

UNIVERSITE LUMIERE LYON 2

Institut de Psychologie

THESE Pour l'obtention du grade de Docteur de l'Université Lumière Lyon 2

Discipline : Psychologie Cognitive

Présentée et soutenue publiquement par

Frédéric DEVINCK

le 19 juin 2003

***Les traitements visuels chez l'homme :
stratégies de classification de la forme***

Co-directeurs de thèse : Dr. Sandrine Delord & Dr. Kenneth Knoblauch

JURY Docteur Sandrine Delord, Maître de Conférence à l'Université Bordeaux II Professeur Olivier Koenig, Professeur à l'Université Lyon II (président du jury) Docteur Kenneth Knoblauch, Chargé de Recherche à l'Inserm Professeur Christian Marendaz, Professeur à l'Université de Grenoble II (rapporteur) Docteur Denis Péliisson, Directeur de Recherche à l'Inserm (examineur) Docteur Simon Thorpe, Directeur de Recherche au CNRS (rapporteur)

Table des matières

Summary . .	1
Résumé .	3
Remerciements . .	5
Liste des abréviations .	7
Introduction . .	9
Dimensions importantes dans la structuration de la forme. .	10
L'image de classification perceptive. . .	11
Organisation de la thèse. .	12
Chapitre I : Partie théorique. .	15
1. Architecture du système visuel .	15
1.1. Arguments physiologiques .	16
1.2. Arguments comportementaux . .	37
2. Le traitement de la surface : une reconstruction ? . .	47
2.1. La surface : un processus de remplissage ? .	47
2.2. Etudes neurophysiologiques du traitement de la surface . .	50
2.3. Arguments comportementaux en faveur de l'existence du processus de remplissage .	63
2.4. Modèle de traitement hiérarchique contour/surface .	71
2.5. Arguments comportementaux en faveur d'un traitement hiérarchique . .	77
2.6. Implications des processus attentionnels . .	89
3. Le bruit visuel : perturbation ou influence sur la prise de décision ? . .	92
3.1. Le signal et le bruit : l'exemple de la théorie de la détection du signal . .	92
3.2. Le phénomène de masquage . .	96
3.3. Le paradigme de classification perceptive .	102
En résumé	110
Chapitre II : Partie expérimentale. . .	111

1. Première série d'expériences : variation de la durée de présentation du stimulus . .	111
1.1. Expérience princeps : durée de présentation intermédiaire (400 ms) .	112
1.2. Expérience préliminaire : détermination de seuil .	112
1.3. Détermination de l'image de classification .	123
1.4. Diminution du temps de présentation . .	141
1.5. Expérience de détermination de seuil pour une durée de 40 ms .	141
1.6. Détermination de l'image de classification pour une durée de 40 ms .	143
1.7. Conclusion .	148
2. Seconde série d'expériences : variation de la taille du signal et du bruit .	149
2.1. Expérience de détermination de seuil pour deux nouvelles tailles du signal .	149
2.2. Détermination de l'image de classification pour différentes variations de taille du signal .	152
2.3. Comparaison des trois tailles du signal .	161
2.4. Expérience de détermination de seuil pour une taille du signal de 3 degrés d'angle visuel et une taille du bruit de 30 degrés d'angle visuel .	164
2.5. Expérience de détermination de l'image de classification pour une taille du signal de 3 degrés et une taille du bruit de 30 degrés . .	165
2.6. Conclusion .	170
3. Troisième série d'expériences : changement de la consigne de détection . .	172
3.1. Expérience de détermination de seuil . .	172
3.2. Expérience de détermination de l'image de classification pour la condition du changement de consigne . .	173
4. Variation de la tâche : discrimination de luminance . .	179
4.1. Expérience de détermination de seuil . .	180
4.2. Expérience de détermination de l'image de classification pour une tâche de discrimination . .	182
5. Réponse de détection par un observateur modèle . .	187
5.1. Introduction . .	187
5.2. Expériences préliminaires . .	189
5.3. Détermination de l'image de classification de l'observateur modèle .	191
Conclusion et Perspectives . .	197

Pertinence de la surface pour la détection. . .	198
Rôle des voies précoces pour le traitement de la surface. .	199
Existe-t-il une contradiction entre la technique de l'image de classification et le paradigme de masquage ? .	200
Perspectives de recherches. . .	201
Bibliographie . .	203
Annexes . .	219
Annexe 1 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour un signal de 3 deg et un temps de présentation de 400 ms. . .	219
Annexe 2 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour un temps de présentation de 40 ms. . .	235
Annexe 3 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour une taille du signal de 1 deg. . .	237
Annexe 4 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour une taille du signal de 6 deg. . .	239
Annexe 5 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour une taille du bruit de 30 deg de l'observateur FD. . .	241
Annexe 6 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour la tâche de détection où la consigne est imposée. .	242
Annexe 7 : Résultats de l'expérience préliminaire pour l'observateur modèle. . .	244

Summary

The thesis examines the role of surface and contour cues in form perception. The classification image paradigm was used to evaluate which features of a signal (a square patch) are used during detection. This technique provides an estimation of the template used by observers to detect a signal. The principle works by comparing visual noise added to signal present and signal absent trials to evaluate what features of the noise correlate best with the observer's perception decision. One obtains an image of the visual cues used by the observer to discriminate presence of the signal. The main results show that observers detect a luminance patch mainly based on the summated luminance within the surface of the stimulus and not by its edges. There is, however, a weak contribution of luminance contrast over a region several times larger than the test stimulus. These results are independent of signal duration and size. The results show that the pixels in the center of the stimulus contributed directly to the task and that edges or filling-in from the edges was not necessary. Subsequently, a model observer was developed to evaluate how the information transmitted along different visual streams might contribute to an observer's performance. Parameters of the receptive field controlling center/surround balance and size were varied. Classification images were constructed on the basis of the model observer judgments. The classification images for balanced center/surround profiles were based on edge and not surface information. Unbalancing the center/surround profiles was sufficient to give a classification image with significant surface information. These results suggest that the surface information could have been transmitted over either pathway as long as the center and surround were unbalanced. These results showed that surface is an early information unlike dominant idea.

Résumé

Selon la théorie prédominante, les contours constituent une information prépondérante encodée par le système visuel, et la surface est une dimension reconstruite à partir des contours. Au cours de la thèse, nous avons voulu tester si cette hypothèse était prise en compte par l'observateur. Pour cela, nous avons utilisé la technique de l'image de classification pour déterminer quelles informations sont utilisées par l'observateur humain pour détecter une forme. Le principe général est qu'une variation aléatoire va permettre de déterminer les dimensions du signal utilisées par l'observateur pour répondre. La tâche de l'observateur est de détecter la présence d'un signal superposé à un bruit. Nous examinons ensuite comment la distribution du bruit influence la réponse de l'observateur. Les résultats montrent que les observateurs utilisent principalement le centre du signal pour répondre. De plus, un faible pourtour sombre entourait cette zone centrale claire. Ces résultats indiquent que les observateurs basaient leur réponse de détection sur la luminance des pixels de la région centrale et non pas sur les contours. Ces résultats ne peuvent pas être expliqués par un mécanisme de reconstruction puisque le centre du signal contribue directement à la détection. Par la suite, un observateur modèle a été développé afin de déterminer comment les voies visuelles contribuent aux performances de l'observateur humain. Deux paramètres étaient manipulés : la taille et l'équilibre centre/pourtour des champs récepteurs. Pour un centre/pourtour équilibré, la réponse était centrée sur les contours du signal. Par contre, la dimension de surface était utilisée lorsque l'antagonisme centre/pourtour était déséquilibré. La taille des champs récepteurs n'était pas un facteur déterminant. Ces résultats suggèrent que la dimension de surface est véhiculée par les voies visuelles précoces.

Remerciements

Je tiens tout d'abord à remercier Sandrine Delord et Ken Knoblauch mes deux directeurs de thèse. Ils m'ont apporté un soutien, une rigueur scientifique, un savoir faire qui m'a permis de mener à bien ce travail et qui ont dû s'armer de patience pour mes nombreuses maladresses.

Je remercie également O. Koenig, C. Marendaz, D. Pélisson et S. Thorpe d'avoir accepté de faire partie du jury de cette thèse.

Je remercie chaleureusement les amis qui m'ont accompagné quotidiennement, qui ont subi mes nombreux grognements contre les machines et pour tous les moments merveilleux que j'ai pu passer avec eux : Gaël (pour ses sorties sportives), Gaëlle (qui a toujours trouvé des t-shirts plus kitsch que les miens), Valérie, les trois Nathalie, Luc, Bill (pour nos pauses philosophiques), Sabine, Fanny (si son accent n'existait pas il faudrait l'inventer) et Guillemette. Je tiens tout particulièrement à remercier Sandrine qui a passé quatre longues années avec moi et avec laquelle j'ai partagé bien plus qu'un bureau. Et comment remercier Céline, la petite blondinette, complice depuis trop peu de temps, sans évoquer l'aide qu'elle m'a apporté pour la rédaction et son sourire communicatif. Merci à Sandrine B., Marie-Pierre et Rosa pour les bons moments que nous avons passés à Stockholm et par la suite.

Je remercie enfin mes amis qui m'ont soutenue au cours de cette thèse et bien avant : Vincent (mon principal supporter et mon plus grand complice depuis plus de quinze ans !), Pascal (pour toutes les journées et soirées à refaire la vie et le monde), Jibé (pour ses histoires et les rêves qu'il nous fait si bien partagé), Katar et Flo (pour leur tendresse et leur obstination à faire de St Etienne une ville idéale), Marion (un jour nous t'offrirons un costume de fraise Tagada), Bertrand (ta joie de vivre est si communicative). Merci aussi à Claude et à Fred qui sont toujours présents malgré l'éloignement (heureusement que le mail existe).

Mes remerciements les plus aimants vont à ma mère et à ma petite sœur qui m'ont encouragé, et qui sont restées tellement présentes malgré l'éloignement. Je tiens à leur dire toute ma reconnaissance.

A Pénélope, tes encouragements, ta tendresse, ta joie pour les petites choses et les fables m'ont aidé et ont été mes repères tout le temps où j'étais devant un petit écran.

Et enfin, un grand merci à la région Rhône-Alpes qui a financé ce travail.

Liste des abréviations

- CGL : Corps Genouillés Latéraux
- COCE : Craik-O'Brien-Cornsweet effect
- Deg : degrés d'angle visuel
- DoG : Difference of Gaussians
- FS : Fréquences Spatiales
- M : Magnocellulaire
- P : Parvocellulaire
- ROC : Receiver Operating Characteristic curve
- TDS : Théorie de la Détection du Signal

Introduction

La première étape de la perception visuelle est l'activation, par la lumière, de millions de photorécepteurs dans la rétine. La répartition de l'image rétinienne se définit par la distribution de l'activité des photorécepteurs. A partir de cet état initial, on peut être étonné de son aboutissement en une expérience visuelle structurée en scènes et en événements complexes. Pour aider à comprendre ce paradoxe, on propose une distinction entre le codage et la perception. Le codage fait référence à la façon dont une image est analysée, intégrée et transmise tout au long des différentes étapes du système visuel. Par exemple, on peut parler du codage de l'information par les photorécepteurs, mais aussi de la sélectivité à l'orientation des cellules de l'aire visuelle primaire. De ce point de vue, une question fondamentale consiste à déterminer comment un stimulus est représenté dans l'activité des neurones du système visuel (DeCharms & Zador, 2000 ; Knoblauch & Shevell, 2003). A l'inverse, la perception se rapporte à notre expérience actuelle ou phénoménale. Il s'agit d'interpréter les objets du monde environnant en lien avec la mémoire.

La transformation qui s'effectue à partir de récepteurs indépendants en une perception structurée est l'un des principaux objectifs de l'organisation perceptive en vision (Palmer, 2002). De ce point de vue, l'objectif général de la thèse est de mieux connaître l'influence de ce codage sur la perception. Pour cela, nous cherchons à déterminer les indices du stimulus utilisés par un observateur, afin de préciser la nature du codage précoce. En d'autres termes, il s'agit de modéliser les stratégies perceptives que le système visuel met en œuvre pour traiter une information.

Dimensions importantes dans la structuration de la forme.

Dès la rétine, la scène visuelle est décomposée en plusieurs éléments, et l'une des fonctions les plus importantes du système visuel précoce est le regroupement de ces éléments. La façon dont le système regroupe l'ensemble des indices pour obtenir un percept unifié est une question primordiale qui n'a pas, pour le moment, trouvé de réponse satisfaisante (Paradiso, 2000). Toutefois, au cours des différentes étapes de traitement qui relie la stimulation à la structuration de la forme, l'un des premiers aspects, proposé dans la littérature, se centre sur la distinction entre le traitement du contour et le traitement de la surface. De nombreux auteurs (Grossberg & Mingolla, 1985 ; Marr, 1982 ; Palmer, 1999) ont émis l'hypothèse que ces processus formaient l'une des premières étapes de traitement de structuration de la forme.

D'un point de vue classique, le codage des contours par le système visuel permet d'expliquer les mécanismes fondamentaux de la perception de forme. En effet, la plupart des théories ont minimisé le rôle du codage de la surface et ont attribué la perception de la surface à un mécanisme secondaire : le remplissage (Grossberg & Mingola, 1985 ; Macknik, Martinez-Conde & Haglung, 2000 ; Spillman, 1997).

Le rôle important des contours en vision est attesté par de très nombreux phénomènes visuels qui leur sont associés (Ratliff, 1965), la saillance de cette dimension chez les nouveaux nés (Haith, Bergman & Moore, 1977 ; Kessen, Salapatek & Haith, 1972 ; Salapatek & Kessen, 1966) ou encore leur importance dans l'analyse informatique d'image (Geissler, Perry, Super & Gallogly, 2001 ; Marr, 1982). De plus, l'antagonisme centre/pourtour des champs récepteurs de la rétine au cortex visuel accentue l'importance du codage des contours et non pas de la surface (Wilson, Levi, Maffei, Rovamo, & DeValois, 1990). En effet, selon la théorie prédominante, les contours constituent une information prépondérante encodée par le système visuel, et notre perception de la surface des objets ne provient pas directement de la lumière mais d'un processus de remplissage réalisé à partir de la détection des contours (Grossberg & Mingola, 1985 ; Komatsu, Murakami & Kinoshita, 1996, 2002 ; Macknik, Martinez-Conde & Haglung, 2000).

Si les contours sont encodés par le système visuel précoce et que le processus de remplissage permet la perception de la surface, alors il faut s'attendre à ce que l'élément le plus pertinent pour la détection d'une forme soit le contour. Au cours de la thèse, nous avons voulu tester si cette hypothèse générale était effectivement prise en compte par l'observateur. Pour cela, nous avons utilisé une technique de corrélation inverse, l'image de classification, pour déterminer quelles informations sont utilisées par l'observateur humain pour détecter une forme.

L'image de classification perceptive.

Traditionnellement, l'objectif des études en perception est de mesurer et d'expliquer les variations du comportement provoquées par les variations du signal.

Il existe deux principales approches expérimentales en psychophysique. La première démarche consiste à établir une correspondance entre le domaine physique et le domaine sensoriel ; une échelle de mesure est nécessaire pour mesurer les sensations, et s'appuie sur la notion de seuil. Ainsi, pour une dimension donnée du signal, le seuil absolu est la valeur minimale du signal qui provoque une sensation juste perceptible. La théorie des seuils considère que la réponse d'un observateur ne dépend que de la stimulation. Toutefois, la mesure des seuils peut être affecté par un biais de réponse de l'observateur. Le biais fait référence à une tendance de l'observateur d'augmenter ou de diminuer sa tendance à répondre présent ou absent en fonction d'une stratégie. Ceci implique une forte contrainte sur la stabilité et l'homogénéité du critère de réponses des observateurs.

La seconde approche tente d'expliquer et de prédire les effets du bruit sur le signal, c'est notamment le cas de la théorie de la détection du signal (TDS). Le bruit est considéré comme l'ensemble des phénomènes non pertinents qui demeurent en l'absence du signal. Dans la TDS, le bruit n'est pas forcément considéré comme une variation totalement aléatoire. Par exemple, une tâche expérimentale pourrait très bien être de détecter la présence d'un signal parmi un ensemble de distracteurs, le bruit est tout ce qui présenté hormis le signal. L'objectif de la TDS est d'obtenir un critère de réponse de l'observateur en distinguant les essais où seul le bruit est présent et les essais où le signal est superposé au bruit. Le critère de décision permet d'indiquer le niveau de sensation pour lequel le jugement de l'observateur change. De plus, toutes les variations que l'on rajoute sur le signal auront un effet sur la sensibilité de l'observateur calculée par l'indice d' .

Le paradigme de l'image de classification est une nouvelle démarche qui a été développé par Ahumada et al. (Ahumada, 1996 ; Ahumada & Lovell, 1971 ; Ahumada, Marken & Sandusky, 1975). Le principe général est qu'une variation aléatoire va permettre de déterminer les dimensions du signal utilisées par l'observateur pour répondre à une tâche. Cette méthode est radicalement différente des techniques psychophysiques traditionnelles puisqu'ici on introduit un facteur d'incertitude : un bruit. Dans nos expériences, le bruit se définit par une matrice de plusieurs petits carrés dont la luminance varie aléatoirement. La tâche de l'observateur est de détecter la présence ou non d'un signal qui est superposé à un bruit, le signal est présent dans la moitié des essais. Nous examinons ensuite comment la distribution du bruit influence la réponse de l'observateur tout au long des essais. Ceci est réalisé par la corrélation entre les perturbations aléatoires et les réponses de l'observateur à chaque essai. Lors des essais où l'observateur classe le stimulus comme présent, nous supposons que le bruit contient les traits qui renforcent ceux utilisés par l'observateur pour donner la réponse « signal présent ». Le résultat est une image qui illustre l'influence relative de chaque pixel sur la

réponse de l'observateur. Nous obtenons ainsi des indications sur la façon dont le système visuel traite un signal grâce à l'analyse du bruit. En d'autres termes, ce paradigme fournit le **prototype** utilisé par l'observateur pour donner sa réponse.

Pour résumer, nous savons qu'il existe un biais de réponse dans la théorie des seuils qui conduit à contaminer le critère de réponse de l'observateur. La TDS permet d'estimer le critère de décision de l'observateur, ce qui permet de vérifier si la stratégie de l'observateur change pendant l'expérience. Enfin, l'image de classification fait partie intégrante du critère de décision, mais ici l'avantage de cette nouvelle démarche est de pouvoir apprécier la distribution sur la dimension évaluée du critère.

La technique de l'image de classification perceptive a déjà été appliquée à des tâches et des domaines variés comme l'audition (Ahumada & Lovell, 1971 ; Ahumada, Marken & Sandusky, 1975), l'acuité de Vernier (Ahumada, 1996 ; Ahumada & Beard, 1998), la discrimination de lettres (Watson & Rosenholtz, 1998), la détection temporelle (Thomas & Knoblauch, 1998), les contours illusoires (Gold, Murray, Bennett & Sekuler, 2000) ou encore la reconnaissance des visages (Sekuler, Gold, Gaspard & Bennett, 2001). Elle va nous permettre de modéliser la représentation visuelle utilisée par l'observateur dans la structuration de la forme.

Notre travail consiste à examiner dans quelle mesure les prototypes utilisés par l'observateur sont compatibles avec les dimensions structurales classiquement admises : une forme se définit par ses contours plus sa surface. La technique que nous utilisons suppose que les zones du signal préservées dans l'image de classification sont les informations sur lesquelles l'observateur fonde sa décision. Selon l'hypothèse la plus répandue, les contours d'une forme visuelle sont encodés précocement par le système, et la surface serait une dimension secondaire, reconstruite ultérieurement par un processus de remplissage des contours précédemment traités. Dans ce cas, la dimension de contour est une caractéristique primordiale pour l'observateur, voire la seule. Nous pouvons poser l'hypothèse suivante : l'image de classification comporte uniquement les contours (i.e. la dimension qui est encodée par les neurones).

Organisation de la thèse.

Le premier chapitre de cette thèse est consacré à des éléments théoriques en rapport avec la problématique étudiée. Ce chapitre théorique s'articule en trois sections. La première traite de l'intégration des contours par le système visuel, et est composée de deux parties. La première partie décrit essentiellement l'architecture physiologique, et aborde notamment la notion de champs récepteurs, le modèle hiérarchique de Hubel et Wiesel (1962, 1968), ainsi que des éléments plus récents permettant une meilleure compréhension du codage de l'information par notre système. La deuxième partie présente plus particulièrement les études comportementales. Cette partie permettra une comparaison entre les données physiologiques et la perception.

La seconde section présentera tout d'abord une première partie qui explique pourquoi

la dimension de surface a été assimilée au processus de remplissage. Dans la deuxième partie, nous donnerons une explication physiologique du remplissage, et des résultats qui montrent une dissociation entre des processus précoces et tardifs pour ce processus. Nous présenterons également un ensemble d'études portant sur la réponse des neurones à une surface uniforme, ce qui nous conduira à distinguer le traitement du remplissage et le codage de la surface. La troisième partie sera consacrée à la mise en évidence du phénomène de remplissage d'un point de vue comportemental pour des phénomènes naturels comme la tâche aveugle. Dans une quatrième partie, nous décrirons les deux principaux modèles d'un traitement hiérarchique qui ont discuté du rôle des contours et de la surface : le modèle de Marr (1982) et celui de Grossberg et de ses collaborateurs (Grossberg, 2000 ; Grossberg & Mingolla, 1985 ; Grossberg, Mingolla & Ross, 1997 ; Grossberg & Todorovic, 1988). La cinquième partie se centrera sur une revue des données comportementales relatives au traitement hiérarchique pour une forme visuelle. L'objectif est de déterminer si le remplissage explique le traitement de la surface. Enfin, la sixième partie sera consacrée aux processus attentionnels impliqués dans le remplissage et le traitement de la surface.

La troisième section comporte trois parties qui sont centrées sur la notion du signal et du bruit dans la détection. La première partie sera consacrée à la théorie de la détection du signal dans laquelle le rôle du signal et du bruit sont explicités. Dans la seconde partie, nous présenterons le paradigme de masquage (ou comment le bruit perturbe la perception du signal) et une dernière partie présentera le paradigme de classification perceptive (ou comment l'analyse du bruit nous permet d'étudier la perception).

Le deuxième chapitre de cette thèse est réservé à la description de la méthode expérimentale utilisée, des résultats obtenus et de discussions relatives à ces différents résultats. Notre recherche tente de déterminer quel type d'information visuelle, le contour ou la surface, est pertinent pour le traitement selon la durée de présentation du signal, la taille du signal et le niveau de traitement exigé par la tâche (détection/discrimination). En outre, la modélisation d'un observateur modèle, basée sur les caractéristiques d'intégration d'un signal par les champs récepteurs (variant en taille et par le déséquilibre centre/pourtour), a été réalisée afin de connaître les propriétés des canaux de traitement précoce qui sont susceptibles d'influencer la réponse de l'observateur humain. Ceci est présenté dans le dernier chapitre. Une conclusion générale mettant en avant les apports de cette recherche terminera ce travail.

Chapitre I : Partie théorique.

1. Architecture du système visuel

Une question à la fois simple et fondamentale en perception visuelle est de savoir comment une information visuelle est intégrée en un tout unique alors que le système visuel commence par séparer une image en plusieurs composantes individuelles. Cette décomposition de l'information s'effectue dès la rétine, et permet de distinguer la notion de codage et de perception. Ce premier chapitre traitera plus particulièrement de la notion de codage de l'information visuelle. Du point de vue physiologique, le codage de la dimension de contour est considéré traditionnellement comme l'un des premiers mécanismes de la perception de forme (Hess & Field, 1999 ; Ratliff, 1965). Nous aborderons dans cette partie une revue des données qui ont amené à considérer les contours comme l'une des premières étapes de traitement d'un stimulus.

La limite entre deux surfaces est définie par des différences de luminance, de couleur ou de texture. Malgré cette hétérogénéité, notre système visuel est capable d'intégrer ces différentes dimensions pour obtenir un percept unifié. L'intégration des contours est donc une étape essentielle dans le traitement d'une forme visuelle. La première partie de ce chapitre abordera l'extraction des contours d'un point de vue physiologique et la notion de champ récepteur sera particulièrement développée. Le modèle hiérarchique de Hubel et

Wiesel (1962) sera présenté, des données plus récentes seront également abordées. La seconde partie de ce chapitre s'intéressera à la confrontation des données physiologiques avec les études comportementales.

1.1. Arguments physiologiques

Depuis les années 60, de nombreuses connaissances ont été accumulées sur la définition et l'extraction des primitives visuelles, comme la sensibilité à l'orientation ou à la taille. Néanmoins, nos connaissances sont limitées lorsque nous essayons de comprendre comment ces éléments, traités par des analyseurs locaux, sont assemblés pour permettre une représentation unifiée.

Les structures susceptibles de fournir un support neurophysiologique à l'extraction des contours d'une forme se situent principalement dans le cortex visuel primaire. Les structures antérieures (rétine et corps genouillés latéraux) conditionnent et élaborent l'organisation fonctionnelle de l'aire visuelle primaire (ou V1).

1.1.1. Distinction entre le système photopique et scotopique

Chaque signal visuel est tout d'abord échantillonné précocement par les photorécepteurs de la rétine : les cônes et les bâtonnets. Les bâtonnets permettent la vision nocturne (scotopique), alors que les cônes sont impliqués dans la vision diurne (photopique). Ces deux types de photorécepteurs se distinguent également par leur emplacement dans la rétine. Les bâtonnets sont situés à la périphérie de la rétine, tandis que les cônes sont principalement centrés autour de la fovéa (région centrale de la rétine). Cette spécificité permet au système visuel de coder des images avec des niveaux d'éclairage très variés (Bullier, 1998).

Le système scotopique, qui dépend uniquement des bâtonnets, ne permet pas un codage des différences spectrales. Seul le système photopique, qui implique le fonctionnement des cônes, permet de discriminer les longueurs d'onde¹ qui est un préalable pour la vision des couleurs. Il existe trois types de cônes qui sont spécifiques à des courtes (cônes S), moyennes (cônes M) et grandes (cônes L) longueurs d'onde. Les cônes L et M sont impliqués dans le codage de la luminance et des différences chromatiques, tandis que les cônes S n'interviennent que dans le codage chromatique.

Après cette première étape, les cellules bipolaires acheminent l'information aux cellules ganglionnaires. Les cellules horizontales et amacriennes permettent un contact transversal. Les cellules ganglionnaires forment la dernière couche de la rétine et répondent aux variations de luminance ou aux aspects chromatiques. Une représentation de cette organisation physiologique de la rétine est présentée à la Figure 1.

¹ On désigne sous le nom de « lumière » le domaine très restreint des ondes électromagnétiques auxquelles le système visuel humain est sensible. Il existe d'autres radiations électromagnétiques que la lumière, toutes caractérisées par la longueur d'onde. La lumière n'en constitue qu'une bande très étroite, appelée spectre visible (compris entre 400 et 700 nm).

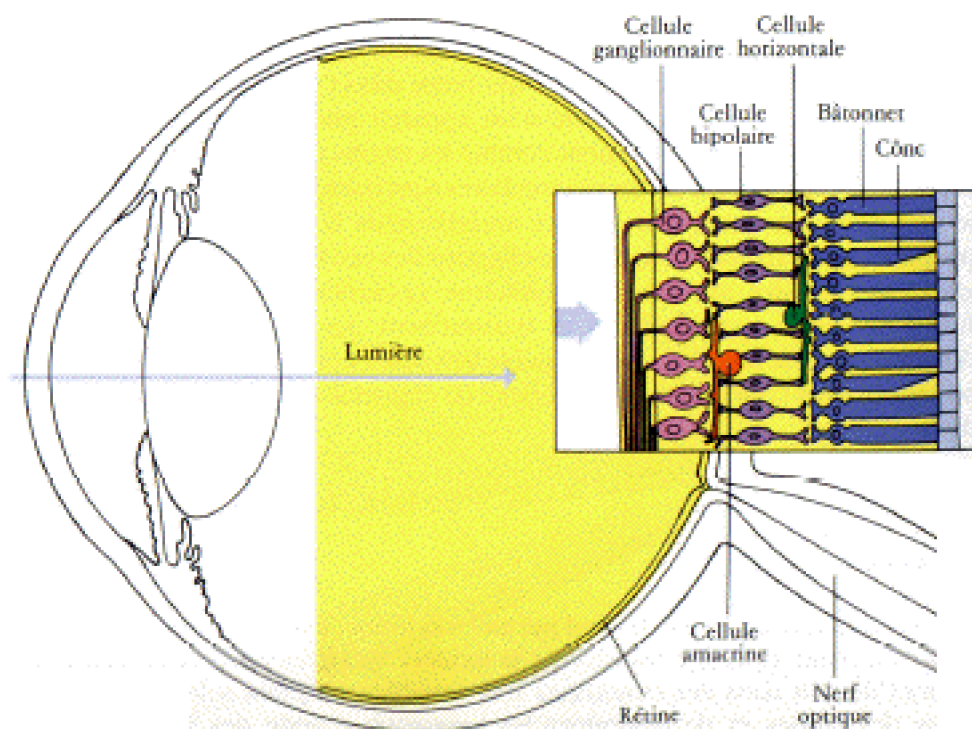


Figure 1 : Organisation physiologique de la rétine chez l'homme (d'après l'illustration de Hubel, 1994, p. 44).

Une distinction majeure entre le système photopique et scotopique est le nombre de photorécepteurs qui communiquent avec une seule cellule ganglionnaire (dans un souci de simplification, nous ne parlerons pas du rôle des cellules intermédiaires). En effet, une cellule ganglionnaire est reliée à plus de bâtonnets que de cônes. Le système scotopique possède donc une résolution spatiale plus faible que le système photopique : il résume davantage l'information qui lui est envoyée et manifeste ainsi ce qu'on appelle une plus grande « sommation spatiale ». En d'autres termes, la sommation spatiale correspond à la capacité de la rétine à additionner des quanta de lumière sur une certaine aire. Par exemple, supposons qu'une cellule ganglionnaire signale la présence d'un stimulus si un minimum de 10 quanta est absorbé. Deux points lumineux qui sont séparés par une certaine distance, et qui n'émettent que 5 quanta chacun, sont présentés. Imaginons que cette distance soit telle que les deux points activent deux bâtonnets qui transmettent l'information à une seule et même cellule ganglionnaire. À eux deux, ils produisent un total d'absorption de 10 quanta. La cellule ganglionnaire va donc signaler la présence d'un seul point lumineux, en d'autres termes la cellule ne fait pas la différence entre un point et deux points. Ainsi, la « sommation spatiale » du système scotopique produit une excellente sensibilité (une faible luminance peut être perçue), mais une faible résolution spatiale (le signal double est perçu comme un signal unique). Considérons maintenant le système photopique : dans des conditions identiques, la distance entre les points étant toujours la même, mais elle est cette fois-ci suffisante pour que les cônes concernés se projettent sur deux cellules ganglionnaires. Puisque chaque cellule ganglionnaire ne reçoit que 5 quanta, et pas 10, aucun signal ne sera perçu. Le système photopique manifeste

ainsi une mauvaise sensibilité. Par contre si nous doublons le nombre de quanta contenu dans un seul point lumineux, alors les deux cellules ganglionnaires, dans le cas du système photopique, signaleront la présence des deux points. Pour la condition scotopique, puisque les bâtonnets convergent vers une seule cellule ganglionnaire, les deux points lumineux ne seront toujours pas distingués, mais un seul point sera perçu (Schwartz, 1999).

Pour résumer, le système scotopique montre une excellente sommation spatiale, ce qui contribue à une sensibilité élevée, mais possède une faible résolution. Par contre, le système photopique a des capacités de sommation spatiale faible et donc une faible sensibilité, mais est au contraire capable d'une résolution spatiale élevée.

1.1.2. Propriétés des champs récepteurs des cellules de la rétine

Chaque cellule ganglionnaire de la rétine signale par une augmentation de sa fréquence de décharge l'apparition d'un contraste chromatique ou de luminance dans une petite fenêtre du champ visuel. Cette région d'activation est appelée le champ récepteur du neurone (Bullier, 2002 ; Hubel, 1994 ; Fitzpatrick, 2000). Il s'agit de la zone du champ visuel qui influence l'activité d'une cellule. Chez le primate, il existe une vingtaine de cellules ganglionnaires différentes (Masland, 2001 ; Rodieck, 1998). Parmi toutes ces cellules, trois grandes catégories semblent être les plus importantes dans la voie rétino-géniculée-corticale (Knoblauch, 1999) : les cellules alpha (appelées également en parasol, de grandes tailles), les cellules bêta (ou naines, de petites tailles) et les cellules gamma (ou bi-stratifiées). Chez le primate, les cellules en parasol représentent environ 10 % de la population des cellules ganglionnaires, les cellules naines, les plus nombreuses de toutes, représentent 70 % du total (Lee, 1996), et enfin les cellules bi-stratifiées, les moins nombreuses, représentent entre 5 et 10 % selon les espèces ; par conséquent ces dernières ont été les moins étudiées (Dacey & Peterson, 1992 ; Watanabe & Rodieck, 1989).

Le champ récepteur d'une cellule ganglionnaire comprend deux régions concentriques : le centre qui est la région la plus sensible et le pourtour qui a une action antagoniste à celle du centre. Deux types fonctionnels de neurones sont distingués : les neurones à *centre-on* et à *centre-off*. L'activité des neurones à *centre-on* s'accroît lorsque le contraste de luminance augmente entre le centre et le pourtour du champ récepteur. Ainsi, une cellule à *centre-on* décharge avec une fréquence plus élevée lorsqu'une tache de lumière est présentée précisément au centre du champ récepteur et que son pourtour n'est pas éclairé (Figure 2b). La réponse de cette cellule est faible ou nulle lorsque son champ récepteur est totalement éclairé (Figure 2d). Réciproquement, ce type de cellule diminuera son activité spontanée lorsqu'un anneau lumineux est présenté dans le pourtour du champ récepteur (Figure 2e). Les cellules à *centre-off* répondent de façon opposée aux stimulations lumineuses. L'organisation des champs récepteurs permet donc aux cellules de répondre au contraste et non pas à une luminance absolue.

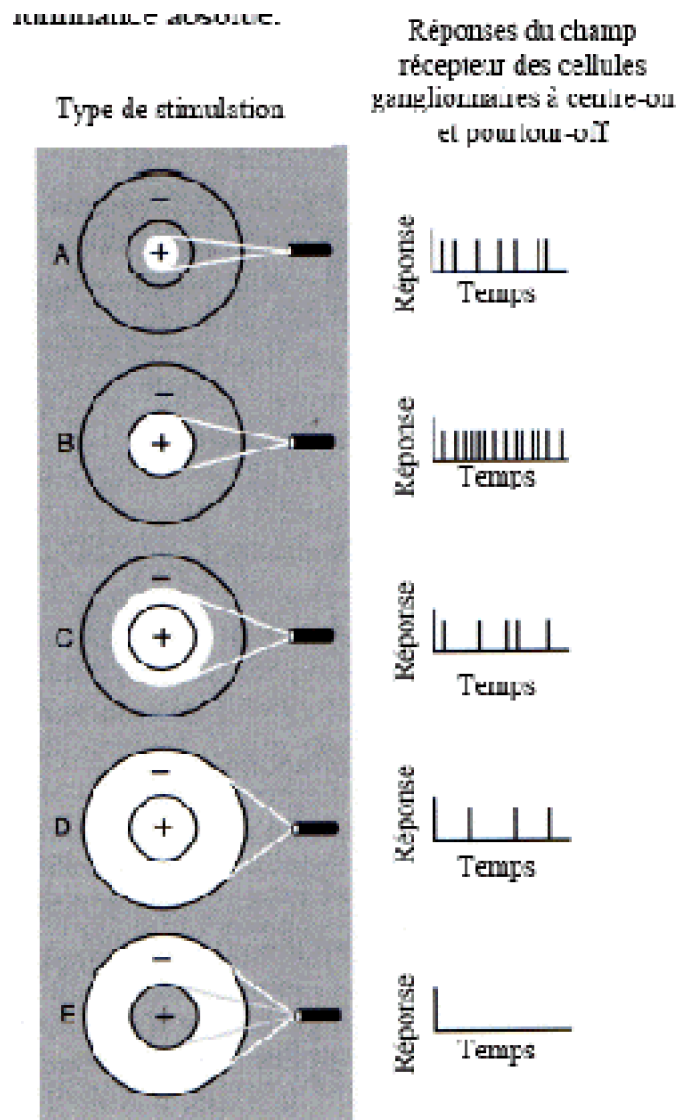


Figure 2 : Champ récepteur d'une cellule ganglionnaire à centre-on et à pourtour-off. Les figures de droite illustrent les différentes réponses neuronales (en décharge/sec) qui correspondent à la présentation d'une tache lumineuse dans le champ récepteur (sur la gauche) (d'après Palmer, 1999).

Le centre et le pourtour ont été modélisés comme deux mécanismes antagonistes avec une sensibilité proche d'une fonction gaussienne pour chacun d'entre eux. Le profil de la zone centrale est étroit et élevé, tandis que le profil de la zone du pourtour est étendu spatialement et faible (un exemple, au trait fin, est donné à la Figure 3). Le centre et le pourtour interagissent, et leur combinaison produit une forme ressemblant à un chapeau mexicain, c'est le profil du champ récepteur des cellules ganglionnaires (Bullier, 2002 ; Rodieck, 1998). Ce profil est obtenu par la soustraction entre les deux fonctions gaussiennes, cette représentation des champs récepteurs a été appelée « Difference of Gaussian » ou DoG (Rodieck, 1965 ; Enroth-Cugell & Robson, 1966). La Figure 4 illustre en gras ce profil. Dans ce modèle, la réponse neuronale dépend de deux mécanismes indépendants : un centre opposé à un pourtour. La fonction DoG permet de faire des

prédictions sur la réponse des neurones sur la base d'un calcul qui ne requière que quatre paramètres, à savoir l'amplitude et la taille du centre et du pourtour des champs récepteurs.

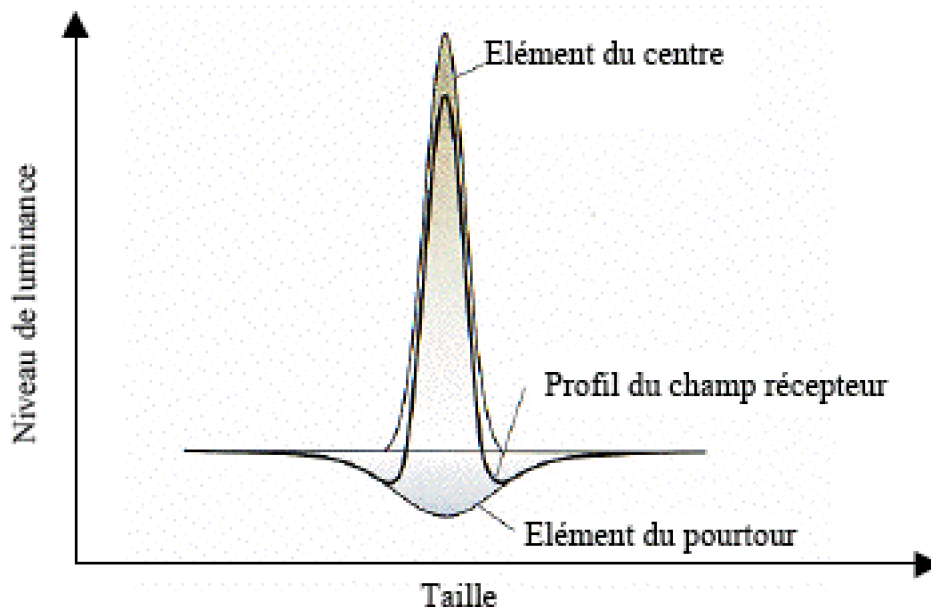


Figure 3 : Profil unidimensionnel du centre et du pourtour d'une cellule ganglionnaire à centre-on. Le champ récepteur de ces cellules est circulaire. La sensibilité spatiale du centre et du pourtour est une fonction gaussienne et est indiquée par un trait fin. Le profil du champ récepteur d'une cellule ganglionnaire est indiqué par un trait gras et est obtenu par l'addition des éléments du centre et du pourtour (l'illustration est issue de Rodieck, 1998, pp 337).

La taille des champs récepteurs dépend du type de cellule ganglionnaire mais également de l'excentricité rétinienne : plus les cellules ganglionnaires sont éloignées de la périphérie et plus la taille de leur champ récepteur est importante (Rodieck, 1998). Par contre, la différence de taille entre les cellules en parasol et naines varie en fonction de l'excentricité. En effet, le ratio est de 3/1 dans la rétine périphérique et de 10/1 à 3 deg d'excentricité (Dacey & Petersen, 1992). De plus, le diamètre de la zone centrale n'est pas équivalent à celui du pourtour avec une taille plus importante pour le pourtour (Dacey, Packer, Diller, Brainard, Peterson & Lee, 2000). Par ailleurs, le diamètre d'un champ récepteur à centre-on est de 30 à 50 % plus large qu'un centre-off (Dacey & Petersen, 1992).

La région centrale du champ récepteur des cellules en parasol somme les signaux des cônes L et M. En rappelant que les cônes M et L sont répartis de manière aléatoire sur la rétine, ces cellules semblent plus adaptées à transmettre les variations de luminance d'une image. Par ailleurs, les cellules naines possèdent non seulement un antagonisme spatial mais aussi spectral. En effet, la réponse du centre des champs récepteurs dépend de l'activation des cônes de type M ou L. A l'inverse, le pourtour est activé par un mélange aléatoire comprenant les signaux des cônes L et M. Ces cellules peuvent aussi être classées selon leurs réponses du centre de type *on* et *off*, ce qui

donne ainsi $-M+L$ et $-L+M$ pour les cellules *off*. En ce qui concerne les cellules bi-stratifiées, les signaux de la couche *on* viennent des cônes S, ceux de la couche *off* des cônes M/L. L'organisation du champ récepteur des cellules bi-stratifiées indiquent que le champ *on* des cônes S est coextensif avec le champ *off* des cônes M/L. Par conséquent, le signal véhiculé par ces cellules est purement chromatique de type S-L/M, sans antagonisme spatial (Knoblauch, 1999 ; Rodieck, 1998). Toutefois, nous ne développerons pas une discussion sur le codage chromatique, puisque notre thèse s'inscrit dans une perspective achromatique.

Il existe 120 millions de bâtonnets, 6 millions de cônes et 1,5 millions de cellules ganglionnaires (Schwartz, 1999 ; Wandell, 1995). L'organisation en champs récepteurs permet l'intégration d'un stimulus initialement échantillonné puis compressé. Les cellules ganglionnaires signalent en effet le résultat d'une comparaison entre la quantité de lumière atteignant une région de la rétine et la quantité de lumière qui atteint le pourtour de cette région. Les champs récepteurs des cellules naines et en parasol fait que ces derniers répondent de façon préférentielle au contraste et pas à une luminance absolue. Ceci implique que l'information de surface est perdue : les cellules répondent si un contour (un contraste noir/blanc) correctement placé atteint leur champ récepteur (la limite entre deux formes est indiquée par une différence de contraste ; Hubel, 1994 ; Schwartz, 1999 ; Wandell, 1995). Les bords d'une forme visuelle sont donc encodés dès les premiers niveaux de traitement.

Le fait que la cellule ne réagisse qu'aux différences locales d'intensité est étonnant, car lorsque l'on regarde une forme lumineuse et uniforme, l'intérieur nous semble aussi clair que le bord. Or, nous avons vu que les cellules ganglionnaires sont activées uniquement par les bords de la forme, par conséquent, si le traitement visuel s'arrêtait là, notre perception porterait exclusivement sur les contours. Le chapitre 2 présentera comment les modèles résolvent ce problème en postulant que la surface est reconstruite.

1.1.3. Rôle de la voie Magnocellulaire et Parvocellulaire dans l'intégration d'une forme

Les axones des cellules ganglionnaires convergent à la surface de la rétine pour former le nerf optique qui se projette principalement sur les corps genouillés latéraux dont la structure est organisée en six couches. L'organisation des champs récepteurs du CGL est identique à celle des cellules ganglionnaires. Mais, la taille des champs récepteurs, à partir de la rétine, augmente au fur et à mesure des différentes étapes de traitement (CGL, V1, V2, etc...).

Les cellules en parasol se projettent dans les couches magnocellulaires (les deux couches inférieures), les cellules naines dans les couches parvocellulaires (les quatre couches supérieures), et les cellules bi-stratifiées dans les inter-couches koniocellulaires. Ces destinations anatomiques ont donné leur nom à trois voies visuelles : la voie magnocellulaire, la voie parvocellulaire et la voie koniocellulaire (Lee, 1996). Nous nous intéresserons plus particulièrement aux voies magnocellulaire et parvocellulaire car elles sont dépendantes du contraste de luminance. Les cellules bi-stratifiées, qui forment le trajet koniocellulaire, ont des champs récepteurs de grande taille, et répondent aux

fréquences spatiales basses. Mais, la voie K n'est pas sensible aux stimulations achromatiques (Hendry & Reid, 2000 ; Kaplan, 2003). C'est pour cette raison que cette voie ne sera pas développée ici.

Tout signal visuel est donc précocement analysé par ces deux systèmes qui ont des propriétés distinctes et complémentaires. Le système magnocellulaire (ou voie M) répond à des hautes fréquences temporelles (image en déplacement rapide sur la rétine ou papillotement rapide de tâche lumineuse) et aux basses fréquences spatiales² (l'exemple d'une image filtrée est donnée Figure 4). Ce dernier répond très rapidement, 50 ms après l'apparition et à l'extinction d'un signal visuel. En d'autres termes, cette voie permet de signaler efficacement les changements brusques et rapides de luminance. Le système parvocellulaire (ou voie P) possède au contraire une résolution temporelle basse et répond aux fréquences spatiales élevées. Les cellules de cette voie répondent préférentiellement à un contraste stable dans leur champ récepteur. Cette voie est activée plus tardivement (70 ms après l'apparition du signal), et véhicule des informations chromatiques et des détails de l'image, son contour par exemple (Kaplan, 2003).



Figure 4 : Deux filtres ont été appliqués à une image préservant soit les FS basse, soit les FS élevées. La photo de Groucho Marx (sur la gauche) a été filtrée en intégrant seulement les fréquences spatiales basses (au milieu), puis seulement les fréquences spatiales élevées (sur la droite). Les fréquences spatiales basses véhiculent des informations globales comme la distinction entre les grandes parties claires et sombres de l'image, alors que les fréquences spatiales élevées transmettent une information concernant les contrastes locaux aux frontières de l'objet (l'exemple est issu de Palmer, 1999, p. 163).

Les différences de sensibilité spatiale et temporelle, des voies M et P, ont été démontrées dans une série d'études où des inhibitions spécifiques aux deux voies étaient réalisées. Merigan et al. (1990, 1991) ont réalisé une lésion partielle dans la voie M ou P du CGL de deux singes, et une étude comportementale a ensuite été effectuée. Merigan et Maunsell (1990) ont inhibé spécifiquement la voie M. Dans ce cas, les deux animaux ne répondaient pas à des fréquences temporelles élevées (10 Hz) lorsque ces dernières étaient présentées dans le champ visuel affecté par la lésion. Par ailleurs, des réponses correctes étaient enregistrées lorsqu'un réseau de 2 cdp était présenté (correspondant à

² Un réseau sinusoïdal a deux dimensions : 1) une fréquence spatiale qui désigne le nombre de périodes (alternance noir/blanc) contenu dans 1 degré d'angle visuel (exprimé en cycle par degré ou cpd), et 2) une fréquence temporelle qui est définie comme une variation périodique au cours du temps, cette variation périodique est exprimée en nombre de cycles par seconde (Hertz ou Hz).

une fréquence spatiale élevée). Merigan, Katz et Maunsell (1991) ont par la suite montré qu'une lésion de la voie P réduisait considérablement les réponses comportementales pour un réseau de 2 cpd. De même, les réponses n'étaient pas affectées par la présentation de fréquences temporelles élevées (10 Hz). Allison, Melzer, Ding, Bonds et Casagrande (1999) ont montré que les réponses des cellules simples de V1 aux contrastes étaient affectées par une inactivation de la voie M ou P. Une inhibition de la voie M dans le CGL réduisait la réponse des neurones pour l'ensemble des niveaux de contrastes. Par contre, l'inhibition de la voie P n'avait aucun effet sur des réseaux dont le contraste était inférieur à 20%. Ces deux voies agissent de façon parallèle mais elles sont confondues à la sortie de V1.

Les premières études physiologiques ont montré que les cellules de la voie M étaient beaucoup moins représentées autour de la fovéa, comparativement aux cellules de la voie P. Toutefois, des études plus récentes ont montré que le ratio des cellules naines et en parasol était identique dans la rétine du singe (Perry, Oehler & Cowey, 1984 ; Silveira & Perry, 1991). Ces résultats mettent ainsi en doute la sélectivité des voies M et P pour des stimulations où l'excentricité du signal varie. Toutefois, une étude de Dacey et Petersen (1992), chez l'observateur humain, n'a pas corroboré ces résultats. En effet, ils ont observé que le ratio des voies P/M dans la rétine était de 30/1.

Une distinction supplémentaire entre les deux systèmes concerne la taille des champs récepteurs³. En effet, le diamètre des champs récepteurs augmente entre la fovéa et la périphérie pour les deux voies, mais quelle que soit l'excentricité, le diamètre des champs récepteurs des cellules magnocellulaires est toujours plus large que celui des champs récepteurs des cellules parvocellulaires (Croner & Kaplan, 1995). Cette différence de taille entre les deux voies a une conséquence importante, puisque la sensibilité au contraste des deux systèmes n'est pas la même. Ainsi, lorsqu'on enregistre la réponse des deux types de cellules pour des contrastes de luminance qui augmentent pas à pas, les résultats montrent une réponse élevée des cellules magnocellulaires pour des contrastes faibles et une réponse faible des cellules parvocellulaires pour tous les niveaux de contraste (Kaplan & Shapley, 1986). Par ailleurs, le gain de contraste⁴ des cellules de la voie M augmente rapidement pour les premiers niveaux de contraste, les cellules sont ensuite saturées pour un contraste excédant 15-20 %. La réponse des cellules de la voie P, au contraire, augmente de façon graduelle. Il est tout à fait possible que les cellules de la voie M soit déterminées par deux mécanismes : un pour les plus petits niveaux de contraste et l'autre pour des niveaux de contraste élevés (Kaplan, 2003). La Figure 5 montre la réponse des deux populations de neurones selon le niveau de contraste. De la même façon que pour les canaux M et P, Lee, Pokorny, Smith, Martin et Valberg (1990) avaient observé des réponses aux contrastes différentes pour les cellules ganglionnaires en parasol et naines de la rétine.

³ Les propriétés des champs récepteurs des cellules du CGL sont similaires à celles des cellules ganglionnaires avec une organisation centre/pourtour antagoniste.

⁴ Le gain de contraste se définit comme le changement de la réponse lorsque le contraste varie.

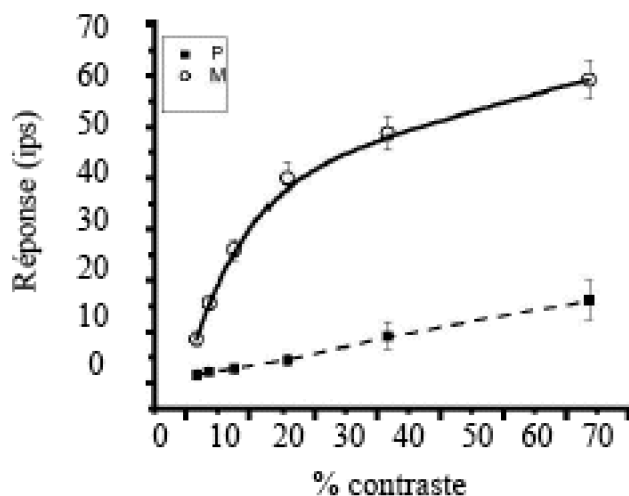


Figure 5 : La réponse des cellules magnocellulaires (M, trait plein) et parvocellulaires (P, trait en pointillé) est présentée en fonction du contraste. Les réponses des cellules (formulée en « impulse per second » ou ips) magnocellulaires sont plus élevées que celles des cellules parvocellulaires pour les plus faibles niveaux de contraste, mais sont de pente similaires pour des niveaux de contraste élevés (d'après Kaplan, 2003).

1.1.4. Éléments de neurophysiologie du cortex visuel

Une présentation succincte de l'organisation de V1 sera tout d'abord réalisée. Par la suite, cette partie définira plus particulièrement l'organisation des champs récepteurs des cellules de l'aire visuelle primaire. Le modèle hiérarchique proposé par Hubel et Wiesel (1962, 1968), qui met l'accent sur les traitements ascendants, sera particulièrement développé. Les modifications, qui ont été apportées sur la structure du système visuel, seront présentées au cours des deux parties suivantes.

1.1.4.1. Organisation modulaire de V1

L'aire V1 est divisée en un certain nombre de couches parallèle à la surface du cortex : 1, 2/3, 4B, 4C, 5 et 6. Ces différentes couches sont présentées à la Figure 6. La couche 4C reçoit les informations issues directement du CGL. Elle est séparée en deux. En effet, la voie P se projette sur la couche 4C₁, alors que la voie M converge vers la couche 4C₂. L'information est ensuite transmise aux couches 2/3 et 4B qui distribueront l'information aux autres aires visuelles. À l'inverse, les couches 5 et 6 se projettent dans le CGL mais aussi dans le colliculus supérieur. Les couches 2 et 3 reçoivent les afférents de la voie K au niveau des blobs. Contrairement à la plupart des cellules de l'aire V1, les champs récepteurs des blobs sont insensibles à l'orientation et à la direction du mouvement (Bullier, 1998, 2001, 2002). En revanche, ces cellules apparaissent plus sensibles aux basses fréquences spatiales (Kaplan, 2003). La Figure 6 présente le trajet réalisé par les voies M et P, et nous permet d'avoir une vue d'ensemble des connexions qui existent dans l'aire V1 et des connexions extrinsèques, entre V1, V2, mais aussi V4 et MT. La

Figure 6 permet également de montrer qu'il existe deux grands types de voies : l'une reliée au cortex pariétal (voie dorsale) et l'autre reliée au cortex inférotemporal (voie ventrale). Nous pouvons ainsi constater que les afférents de la voie dorsale sont issus des cellules de type M, tandis que la voie ventrale est activée par les cellules M, P et K (Bullier, 1998 ; 2002).

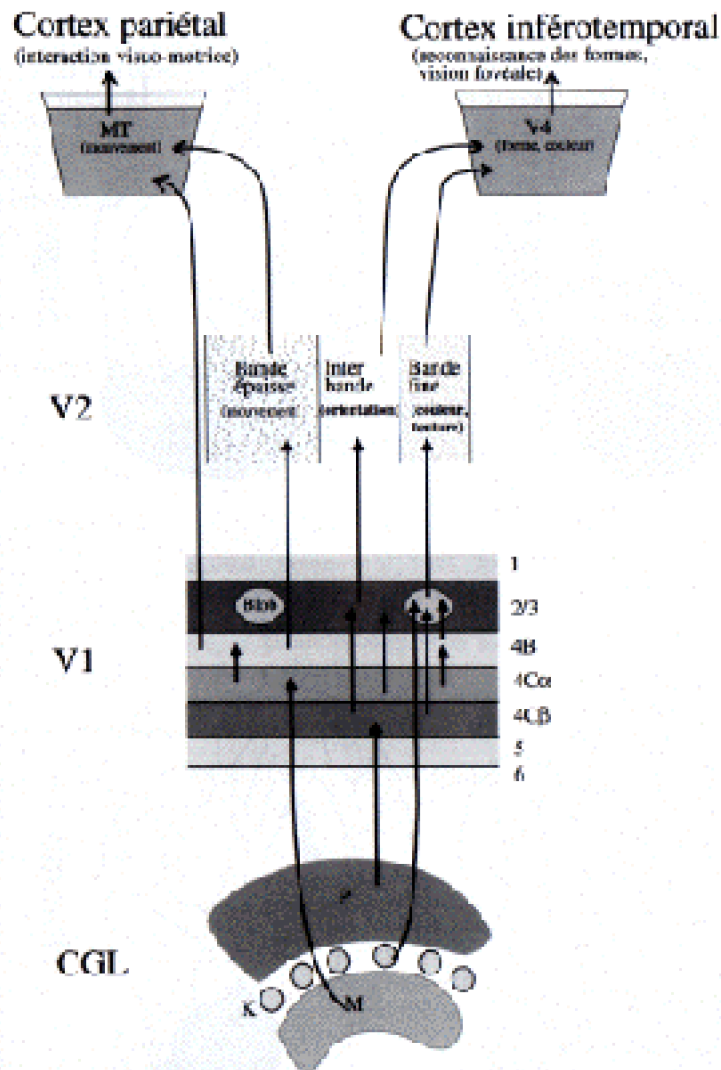


Figure 6 : Les flèches représentent les connexions ascendantes entre le CGL, l'aire V1 et l'aire V2, mais aussi avec l'aire MT et V4 (d'après Bullier, 1998).

1.1.4.2. Propriétés des champs récepteurs de V1

Une distinction existe entre les champs récepteurs des cellules corticales et les champs récepteurs des cellules ganglionnaires ; en effet, la majorité des cellules corticales répondent à des lignes orientées : ce sont des cellules spécifiques à une orientation donnée. Par exemple, une cellule corticale ne sera activée par une forme que si la limite de cette forme coupe son champ récepteur avec l'orientation appropriée. De telle sorte

que les neurones sensibles aux autres orientations dans la même région ne répondront pas ou moins. De même, l'activité des neurones diminuera fortement lorsque l'ensemble du champ récepteur est éclairé, comme pour les cellules ganglionnaires. L'information concernant l'orientation des contours de l'objet est codée spécifiquement dans V1. Deux principaux types de cellules sont localisés dans l'aire V1 : les cellules simples et les cellules complexes. Dans le modèle hiérarchique de Hubel et Wiesel (1962, 1968), les cellules simples sont situées en amont des cellules complexes.

1.1.4.3. Propriétés des cellules simples

Les cellules simples possèdent, comme les cellules ganglionnaires, un champ récepteur comportant des régions excitatrice et inhibitrice distinctes l'une de l'autre. Toutefois, l'organisation de ces cellules est différente : leurs champs récepteurs sont allongés, et sont divisés en deux ou trois parties inhibitrices ou excitatrices. Dans le premier cas, les champs récepteurs des cellules simples comporte une bande centrale excitatrice (ou inhibitrice) qui est entourée par deux régions inhibitrices (ou excitatrices). Le profil spatial des champs récepteurs des cellules simples est alors décrit comme une fonction Gabor⁵ 2D optimisée (Ringach, 2002). Un exemple de champ récepteur est présenté à la Figure 7. Dans le deuxième cas, les champs récepteurs sont divisés en deux parties : une bande excitatrice et une bande inhibitrice.

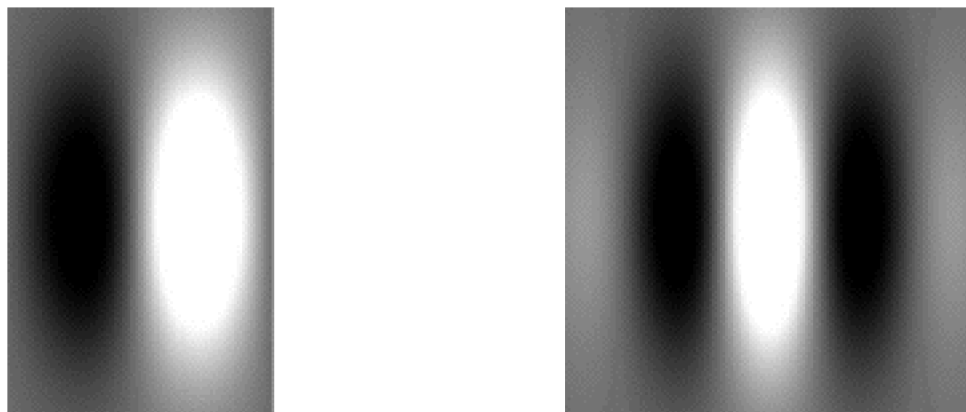


Figure 7 : Profil des champs récepteurs de l'aire V1. La zone claire représente la partie excitatrice, et la zone noire la partie inhibitrice. La figure de gauche présente un champ récepteur avec un profil de fonction Gabor optimisé, et la figure de droite un profil de champ récepteur dont le profil a deux composants.

De même que pour les champs récepteurs du CGL ou des cellules ganglionnaires, plus une stimulation remplit une région donnée, et plus l'excitation ou l'inhibition est intense. De plus, il n'y a pas de réponse quand les deux régions antagonistes sont stimulées simultanément. La stimulation la plus forte correspond à une bande allongée, orientée et positionnée de telle façon à ce qu'elle recouvre entièrement la région excitatrice. La réponse diminue lorsque l'orientation de la stimulation change puisque

⁵ Une fonction Gabor est obtenue par la multiplication entre une sinusoïde et une fonction gaussienne.

cette stimulation recouvre alors de plus en plus la partie inhibitrice. La spécificité de ces cellules est donc d'être hautement sensible aux contours d'une forme. Une hypothèse a été formulée pour déterminer la façon dont les cellules corticales sont connectées pour engendrer des champs récepteurs comme ceux des cellules simples. Les champs récepteurs des cellules simples seraient formés à partir du recouvrement de plusieurs champs récepteurs, du CGL, alignés dans la même direction. A partir des travaux sur les cellules simples, Hubel et Wiesel (1968) ont proposé une architecture en colonne du cortex visuel. Cette architecture est composée de plusieurs petites unités corticales appelées les hypercolonnes. Ces unités seraient localisées sur la surface du cortex pour une taille de 1.5 mm (Gilbert & Wiesel, 1979). Dans chaque hypercolonne, une progression régulière et unidimensionnelle représente la sensibilité des cellules à l'orientation. Un exemple est présenté à la Figure 8. Ces modules analysent une région particulière de l'espace, et la succession de ces modules est arrangée de façon rétinotopique.

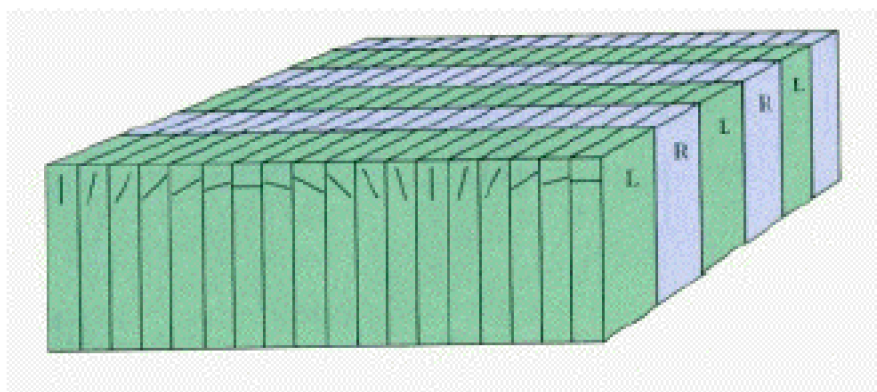


Figure 8 : Modèle schématique de l'organisation en colonne de V1. Les parties claires et foncées distinguent les colonnes à dominance oculaire (œil gauche et œil droit). A l'intérieur de chaque région de dominance oculaire, des colonnes de cellules accordées à différentes orientations sont disposées de façon orthogonale. Chaque région répond à une région particulière du champ visuel.

1.1.4.4. Propriétés des cellules complexes

Selon le modèle de Hubel et Wiesel (1962, 1968), les cellules complexes ont le rôle suivant. Comme les cellules simples, elles répondent à des barres ayant une orientation spécifique. Mais, les cellules complexes ont un champ récepteur uniforme, c'est-à-dire sans régions excitatrice et inhibitrice distinctes bien qu'elle code un contraste. Ces cellules sont fortement activées lorsqu'un signal balaie le champ récepteur avec une orientation correcte. Cette propriété les différencie principalement des cellules simples où le signal doit être fixe et correctement orienté. Hubel et Wiesel (1962) ont formulé l'idée que les champs récepteurs des cellules simples sont formés à partir des champs récepteurs du CGL. De la même façon, l'émergence des cellules complexes est le résultat de l'activité des cellules simples.

Hubel et Wiesel (1962) ont également décrit une troisième catégorie à savoir les cellules hypercomplexes. Les propriétés des champs récepteurs de ces cellules sont identiques à celles des cellules complexes. Les cellules simples et complexes additionnent les stimulations : plus la barre est longue et plus la réponse de ces cellules est importante, et ce, jusqu'au moment où la barre remplit entièrement la longueur du champ récepteur, au-delà de cette limite leur réponse ne varie plus. Par contre, la réponse des cellules hypercomplexes, appelées également cellules à inhibition terminale, augmente avec la longueur de la barre jusqu'à une limite donnée, puis l'activité diminue lorsque la barre dépasse cette limite. En réalité, cette propriété concerne aussi bien les cellules simples que les cellules complexes. Les cellules qui possèdent cette propriété sont appelées des détecteurs de terminaisons (« end-stopped cells ») ou encore cellules hypercomplexes de type I, s'il s'agit des cellules simples, et cellules hypercomplexes de type II s'il s'agit de cellules complexes (Palmer, 1999).

1.1.4.5. Et au delà de V1 ?

Ces premières recherches en physiologie ont montré que l'aire visuelle primaire traite les éléments locaux. Toutefois, comme le codage des bords est relativement rudimentaire, nous aborderons dans cette partie, le rôle des autres aires visuelles dans l'intégration des contours. Nous nous intéresserons plus particulièrement aux aires constituant la voie ventrale (Ungerleider & Mishkin, 1982). Cette voie part du cortex occipital (V1, V2, V3, V4) pour aboutir dans le cortex inférieur temporal (un exemple est présenté à la Figure 6). Les cellules du cortex inférieur temporal, qui possèdent des champs récepteurs larges, répondent à des informations très élaborées. Elles semblent avoir une fonction intégrative des propriétés des objets telles la forme, la couleur et la texture. Le système ventral apparaît donc être crucial pour l'identification visuelle des objets.

Une étude de Schwartz, Desimone, Albright et Gross (1983) conforte l'hypothèse selon laquelle V1 ne serait pas la seule aire corticale à être sensible aux contours d'une forme. Schwartz et al. (1983) ont enregistré l'activité des cellules de l'aire temporelle inférieure en testant plus particulièrement la complexité des contours. Ce type de contour était obtenu par l'utilisation de descripteur de Fourier sur le bord d'un cercle. Leur recherche a montré que ces cellules traitent sélectivement les fréquences spatiales relatives au contour du cercle et, leur activité ne varie pas avec la taille du signal. En d'autres termes, les neurones sont sensibles à des contours de plus en plus complexes au cours du traitement. Pasupathy et Connor (1999) ont également montré que les neurones de V4 étaient plus sensibles aux caractéristiques d'un contour (des angles ou des courbes pointés vers l'extérieur de façon convexe) que d'une simple barre lumineuse. Par ailleurs, Pasupathy et Connor (1999) ont observé une sensibilité de ces neurones à l'orientation. Ce résultat indique que l'aire V4 est une étape intermédiaire dans la reconnaissance d'une forme plus complexe. Ces deux études mettent en évidence des mécanismes de plus en plus élaborés, pour l'intégration des contours, au fur et à mesure du traitement cortical.

1.1.4.6. Intégration des contours illusoirs par le cortex visuel

Une différence de luminance présente sur le bord de la forme est nécessaire pour la détection des contours, mais le phénomène des contours illusoires indique qu'une discontinuité de luminance n'est pas nécessaire pour détecter des bords. Nous aborderons ici la façon dont les contours illusoires sont encodés par le système visuel.

Des études physiologiques ont étudié particulièrement le traitement des contours en enregistrant les réponses des cellules des aires V1 et V2 du singe après la présentation de bords continus (des barres) et de bords discontinus composés par des segments alignés formant un contour illusoire (Peterhan & von der Heydt, 1989 ; von der Heydt & Peterhans, 1989). Les résultats ont indiqué que les cellules sensibles à l'orientation d'une barre, dans l'aire V1, sont activées par les bords continus mais ne répondent pas aux bords discontinus. Par ailleurs, les auteurs ont effectué une brèche, d'une taille de 2 à 8 min, au sein d'un contour continu. Aucune réponse des neurones n'a été enregistrée, ce qui signifie que les discontinuités étaient négligées dans le cortex visuel primaire. Néanmoins, 40 % des neurones enregistrés dans V2 répondent aux bords continus comme à des segments discontinus formant un contour illusoire. Ces données indiquent donc que l'aire V2 intègre les éléments locaux concernant l'intégration des contours illusoires. Selon von der Heydt et Peterhans (1989), cette intégration est très rapide puisqu'elle se produit 70 ms après la présentation du signal.

Pour expliquer le traitement des contours, von der Heydt et Peterhans (1989) ont proposé l'existence de neurones qu'ils ont appelé les « cellules contours ». Les « cellules contours » de V2 sont basés sur un mécanisme de sommation des signaux provenant : 1) des cellules sensibles à l'orientation, depuis les cellules simples ou complexes de V1, et 2) des cellules à inhibition terminale. Les « cellules contours » seraient situés dans l'aire V2. Ces cellules seraient activées lorsque les fins de lignes sont placées dans la région excitatrice des champs récepteurs à inhibition terminale. La réponse des « neurones contours » est réduite lorsque des fins de lignes sont placées dans la région inhibitrice des cellules à inhibition terminale. La théorie des « cellules contours » dans V2 n'implique pas une explication complète de l'intégration des contours, mais un mécanisme possible pour connaître la façon dont l'activité des cellules de V2 est dérivée des réponses de V1 (von der Heydt, 1995).

Par la suite, ces auteurs ont montré que les neurones possédant ces capacités d'intégration étaient sensibles à la colinéarité des éléments locaux du contour. Peterhans et von der Heydt (1991) ont effectué une autre série d'expériences dans laquelle, des points étaient présentés sur un axe horizontal et l'écartement des points était manipulé par rapport à cet axe. Les résultats ont montré que la réponse des neurones de V2 diminuait en fonction de l'écartement des points. Un écart de deux minutes était suffisant pour obtenir une diminution de la réponse des neurones. Le mécanisme des « cellules contours » permet d'expliquer la modulation des réponses en fonction de l'alignement. En effet, le mécanisme de sommation serait optimisé lorsqu'un grand nombre de point tombe dans la région excitatrice des champs récepteurs des cellules à inhibition terminale. Par conséquent, plus les points sont positionnés sur les détecteurs de terminaison, et plus l'activité des « cellules contours » sera réduite (von der Heydt, 1995).

Cet effet apparaît également dans des études comportementales chez l'observateur humain. Dans l'expérience de Zucker et Davies (1988), les stimuli étaient composés de

plusieurs barres qui formaient un soleil. Cette illusion est communément appelée la figure d'Erhenstein. Les barres étaient continues ou en pointillées et l'espacement entre les points était également manipulé. Dans cette illusion, un rehaussement de la luminosité au centre de la figure est constaté. La tâche des observateurs était de détecter le rehaussement ou non de la luminosité. Les résultats indiquaient que la région centrale apparaissait plus claire que le fond jusqu'à un espacement de 6 min entre les segments.

De nombreuses études neurophysiologiques ont montré que les aires V1 et V2 étaient activées après une présentation de contours illusoires (Grosf, Shapley & Hawken, 1993). De plus, les réponses étaient plus importantes dans V2 que dans V1 (von der Heydt & Peterhans, 1989 ; Lee & Nguyen, 2001 ; Sheth, Sharma, Rao & Sur, 1996). La réponse des neurones aux contours illusoires a permis à von der Heydt et Peterhans (1989) de proposer un modèle sur l'intégration des contours par le cortex visuel. Toutefois les études physiologiques, traitant des contours illusoires, n'ont jamais analysé le rôle des activations des aires visuelles au-delà de V2. Pour cette raison, Mendola, Dale, Fischl, Liu et Tootell (1999) ont réalisé une étude en IRMf du cortex visuel chez l'observateur humain durant la perception de plusieurs contours illusoires et réels. Cette étude permet de répondre à la question suivante : est ce que les aires visuelles impliquées dans le traitement des contours illusoires et réels sont les mêmes ? Deux types de contours illusoires ont été utilisés : un carré de Kanisza et un ensemble de lignes horizontales où certains éléments du centre sont déplacés afin de former un carré. Une comparaison était ensuite réalisée avec une condition qui n'induisait plus de contours illusoires, c'est-à-dire que pour le carré de Kanisza, les pac-men étaient retournés vers l'extérieur, et la seconde illusion ne comprenait que des lignes horizontales. Le contour réel était représenté par un disque plein dont la luminance était supérieure à celle du fond. Les deux types de contours illusoires engendraient une forte activité neuronale dans des aires de haut niveau comme V3A, V7, V4v et V8. Par contre, une faible activité était enregistrée dans les aires V1 et V2 pour le carré de Kanisza. Par ailleurs, la présentation d'un disque uniforme indiquait une activité équivalente entre les différentes aires visuelles. Mendola et al. (1999) supposent que si un enregistrement neuronal est réalisé chez le singe pour des régions situées au-delà de V2, la réponse devrait être plus importante. Hirsch, DeLaPaz, Relkin, Victor, Kim, Li, Borden, Rubin et Shapley (1995) ont également réalisé une étude en IRMf pour une figure à contour illusoire (le carré de Kanisza). Les résultats ont montré des activations spécifiques du cortex extra strié, plus particulièrement dans l'hémisphère droit, quand des formes illusoires étaient perçues, par rapport à des conditions où des contours réels étaient présentés. Pour résumer, l'ensemble des études en IRMf implique que : 1) certaines aires sont associées de façon plus spécifique dans la perception des figures illusoires, et 2) l'activité est plus importante au-delà de V2. L'utilisation de figures à contour illusoire semble être controversée puisque ce type de stimuli n'impliquent pas un traitement équivalent à celui des formes naturelles de la part du système visuel.

1.1.5. La théorie des filtres

Les neurones du cortex visuel sont sensibles aux « barres » d'une orientation spécifique. Ceci implique que l'aire V1 détecte les contours des objets, et à partir de cette information Hubel et Wiesel (1962, 1968) ont émis l'hypothèse qu'un percept pouvait être construit.

Toutefois, DeValois, Albrecht et Thorell (1982) ont présenté une approche alternative du traitement cortical. Ces auteurs ont étudié les propriétés des champs récepteurs des cellules corticales en utilisant des réseaux sinusoïdaux au lieu de barres lumineuses. En effet, les travaux précurseurs de Campbell et Robson (1968) ont montré une sensibilité de l'observateur pour différentes fréquences spatiales. Les résultats de DeValois et al. (1982) ont indiqué que les cellules corticales étaient activées par ces réseaux sinusoïdaux mais avec des fréquences spatiales spécifiques. Par exemple, une cellule répondait à une fréquence spatiale de 3 cpd, alors qu'une autre cellule était activée par une fréquence spatiale de 6 cpd. DeValois et al. (1982) ont enregistré à la fois des cellules simples et complexes, ces cellules étaient accordées aux fréquences spatiales basses, moyennes et hautes.

Cette sensibilité aux fréquences spatiales a conduit de nombreux chercheurs à utiliser ce type de stimuli dans leurs expériences. Par exemple, dans une étude en IRMf, Boynton, Demb, Glover et Heeger (1999) ont montré une activation dans l'aire V1 pour une tâche de discrimination de contraste entre deux fréquences spatiales. Les deux fréquences spatiales étaient présentées à la suite, et l'observateur devait indiquer lequel des deux stimuli avait le niveau de contraste le plus élevé. Le niveau d'activité de l'aire V1 était de plus en plus important en fonction de l'augmentation du niveau de contraste entre ces deux stimuli. Par ailleurs, une similitude a été constatée entre le niveau d'activité de V1 et le seuil des réponses pour chacun des niveaux de contraste.

La théorie de Hubel et Wiesel (1962, 1968) propose que le codage d'une barre reflète en quelque sorte sa perception, mais la théorie des filtres montre que cela n'est pas aussi simple. L'opposition entre ces deux théories réside dans le fait que le modèle de Hubel et Wiesel (1962, 1968) avait un objectif qui était la représentation d'un contour par le système visuel, tandis que la théorie des filtres n'a pas d'objectif direct puisqu'un réseau sinusoïdal n'a pas de sens pour une forme. La théorie des filtres a permis d'apporter un nouveau point de vue : est-il pertinent d'envisager les neurones par des détecteurs de caractéristiques spatiales comme des traits dans V1 ? Les implications de ces deux théories sont différentes, la théorie des filtres suggère que les cellules de V1 ne sont pas des détecteurs de traits, comme des lignes ou des bords, mais plutôt des filtres qui décomposent l'image en primitives. Ce point de vue n'exclut pas l'existence de détecteurs de contours ou de lignes. Mais ces détecteurs seraient situés à un plus haut niveau, où les filtres seraient combinés pour déterminer les contours d'une forme. Par conséquent, la théorie des filtres est compatible avec la théorie des détecteurs de bord ou de ligne, mais avec la nuance que ces détecteurs ne seraient pas localisés dans V1 (Palmer, 1999).

Par ailleurs, la théorie des filtres ne remet pas en cause l'existence de la sélectivité à l'orientation des cellules de V1 (Hubel & Wiesel, 1968), ainsi que la participation de cette aire corticale dans la perception (Paradiso, 2002). En effet, DeValois et DeValois (1988) ont mis en évidence l'existence d'un lien entre les caractéristiques de sélectivité à l'orientation et celles aux fréquences spatiales. Une dimension supplémentaire a été ajoutée à l'architecture en hypercolonne décrite par Hubel et Wiesel (1968). De façon schématique, l'organisation de chaque hypercolonne se déroule de la façon suivante : des colonnes de cellules accordées à différentes fréquences spatiales et d'orientation sont disposées de façon orthogonale afin que chaque combinaison de fréquence spatiale et

d'orientation soit représentée dans une hypercolonne.

Les cellules de V1 sont sensibles aux fréquences spatiales d'une image, mais l'analyse effectuée par les champs récepteurs est essentiellement locale puisque la taille des champs récepteurs limite leur activité à de petites régions du champ visuel. Cet argument restreint l'utilisation de réseaux sinusoïdaux dont l'extension est théoriquement infinie selon le théorème de Fourier. Néanmoins, l'analyse des fréquences spatiales peut être réalisée en utilisant un signal de luminance sinusoïdale dont le contraste (l'amplitude) diminue en fonction de l'éloignement du centre du réseau. Ce signal est appelé une fonction Gabor comme nous l'avons discuté plus haut, un exemple est présenté à la Figure 9. Les cellules simples de V1 sont activées par la présentation de ce type de stimulus (Palmer, 1999). Une hypothèse prédominante suppose que les premières représentations corticales de l'image visuelle consistent en une décomposition par des filtres Gabor. Ces filtres sont localisés spatialement et répondent à un signal provenant d'une région spécifique de la rétine. Toutes ces raisons ont conduit à remplacer l'utilisation des réseaux sinusoïdaux par des fonctions Gabor.



Figure 9 : Exemples de signaux Gabor. La figure de gauche présente un signal Gabor dont la fréquence spatiale est basse, et la figure de droite montre un signal Gabor avec une fréquence spatiale élevée.

1.1.6. Que reste-t-il du modèle de Hubel et Wiesel aujourd'hui ?

1.1.6.1. Peut-on vraiment considérer le traitement visuel comme hiérarchique ?

Les travaux de Hubel et Wiesel (1962, 1968) privilégient un traitement hiérarchique. Leur modèle propose que les propriétés des champs récepteurs, pour un niveau donné, seraient la conséquence de l'arrangement des réponses données par les neurones des couches inférieures de la hiérarchie. Par conséquent, au cours des premières étapes du traitement visuel, les champs récepteurs se spécialisent dans le traitement de l'information. Par exemple, les cônes répondent à la luminance, les cellules ganglionnaires à un contraste de luminance, les cellules simples à un contraste orienté, et les cellules complexes à un contraste orienté en mouvement. De façon plus spécifique, les neurones du CGL convergent vers la couche 4 (de V1) pour former les cellules simples, et ensuite les neurones de la couche 4 activent la couche 3 (de V1) pour former les cellules complexes (Hubel & Wiesel, 1962). Une illustration de ce modèle est

présentée dans la Figure 10(a).

Ce modèle hiérarchique donne une importance aux connexions ascendantes (« feedforward »), mais cette conception ne permet pas d'expliquer pourquoi dans la couche 4 du chat, 95 % des synapses d'une cellule simple proviennent d'autres neurones corticaux. Réciproquement, les connexions excitatrices de la couche 3 de V1, qui est l'étape suivante dans le modèle hiérarchique, ne proviennent que de 20 % de la couche 4. Le modèle purement ascendant de Hubel et Wiesel ignorait l'existence de connexions locales pour une même couche, à savoir les connexions horizontales, mais également descendantes entre les couches (Gilbert, 1995). En effet, de nombreuses connexions horizontales excitatrices et inhibitrices existent au niveau des couches 3 et 4, et ces cellules interagissent entre elles (Fitzpatrick, Lund & Blasdel, 1985 ; Miller, Pinto & Simons, 2001). Par ailleurs, des connexions descendantes relient la couche 3 à la couche 4 (Martin, 2002). Ces connexions descendantes sont plus nombreuses que les connexions ascendantes. En effet, leur organisation est telle qu'un neurone reçoit uniquement une information depuis les neurones afférents. En revanche, les connexions descendantes sont organisées afin de transmettre une information depuis un neurone cible vers plusieurs neurones d'un niveau inférieur (Bullier, 1998). La Figure 10(b) présente les différentes modifications qui ont été apportées au modèle de Hubel et Wiesel (1962). Par conséquent, la structure de V1 ne peut pas se limiter à un modèle hiérarchique, puisque seul un nombre restreint de cellules corticales permettent ce traitement (Martin, 2002 ; Somers, Nelson & Sur, 1995). Cette idée ne concerne pas seulement l'aire V1, mais s'étend à l'ensemble des aires visuelles. En effet, les connexions entre les aires visuelles sont établies à travers deux types de projections. Une projection ascendante transmet l'information d'une aire inférieure dans la hiérarchie à une aire supérieure. Une projection descendante effectue le chemin inverse, c'est-à-dire, qu'elle transmet l'information d'une aire hiérarchique supérieure à une aire inférieure.

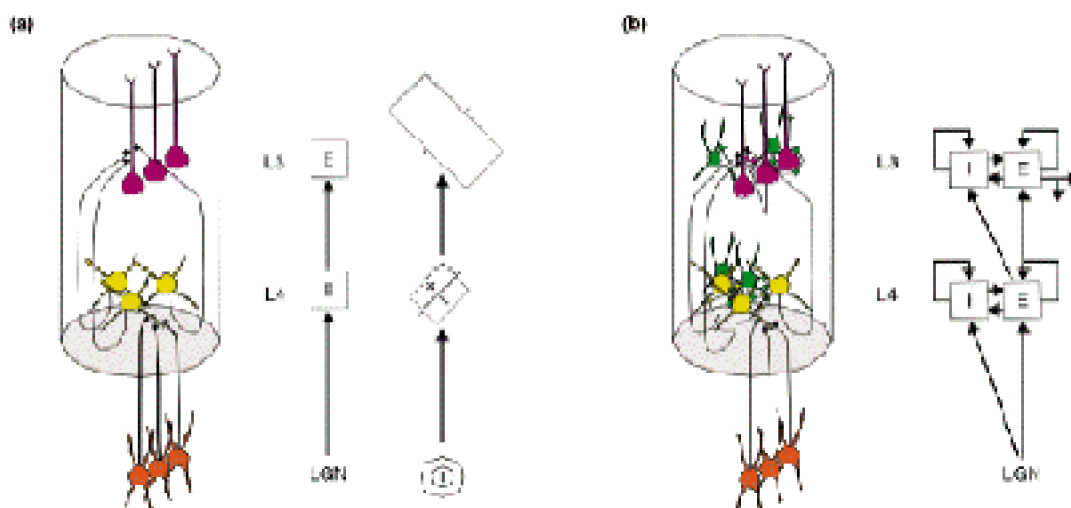


Figure 10 : La figure de gauche présente le modèle hiérarchique de Hubel et Wiesel, où un traitement purement ascendant s'effectue depuis le CGL jusque dans la couche 3 de V1.

La figure de droite présente un modèle alternatif dans lequel les couches 3 et 4 contiennent des neurones excitateurs et inhibiteurs qui sont connectés de toutes les façons

possibles (issu de Martin, 2002, p. 419).

D'un point de vue classique, les champs récepteurs de V1 sont stimulés spécifiquement par l'orientation d'un signal. Les travaux de Hubel et Wiesel (1962) ont eu des conséquences importantes sur la façon de considérer le fonctionnement du système visuel. Ainsi, les aires visuelles étaient considérées comme étant sensibles à différents éléments, par conséquent, possédaient des fonctions spécifiques. Par exemple, l'aire MT était spécifique au mouvement tandis que V4 était l'aire de la couleur. Une organisation de type modulaire était établie de façon stricte (Felleman & Van Essen, 1991 ; Livingstone & Hubel, 1988). La difficulté majeure de ce type d'organisation réside dans le traitement temporel d'une image. En effet, il existe une contradiction entre la rapidité à laquelle le système visuel traite une information (Thorpe, Fize & Marlot, 1996) et la lenteur d'activation des neurones (Nowak & Bullier, 1997).

La rapidité du système visuel pourrait résider dans le fait que l'information en provenance de la rétine ne se propage pas comme une vague qui escalade les différents niveaux mais plutôt comme trois vagues successives. Le décalage temporel entre les trois vagues successives d'activation (voie M, P et K) crée un fonctionnement asynchrone que le système pourrait utiliser. D'après Bullier (1998, 2001), ces différences de temps de réponses jouent un rôle important dans le traitement de l'information par le cortex cérébral. Notamment, un modèle par itération basé sur l'activation des neurones M, P et K à différents moments permettrait d'expliquer la vitesse importante à laquelle un stimulus est traité. Ce modèle remettrait en question l'hypothèse d'une organisation hiérarchique basée principalement sur des connexions ascendantes (Hubel & Wiesel, 1962 ; Livingstone & Hubel, 1988) et mettrait l'accent sur les connexions descendantes caractérisées par un décalage temporel à des niveaux spatiaux différents selon la catégorie de neurones de relais.

1.1.6.2 .Influence contextuelle sur l'activité des champs récepteurs

Cette partie a pour objectif de montrer que les champs récepteurs ne se limitent pas à une simple transformation des entrées en une sortie, mais, au contraire ce serait une structure plus dynamique (Wörgötter et Eysel, 2000).

De nombreuses études ont indiqué que la sélectivité à l'orientation des cellules corticales, est influencée par le contexte de la stimulation. Dans ces expériences, lorsqu'une barre lumineuse est présentée dans le champ récepteur d'un neurone avec la bonne orientation, et que simultanément apparaissait un autre stimulus positionné en dehors du champ récepteur, la réponse de la cellule pouvait être inhibée ou augmentée. Si une ligne verticale est entourée par des stimuli identiques, alors la réponse de la cellule diminue. Par contre, si cette ligne est entourée par d'autres stimuli dont l'orientation est différente (ligne verticale), la réponse de la cellule augmente (Gilbert & Wiesel, 1990 ; Knierim & Van Essen, 1992 ; Sillito & Jones, 1996). Par ailleurs, Levitt et Lund (1997) ont montré que cet effet était accentué selon le niveau de contraste entre la cible et le pourtour. La cible était composée d'un carré et le pourtour était également un carré mais de taille plus importante. Ces deux carrés étaient constitués de lignes dont l'orientation était différente dans le centre et le pourtour. Le contraste dépendait de la fréquence temporelle des barres qui composaient ces deux stimuli. Dans le cas où le contraste de la

cible était élevé, la réponse des neurones était supérieure, comparativement à une condition où la cible était présentée seule. Par contre, un faible contraste du signal engendrait une diminution de la réponse des cellules par rapport à la condition contrôle. Un effet identique a été mis en évidence avec l'utilisation de réseaux sinusoïdaux (Ohtani, Okamura, Yoshida, Toyama, & Ejima, 2002). Pour Lamme et Spekreijse (2000), ces effets contextuels permettraient d'expliquer des phénomènes attentionnels comme l'effet « pop out ». En effet, lorsque la cible est entourée de signaux dont l'orientation est aléatoire, alors la réponse des cellules n'est pas différente de la condition où la cible est entourée de signaux avec la même orientation (Kapadia, Ito, Gilbert, & Westheimer, 1995). Dans ces deux cas, il n'y a pas d'effet « pop out », puisque la cible ne se distingue pas des éléments du pourtour par une orientation particulière (Lamme & Spekreijse, 2000 ; Li, 1999). Kastner, Nothdurft et Pigorev (1997) ont répliqué l'expérience de Knierim et van Essen (1992) en utilisant cette fois des barres en mouvement.

1.1.6.3. Propriétés dynamiques des champs récepteurs

Les expériences que nous venons de présenter dans partie précédente montre l'existence d'interactions entre les neurones. Nous nous intéresserons ici à la structure même du champ récepteur. La question est ici de déterminer si la taille d'un champ récepteur est fixe ou, au contraire, évolue en fonction de la stimulation. Sceniak, Ringach, Hawken et Shapley (1999) ont enregistré la réponse de 85 cellules de V1 dans une expérience de sommation spatiale⁶, en utilisant un réseau sinusoïdal comme stimulus. Les fréquences temporelles et l'orientation étaient spécifiques à chaque cellule et seul le contraste du réseau variait. L'activité des cellules indiquait que l'augmentation du contraste du stimulus diminuait la taille du centre des champs récepteurs, par ailleurs, aucun effet n'a été constaté pour le pourtour antagoniste. Ces données suggèrent que les variations de contraste implique un changement de la polarité entre les cellules corticales. Selon Sceniak et al. (1999), ce mécanisme permettrait une meilleure détection des contours lorsque le signal est faible. Kapadia, Westheimer et Gilbert (1999) ont également réalisé une expérience de sommation spatiale en utilisant une barre lumineuse. Les résultats ont aussi indiqué que la taille des champs récepteurs n'était pas fixe, mais dépendait du contraste du signal. La taille du champ récepteur était 4 fois plus importante pour un stimulus de faible contraste par rapport à un stimulus ayant un contraste élevé. Cette dynamique des champs récepteurs a également été mis en évidence dans le cas de lésion cérébrale dans laquelle la plasticité cérébrale permet une augmentation de la taille des champs récepteurs. Cette question sera traitée dans une partie ultérieure (concernant le remplissage des scotomes).

1.1.6.4 Intégration des contours en dehors de la limite des champs récepteurs

Un champ récepteur n'est pas suffisant pour intégrer le contour d'une forme, l'objectif de cette partie est de montrer l'importance de l'aire visuelle primaire pour intégrer un contour dont la taille est plus grande qu'un champ récepteur. La théorie des champs récepteurs

⁶ Le principe des expériences de sommation spatiale sera traité dans une partie ultérieure.

classique (Hubel & Wiesel, 1962, 1968), décrite dans les parties précédentes, supposait que l'intégration d'une forme serait réalisée par des mécanismes de haut niveau. Par ailleurs, la conception de Hubel et Wiesel (1962) ne permettait pas d'expliquer à elle seule l'intégration des contours une fois la limite spatiale des champs récepteurs dépassée. Mais, les expériences liées aux effets contextuels indiquent que les cellules du cortex sont capables d'intégrer les informations à travers des grandes parties du champ visuel (Gilbert, 1995). L'activité des cellules est influencée par un signal situé en dehors de son champ récepteur, ce qui est cohérent avec l'existence des connexions horizontales « long-range » ($< 0.2-3\text{deg}$). En effet, ces projections relient préférentiellement des cellules codant les mêmes orientations et inhibent celles qui ont une orientation orthogonale (Gilbert, 1995 ; Gilbert & Wiesel, 1989 ; Malach, Amir, Harel & Grinvald, 1993). Kapadia, Westheimer, et Gilbert (2001) ont utilisé des barres qui étaient alignées dans un même axe ou parallèle. La réponse des neurones à une barre était plus élevée lorsque ces dernières étaient alignées que parallèles. La réponse des neurones diminuait pour une différence d'orientation supérieure à 20 deg.

Par ailleurs, Polat, Mizobe, Pettet, Kasamatsu, et Norcia (1998) ont montré que ces résultats variaient selon le niveau de contraste entre la cible et les signaux du pourtour. Dans leur expérience, Polat et al. (1998) ont utilisé des signaux Gabor pour enregistrer la réponse des neurones de V1, dont le champ récepteur était positionné sur l'un de ces signaux (la cible). Le contraste de luminance de la cible Gabor était manipulé, et le niveau de luminance des signaux Gabor du pourtour était constant. Les signaux Gabor du pourtour étaient situés à une distance équivalente à 2-3 fois le diamètre du champ récepteur correspondant aux cellules enregistrées. Lorsque la cible était présentée seule, l'enregistrement des neurones montrait que plus le contraste de la cible augmentait, plus la réponse des neurones était importante. Par contre, lorsque deux signaux Gabor étaient présentés de façon colinéaire à la cible, la réponse des cellules était facilitée pour les niveaux faibles de contraste de la cible, et réduite pour des niveaux de contraste élevés. Lorsque les deux signaux Gabor étaient présentés de façon orthogonale à la cible, les effets inhibiteurs augmentaient et les effets facilitateurs diminuaient. Chavane, Monier, Bringuier, Baudot, Borg-Graham, Lorenceau, et Frégnac (2000) ont également mis en évidence l'importance de la colinéarité des signaux Gabor pour l'aire V1 du singe. De plus, Ghose, et Maunsell (2002) ont montré que les réponses neuronales étaient plus rapides lorsque l'orientation des signaux Gabor était proche. Ainsi, les réponses neuronales dépendent de l'organisation d'ensemble du stimulus (Kasamatsu, Polat, Pettet, & Norcia, 2001). Par ailleurs une étude récente en IRMf a permis de montrer des signaux Gabor colinéaires activaient non seulement V1 mais aussi le cortex occipital complexe (Altmann, Bühlhoff, & Kourtzi, 2002), cette aire est fortement impliquée dans la perception d'objet (Kourtzi & Kanwisher, 2000).

Ces résultats confortent les résultats précédents concernant l'influence contextuelle, en l'élargissant aux signaux placés sur un même axe (Lamme & Spekreijse, 2000). Cette dynamique des champs récepteurs semble indiquer que les mécanismes nerveux précoces du système visuel sont capables d'encoder et d'intégrer un contour en dehors de la limite du champ récepteur classique.

Les effets contextuels montrent que V1 peut intégrer des contours sur des grandes

parties du champ visuel. Une autre possibilité, et sur des régions plus éloignées encore, est fournie par la théorie de la synchronisation. Les neurones, qui répondent à une même image, sont activés simultanément alors que les neurones qui répondent à une image différente ne devraient pas répondre de façon synchrone. Le principe de la synchronisation permettrait de favoriser l'intégration des contours, lorsque les éléments d'une même forme sont liés du point de vue perceptif (Roelfsema & Singer, 1998 ; Singer, 1999). Cette hypothèse a été validée par une expérience dans laquelle des neurones, distants les uns des autres, étaient activés simultanément (synchrone) lorsqu'ils répondaient à un contour continu (Singer & Gray, 1995). Par contre, l'activation de ces mêmes neurones était réduite quand on réalisait une segmentation de ce contour (Gray, König, Engel & Singer, 1989), ce qui suggère que la synchronie est seulement possible pour des formes, dont les contours sont continus.

1.2. Arguments comportementaux

Au cours de l'introduction nous avons vu qu'il existait une distinction entre le codage neuronal et la perception. Il nous semble donc pertinent de tenter de faire un lien entre les études physiologiques et les études comportementales qui s'interrogent sur des fonctions visuelles identiques (par exemple le contraste). Nous aborderons essentiellement des études psychophysiques qui ont étudié le mécanisme de sommation spatiale de la rétine, la façon de distinguer les voies M et P, et des expériences portant sur l'intégration des contours qui se centrent plus particulièrement sur le cortex strié.

1.2.1. La loi de Ricco

Les différences physiologiques qui existent entre le système photopique et scotopique entraîne des mécanismes distincts de sommation spatiale. En psychophysique, ce processus est classiquement étudié par l'expérience suivante : un point d'une luminance donnée est présenté à un observateur. La taille et la luminance de ce point sont manipulées et la tâche de l'observateur est de dire s'il perçoit ce point ou non. Le résultat classique est que pour des tailles inférieures à une surface critique, le seuil de luminance reste insensible à l'agrandissement de la taille du stimulus. Au-delà de cette surface le seuil de luminance augmente avec la taille du stimulus. Cet effet est communément appelé loi de Ricco. La taille critique de Ricco du système photopique est plus petite que celle du système scotopique, ce qui reflète les capacités réduites de sommation spatiale, c'est-à-dire la plus grande résolution, du système photopique (Barlow, 1958). Pour des conditions scotopiques, la taille critique est de 10 min.

1.2.2. Comment distinguer les canaux M et P chez l'observateur humain ?

Les données physiologiques de Kaplan et Shapley (1986) ont permis de déterminer des caractéristiques de contraste différentes pour la voie M et P. En effet, nous avons vu que la taille des champs récepteurs de ces canaux était différente, par conséquent les contours sont encodés sur différentes échelles spatiales. L'objectif de Smith et Pokorny (1997) était de distinguer ces deux voies de façon comportementale en s'appuyant sur

des techniques psychophysiques chez l'observateur humain. La sensibilité au contraste des canaux M et P est développée car le paradigme qui privilégie la voie M sera utilisé dans l'une de nos expériences.

Pour cela, Smith et Pokorny (1997) ont utilisé trois paradigmes différents. Dans ces trois paradigmes, quatre carrés étaient présentés sur un fond gris moyen, et la tâche consistait à indiquer le carré dont la luminance était inférieure ou supérieure aux trois autres carrés. Dans la condition « pulse », les carrés apparaissaient brièvement. Dans le second paradigme dont la condition était « steady-pedestal », quatre carrés de luminance identique étaient présentés, puis la luminance de l'un de ces carrés diminuait ou augmentait. Pour la condition « pedestal- Δ -pedestal », les quatre carrés avaient la même luminance, puis la luminance de ces carrés changeait mais l'un de ces éléments changeait différemment des trois autres. Le seuil de luminance (le plus petit changement de luminance détecté) était relevé pour chaque observateur et pour différents niveaux de luminance des trois carrés de manière à calculer des courbes de contraste.

Afin de déterminer la sensibilité des trois paradigmes à mettre sélectivement en œuvre la voie M ou P, une comparaison a été réalisée avec les données de Kaplan et Shapley (1986). Les courbes de la Figure 5 indiquent que la réponse des cellules de la voie M augmentait rapidement avec l'augmentation du contraste ; de même la réponse des cellules de la voie P augmentait de façon linéaire et fixe. Ces courbes sont principalement caractérisées par leur pourcentage de gain de contraste. Ce pourcentage est obtenu par la formule suivante : $(R_{\max} / C_{\text{sat}}) / 100$, dans laquelle R_{\max} est l'amplitude de réponse maximale des neurones (en impulsion par seconde ou ips), et C_{sat} est le contraste pour lequel l'amplitude de la réponse est la moitié de R_{\max} . Kaplan et Shapley (1986) ont montré que le pourcentage de gain de contraste était inférieur à 1 (compris entre 0.15 et 0.5) pour la voie P, et supérieur à 1 (compris entre 5 et 8) pour la voie M. Smith et Pokorny (1997) ont effectué une comparaison entre les pentes obtenues à partir des trois paradigmes, en remplaçant le nombre d'impulsion par seconde par le seuil de luminance des observateurs, avec les pentes obtenues à partir de l'expérience de Kaplan et Shapley (1986), ce qui permettait de prédire directement la participation de la voie M ou P dans les trois paradigmes.

Les résultats ont indiqué que le pourcentage de gain de contraste de la condition « pulse paradigm » était spécifique à la voie P. Par ailleurs, les résultats de la condition « steady-pedestal » impliquaient une participation de la voie M. Cette étude a permis de spécifier une participation de la voie M et P à l'aide de différentes tâches de discrimination, indépendamment de tout enregistrement neuronal. Par la suite, ces deux paradigmes ont été utilisés pour mettre en évidence un déficit important de la voie M chez des patients atteints d'une maladie dégénérative de la rétine : la rétinite pigmentosa⁷ (Alexander, Pokorny, Smith, Fishman & Barnes, 2001). Cette technique permettrait également d'étudier de façon comportementale l'implication des voies M et P pour différentes pathologies ; par exemple, il est proposé qu'un certain type de dyslexie soit spécifiquement associé à un trouble de la voie M (Schwartz, 1999 ; Stuart, McAnally &

⁷ Cette maladie se caractérise par une perte progressive de la vision périphérique et une capacité lente à s'adapter à des conditions où la lumière est faible (Rodieck, 1998).

Castles, 2001).

1.2.3. Intégration des contours sur des petites parties du champ visuel

La partie physiologique a permis de montrer que l'intégration des contours d'une forme nécessite des interactions entre les champs récepteurs. En effet, la taille d'un champ récepteur est bien trop petite pour pouvoir coder à lui seul les bords d'une forme. La question est donc de déterminer si l'existence de ces connexions reflète l'intégration des contours qui est réalisée par l'observateur humain. Nous aborderons au cours de cette partie la nature des relations entre des signaux Gabor sur des petites régions du champ visuel. L'intérêt suscité par ce thème montre l'importance accordée au codage et à la perception des contours.

Les interactions latérales entre les filtres spatiaux ont été étudiées dans un paradigme de masquage par Polat et Sagi (1993, 1994a, 1994b). Dans ces expériences, le signal était composé d'une fonction Gabor présentée en fovéa et le masque était constitué par deux autres fonctions Gabor qui étaient placées de chaque côté du signal. Les deux masques Gabor avaient un contraste identique mais plus élevé que celui de la cible. La cible et le masque avait la même orientation (horizontale, verticale et en diagonale), mais la distance les séparant variait. Un essai se déroulait de la façon suivante : dans un premier temps, le masque était présenté seul, il était suivi, après un intervalle, du masque et de la cible ; l'ordre de présentation des stimuli était contrebalancé. L'observateur devait indiquer si le signal était présent dans le premier ou le second stimulus. Les résultats ont montré un effet suppressif du masque sur la cible, puisque le seuil était plus élevé pour des distances courtes. Réciproquement, le seuil était plus bas pour des distances importantes entre la cible et le masque, et cette facilitation du traitement de la cible par le masque se maintenait au delà de 0.9 deg. Par ailleurs, les cibles horizontales (c'est-à-dire colinéaires au masque) étaient détectées plus facilement. Polat et Tyler (1999) ont également différencié la forme des fonctions Gabor. Dans leur expérience, les stimuli étaient présentés sous une forme circulaire ou allongée. La méthode était identique à celle utilisée dans l'expérience de Polat et Sagi (1993) et les résultats ont indiqué que la sensibilité était grande lorsque les fonctions Gabor étaient allongées. Cette dernière étude montre que la colinéarité des stimuli est importante. Par ailleurs, une diminution des performances était observée pour des stimuli non colinéaires (Vidnyanszky, Pappathomas & Julesz, 2001), ce résultat était identique à celui de Knierim et van Essen (1992).

Ces résultats sont compatibles avec les propriétés des connexions horizontales « long-range » dans le cortex visuel. En effet, ces connexions permettent une transmission de l'information entre des cellules de la même couche, en fonction de l'input reçu et un effet inhibiteur ou excitateur est obtenu (Hirsch & Gilbert, 1991). Par ailleurs, Polat et Sagi (1994) ont émis l'hypothèse de l'existence d'une chaîne d'interaction entre les neurones, nécessaire pour expliquer un seuil constant au delà de 0.9 deg. Cette hypothèse sera développée dans la prochaine partie.

1.2.4. Intégration des contours sur de grandes régions du champ visuel

Les études précédentes utilisaient des stimuli colinéaires dont l'étendue n'excédait pas 3 deg. Dans cette partie, nous aborderons la façon dont le système perceptif traite les contours pour des distances plus importantes. L'objectif est de montrer comment les contours sont intégrés par le système visuel précoce, le nombre important de ces recherches indique, là encore, la place prépondérante qu'occupe la dimension des contours en vision.

Des études (Kapadia, Westheimer & Gilbert, 2001 ; Polat, Mizobe, Pettet, Kasamatsu & Norcia, 1998) ont montré l'importance de signaux colinéaires sur une toute petite région du champ visuel (équivalent à la taille de trois champs récepteurs mis l'un à côté de l'autre). A notre connaissance, les études physiologiques se sont surtout intéressées à un traitement des stimuli colinéaires sur de petites parties du champ visuel. Par ailleurs, les connexions horizontales « long range » dans V1 s'étendent sur des distances plus importantes (Gilbert, 1995 ; Gilbert & Wiesel, 1989) que les distances entre les signaux Gabor utilisés par Polat et al. (1998). Il est donc tout à fait possible que les résultats de Polat et al. (1998) soient généralisables sur des régions plus importantes. Cette partie a pour objectif de développer la façon dont les contours sont intégrés par le système visuel pour des tailles supérieures à un seul champ récepteur.

1.2.4.1 . Principe du champ d'association

Field, Hayes et Hess (1993) ont proposé la notion de champ d'association afin d'expliquer l'intégration des contours sur des grandes régions du champ visuel. Ce concept se définit par des connexions latérales entre des cellules sensibles à la même orientation ou avec une orientation voisine. Les stimuli étaient composés de plusieurs fonctions Gabor dont l'orientation était manipulée pour être aléatoire ou, avoir la forme d'une courbe parmi d'autres signaux Gabor dont l'orientation était aléatoire. Ces deux types de stimuli étaient présentés l'un après l'autre (l'ordre de présentation était contrebalancé). La tâche des observateurs était de déterminer dans quel intervalle la courbe était présentée. Les observateurs détectaient plus facilement la cible lorsque l'orientation des signaux Gabor formant la courbe variait peu. De plus, le pourcentage de détection correcte diminuait de façon continue lorsque la différence d'orientation entre les signaux Gabor était importante. Au-delà d'une orientation de 30 deg, le pourcentage de détection était inférieur à 50%. Ce résultat est illustré dans la Figure 11 dans laquelle un lien est possible entre deux Gabor (traits pleins) ou non (traits en pointillés). Ce résultat suggère que les cellules sensibles à la même orientation ou à une orientation voisine sont connectées pour la perception des contours. Le principe du champ d'association indique que des connexions horizontales entre des neurones sensibles à des orientations voisines permettent d'intégrer des contours sur des grandes parties du champ visuel.

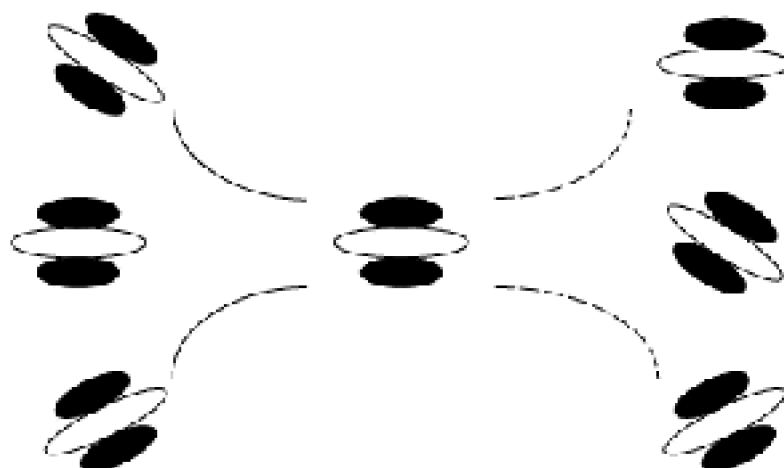


Figure 11 : Modèle schématique de la notion de champ d'association. Le trait continu indique que les cellules sensibles à une orientation voisine (< 30 deg) sont préférentiellement activées. L'inverse est illustré par les traits en pointillé.

Hess et Dakin (1997) ont testé si la règle du champ d'association était similaire en vision centrale et périphérique. Une courbe similaire aux précédentes était présentée en fovéa, ou à différentes excentricités. Les résultats montraient que les observateurs détectaient plus facilement la courbe apparaissant en fovéa, et que la détection diminuait quand l'excentricité augmentait jusqu'à 10 deg mais que, au-delà de 10 deg les performances ne variaient plus. Les mécanismes d'intégration de contour sont donc spécifiques de la région centrale du champ visuel. Par la suite, Hess et Dakin (1997) ont manipulé la phase des signaux Gabor (identique ou alterné), avec une présentation en fovéa ou à 20 deg d'excentricité (par rapport à un point de fixation situé dans la fovéa). Les performances diminuaient lorsque la phase était alternée et cette différence était plus importante lorsque les stimuli étaient présentés dans le champ visuel périphérique. Pour les auteurs, ce résultat indique que le champ visuel périphérique ne peut pas utiliser plus d'une opération de filtrage à la fois. On sait depuis longtemps que les fonctions visuelles dans la périphérie sont différentes de celles de la fovéa pour de nombreux aspects, et notamment l'incertitude spatiale. Hess et Dakin (1999) ont testé si ce facteur pouvait rendre compte des effets précédents. L'incertitude spatiale était manipulée dans une tâche de discrimination. Deux stimuli étaient présentés l'un après l'autre et ces stimuli étaient composés de signaux Gabor alignés, ou de signaux Gabor dont la position était incertaine le long d'un axe (selon une répartition 2D gaussienne). L'incertitude spatiale⁸ est plus élevée dans la périphérie, donc la performance pour les Gabors « incertains » devrait être améliorée en périphérie et non pas en fovéa. Les résultats ont montré que plus le sigma de la fonction gaussienne était important (plus l'incertitude était grande) et plus la discrimination du contour se détériorait. Par ailleurs, les performances étaient toujours inférieures dans la périphérie que dans la fovéa. Ce résultat indique que l'incertitude spatiale n'a pas d'impact direct sur la réponse des observateurs. Ce facteur ne peut donc pas expliquer la diminution de performance observée dans le champ visuel

⁸ L'incertitude spatiale correspond au degré d'imprécision pour lequel nous pouvons localiser un signal.

périphérique. L'ensemble de ces résultats semblent donc indiquer que l'intégration des contours n'est pas possible en périphérie. Li et Gilbert (2002) ont réalisé une expérience similaire à celle de Hess et al. (1993) mais les stimuli étaient composés de segments de ligne. De plus, l'espacement entre les segments et le nombre de signaux colinéaires (la taille du contour) étaient manipulés. Les résultats indiquaient que la perception d'un contour continue diminuait lorsque l'espacement entre les segments augmentait et ce quel que soit le nombre d'éléments formant le contour. La détection des signaux colinéaires était presque nulle pour un espacement supérieur à 2 deg. Cette taille est comprise dans la zone d'interaction des connexions horizontales long-range. Cette étude indique que l'intégration des éléments d'un contour est effectué pour des grandes tailles tant que l'espace entre les éléments reste en dessous d'une certaine distance.

Enfin, pour des stimuli plus complexes, Saarinen et Levi (2001) ont répliqué les résultats obtenus par Field et al. (1993) en utilisant un autre paradigme. Dans cette expérience, des éléments Gabor formaient un C et cette forme globale avaient 4 orientations différentes sur un fond gris moyen. L'observateur devait indiquer l'orientation globale de cette forme. Trois conditions étaient utilisées : dans la première condition, les segments Gabor étaient alignés avec le contour de la forme en C ; dans la seconde condition, l'orientation des segments était orthogonale à la forme ; et dans la dernière condition, les segments étaient alignés et orthogonaux. Le seuil des observateurs était plus bas dans la condition où les segments étaient alignés par rapport à la condition où les segments étaient orthogonaux. Ceci implique que des éléments locaux correctement orientés, facilitent l'identification de l'orientation d'une forme globale. Ce résultat peut être expliqué par le fait que les interactions locales permettent la diffusion d'une activité neuronale le long du contour.

En résumé, la présentation de l'ensemble de ces études a permis d'étendre les recherches de Polat et Sagi (1993, 1994) sur des grandes régions du champ visuel. D'un point de vue neurophysiologique, les connexions horizontales « long-range » permettraient d'expliquer les résultats obtenus avec la notion de champ d'association (Kovacs, 1996).

1.2.4.2. Influence de la forme des contours sur la notion de champ d'association

L'ensemble des expériences physiologiques et comportementales manifeste une grande cohérence entre les données portant sur l'intégration des contours. Toutefois, les auteurs des études n'utilisent pas de signaux Gabor colinéaires représentant une forme. Cette partie a donc pour objectif de tester ce type de stimulus.

Ainsi, Kovacs (1996) suggère que le principe des connexions horizontales « long-range » ne serait pas une condition suffisante à elle seule pour expliquer l'intégration des contours d'une forme. En effet, des processus comme notre expérience passée influencent les connexions entre deux neurones. Kovacs et Julesz (1993) ont étudié cette question en reprenant le même paradigme que celui de Field et al. (1993), mais une condition dans laquelle la courbe était fermée, a été ajoutée (elle avait ainsi la forme d'un cercle). Les résultats ont montré une différence significative entre des contours

fermés ou ouverts : un pourcentage plus élevé de détection correcte était observé pour des cercles. Ce résultat démontre que la sensibilité augmente grâce au processus de fermeture, et non pas par une autre propriété du stimulus. Le principe de fermeture influencerait le champ d'association.

Ces résultats ont amené Kovac et Julesz (1993) à conclure que les éléments Gabor composant le cercle paraissaient avoir un contraste plus élevé, ce qui faciliterait leur perception. Hess, Dakin et Field (1998) ont testé cette hypothèse en réalisant une tâche où les observateurs ajustaient le contraste d'un signal Gabor avec celui des éléments présentés dans le fond. Le signal Gabor était placé parmi des éléments Gabor alignés ou sans aucun lien. Mais, quelle que soit la condition de présentation des éléments, les observateurs ne surestimaient pas le contraste du signal. Les résultats n'ont pas confirmé l'hypothèse selon laquelle le contraste des éléments serait perçu plus élevé lorsque ces derniers forment un contour par rapport à une condition où ils n'en forment pas. Toutefois, Hess et al. (1998) n'ont pas utilisé de condition dans laquelle les éléments étaient regroupés en un cercle, ce qui ne permet pas de comparer directement leur résultats avec l'hypothèse formulée par Kovac et Julesz (1993).

Par la suite, Pettet, McKee et Grzywacz (1998) ont déterminé spécifiquement les conditions pour lesquelles le mécanisme de fermeture facilitait la détection (Kovacs & Julesz, 1993). Pour l'ensemble des expériences, deux ensembles de stimuli, composés d'éléments Gabor, étaient présentés de chaque côté de l'écran. Les observateurs déterminaient le côté où le contour était situé. Cette expérience a permis de répliquer les résultats de Kovacs et Julesz (1993) en montrant qu'un contour fermé était plus facilement détectable qu'un contour aligné. Ceci restait vrai même si les performances diminuaient quand le nombre de distracteurs augmentait (les signaux Gabor composant le fond). Field et al. (1993) ont montré que plus un contour était arrondi et plus sa visibilité diminuait, par conséquent, des modifications au sein de la courbe d'un contour fermé pourrait influencer sa détection. Pour tester cette hypothèse, trois contours fermés étaient comparés : la première forme était un contour circulaire, la seconde forme se rapprochait d'un croissant de lune, la troisième forme était similaire à un haricot sec. Parmi ces trois formes, le contour circulaire était le plus facilement identifiable et la forme du croissant de lune était le plus difficile à identifier, la forme des contours en haricot ayant une visibilité intermédiaire. En outre, les performances obtenues dans les conditions « croissant de lune » et « haricot » étaient comparables à celles obtenus dans une condition contrôle où le contour n'était pas fermé. Le principe de fermeture facilite donc la visibilité des contours seulement lorsque certaines contraintes géométriques sont réunies. Ainsi, ce mécanisme ne suffit pas, à lui seul, à améliorer la détection des contours. Kovac et Julesz (1994) ont obtenu des résultats similaires à ceux de Pettet et al. (1998) lorsque des éléments Gabor alignés en cercle étaient comparés à une condition où une forme ovale était utilisée.

Les études de psychophysiques et neurophysiologiques ont permis de montrer que l'orientation, entre les éléments constituant un contour, était un facteur important. Si ce phénomène est vrai, alors la forme du contour devrait affecter les interactions entre les unités locales. En d'autres termes, Pettet (1999) avance l'idée que la forme globale influencerait la réponse de détection. Les stimuli étaient composés de fonctions Gabor qui étaient présentées à droite et à gauche de l'écran. Conformément aux données de Field

et al. (1993), l'orientation des Gabor était toujours inférieure à 30 deg. Parmi ces stimuli, certaines fonctions Gabor étaient regroupées pour former une courbe et étaient présentées à droite ou à gauche de l'écran. L'observateur indiquait la position de la courbe. Il a également manipulé la courbure du signal qui allait d'un demi-cercle à un serpent. Pettet (1999) a observé que plus le signal était courbé et plus il était facile à détecter. L'expérience de Pettet (1999) a mis en avant des différences de détection alors que l'orientation des éléments Gabor était inférieure à 30 deg. Pour Hess et Field (1999), ceci indique que la notion de champ d'association n'est pas suffisante pour expliquer le processus d'intégration des contours. Par ailleurs, les résultats de Kovacs et Julesz (1993) montrent la nécessité d'aller au-delà des connexions latérales « long-range » pour expliquer le traitement des contours.

Par conséquent, il est tout à fait possible que des mécanismes intermédiaires existent et permettent l'intégration des contours. Ainsi, Ullman (1984) suggérait l'existence de routines visuelles qui établissent des relations entre les premiers niveaux de traitement, qui ne sont pas réalisées explicitement, et des représentations de haut niveau. Les routines visuelles permettraient d'intégrer les informations locales en un tout, ce processus serait un intermédiaire entre les premiers niveaux de traitement et les traitements de haut niveau. Par exemple, ce mécanisme permettrait de relier les différents éléments locaux en un contour unique. Néanmoins, dans une étude neuropsychologique récente, Giersch, Humphreys, Boucart et Kovacs (2000) ont rapporté le cas d'un patient agnosique (HJA) dont la lésion est principalement située dans l'aire occipitale mais sans que l'aire V1 ne soit atteinte. La méthode était en tout point identique à celle utilisée dans Kovacs et Julesz (1993). Les résultats n'ont montré aucune différence entre la sensibilité du patient et des observateurs contrôle. Ce patient n'avait aucun déficit particulier dans le traitement des contours, ce qui indiquerait que l'aire V1 est suffisante pour intégrer les contours d'une forme.

De nombreuses études comportementales (Kovacs & Julesz, 1993, 1994 ; Pettet, 1999 ; Pettet et al. 1998) ont permis de montrer que les connexions horizontales n'étaient pas suffisantes à elles seules pour déterminer la perception des contours sur des grandes régions du champ visuel. Une étude récente de Angelucci, Levitt, Walton, Hupé, Bullier et Lund (2002) suggèrent une implication des processus « feedback » dans le traitement des contours dans V1 à partir du cortex extrastrié.

Boucart, Delord et Giersh (1994) ont montré que les effets de colinéarité étaient généralisables à des stimuli plus complexes que des éléments Gabor. Les stimuli étaient des dessins d'objet composés de segments de lignes. L'espacement et la colinéarité entre les segments étaient manipulés. L'expérience se déroulait de la façon suivante : une forme apparaissait au centre de l'écran, elle était suivie d'une cible (identique en tout point à la figure de référence) et d'un distracteur (l'objet est identique à la figure de référence, mais l'orientation est l'image en miroir). La tâche des observateurs était de préciser laquelle des deux formes était identique à la forme précédente. La figure de référence était constituée d'un dessin au contour complet ou segmenté. Cette manipulation supplémentaire permettait de vérifier si la présentation préalable d'une forme complète facilitait le traitement d'une forme incomplète. Les résultats indiquaient une diminution des performances lorsque l'espacement entre les segments était supérieur à 10 min et pour

des stimuli non colinéaires. De plus, les temps de réponse étaient plus longs et les taux d'erreur plus faibles lorsque la figure de référence était un dessin au contour complet. Ainsi, l'identification de la figure complète facilite le traitement, et l'accès à l'identité de l'objet ralentit le temps de réponse. Par ailleurs, Boucart, Grainger et Ferrand (1995) ont cherché à déterminer l'importance des caractéristiques de contour sur l'identification d'objet dans une tâche d'amorçage. L'amorce était composée de dessin avec des parties de contour supprimées, et la cible était toujours la version complète du dessin présentée en amorce. Un masque, composé de point noir et blanc aléatoire, était intercalé entre l'amorce et la cible. Boucart et al. (1995) ont tout d'abord réalisé une expérience dans laquelle le contour était constitué uniquement par les coins ou par les segments de ligne de l'objet. Les résultats indiquaient que la facilitation due à une pré-activation en mémoire, était équivalente pour les deux versions de stimuli. De plus, lorsqu'aucune amorce n'était présentée, le temps de dénomination de l'objet est plus long par rapport à une condition avec une amorce de dessin segmenté. L'identification est facilitée par le degré de structuration locale. Dans une dernière expérience, les auteurs ont remplacé les coins bi-dimensionnels par des coins tri-dimensionnels. Ces stimuli étaient comparés aux dessins constitués par des segments de lignes. Dans ce cas, la dénomination d'objet était facilitée pour des contours composés de coins tri-dimensionnels. Il semble donc que le facteur 3D transmet plus d'information que des segments de lignes sur l'arrangement spatial entre les différentes parties de l'objet.

1.2.4.3. Un aspect méconnu dans l'intégration des contours : le facteur temporel

Les études, qui ont été présentées dans les parties précédentes, intégraient uniquement des paramètres qui limitaient la sensibilité à des images statiques. Toutefois, des études récentes ont examiné le rôle des aspects temporels dans l'intégration des contours.

D'un point de vue comportemental, l'aspect temporel de l'intégration des contours a été étudié récemment par Usher et Donnelly (1998). Dans leur expérience, une tâche de détection était réalisée dans laquelle des segments de lignes colinéaires étaient affichés brièvement de façon synchrone ou asynchrone par rapport à des segments du fond. Dans la condition synchrone, les éléments du contour et du fond étaient présentés simultanément. Tandis que dans la condition asynchrone, les éléments du contour étaient présentés avant les éléments du fond. Les observateurs avaient des performances supérieures dans la condition asynchrone que dans la condition synchrone. Une seconde condition asynchrone a été testée pour montrer que les effets obtenus n'étaient pas la conséquence d'un traitement implicite. Dans cette condition, l'orientation des segments était aléatoire, et les résultats ont indiqué une baisse significative des performances. Cette étude montre que l'information temporelle interagit avec le traitement spatial et que, par conséquent, ces deux informations ne peuvent donc pas être utilisées indépendamment l'une de l'autre. Beaudot (2002) a répliqué cette expérience en utilisant des signaux Gabor à la place de segments de lignes. Il a observé que l'effet obtenu par Usher et Donnelly (1998) dépendait de l'ordre de présentation entre le signal et le fond. En effet, un résultat identique était observé uniquement lorsque les signaux Gabor du contour précédaient les éléments Gabor du fond. Pour Beaudot (2002), l'utilisation de stimuli

simples, comme des barres lumineuses, ne serait pas une condition suffisante pour déterminer le codage neuronal impliqué dans l'intégration des contours. Les expériences de Usher et Donnelly (1998) et de Beaudot (2002) indiquent qu'une présentation synchrone ou asynchrone n'est pas suffisante pour étudier l'aspect temporel dans l'intégration des contours (Farid, 2002).

Néanmoins, d'autres auteurs se sont intéressés aux traitements temporels en dehors du principe d'asynchronie. Bex, Simmers et Dakin (2001) ont étudié les aspects temporels dans l'intégration des contours en utilisant des signaux Gabor et en manipulant la fréquence temporelle (de 2 à 10 Hz) pour deux types différents de contours : la phase des éléments Gabor était soit dans le même alignement que le contour, soit opposée (comme les barreaux d'une échelle). La tâche des observateurs était de détecter la présence ou non du contour. Les résultats ont indiqué que les éléments correctement alignés étaient plus facilement détectables que des éléments dont la phase était orthogonale par rapport à l'alignement. Quelle que soit l'orientation des éléments Gabor, le seuil de détection était plus élevé à partir du moment où la fréquence temporelle n'était plus statique, puis le seuil des observateurs diminuait à partir de 2 Hz, mais il était toujours supérieur à celui de la condition contrôle où la fréquence temporelle ne variait pas.

Hess, Beaudot et Mullen (2001) ont manipulé différents facteurs comme les fréquences temporelles, l'orientation entre les éléments Gabor du contour ou encore le contraste des éléments formant le contour. La méthode utilisée était en tout point identique à celle de Field, Hayes et Hess (1993). Les résultats étaient similaires à ceux de Field et al. (1993) : plus l'orientation entre les éléments était importante, et moins les performances de détection étaient élevées. Mais ici, Hess et al. (2001) ont montré que les performances diminuaient avec l'augmentation de la fréquence temporelle. Une distinction était observée puisque le pourcentage de détection correcte diminuait entre 6 et 12 Hz lorsque les éléments du contour étaient alignés, alors que la diminution des performances s'observait dès 1-2 Hz pour des contours arrondis. Par ailleurs, le contraste du contour n'avait peu ou pas d'influence sur les performances des observateurs.

Des critiques ont été formulées à propos du rôle joué par l'asynchronie dans le traitement des contours. Toutefois, ces critiques ne remettent pas en question l'importance des fréquences temporelles dans l'intégration des contours. Cet aspect mériterait d'être pris en considération dans les modèles du traitement des contours (Grossberg & Mingolla, 1985 ; Yen & Finkel, 1996).

Le codage physiologique d'une information visuelle est déterminé par l'antagonisme centre/pourtour des champs récepteurs de la rétine jusque dans le cortex visuel. Ce modèle de l'intégration d'un stimulus par les champs récepteurs souligne le fait que la transmission de l'information est réalisée à partir des bords et non pas à partir d'une surface uniforme. Malgré les critiques qui ont été formulées aux modèles de Hubel et Wiesel (1962, 1968), la sensibilité des champs récepteurs à un contraste n'a jamais été remise en question. Au cours de cette partie, nous nous sommes intéressés à l'intégration des contours d'un point de vue physiologique mais aussi comportemental sur des petites et des grandes parties du champ visuel. Nous avons vu qu'il existait de nombreuses similarités entre les études psychophysiques et physiologiques concernant le traitement des contours. La cohérence de ces études indique que les bords constituent une

information primordiale pour le traitement d'une forme. La question que nous aborderons dans la prochaine partie est de savoir comment l'information de surface est intégrée et comment cette dimension est traitée de façon perceptive, puisque c'est une dimension, selon ces théories, qui n'est pas codée directement par le système visuel.

2. Le traitement de la surface : une reconstruction ?

Nous venons de présenter un ensemble d'argument physiologique et psychophysique qui montraient que la première analyse visuelle consiste en une décomposition de l'image en contrastes locaux. La question est donc de savoir si la dimension de surface est perdue ou non. La plupart des modèles de reconnaissance visuelle proposent que l'information de surface soit reconstruite dans une deuxième étape à partir de l'intégration des contours (von der Heydt, et al., 1984 ; Grossberg & Mingolla, 1985 ; Marr, 1982 ; pour une revue voir Palmer, 1999). Notre système visuel code une information basée sur les bords d'une forme, mais notre perception nécessite l'intégration de la dimension de surface. Notre objectif sera de déterminer comment notre système perceptif, d'un point de vue physiologique et comportemental, traite l'information de surface.

La première partie de ce chapitre indiquera pourquoi de nombreux auteurs ont rapidement considéré que la dimension de surface était purement secondaire, et qu'elle résultait d'un processus dit de remplissage. Selon cette théorie, les contours seraient extraits dans un premier temps, puis la surface contenue par les contours serait remplie. Nous présenterons dans une seconde partie des éléments physiologiques qui sous-tendent la notion de remplissage pour des phénomènes comme la tache aveugle. Nous présenterons également des études qui enregistrent directement l'activité des neurones lorsque leur champ récepteur est positionné sur une surface uniforme. La question qui sera ainsi abordée est de déterminer si le remplissage peut expliquer le codage de la surface. En d'autres termes, l'objectif de cette partie physiologique est de déterminer s'il existe un remplissage à partir des contours et s'il existe également des cellules qui répondent spécifiquement à la luminance au centre d'une forme.

Par la suite, nous aborderons, dans une troisième partie, les études comportementales qui ont mis en évidence l'existence du remplissage pour des phénomènes comme la tache aveugle, les scotomes ou encore les scotomes artificiels. La quatrième partie traitera de deux principaux modèles de traitement hiérarchique du contour et de la surface. Puis, la cinquième partie présentera un ensemble d'études comportementales en faveur d'un traitement hiérarchique. Enfin, nous terminerons ce chapitre en décrivant les processus attentionnels impliqués dans le traitement de la surface.

2.1. La surface : un processus de remplissage ?

De nombreux auteurs (Gerrits & Vendrik, 1970 ; Grossberg & Mingolla, 1985 ; Spillman, 1997 ; Walls, 1954) ont considéré que la surface était traitée par l'intermédiaire du

processus de remplissage. En d'autres termes, la surface n'est pas encodée directement mais elle est reconstruite à partir des contours. L'objectif de cette partie est de montrer ce qui a amené de nombreux auteurs à considérer le traitement de la surface comme un processus de remplissage.

Depuis plusieurs siècles (Mariotte, 1668 ; cité par Wade, 1998), nous connaissons l'existence d'une curiosité dans la conception anatomique de l'œil. Cette curiosité concerne une petite zone de la rétine sur laquelle les axones des cellules ganglionnaires convergent pour former le nerf optique. Cette région de la rétine est appelée la papille, et la zone du champ visuel correspondante est connue sous le nom de tache aveugle, appelée aussi disque optique. Comme cette région ne comporte pas de photorécepteur, elle est insensible. La Figure 12 à gauche indique l'emplacement de la tache aveugle dans la rétine (situé à 13-15 deg de la fovéa). De plus, la Figure 12 à droite montre que cette région est plus large que la fovéa. De façon plus précise, la longueur de la tache aveugle est approximativement de 6 deg pour une largeur de 4,5 deg (Churchland & Ramachandran, 1996).

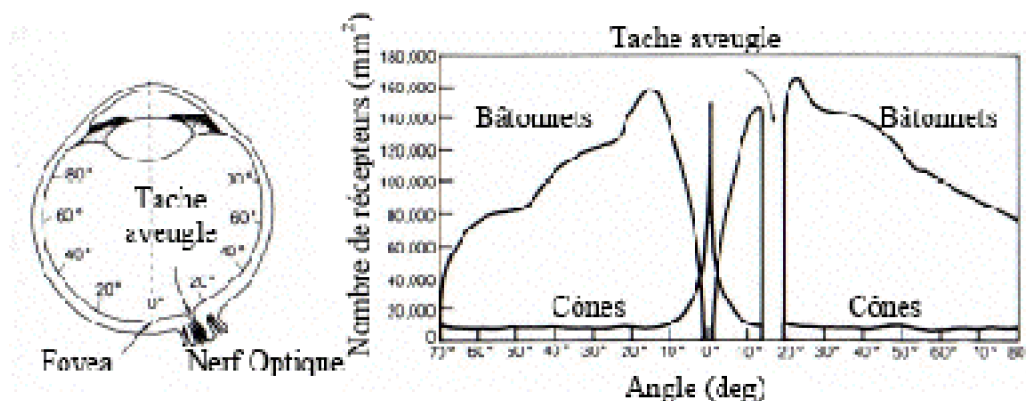


Figure 12 : L'œil humain est représenté sur la gauche. La tache aveugle est la région de la rétine où les cellules ganglionnaires se regroupent pour se projeter ensuite sur le CGL. La figure de droite représente la densité des cônes qui est plus élevée dans la fovéa et qui diminue brutalement dans le champ visuel périphérique. À l'inverse, la densité des bâtonnets est plus importante dans la région qui entoure la fovéa puis diminue graduellement dans le champ visuel périphérique. Nous pouvons noter sur ces deux figures que la tache aveugle se situe dans le champ visuel périphérique et que sa taille est plus grande que la fovéa (tiré de Palmer, 1999, p. 36).

Cependant, nous n'avons pas conscience de la présence d'une tache « noire » (ou vide) dans notre champ visuel. Il existe deux raisons pour expliquer ce phénomène. La première explication repose sur le fait que les récepteurs d'un œil enregistrent les informations manquantes du deuxième œil qui sont situées dans la région de la tache aveugle. Cette explication n'est valable que pour une condition binoculaire. En effet, la région corticale qui correspond à la tache aveugle de l'œil gauche controlatéral correspond à l'aire de la rétine de l'œil droit ipsilatéral. La Figure 13 permet de suivre le cheminement d'un stimulus qui serait positionné dans la région de la tache aveugle.

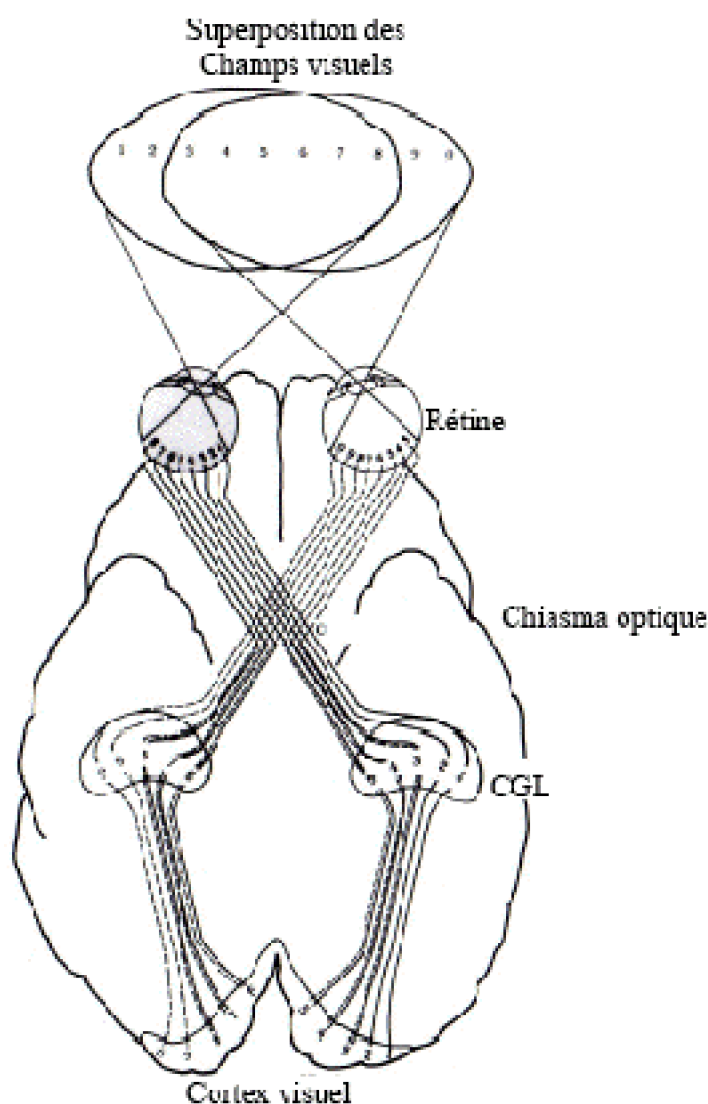


Figure 13 : Le schéma illustre les projections de la rétine jusqu'au cortex visuel, en précisant quelle partie du champ visuel est représenté dans ces différentes aires. Par exemple, l'hémichamp gauche se projette dans la partie controlatérale droite du CGL et se projette ensuite dans la partie droite du cortex visuel. La tache aveugle de l'œil gauche correspond approximativement à la région 3, où il y a une superposition des champs visuels. En suivant la voie de la région 3 depuis le champ visuel jusqu'au cortex, nous pouvons suivre le chemin d'un stimulus qui serait placé dans la région de la tache aveugle de l'œil gauche (tiré de Churchland et Ramachandran, 1996, p. 150).

La seconde raison est que notre système perceptif remplit cette zone du champ visuel avec les qualités appropriées. Par exemple, si la zone du champ visuel correspondant à la tache aveugle est entourée par une surface uniforme, elle est « remplie » par l'information environnante. Le remplissage suppose qu'il y a une propagation de l'information à partir de l'environnement jusqu'au centre de la zone de la tache aveugle. Il doit donc exister une différence temporelle dans le traitement de l'information entre la région située sur le bord, qui est réalisé en premier, et la région située

au centre, qui est effectué en dernier, de la tache aveugle. Des hypothèses de nature physiologique ont été émises sur le fonctionnement du remplissage, et seront exposées dans une partie ultérieure. Ce phénomène a également été mis en évidence dans la littérature clinique, en cas de scotome⁹. Les conséquences de ce traumatisme sont proches de la tache aveugle du nerf optique puisqu'une ou plusieurs parties du champ visuel sont aveugles. Toutefois, les patients ont au départ la sensation qu'une tache noire est présente dans leur champ visuel mais cette impression disparaît très rapidement (Palmer, 1999). Ce traumatisme est le résultat d'une lésion de la rétine ou du cortex visuel.

Pour Walls (1954), un certain nombre de phénomènes perceptifs, comme la tache aveugle, les scotomes ou encore l'effet Troxler¹⁰, seraient issus d'un même et unique processus : le remplissage. Selon plusieurs auteurs (Gerrits & Vendrik, 1970 ; Walls, 1954), ce processus ferait partie intégrante du fonctionnement normal du système visuel.

Un point important est que le processus de remplissage permet d'avoir une expérience conforme du monde visuel alors qu'il n'y a aucune activité neuronale provenant de la rétine ou de la zone endommagée. Il est classiquement admis que l'activité des cellules est faible lorsque les champs récepteurs sont positionnés sur une surface uniforme. Par conséquent, il existe une similitude entre la dimension de surface, qui n'est codée par la rétine, et le processus de remplissage qui permet de combler l'information manquante. Ce phénomène a donc été considéré comme une explication possible de l'intégration par le système visuel de la dimension de surface (Durgin, Tripathy & Levi, 1995 ; Grossberg & Mingolla, 1985). Plusieurs auteurs (Gerrits & Vendrik, 1970 ; Palmer, 1999) ont émis l'hypothèse que ce phénomène serait la conséquence d'un traitement neuronal de haut niveau.

2.2. Etudes neurophysiologiques du traitement de la surface

2.2.1. Le remplissage : un processus neuronal ou cognitif ?

Nous avons vu précédemment que le processus de remplissage se présentait comme une possibilité pour expliquer le traitement de la surface. Une question a été de déterminer la nature passive ou active du processus de remplissage. En effet, le débat portait sur l'implication directe d'un traitement neuronal réalisé par le système visuel. Les auteurs en faveur d'un remplissage passif (Dennett, 1991) supposent que la tache aveugle ou le scotome sont simplement ignorés de la même façon que les régions en dehors du champ visuel (située derrière la tête, par exemple). À l'inverse, de nombreux auteurs (Churchland & Ramachandran, 1996 ; DeWeerd & Spillman, 2003 ; DeWeerd, Gattass, Desimone & Ungerleider, 1995 ; Gilbert, 1996 ; Spillman & Werner, 1996)

⁹ Quand une lésion détruit les cellules répondant à une région particulière de l'espace dans la rétine ou le cortex visuel, cette région devient aveugle. On dit que la patient présente un scotome (i.e. cette région aveugle).

¹⁰ L'effet Troxler est le suivant : le regard de l'observateur doit être fixe puisqu'on lui présente un petit stimulus dans la périphérie du champ visuel. Après quel que instants, le stimulus disparaît et est remplacé par ce qui l'entoure.

supposent l'existence d'un remplissage actif, pour lequel, une représentation est construite à partir d'une stimulation visuelle entourant la région de la tache aveugle ou du scotome. L'une des questions est donc de déterminer la nature du processus de remplissage. Les expériences qui seront présentées dans cette partie suggèrent contrairement à Dennett (1991) l'existence d'un remplissage neuronal. Ces arguments sont principalement basés sur des données empiriques et non pas purement théoriques.

L'explication neuronale du processus de remplissage est basée sur l'existence des connexions horizontales « long-range » entre les neurones (Gilbert, 1996). Ces connexions ont été mises en évidence chez le chat (Lund, Yoshioka & Levitt, 1993) et le singe (Gilbert & Wiesel, 1992). L'étendue des axones pour ce type de connexion est importante et peut aller jusqu'à 6-8 mm sur le cortex (Gilbert, Das, Ito, Kapadia & Westheimer, 1996). Dans le premier chapitre, nous avons vu que la distance entre deux hypercolonnes est de 1.5 mm, ce qui représente la distance corticale pour que deux champs récepteurs ne soient pas superposés. Par conséquent, la distance couverte par les connexions horizontales permet de relier des cellules qui sont séparées par un intervalle équivalent à la taille de plusieurs champs récepteurs. Cet état physiologique suggère que ces connexions peuvent transmettre une information depuis les contours jusqu'au centre d'une figure (DeWeerd & Spillman, 2003).

Les connexions horizontales supposent que lorsque des neurones d'une même couche sont activés, ces derniers peuvent activer ou inhiber les neurones proches spatialement. Pour illustrer cette idée, prenons l'exemple de trois neurones corticaux : A, B et C. A la suite d'un scotome, le neurone B ne reçoit plus d'afférence des couches inférieures, mais continue à recevoir celles des neurones A et C, qui sont eux activés par les couches inférieures. On peut appliquer ce raisonnement au remplissage : alors que B ne reçoit aucune projection, ce neurone transmettra la même information que A et C, et par conséquent, l'observateur n'aura aucune lacune (Spillman & Werner, 1996). La Figure 14 illustre cette notion.

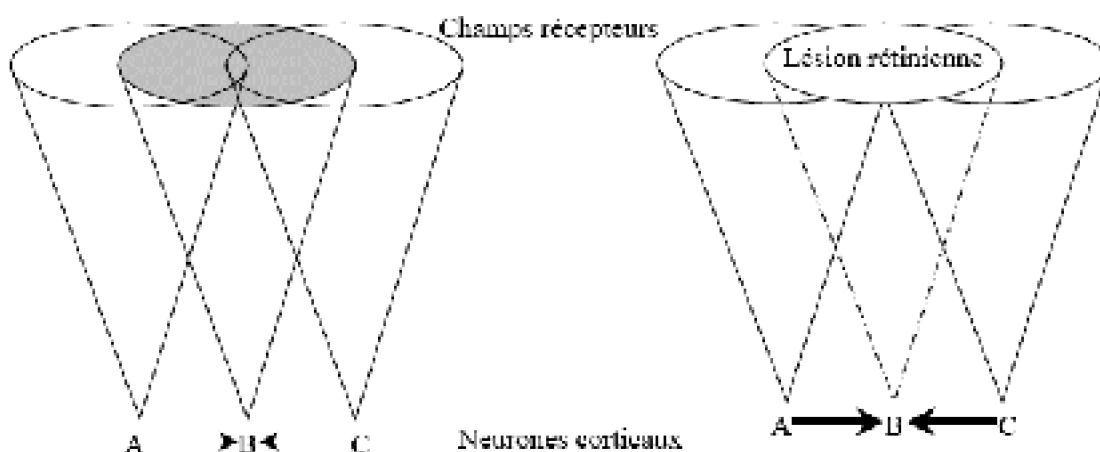


Figure 14 : Les ellipses représentent les champs récepteurs des cellules corticales A, B et C. Les flèches illustrent les connexions horizontales depuis les cellules A et C en direction de la cellule B. La figure de gauche présente une condition normale d'illumination pour laquelle les connexions horizontales sont relativement faibles. La figure de droite montre

une condition pour laquelle le neurone B ne reçoit plus d'activation à la suite d'une lésion de la rétine. Le neurone B est toutefois réactivé par l'intermédiaire des connexions horizontales A et C.

Les différentes parties, qui seront présentées par la suite, permettront de vérifier l'implication des connexions « long-range » pour le remplissage, mais également de déterminer son niveau de traitement (précoce ou tardif). Nous nous interrogerons ensuite sur les aspects physiologiques de ce processus pour le traitement de la surface.

2.2.2. Le remplissage est-il un processus précoce ou tardif ?

2.2.2.1. Arguments expérimentaux d'un processus tardif de remplissage

Les études princeps en faveur d'un traitement tardif du remplissage utilisaient un « scotome artificiel » : un carré de luminance homogène, excentré par rapport à un point de fixation présenté au préalable, apparaissait d'abord dans une texture donnée (une disposition aléatoire de petit carré blanc sur fond noir), et, après un bref délai, le carré était rempli phénoménalement par la texture environnante. Les résultats comportementaux ont montré qu'après ce remplissage, le carré n'était plus perçu, bien que sa luminance soit toujours différente du fond. Ce phénomène a été observé chez l'homme (Ramachandran & Gregory, 1991) et le singe (DeWeerd, Gattass, Desimone & Ungerleider, 1995). Les singes ne signalaient pas comportementalement le remplissage, seule la réponse des neurones était enregistrée. DeWeerd et al. (1995) ont enregistré dans différentes aires corticales : V1, V2 et V3 l'activité des neurones ayant un champ récepteur dans la zone du carré qui devait disparaître. Pour déterminer le moment où le processus de remplissage se produisait, une expérience contrôle a été effectuée dans laquelle aucun carré n'était présenté dans la texture. Les résultats ont indiqué une augmentation de l'activité des neurones, par rapport à la condition contrôle, après une durée de présentation de 6.2 à 9.8 secondes. Pour DeWeerd et al. (1995), cette augmentation de l'activité, au cours de cette période, correspondait au processus de remplissage. En effet, les auteurs ont effectué une expérience chez des observateurs humains avec des stimuli identiques et la disparition du carré se produisait avec la même latence. Les résultats de l'enregistrement des neurones pour différentes couches corticales, ont montré que la présentation des stimuli s'accompagnait d'une faible activité neuronale dans l'aire visuelle primaire (V1), mais d'une activation plus élevée dans les régions ultérieures (V2 et surtout V3). Par ailleurs, les données observées dans V1 étaient d'autant plus restreintes que seul quatre neurones sur vingt étaient activés pour une excentricité de 24 deg et aucun pour une excentricité de 4.5 deg, à partir d'un point de fixation. Cette étude physiologique situe le traitement du remplissage à un niveau élevé dans le traitement cortical d'un stimulus.

2.2.2.2. L'exemple de la tache aveugle : un argument en faveur d'un traitement précoce

L'expérience de DeWeerd et al. (1995) a montré que le processus de remplissage n'était pas prévalent dans l'aire V1, mais nécessitait une étape dans les aires V2 et surtout V3.

Les deux premières parties tenteront au contraire de montrer la précocité du processus de remplissage pour la tache aveugle et les scotomes.

Fiorani, Rosa, Gattass et Rocha-Miranda (1992) ont été les premiers à poser la question suivante : quelle est la réponse des cellules de V1 correspondant à la tache aveugle de l'œil droit lorsque l'œil gauche est fermé et qu'un stimulus est présenté à l'œil droit ? Pour cela, ils ont utilisé une stimulation visuelle dont la taille était supérieure à la région de la tache aveugle. Le stimulus était composé d'une barre noire qui traversait la tache aveugle, ce qui correspondait à un processus de remplissage plus naturel (puisque l'expérience est réalisée dans la tache aveugle). Les auteurs ont enregistré la réponse des neurones de l'aire V1 à cette stimulation, dont les champs récepteurs étaient situés à l'intérieur de la tache aveugle (la région est souvent appelée « optic disk cortex »). Lorsque les conditions expérimentales impliquaient un remplissage (i.e. les deux extrémités de la barre étaient situées hors de la tache, mais son milieu tombait à l'intérieur), l'activité des neurones dans l'aire V1 était élevée. En revanche, lorsque la taille de la barre était diminuée de moitié, de manière à ce que l'une des extrémités aboutisse dans la tache aveugle (et qu'aucun remplissage ne soit nécessaire), une réponse plus faible était enregistrée. De plus, lorsqu'un masque uniforme était placé à l'intérieur de la tache aveugle, Fiorani et al. (1992) ont obtenu des réponses identiques à la condition où aucun masque n'était présenté, ce qui indique qu'aucune information n'était transmise par le masque. Par ailleurs, l'utilisation d'un masque plus important que la taille de la tache aveugle empêchait la stimulation des neurones de V1. L'implication de l'aire visuelle primaire témoigne de la précocité du processus de remplissage. Pour les auteurs, les résultats obtenus mettent en jeu des mécanismes différents de ceux trouvés par Peterhans et von der Heydt (1989). Ces derniers ont montré une réponse des neurones de l'aire V2, alors que la réponse est présente dès V1 chez Fiorani et al. (1992). La différence existant entre ces deux études est que les contours illusoire sont perçus avec les mouvements oculaires tandis que le phénomène de complétion nécessite une fixation. Tootell, Hadjikhani, Vanduffel, Liu, Mendola, Sereno et Dale (1998) ont complété cette étude chez l'observateur humain avec une expérience en IRMf. Un signal était présenté dans la tache aveugle, de façon monoculaire. Les résultats ont montré une activation de la couche 4 de V1. Par contre, aucune activation n'était constatée au-delà de V1, et aussi lorsque le signal était présenté de façon binoculaire.

Komatsu, Murakami et Kinoshita (1996) ont utilisé une surface uniforme (un carré) dans la zone du champ visuel correspondant à la tache aveugle. Les auteurs ont examiné l'activité de 85 neurones dans V1 susceptible de répondre à cette zone. Les résultats de l'enregistrement ont indiqué que quarante d'entre eux étaient activés par un signal plus large que la tache aveugle. Ceci indique que même s'il n'y a pas d'input dans la tache aveugle, certains neurones sont actifs lorsque la taille d'un signal est plus grande que la tache aveugle. Cette dernière expérience démontre l'existence du processus de remplissage pour le traitement des surfaces recouvrant la tache aveugle. Par la suite, Komatsu, Kinoshita et Murakami (2000) ont répliqué cette expérience en montrant que l'activité neuronale (24 % des neurones enregistrés) était beaucoup plus importante dans la couche 6 de V1 que dans les autres couches. Un résultat identique était observé lorsque la surface uniforme était présentée autour de la tache aveugle ou en dehors de

cette zone. Les projections de la couche 6 de V1 vont essentiellement vers le CGL (Fitzpatrick, Usrey, Schofield & Einstein, 1994). Les auteurs émettent l'hypothèse d'un processus de « feedback » vers le CGL qui modulerait les réponses visuelles en augmentant l'activité des neurones de V1 à une surface (Silito, Jones, Gerstein & West, 1994).

Les études présentées dans cette partie ont montré la précocité du processus de remplissage dans un cadre plus naturel, que celui des scotomes artificiels, utilisé par DeWeerd et al. (1995).

2.2.2.3. L'exemple des scotomes : un second argument en faveur de la précocité du remplissage

Pettet et Gilbert (1992) se sont intéressés à la réorganisation des champs récepteurs à la suite d'un scotome induit par un impact au laser sur l'aire V1 du chat. Pettet et Gilbert (1992) ont enregistré la réponse des neurones avant et après l'impact au laser. Leurs résultats ont indiqué une expansion de la taille des champs récepteurs à la suite du scotome. Les auteurs suggéraient que l'augmentation de la taille des champs récepteurs pouvait expliquer le phénomène du remplissage des scotomes. Une réorganisation des champs récepteurs de V1 a également été démontrée à la suite d'une lésion de la rétine (Chino, Kaas, Smith III, Langston & Cheng, 1992). Ces résultats ont néanmoins été controversés par une étude dans laquelle cet effet n'était pas répliqué. Par contre, l'activité des neurones augmentait considérablement autour de la lésion (DeAngelis, Anzai, Ohzawa & Freeman, 1995). Eysel et Schweigart (1999) ont récemment complété ces études. Les auteurs ont induit un scotome par une lésion corticale chez le chat. Les auteurs estimaient ensuite la taille des champs récepteurs corticaux situés au bord de la lésion, 2 jours puis 2 mois après l'opération. Les résultats ont indiqué que la taille des champs récepteurs des cellules, situés au bord de la lésion, ne variait pas immédiatement après la lésion mais augmentait 2 mois après l'opération. Pour les auteurs, ceci permettrait de réduire l'étendue de la taille du scotome, puisque l'augmentation de la taille des champs récepteurs permettrait de diminuer le temps de transmission de l'information pour les connexions horizontales. Murakami, Komatsu et Kinoshita (1997) ont étendu ces recherches chez le singe pour lesquels des résultats similaires ont été montrés. Pour Gilbert (1995), l'expansion des champs récepteurs pourrait expliquer le processus de remplissage dans le cadre des scotomes artificiels (DeWeerd, Gattass, Desimone & Ungerleider, 1995).

Les études, qui ont utilisé un signal positionné autour de la tache aveugle ou des scotomes, étendent à une structure visuelle encore plus précoce les résultats de De Weerd, Gattass, Desimone et Ungerleider (1995) : ces derniers ont montré une activation des neurones dans les aires V2 et V3 (mais pas dans V1) pour le traitement du remplissage lorsqu'un scotome artificiel était utilisé.

2.2.3. Le remplissage permet-il d'expliquer le traitement de la surface ?

Le remplissage est un processus qui permet de combler une information manquante dans le champ visuel, et certain de ces phénomènes ont été décrits dans la partie précédente.

La plupart des théories ont donc supposé que le traitement de la surface pouvait être expliqué par ce processus. Dans cette partie, deux questions seront abordées : 1) quelle est la réponse des neurones du cortex visuel lorsque leurs champs récepteurs sont situés à l'intérieur d'une surface ? 2) est-ce que le remplissage est impliqué directement dans le traitement de la surface ? En effet, mettre en évidence un remplissage pour les scotomes ou la tache aveugle ne signifie pas que ce processus soit lié au traitement de la surface. Pour cela, nous nous intéresserons plus particulièrement à la réponse des neurones du cortex visuel pour une surface uniforme ou d'une texture. La réponse des neurones sera ainsi directement étudiée en dehors de tous phénomènes perceptifs impliquant un remplissage. Cette question sera traitée au cours des deux dernières parties. En effet, la première partie sera consacrée à une mise en évidence expérimentale du traitement hiérarchique entre le contour et la surface.

2.2.3.1. Mise en évidence d'un traitement hiérarchique contour/surface

Une recherche est venue compléter l'importance accordée au traitement hiérarchique du contour et de la surface (Macknik, Martinez-Conde, & Haglund, 2000). L'objectif de Macknik et al. (2000) était de déterminer la dimension d'un stimulus qui engendrait la plus forte réponse neuronale et perceptive. Pour cela, les auteurs ont combiné des méthodes psychophysiques (un paradigme de masquage) et électrophysiologiques (l'image optique). Cette dernière technique permet de mesurer l'activité engendrée par un signal à la surface du cortex visuel. L'une des expériences est une tâche de discrimination dans laquelle deux barres de longueurs différentes étaient présentées à droite et à gauche de l'écran. L'observateur devait indiquer quelle barre était la plus grande. Un masque rectangulaire dont la largeur variait (de 0 à 4 deg) était superposé aux deux cibles et les frontières des masques étaient éloignées ou proches des cibles. Les résultats ont montré que plus les frontières des masques étaient proches des cibles, plus la tâche de discrimination était difficile à réaliser. Il semble donc que les contours des masques transmettent une plus grande inhibition que leur surface. Lorsque Macknik et al. (2000) ne présentaient qu'une seule barre étroite lumineuse à des singes, les résultats de l'image optique ont montré qu'une seule rayure émergeait dans l'aire V1 (cf. Figure 15a). Par contre, lorsque la barre s'élargissait, l'image optique se séparait en deux rayons d'activité correspondant aux contours du stimulus (cf. Figure 15b). Ces résultats qui mettent en évidence l'importance des contours du signal visuel dans la perception, suggèrent donc que le traitement de la surface est la conséquence d'un processus plus tardif. La critique qui peut être formulée à cette expérience est que les auteurs n'utilisent pas de surface uniforme dans leur expérience, et nous verrons dans la prochaine partie que cela entraîne des résultats assez différents.

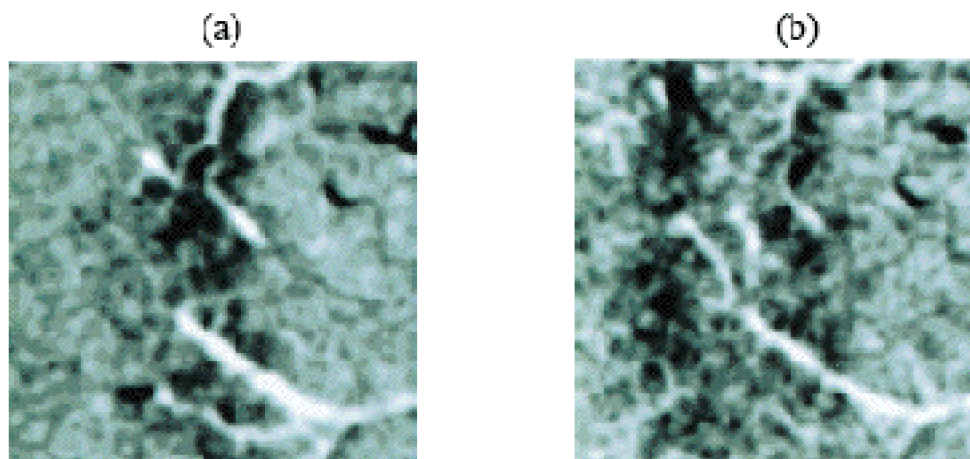


Figure 15 : Résultats obtenus par la technique de l'image optique dans l'expérience de Macknik, Martinez-Conde et Haglund (2000). La figure de gauche montre l'activité de l'aire V1 lorsqu'une seule barre étroite est présentée, et la figure de droite quand la barre s'élargissait.

2.2.3.2. Quelle est la réponse des neurones du cortex visuel à une surface uniforme ?

Les expériences précédentes ont permis de montrer des activations des neurones de V1 lorsqu'un signal est présenté dans une région du champ visuel recouvrant la tache aveugle (Fiorani et al., 1992) ou les scotomes (Pettet & Gilbert, 1992). Ces résultats permettent de conforter l'hypothèse de Spillman et Werner (1996) selon laquelle les connexions horizontales seraient fortement impliquées dans le processus de remplissage. Au début de ce chapitre, nous avons vu que de nombreux auteurs (Durgin, Tripathy & Levi, 1995 ; Gerrits & Vendrik, 1970 ; Grossberg & Mingolla, 1985 ; Walls, 1954) émettaient l'hypothèse selon laquelle ce processus expliquerait le codage de la surface. L'objectif est ici de déterminer si le remplissage participe directement au traitement de la surface.

La luminosité ¹¹ d'une surface peut facilement être manipulée par rapport au niveau réel de luminance par une technique d'induction. Par exemple, lorsqu'un carré gris moyen est entouré d'un fond noir, le carré paraît plus clair qu'il ne l'est quand il est perçu de façon isolée, et inversement lorsqu'un fond plus clair est présenté. Un exemple est montré à la Figure 16.

¹¹ La luminosité est l'impression subjective du niveau lumineux perçu. Elle varie du sombre au clair et n'a pas d'unité spécifique. A l'inverse, la luminance correspond précisément à la puissance lumineuse émise par une source.

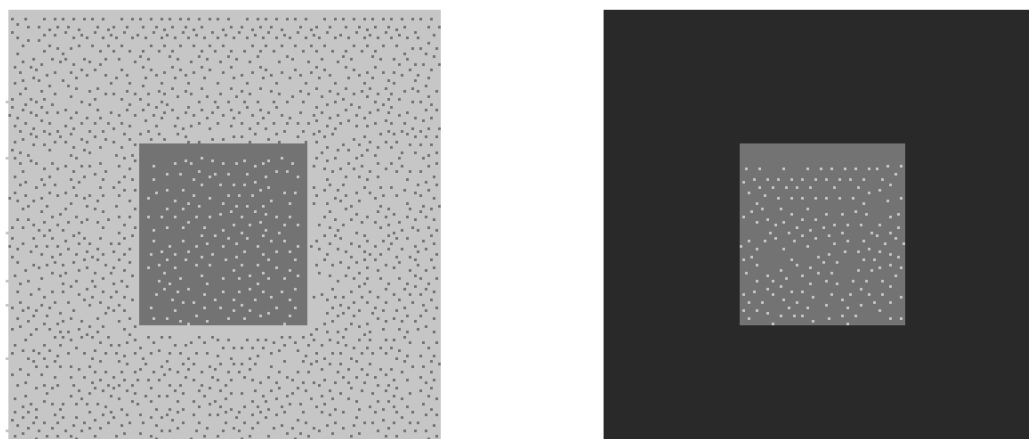


Figure 16 : Bien que les deux carrés du centre ont une luminance identique, la luminosité est différente. Ce phénomène est communément appelé : *contraste simultané*.

Rossi, Rittenhouse et Paradiso (1996) ont utilisé cette technique d'induction mais avec des stimuli différents. Dans leur condition « induction », une surface centrale était entourée par deux autres régions (à droite et à gauche) dont la luminance variait dans le temps du clair au sombre. Dans une autre condition, les deux surfaces étaient statiques avec un niveau gris moyen, mais alors, c'est la luminance de l'aire centrale qui variait du clair au sombre. Une condition contrôle était utilisée dans laquelle aucun autre stimulus n'était présenté, mais la luminance des deux surfaces adjacentes variait du noir jusqu'à un niveau de gris moyen. Les auteurs enregistraient chez le chat les neurones de l'aire visuelle primaire qui avaient leurs champs récepteurs situés dans la surface centrale. Les résultats ont indiqué que dans la condition d'induction, 75 % des neurones (120 sur 160) étaient activés par les variations de luminance du pourtour, et ce malgré le fait que les variations de luminance se produisaient en dehors du champ récepteur enregistré. La variation de luminance de la région centrale produisait aussi une réponse, mais qui était plus importante lorsque la luminance était claire. Dans la condition où l'aire centrale était absente, la décharge des neurones était équivalente à une décharge spontanée. Cette recherche montre un résultat très important pour la question du traitement de la surface, à savoir que la luminosité est traitée à partir de V1. Mais ici, les neurones n'agissent pas forcément comme des détecteurs de luminance puisque le pourtour peut être la dimension qui contribue à leur activité. Un résultat identique a été montré lorsque le champ récepteur des neurones enregistrés était situé dans un Mondrian¹² (MacEvoy & Paradiso, 2001). Ce résultat a été répliqué, par Kinoshita et Komatsu (2001), chez le singe avec une procédure simplifiée : la luminance du fond était fixe, et, seule la luminance du carré variait.

Rossi et Paradiso (1999) ont voulu déterminer si les réponses neuronales observées dans l'étude précédente, se produisaient uniquement dans le cortex ou si des réponses similaires étaient observées à une étape plus précoce du traitement visuel. Avec les

¹² Un Mondrian est composé de carré et de rectangle de différentes couleurs et superposé les uns aux autres. Le nom de ce stimulus provient du peintre hollandais qui a réalisé des tableaux de ce type.

mêmes stimuli que dans l'étude de Rossi et al. (1996), ils ont examiné la réponse de 75 neurones du CGL. Lorsque la luminance variait en dehors de l'aire centrale, la réponse des neurones n'était pas aussi élevée que dans le cortex, puisque 42 neurones (soit seulement 56 %) produisaient une activation significative. Par ailleurs, ils ont utilisé la même procédure expérimentale, mais cette fois pour enregistrer 33 cellules ganglionnaires dans la rétine grâce à un traceur optique. Mais, dans ce cas, aucune activité n'était constatée. Cette étude suggère donc que l'information de surface est intégrée à travers le champ visuel au niveau du cortex, voire dès le CGL, puisque la perception de la luminance est influencée par les variations de luminance des surfaces voisines, même éloignées.

Notre question est de déterminer l'implication du processus de remplissage dans le traitement de la surface. Pour cela, la taille du signal était manipulée. Selon le raisonnement suivant : la réponse des neurones dont le champ récepteur est situé au centre de la figure, devrait diminuer avec l'augmentation de la taille de la stimulation. En effet, augmenter la taille du signal revient à ralentir la propagation de l'information le long des connexions horizontales puisque la distance à parcourir augmente. Komatsu, Murakami et Kinoshita (1996) ont enregistré chez le macaque 29 neurones situés dans l'aire V1 pendant que plusieurs carrés de tailles différentes (plus petits ou plus larges que la taille des champs récepteurs, de 0.3 à 6.7 deg) étaient présentés. Les résultats ont montré un déclin de la réponse lorsque les stimuli devenaient de plus en plus larges. En d'autres termes, plus le contour s'éloigne, moins il y a de remplissage, sauf pour 1/3 des neurones pour lesquels l'activité restait constante quelle que soit la taille (ces neurones étaient situés à des profondeurs différentes dans le cortex). Par la suite, MacEvoy, Kim et Paradiso (1998) ont voulu déterminer la taille critique au-delà de laquelle l'information de surface n'influencerait plus les réponses des cellules dont le champ récepteur est situé à l'intérieur de la surface. Pour cela, ils ont utilisé un disque dont la taille augmentait, et le champ récepteur était positionné au centre. Les résultats ont indiqué qu'une majorité des neurones (72 %) étaient activés par la présentation d'un disque lumineux, même quand les contours de ce disque étaient placés en dehors du champ récepteur de la cellule. L'augmentation de la taille du disque avait pour effet principal de diminuer la réponse des neurones jusqu'à une taille critique (8 deg) pour laquelle les cellules n'étaient plus activées. Par ailleurs, les auteurs ont également constaté d'autres types de réponses neuronales. Dans certains cas, l'activité neuronale diminuait avec l'augmentation de la taille du disque, mais contrairement au résultat précédent, ces cellules restaient toujours actives. Un troisième résultat indiquait que l'activité des neurones augmentait avec la taille du disque, pour des tailles allant jusqu'à 30 deg. Par la suite, MacEvoy et al. (1998) ont remplacé le disque par un réseau sinusoïdal, dont l'utilisation est plus conventionnelle. Les résultats ont indiqué que seule la moitié des neurones étaient affectés par la variation de taille, mais pour ces neurones, les résultats obtenus étaient similaires à la condition où un disque uniforme était utilisé. L'hétérogénéité de ces résultats semble indiquer que chaque cellule n'intègre pas simultanément la surface. L'hypothèse du remplissage est ici controversée puisque malgré les grandes tailles du signal, certains neurones sont toujours activés (Komatsu, Murakami & Kinoshita, 1996 ; MacEvoy, Kim & Paradiso, 1998). Ces résultats montrent que le phénomène de remplissage n'est pas suffisant à lui seul pour rendre compte de la perception d'une surface. Ces résultats sont généralisables

à l'observateur humain, car la perception de la luminance est similaire à celle du macaque (e.g. Huang, MacEvoy & Paradiso, 2002).

Ces études (Rossi & Paradiso, 1999 ; Rossi, Rittenhouse & Paradiso, 1996) sont proches des effets contextuels, déjà cités dans le premier chapitre à la différence près que la cible et le contexte sont non plus des barres mais des surfaces uniformes. L'utilisation des effets contextuels est en effet la seule méthode capable de rendre compte du traitement de la surface, puisqu'en effet, une surface n'est jamais perçue de façon isolée. Ces expériences ont montré que les cellules de V1 sont activées par une surface uniforme.

Pour tester l'hypothèse du rôle du remplissage dans le codage de la surface, de nombreux auteurs ont manipulé la taille du signal (Komatsu, Murakami & Kinoshita, 1996 ; MacEvoy, Kim & Paradiso, 1998). Ainsi, lorsque l'on enregistre des cellules dont le champ récepteur est positionné au centre de la figure, et si la taille du signal augmente alors on pourrait s'attendre à ce que la réponse de ces cellules diminue. Toutefois, ce type de réponse n'est pas observé pour l'ensemble des neurones, en effet certaines cellules maintiennent une activité constante. Par conséquent, le processus de remplissage dans le traitement de la surface semble être fortement mis en jeu, mais les variations d'activité des cellules ne permettent pas d'affirmer que ce processus définit totalement le traitement de la surface (Gilbert, 1996).

Les expériences qui ont été développées dans les paragraphes précédents (Rossi & Paradiso, 1999 ; Rossi, Rittenhouse & Paradiso, 1996) repose sur un contraste entre deux surfaces. Ces études ne déterminent pas l'influence des contours dans la réponse de V1 à la luminosité d'une surface. Hung, Ramsden, Chen et Roe (2001) ont étudié l'implication des contours dans la perception de luminance. Les auteurs se sont intéressés à l'activité de la surface dans le cas de l'effet Craik-O'Brien-Cornsweet (appelé également effet COCE). Cet effet est considéré comme une illusion où la luminosité d'une région est affectée par la discontinuité de luminance au bord de cette région (un exemple est donné à la Figure 17). Dans ce cas, il n'y a pas de contraste entre les surfaces, car la différence de luminosité est interprétée à partir du changement produit par les bords. Hung et al. (2001) ont également utilisé deux autres types de stimuli : deux surfaces contiguës dont la luminance changeait réellement, et deux surfaces qui étaient équiluminantes, avec un profil de luminance aux bords qui augmentait ou diminuait (« Narrow Real Stimulus »). Ce dernier stimulus se distingue du COCE par le fait que les variations de luminance aux bords des deux surfaces ne se font pas graduellement, ce qui a pour effet d'inverser la luminosité par rapport au COCE. Les profils de luminance et de luminosité des stimuli sont présentés à la Figure 17. L'utilisation de ces trois types de stimuli permettrait aux auteurs de déterminer la contribution des contours dans la perception de luminance.

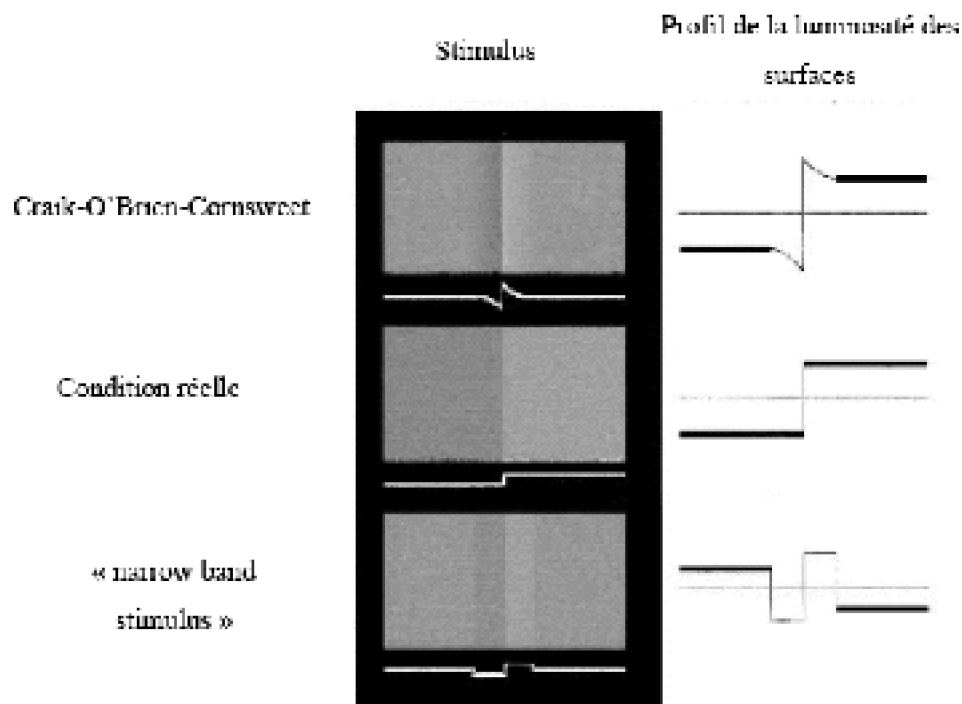


Figure 17 : La figure de gauche présente la luminosité des stimuli réels et illusoires, utilisés dans l'expérience de Hung, Ramsden, Chen et Roe (2001), avec leur profil de luminance en dessous. Sur la figure de droite, le profil de la luminosité des surfaces est représenté en gras.

Hung et al. (2001) ont utilisé la technique d'image optique, mais ils ont également effectué un enregistrement de la réponse des neurones dans les aires V1 et V2 du chat. Pour s'assurer que les neurones ne répondaient pas à un contraste, le contour, qui sépare les deux surfaces, était placé à une distance minimale de 1 deg du champ récepteur. Les résultats obtenus avec la méthode de l'image optique, ont indiqué une activité plus importante dans V2 que dans V1 pour les trois types de stimuli. De plus, une forte activité était notée pour les stimuli dont la luminance changeait réellement. Par contre, l'activité neuronale était plus faible que celle que soit l'aire corticale pour l'effet COCE et le « Narrow Real Stimulus ». Toutefois, ces résultats sont à relativiser puisque les enregistrements ont été réalisés chez le chat et non chez le primate. En effet, il existe des différences fondamentales entre le primate et le chat, notamment il existe des arguments pour supposer qu'il n'y a pas de voie parvocellulaire chez le chat (Rodieck, 1998).

Les études précédentes (Rossi & Paradiso, 1999 ; Rossi, Rittenhouse & Paradiso, 1996) ont montré que les neurones du cortex visuel primaire répondaient à un contraste de luminance entre deux surfaces. Mais, l'origine de cette réponse est incertaine : est ce que les neurones sont activés via les bords de la stimulation ou bien directement à partir des caractéristiques de la surface ? Cette réponse n'a pas trouvé de réponse totalement satisfaisante. Par contre, dans l'effet COCE, le contraste de surface est absent et le changement de luminosité est le résultat d'un contraste. Les résultats de Hung et al . (2001) semblent être plus robustes dans V2 que V1. Ainsi, il faudrait attendre V2 pour que

la luminosité d'une surface soit induite depuis un contour. De plus, les auteurs suggèrent que la perception de la luminosité est encodée par l'aire V1 pour un contraste entre deux surface, et par l'aire V2 pour un contraste induit à partir des bords. En d'autres termes, cette partie nous a permis de montrer qu'il existait des cellules qui sont sensibles à la luminance d'une surface mais cette réponse peut être accentuée à partir du moment où il y a une induction de la luminosité depuis les contours.

2.2.3.3. La texture d'une surface est-elle envisagée différemment ?

Les effets contextuels qui influencent l'activité des neurones, dont le champ récepteur est situé sur une surface uniforme, ont également été démontrés lorsque la surface est constituée d'une texture et non pas d'une luminance uniforme (Lamme, 1995). Lamme, Rodriguez-Rodriguez et Spekreijse (1999) ont testé directement l'hypothèse selon laquelle la détection des contours est suivie par un processus de remplissage de la surface. Les stimuli étaient composés d'une figure où un carré et un fond se distinguait, sur la base de l'orientation des lignes qui formaient leur texture (un exemple est présenté à la Figure 18). Au cours de l'expérience, le carré se déplaçait verticalement afin d'être à l'intérieur, sur les bords ou à l'extérieur d'un champ récepteur (15 positions ont été enregistrées). Pour l'analyse des résultats, les auteurs ont principalement examiné les caractéristiques temporelles de l'activité des neurones, chez un singe, après la présentation du carré. Les résultats ont indiqué qu'une faible activité se produisait entre 70 et 80 msec après la stimulation pour les contours et la surface. Ensuite, entre 90 et 100 msec, la réponse obtenue pour les bords augmentait, et l'activité pour la surface était toujours la même. La réponse des neurones pour les contours du carré était maximale à partir de 115-125 msec et, c'est à partir de 150-160 msec que l'activité augmentait pour la surface. Par la suite, les variations enregistrées diminuaient lentement mais de façon uniforme, c'est-à-dire que la réponse aux contours était toujours plus importante par rapport à la surface et ce jusqu'à la fin (420-430 msec). En d'autres termes, une réponse neuronale est donnée très rapidement aussi bien pour les contours que pour la surface, et immédiatement après, une augmentation de l'activité est enregistrée pour les contours, et plus tardivement pour la surface. Les résultats sont présentés à la Figure 18. La première étape de l'activité neuronale impliquerait un traitement précoce de la surface et ce, de façon parallèle aux traitements des contours. Toutefois, pour Lamme et al. (1999), cette première étape de réponse serait générée par un traitement ascendant du CGL en direction de V1, ce qui permettrait de déterminer les réponses ultérieures. Un résultat similaire était également trouvé pour des tailles importantes (12 deg) de la figure centrale (Zipser, Lamme & Schiller, 1996). Cette dernière expérience est incompatible avec l'idée d'un processus de remplissage, puisqu'en effet, dans les expériences de Lamme et al. (Zipser et al., 1996 ; Lamme et al., 1999), les temps de réponse des neurones ne dépendent pas de la taille de l'aire centrale. En effet, le processus de remplissage implique l'idée d'une propagation progressive de l'information, si l'implication du remplissage avait été vérifié alors la réponse des neurones aurait dû diminuer au fur et à mesure que le champ récepteur placé sur la surface s'éloigne du bord, ce qui n'est pas constaté. Les résultats de Zipser, Lamme et Schiller (1996) ont toutefois été controversés par Rossi, Desimone et Ungerleider (2001) qui ont observé, dans les mêmes conditions,

une diminution de l'activité des neurones lorsque le champ récepteur était positionné au centre du signal et pour des tailles importantes du signal (de 4 à 8 deg).

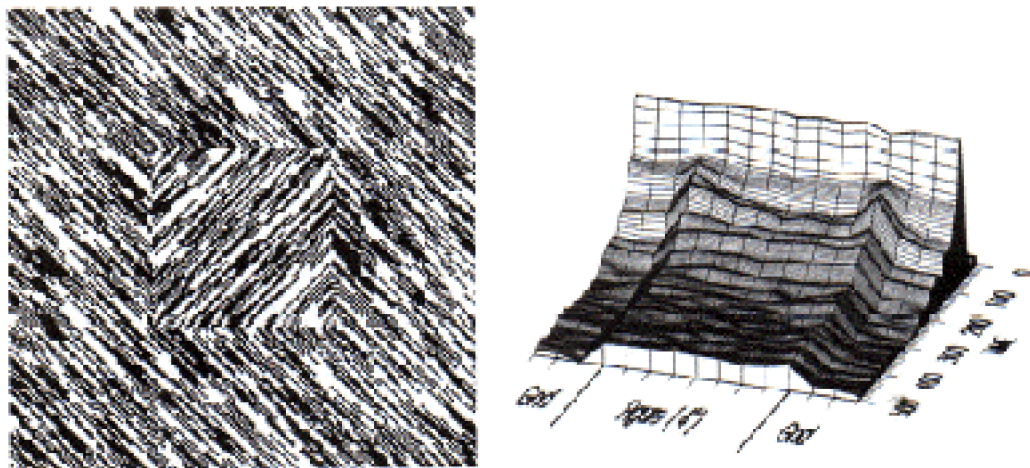


Figure 18 : La figure de gauche représente le type de stimulus utilisé pendant les expériences de Lamme, Rodriguez-Rodriguez et Spekreijse (1999). La figure de droite montre les résultats obtenus dans leurs expériences en fonction du temps d'enregistrement et de la position des champs récepteurs.

Les études, qui ont utilisé une surface uniforme ou une texture, indiquent que la surface est traitée dès l'aire V1. Rossi et Paradiso ont également montré que certaines cellules du CGL étaient activées par la dimension de surface. Néanmoins, pour des structures encore plus précoces, comme la rétine, ce traitement ne se retrouvait pas (Rossi & Paradiso, 1999). Par ailleurs, les résultats de Komatsu et al. (1996), de Mc Evoy et al. (1998), mais aussi de Lamme et al. (Lamme, Rodriguez-Rodriguez & Spekreijse, 1999 ; Zipser, Lamme & Schiller, 1996), concernant la taille de leur stimuli, sont en contradiction avec l'hypothèse du phénomène de remplissage.

La représentation de l'information de surface dans l'aire visuelle primaire est surprenante puisque, depuis les travaux précurseurs de Hubel et Wiesel (1962), l'aire V1 avait surtout été étudiée en terme de réponse à des contours lumineux ou colorés. En d'autres termes, on supposait que seuls les neurones dont les champ récepteurs sont situés aux frontières de la surface, peuvent être activés dans le cortex visuel. La perception de la surface serait la conséquence du codage des contours. Mais, de nombreuses expériences plus récentes (Komatsu, Kinoshita & Murakami, 1996 ; Lamme, Rodriguez-Rodriguez & Spekreijse, 1999 ; Rossi & Paradiso, 1999 ; Rossi, Rittenhouse & Paradiso, 1996) ont mis en évidence une activité neuronale dans V1 non seulement au bord d'un stimulus mais également à l'intérieur de celui-ci. Par ailleurs, l'expérience de Hung, Ramsden, Chen et Roe (2001) a permis de préciser que la luminosité d'une surface à partir des bords est essentiellement représentée dans l'aire V2. De plus, Lamme, Rodriguez-Rodriguez et Spekreijse (1999) ont mis en évidence l'existence d'un traitement hiérarchique depuis les contours vers l'intérieur de la surface. Dans cette étude, différents niveaux d'activité ont été observés dans le temps, et il est tout à fait possible que ces

variations de réponse soient modulées par des processus « feed-back » provenant du cortex extra-strié. Ces connexions « feed-back » permettraient d'augmenter l'activité de V1 au cours du temps, pour des neurones dont le champ récepteur est situé à l'intérieur d'une surface (Lamme & Roelfsema, 2000 ; Lamme, Supèr & Spekreijse, 1998). Dans la continuité de cette idée, les travaux de Hung et al. (2001) permettent de compléter l'hypothèse de l'existence d'un processus « feed-back » depuis le cortex extra-strié pour le traitement de la surface à partir des contours.

Les premières études que nous avons présenté dans cette partie démontraient l'existence précoce du processus de remplissage (Fiorani et al., 1992 ; Gilbert & Wiesel, 1992). Par la suite nous avons voulu déterminer le rôle de ce mécanisme dans la perception de normale de la surface. L'ensemble des études portant sur des stimuli uniformes ou composés d'une texture ont permis d'indiquer l'importance accordée à la dimension de surface par V1. Toutefois, ces études montrent que la surface est une dimension qui est reconstruite dans une seconde étape. En effet, l'expérience de Rossi et Paradiso (1999) a montré que les cellules de la rétine n'encodent pas cette dimension, ce qui est cohérent avec les travaux de Hubel et Wiesel (1962, 1968). Néanmoins, cette étude est à notre connaissance la seule à avoir réalisé des enregistrements de la rétine lorsqu'une surface uniforme était présentée. Par contre, ces expériences ne démontrent pas explicitement le rôle du processus de remplissage dans le traitement de la surface. Mais, cela ne veut pas dire que le remplissage ne joue pas un rôle central pour certains phénomènes tels que la tache aveugle ou les scotomes.

2.3. Arguments comportementaux en faveur de l'existence du processus de remplissage

Cette partie a pour objectif de démontrer l'existence comportementale du processus de remplissage dans le cadre de phénomène naturel comme la tache aveugle ou le « scotome bleu », mais également dans le cadre de lésions de la rétine ou du cortex visuel, qui conduisent à l'apparition de scotomes. Par ailleurs, une dernière partie abordera le cas des scotomes artificiels. La partie sur la physiologie du remplissage a montré une différence de traitement entre la tache aveugle et les scotomes artificiels. Aussi, un second objectif sera de déterminer si cette distinction se retrouve dans les études comportementales.

2.3.1. Le processus de remplissage dans le cas de phénomènes naturels

La tache aveugle est la région de la rétine correspondant départ du nerf optique (voir p. 55). Cette région ne comporte pas de photorécepteurs, aucune information n'y est véhiculée et nous devrions pas avoir de sensation dans la région du champ visuel correspondant à cette zone. Dans la vie de tous les jours, nous n'avons pas la sensation de cette tache noire dans notre champ visuel pour deux raisons : 1) la région de la tache aveugle est positionnée de telle façon que les photorécepteurs de l'autre œil enregistrent les informations manquantes dans la partie aveugle, et 2) le système sensoriel remplit cette région aveugle à partir de la région du pourtour. Ainsi, la tache aveugle ne peut être

étudiée qu'en condition monoculaire. En effet, si nous fixons le point de fixation de la Figure 19 avec uniquement notre œil droit et à une certaine distance, les deux barres noires seront perçues comme une seule. De même, pour la Figure 19 le point noir disparaît lorsque ce dernier est placé dans la zone correspondant à la tache aveugle. Dans les deux cas, la perception obtenue s'explique par le fait que le point noir et l'espace blanc sont situés dans la région de la tache aveugle, et un remplissage est alors réalisé par notre système pour combler l'information manquante. Dans les deux cas que nous avons présenté, le processus de remplissage semble être instantané (Palmer, 1999 ; Ramachandran, 1992a, b). Par contre, si l'on place une seule barre à proximité de la tache aveugle, la taille de la barre n'augmentera pas.



Figure 19 : Mise en évidence de la tache aveugle de façon comportementale. Dans cette expérience, le point de fixation de la figure du haut ou du bas doit être fixé uniquement par l'œil gauche. La feuille doit être fixée le plus près possible, puis l'observateur doit s'éloigner lentement. A une certaine distance, le point noir de la figure du haut disparaît, et les deux barres de la figure du bas sont perçues comme une seule (l'exemple est tiré de Palmer, 1999 p. 34).

Un processus de remplissage a été mis en évidence de manière comportementale chez le singe (Komatsu & Murakami, 1994). Les auteurs ont tout d'abord déterminé, dans une condition binoculaire, si le singe faisait une différence entre un disque homogène et un anneau dont le centre pouvait être plus grand ou plus petit que la région de la tache aveugle. Les mouvements oculaires étaient enregistrés pour déterminer si le singe apprenait à reconnaître le disque homogène d'un disque non homogène. La procédure était la suivante : après la présentation d'un point de fixation, le disque homogène et l'anneau étaient présentés simultanément à droite et à gauche de ce point. Pour chaque essai, l'un des deux stimuli était placé dans la tache aveugle. Les résultats ont indiqué que le singe arrivait correctement à discriminer la zone homogène par rapport à un anneau dont le centre est plus important que l'aire du champ visuel recouvrant la tache aveugle. De la même façon, il y avait autant de mouvement oculaire à droite et à gauche lorsque la zone centrale de l'anneau était plus petite que la zone du champ visuel de la tache aveugle. Par contre, lorsqu'un œil était recouvert et que l'anneau était présenté dans la tache aveugle de l'œil non recouvert, les singes n'arrivaient pas à le discriminer d'un disque homogène. Ce dernier résultat permet de confirmer que le processus de remplissage s'effectue également chez le singe.

Un second phénomène naturel existe en dehors de la tache aveugle, il s'agit d'une trinitopie de courte longueur d'onde (que l'on trouve dans la littérature sous le nom de « scotome bleu ») (Walls, 1954). Lorsque l'on regarde le ciel par temps clair, on peut apercevoir parfaitement une couleur bleu homogène. La perception d'un bleu homogène

est étonnante lorsque l'on sait qu'il n'existe aucun cône dans la fovéa qui réponde à des courtes longueurs d'onde. Cette organisation physiologique devrait occasionner une lacune au centre de notre champ visuel. L'absence d'information ne nous apparaît pas, ce qui est probablement dû à un processus de remplissage (Gerrits & Vendrick, 1970 ; Spillman & Werner, 1996). Ce phénomène a été mis en évidence de façon comportementale par Magnussen, Spillman, Sturzel et Werner (2001). Un disque uniforme, dont la longueur variait aléatoirement de 410 à 510 nm, était présenté à l'observateur. La tâche consistait à dire si une tache noire était perçue ou non sur une échelle de 1 à 10 (1 correspondant à un niveau où le phénomène n'est pas visible et 10 à un niveau où le phénomène l'est fortement). Dans une seconde expérience, l'observateur déterminait le temps de disparition de la tache. Les résultats ont indiqué que la visibilité de la tache suivait une courbe en U inversé dont le plateau se situait à 450 nm, et une courbe identique était observée avec les temps de remplissage. De plus, le coefficient de corrélation était important entre la visibilité du point noir et le temps que celui-ci mettait à disparaître (de 0.84 à 0.93 selon les observateurs).

2.3.2. Le processus de remplissage dans le cas des scotomes

L'idée d'un traitement de remplissage dans le cas des scotomes a rapidement été émise par Walls (1954). Par la suite, Gerrits et Timmerman (1969) ont étudié les temps de remplissage chez des patients ayant des scotomes rétiniens. Pour cela, les auteurs plaçaient un carré autour de la zone du champ visuel correspondant au scotome, et les patients décrivaient ensuite leur perception. Cette observation montrait que le carré était perçu immédiatement ou dans un délai bref comme étant homogène. Les temps de remplissage dépendent principalement de l'étendue du scotome (Gerrits & Timmerman, 1969).

Ramachandran (1993) a étudié les capacités de remplissage de patients ayant des scotomes corticaux. Lorsque l'auteur présentait une ligne droite dans la zone du champ visuel correspondant au scotome de ces patients, ces derniers avaient l'expérience d'une tache noire qui interrompait la ligne, mais très rapidement ils arrivaient à percevoir cette ligne comme continue. Par la suite, Ramachandran (1993) a présenté deux lignes non-alignées, et les résultats de cette expérience ont indiqué que les patients percevaient une ligne droite. Une dernière expérience consistait à présenter aux patients deux segments de colonnes de X, et les résultats ont montré que les deux segments étaient perçus comme continus. Une présentation des stimuli et de la perception des patients est montrée à la Figure 20. Pour l'auteur, l'expérience visuelle de ces patients ne pouvait pas être la conséquence d'une activité neuronale de la zone corticale correspondant au scotome, mais bien à une activité de haut niveau.

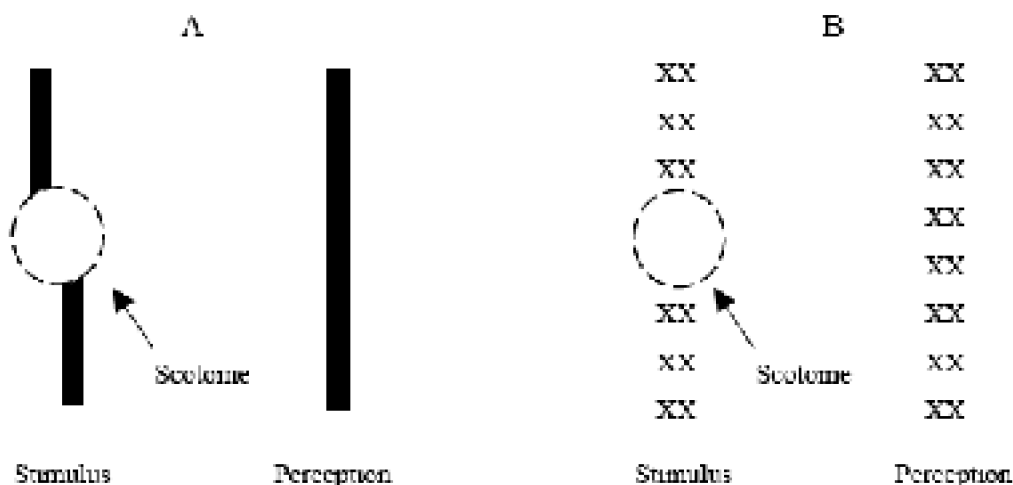


Figure 20 : Résultats du processus de remplissage, chez des patients avec des scotomes corticaux, dans les expériences de Ramachandran (1993). Le cercle en pointillé représente le scotome. Pour chacune des expériences, la figure de gauche indique ce qui est montré aux patients et la figure de droite montre ce qui est perçu. Les segments de lignes non-alignés sont perçu comme continus et droits (A), et la configuration de lignes composée de X est également remplie (B).

Ces expériences montrent le lien étroit qui existe d'un point de vue phénoménologique entre le remplissage pour la tâche aveugle et les scotomes.

2.3.3. Le processus de remplissage dans le cas des scotomes artificiels

Nous aborderons au cours de cette partie le paradigme des scotomes artificiels¹³. Cette technique a été rapidement mise en avant pour étudier spécifiquement le processus de remplissage. Pessoa et Neumann (1998) ont émis l'hypothèse que les scotomes artificiels reflétaient les mécanismes mis en jeu dans la perception de surface normale ou de la tâche aveugle. Toutefois, nous verrons que ce parallèle a été contesté (Dugin, Tripathy & Levi, 1995).

Pour créer un scotome artificiel (e.g. Ramachandran & Gregory, 1991), un carré gris (1.5*1.5°) est présenté sur un bruit dynamique (dont la luminance varie continuellement et de façon aléatoire), à une distance de 6° d'un point de fixation. Après une période brève (en moyenne 5 sec), le carré disparaît et est rempli par le bruit environnant. Le processus de remplissage peut se produire pour différents types d'environnements, comme des lignes horizontales ou en diagonales, mais aussi des lettres et même des caractères chinois (Gyoba, 1997). Il semble que la couleur et la texture soient traitées par des processus de remplissage différents (Ramachandran & Gregory, 1991) : un carré gris était affiché sur un fond rose, avec une texture, des points noirs dont la position variait, superposé au fond rose. Une texture identique était affichée dans le carré gris, mais cette fois les points noirs bougeaient horizontalement. Les résultats ont indiqué que le carré gris

¹³ Le terme de scotome artificiel a finalement été adopté pour décrire la disparition d'un signal lorsque celui est présenté sur un fond bruité ou uniforme, cet effet est identique à l'effet Troxler.

était d'abord rempli par le fond rose puis, dans un deuxième temps, les points noirs à l'intérieur du carré gris disparaissaient, et étaient remplacés par les points noirs dont la localisation changeait continuellement. Ce résultat suggérait l'existence de deux mécanismes de remplissage spécifique à la texture et à la couleur, et l'origine de ces deux mécanismes serait la conséquence de la spécialisation des aires corticales dans le traitement de ces deux dimensions.

Spillman et Kurtenbach (1992) ont suggéré que la vitesse du processus de remplissage des scotomes artificiels était dépendante du fond. La procédure était identique à celle de Gregory et Ramachandran (1991), mais ici, trois fonds différents étaient utilisés : un fond uniforme, un bruit statique et un bruit dynamique (plusieurs fréquences temporelles étaient utilisées pour le bruit). Le temps que mettait le carré pour disparaître était, en moyenne, de 57 sec pour un fond uniforme, de 26 sec pour un bruit statique et de 10 à 16 sec pour un bruit dynamique. Ainsi, plus la fréquence temporelle du bruit est élevée, et plus le carré disparaît rapidement. Spillman et Kurtenbach (1992) ont également utilisé plusieurs niveaux de contrastes pour le carré (blanc, gris moyen et gris foncé), mais aucune influence de ce facteur n'était observée sur les temps de remplissage. La différence de remplissage entre un bruit dynamique et statique a également été démontrée par Ramachandran, Gregory et Aiken (1993) qui étudiaient l'effet de l'excentricité du signal (12, 17.5 et 23 deg). Les résultats ont montré que plus le signal était éloigné du point de fixation, et plus le processus de remplissage s'effectuait rapidement. Cet effet d'excentricité a été répliqué par Lou (1999) et semble indiquer que les scotomes artificiels (ou effet Troxler) sont essentiellement des phénomènes de vision périphérique.

Pour expliquer le phénomène des scotomes artificiels, Spillman et Kurtenbach (1992) ont émis une hypothèse selon laquelle les contours de la cible seraient perturbés par le bruit. En effet, dans un bruit dynamique, les contours d'une cible ne sont pas perçus nettement, mais de nombreuses variations de l'ordre du pixel empêchent une perception distincte, comme les dents d'une scie par exemple. On peut supposer que le signal est alors moins distinct, et qu'il peut alors disparaître plus facilement. Ramachandran, Gregory et Aiken (1993) ont proposé une autre explication : la disparition du signal serait la conséquence d'une « fatigue » des neurones qui sont spécialisés dans l'extraction des bords. Pour les auteurs, cette hypothèse est vérifiée lorsque la disparition du signal est plus rapide pour un bruit dynamique, car, les neurones seraient plus vite « fatigués » par ce type de stimulus plutôt que par un bruit statique.

DeWeerd, Desimone et Ungerleider (1998) ont réalisé une étude paramétrique du processus de remplissage. La tâche était en tout point identique aux études décrites précédemment (Gregory & Ramachandran, 1991 ; Ramachandran et al., 1993 ; Spillman & Kurtenbach, 1992), mais ici le fond était composé uniquement d'une texture dynamique (c'est-à-dire des petits traits blancs qui apparaissaient et disparaissaient aléatoirement sur un fond noir). Les auteurs ont tout d'abord manipulé la taille du signal (un carré gris moyen) qui allait de 0.6 à 5.6 deg. Les résultats ont indiqué que le processus de remplissage était plus important lorsque la taille du signal augmentait. La disparition du signal se faisait sur une période approximative de 3 secondes pour une taille de 0.6 deg, et de 10 à 12 secondes pour une taille de 5.6 deg. De façon réciproque, les temps de

remplissage diminuaient lorsque la taille de la texture augmentait. Une autre expérience a permis de montrer qu'un signal coloré mettait plus de temps à disparaître qu'un signal gris moyen. Par ailleurs, DeWeerd et al. (1998) ont montré qu'un signal fixe permettait un traitement de remplissage plus rapide. Dans cette expérience, les auteurs présentaient un signal carré à 8 deg d'excentricité d'un point de fixation, puis après une seconde, le signal se déplaçait à coté de sa position initiale (de 0.1 à 1.6 deg). Les temps de disparition de la cible étaient plus rapides pour des petits déplacements. Cette expérience montre l'importance des contours dans le traitement de remplissage, puisque le déplacement des contours ralentit le remplissage. Une expérience similaire a été également réalisée avec un signal qui apparaissait et disparaissait rapidement dans la périphérie (Anstis, 1996).

Dans la continuité de cette étude paramétrique, Welchman et Harris (2001) se sont également intéressés aux propriétés physiques de la cible qui influençaient sa disparition. Pour cela, les auteurs ont utilisé différents niveaux de luminance pour la cible et le fond (du noir au blanc). Les temps de disparition de la cible étaient plus longs pour des luminances extrêmes comme le noir et le blanc. Les résultats de cette étude sont toutefois contradictoires avec ceux de Spillman et Kurtenbach (1992), où aucune différence n'était observée entre des luminances extrêmes et un niveau de gris moyen. Par ailleurs, Welchman et Harris (2001) ont montré que le remplissage de la cible était facilité lorsque sa luminance se rapprochait de celle du fond.

Les études qui ont été présentées montrent que la relation entre la cible et le pourtour est un facteur important concernant les temps de remplissage. L'objectif de l'expérience de Sakaguchi (2001) était de manipuler différents facteurs qui constituent un lien entre la cible et le pourtour. Dans la plupart des expériences qui ont été présentées, la cible est toujours un carré gris et le pourtour est souvent un bruit. La tâche de l'observateur était de répondre le plus vite possible sur une touche dès que la cible disparaissait. Dans une première expérience, différents niveaux de luminance étaient utilisés pour la cible et le pourtour. Conformément aux résultats de Welchman et Harris (2001), les résultats ont indiqué que le temps de disparition de la cible était élevé quand le contraste entre la cible et le pourtour était important. Par la suite, afin de connaître les effets de l'orientation sur les capacités de remplissage, la cible et le pourtour étaient composés de réseaux sinusoïdaux orientés différemment. Les résultats ont montré que les temps de remplissage augmentaient lorsque la différence d'orientation entre la cible et le pourtour augmentait. Welchman et Harris (2001) se sont aussi intéressés à l'asymétrie entre la cible et le pourtour dans le domaine du mouvement. Dans leur expérience, le fond était constitué de plusieurs points qui bougeaient de façon aléatoire et la cible était représentée par des points qui se déplaçaient horizontalement dans un espace donné. Plusieurs vitesses de déplacement des points étaient utilisées. Les résultats allaient dans le sens des travaux de Sakaguchi (2001), puisque plus le contraste de mouvement entre la cible et le pourtour était important et plus la cible mettait du temps à disparaître. Ces expériences (Sakaguchi, 1999 ; Welchman & Harris, 2001) mettent en évidence un mécanisme important pour le processus de remplissage : l'asymétrie entre la cible et le pourtour. En effet, lorsque la différence qui existe entre la cible et le pourtour est faible, la visibilité du signal est diminuée, ce qui rend le traitement de remplissage plus rapide.

Une étude récente de Welchman et Harris (2003) est venue contredire l'implication

d'un processus de remplissage dans le cas des scotomes artificiels. Dans leur expérience, les bords de la cible étaient dégradés par une opération de lissage, les observateurs indiquaient à partir de quel moment la cible disparaissait. Les résultats montraient qu'il n'y avait pas de différence entre le temps de disparition de la cible lissée et de la cible non lissée. Ainsi, la dégradation des bords n'influence pas le temps que met une cible à disparaître. Les auteurs suggèrent que la disparition d'une cible serait le résultat d'une adaptation qui ne serait pas spécifique aux contours, et qu'il n'est pas nécessaire de faire l'hypothèse d'un remplissage pour expliquer ce phénomène.

Un second phénomène perceptif est généralement observé dans le cadre des scotomes artificiels. En effet, Ramachandran et Gregory (1991) ont réalisé une expérience identique aux études présentées précédemment, mais cette fois dès que le processus de remplissage se produisait, le bruit était automatiquement remplacé par un fond gris identique au carré. Dans ce cas, le bruit à l'emplacement du carré persistait pendant 2-3 secondes après l'affichage d'un fond gris moyen. Les auteurs ont proposé l'existence d'une représentation neuronale du remplissage qui persisterait quelques secondes après la disparition du phénomène. Cet effet ne se produit pas avec un bruit statique, mais été observé essentiellement avec un bruit dynamique (Ramachandran, Gregory & Aiken, 1993). Reich, Levi et Frishman (2000) ont étudié plus particulièrement ce phénomène en déterminant la taille perçue de cet effet. Pour cela, quatre points étaient placés à l'opposé de l'ancienne position du signal. La taille du scotome artificiel pouvait varier de 2 à 6 deg, et l'observateur jugeait la taille du phénomène perceptif par rapport aux quatre cercles. L'estimation donnée par les observateurs étaient toujours inférieure de 1 deg par rapport à la taille du signal. Ces résultats étaient aussi bien observés en présentation monoculaire ou binoculaire, ce qui est cohérent puisque le scotome artificiel, à l'inverse de la tache aveugle, est binoculaire.

Toutefois, ce second phénomène n'est jamais constaté pour les phénomènes naturels comme la tache aveugle ou les scotomes (Durgin, Tripathy & Levi, 1995). En outre, les temps de remplissage des scotomes artificiels sont beaucoup plus importants (de l'ordre d'une dizaine de secondes) que ceux de la tache aveugle ou des scotomes, qui est instantané. De plus, les études en physiologie ont montré une contradiction dans la réponse des neurones corticaux avec la présentation des scotomes artificiels (DeWeerd, Gattass, Desimone & Ungerleider, 1995) ou de stimuli dans la tache aveugle (Fiorani, Rosa, Gattass & Rocha-Miranda, 1992 ; Komatsu, Murakami & Kinoshita, 1996). La distinction qui est réalisée entre le remplissage naturel (par exemple la tache aveugle) et artificiel (les scotomes artificiels) se retrouve aussi bien dans les études comportementales que physiologiques, ce qui montre bien que les processus de remplissage doivent être différents dans les deux cas.

2.3.4. Problématique du processus de remplissage

2.3.4.1. Problématique des temps de traitement du remplissage

Les résultats des expériences concernant les scotomes artificiels (DeWeerd, Desimone & Ungerleider, 1998 ; Ramachandran & Gregory, 1991 ; Spillman & Kurtenbach, 1992)

montrent que le remplissage perceptif est plus lent que le remplissage, quasi instantané, impliqué dans la perception normale de la surface. Par exemple, les temps de remplissage, dans le cadre des expériences utilisant les scotomes artificiels, oscillent entre 4 et 12 sec selon les paramètres manipulés. Ceci implique un processus relativement lent par rapport à l'estimation effectuée pour une surface normale : Paradiso et Nakayama (1991) avaient estimé un temps de remplissage équivalent à 115 deg/sec. Pour DeWeerd et al. (1998), cette observation laisse entendre que le processus de remplissage est ralenti par un traitement antagoniste qui séparait la région uniforme (la cible) du pourtour. Mais, comme nous venons de l'évoquer, il est tout à fait possible que le phénomène de remplissage, qui se produit pour les scotomes artificiels, ne reflète pas exactement la perception normale de la surface. A l'inverse, cette différence pourrait être liée à l'existence de saccade oculaire durant la fixation, ce qui maintiendrait une activité neuronale le long des bords du signal et retarderait la disparition du signal (Spillman & DeWeerd, 2003).

2.3.4.2. Suppression et substitution : une solution ?

Une question peut alors se poser : est ce que le traitement du remplissage (constaté dans le cadre de la tâche aveugle ou encore dans le traitement d'une surface uniforme) et la disparition d'un signal (un exemple est donné par les images stabilisées ou encore les scotomes artificiels) sont un seul et même mécanisme, ou si l'un n'est pas le pré requis de l'autre ? L'asymétrie qui existe entre les temps de remplissage des scotomes naturels et artificiels ne suggère pas un traitement unitaire, mais plutôt la manifestation de différentes étapes de traitement d'un même mécanisme (Spillman & DeWeerd, 2003).

Selon plusieurs auteurs (Spillman & DeWeerd, 2003 ; Stürzel & Spillman, 2001), le processus de remplissage impliqué dans le cadre des scotomes artificiels résulte d'un traitement en deux étapes. Un premier traitement serait mis en place lentement, dont la durée est déterminée en seconde, durant lequel la séparation figure-fond se détériorerait petit à petit, et ce jusqu'à la suppression des contours du signal. Puis, un second traitement, plus rapide (mesuré en milliseconde), serait effectué dans lequel le fond se diffuserait depuis les contours du signal pour remplir la cible (la substitution). Le temps de disparition du signal est important, dans les expériences se reportant à l'étude des scotomes artificiels, puisqu'il reflète la première étape, la seconde étape se faisant instantanément.

L'hypothèse de suppression et de substitution permettrait d'expliquer pourquoi l'étude physiologique de DeWeerd, Gattass, Desimone et Ungerleider (1995) montrait une forte implication du processus de remplissage dans les aires V2 et V3, par rapport aux études physiologiques portant sur la tâche aveugle et les scotomes qui démontraient au contraire une implication précoce du remplissage. Toutefois, aucune donnée expérimentale n'a permis de valider cette hypothèse. Il est donc nécessaire de faire la distinction entre le remplissage naturel (tâche aveugle, scotome, scotome bleu) et artificiel, car même si ces deux phénomènes semblent similaires dans les effets, les mécanismes qui les sous-tendent n'apparaissent pas comme étant identiques. Par conséquent, le paradigme des scotomes artificiels ne permet pas à lui seul d'étudier l'existence d'un processus de remplissage dans le traitement de la surface.

2.4. Modèle de traitement hiérarchique contour/surface

Cette partie présente les modèles généraux de reconnaissance visuelle qui se fonde sur une conception hiérarchique, celui de Marr (1982) centré sur un point de vue computationnel et celui de Grossberg et Mingolla (1985) centré sur les propriétés physiologiques du système visuel.

2.4.1. L'approche computationnelle de Marr

La théorie de Marr (1982) a dominé pendant de nombreuses années l'approche computationnelle. Selon cette théorie, trois principales étapes permettraient le traitement d'une image en vue de sa reconnaissance. Dans la lignée des physiologistes, les chercheurs dans le domaine de la vision computationnelle ont proposé l'existence d'une première étape d'extraction des contours par le biais de détecteurs de lignes. Pour Marr (1982), cette première étape appelée l'ébauche primitive (« Primal Sketch »), se scinde en deux parties : l'ébauche primitive brute (« Raw Primal Sketches ») et l'ébauche primitive complète (« Full Primal Sketches »). L'ébauche primitive brute permettrait l'extraction de bordures, de barres et de taches par leurs localisations dans l'espace et leurs orientations. L'ébauche primitive complète est une version plus élaborée de l'ébauche primitive brute car elle résulte d'opérations comme le groupement des contours. Ce premier pas permet le traitement des contours d'une image, mais la description d'une image ne se limite pas à cet aspect. Une deuxième étape était ensuite décrite comme une description de la surface et de la forme centrée sur l'observateur. Cette étape était appelée : l'ébauche $2^{1/2}$ (« 2^{1/2} Sketch »), et était obtenue par une analyse de la profondeur et du mouvement, mais également grâce aux structures assemblées dans l'ébauche primitive. Une étape essentielle dans le traitement visuel est la reconnaissance d'objets, et pour connaître à quel objet une forme appartient, une troisième étape était nécessaire : la représentation en modèle 3D (« 3D Model Representation »). Elle correspond à une étape de traitement où des représentations d'objets connus sont activées. La théorie de Marr implique un certain nombre de niveaux de représentations, dont chaque étape a une description symbolique de l'information transmise par le système physiologique.

L'algorithme conçu par Marr et Hildreth (1980) pour générer l'ébauche primitive brute commence par la transformation d'une représentation de l'image en une série de représentations indépendantes à différents niveaux. L'algorithme « zero-crossing » permet de détecter les contours d'une image.

Marr et Hildreth (1980) ont proposé que l'image soit traitée par un ou plusieurs filtres. Ces filtres utilisent des fonctions gaussiennes de différents écart-types (ou largeurs) et amplitudes (ou hauteurs), et le but de cette première opération est de compresser l'information en diminuant la gamme des fréquences spatiales dans l'image. Par exemple, plus un filtre gaussien est large et plus la bande passante transmise sera basse. Cette première opération est notée : $G \cdot I$, où G correspond aux filtres gaussiens et I à l'image. Une seconde opération est réalisée par l'algorithme de Marr et Hildreth (1980) : la localisation des changements d'intensité dans les multiples représentations de l'image, se

fonde sur le calcul de la dérivée seconde d'un filtre gaussien, dont la forme s'apparente à celle d'un champ récepteur, cet opérateur est appelé le ∇^2 (ou Laplacien). Le Laplacien est alors appliqué à chacune des valeurs fournies par les filtres gaussiens. Le résultat correspond à un ensemble de séries de valeurs du Laplacien notées $\nabla^2 G * I$. L'image de départ est montrée dans la Figure 21a, et le résultat de la convolution est présenté dans la Figure 21b. Les valeurs numériques de l'image filtrée par le $\nabla^2 G$ peuvent être positives ou négatives, et le résultat constitue une échelle allant du blanc (positif) au noir (négatif) en passant par plusieurs niveaux de gris. La Figure 21c présente l'image de façon binaire: noire (-1) pour des valeurs de luminance négative et blanc (+1) pour des valeurs de luminance positive. La Figure 21d indique les positions dans l'image où les différents changements d'intensité étaient localisés .

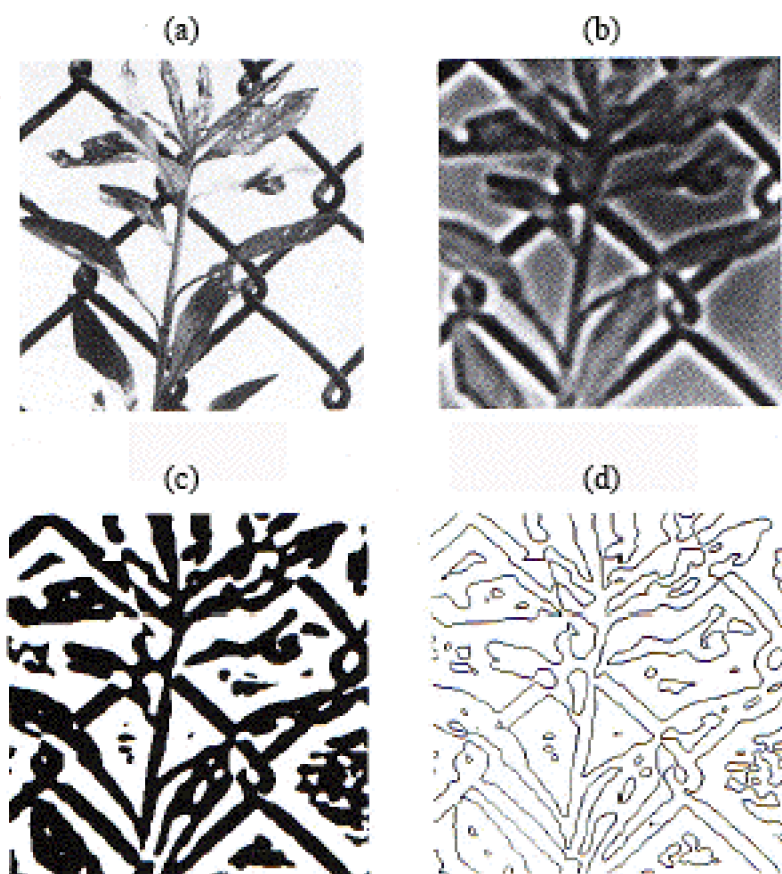


Figure 21 : Détection des contours par l'étape « zero-crossing » de Marr et Hildreth (1980). La figure située en haut à gauche représente l'image naturelle, et les trois autres figures représentent les différentes étapes du « zero-crossing ».

L'ébauche primitive brute est une représentation symbolique de l'image élaborée au moyen de quatre éléments : les segments de bordure, les barres, les terminaisons et les taches, chacun étant doté de propriétés d'orientation, de contraste, de taille et de position. Ces quatre représentations permettent de signaler différents types d'intensités. Un exemple d'ébauche primitive brute est présenté à la Figure 22. Cette ébauche ne décrit pas uniquement les bords de l'objet mais aussi les éléments de texture au sein de la

surface. Les Figures 22a et 22b indiquent l'étape finale du zero-crossing pour deux différentes tailles de $\sigma^2 G$. Les Figures 22c, 22d et 22e montrent respectivement, les taches, les contours et les barres qui sont trouvés par les différentes tailles de filtres lors de l'étape du zero-crossing.

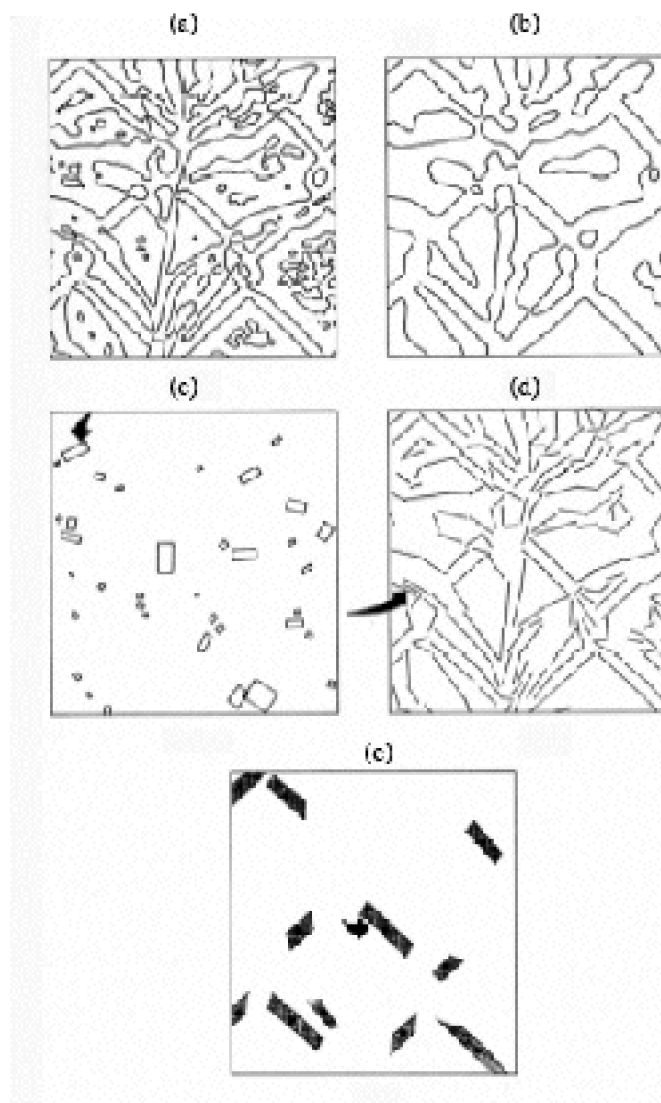


Figure 22 : L'ébauche primitive brute de Marr (1982). L'image de la Figure 1a a été analysée par l'algorithme « zero-crossing » de façon locale (a) et globale (b). En d'autres termes la taille du filtre DoG variait. Ces résultats sont intégrés pour former une carte symbolique de contours orientés (d). D'autres analyses permettent la détection de tache (c) et de barres (e) pour différentes localisations.

L'ébauche primitive complète est dérivée de l'ébauche primitive brute par une opération d'organisation, pour extraire une structure globale à partir de l'ébauche primitive brute qui est désordonnée. L'étape suivante consiste alors à grouper les éléments de l'image en un tout, et à séparer les régions qui diffèrent les unes des autres par leur texture. La première étape de l'ébauche primitive complète est d'attribuer des marques d'emplacement à des structures ou des agrégats de structures issus de l'ébauche visuelle

brute. Puis, ces marques d'emplacement sont à leur tour rassemblées pour former des unités plus larges, et sont définies par la position des taches, par la terminaison d'un bord ou d'une barre. Le rassemblement de ces marques d'emplacement peut s'effectuer par le regroupement des marques proches sur la base de changement de la densité spatiale, par le rassemblement curvilinéaire qui génère des contours en joignant des éléments alignés et proches les uns des autres, et finalement par un rassemblement d'éléments partageant la même orientation. Le rassemblement des marques d'emplacement repose donc sur la proximité locale (des éléments adjacents sont combinés) et la similarité (des éléments orientés de façon similaire sont combinés) ; des considérations plus globales peuvent influencer la détection des structures. En effet, dans le cas d'une courbe, un principe de « fermeture » pourrait permettre à deux segments d'être joints, et ce, même pour des contrastes différents. La théorie de Marr (1982) incorpore la plupart des principes de l'école de la Gestalt.

Marr (Marr, 1982 ; Marr & Nishihara, 1978) considère que le but du traitement visuel précoce est la description d'une surface visible, de telle manière que leurs dispositions et arrangements soient centrés par rapport à l'observateur. Il appelle cette description : l'ébauche en $2^{1/2}$ D. Elle se construit à partir de l'ébauche primitive complète. Au cours de cette étape, l'orientation des surfaces serait traitée de façon précise. Pour Marr (1982), l'ébauche $2^{1/2}$ D est constituée d'une série de représentations de type vectoriel (voir Figure 23) : la longueur du vecteur indique l'oblicité de cette partie de la surface et l'orientation du vecteur représente la direction dans laquelle est inclinée la surface. De plus, la distance par rapport à l'observateur serait indiquée par une échelle scalaire.

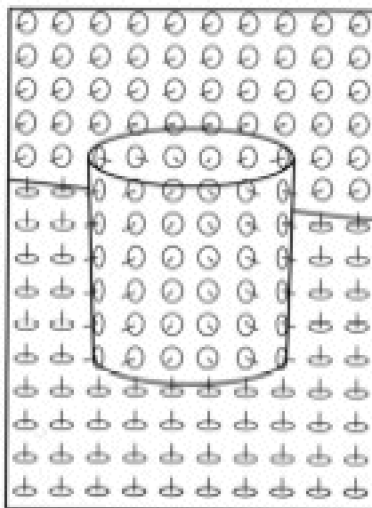


Figure 23 : Ebauche $2^{1/2}$ D d'une tasse. A cette étape, l'image est représentée par les différentes orientations de la surface. Les différentes orientations de la surface sont décrites par des cercles imaginaires et l'inclinaison du vecteur à l'intérieur de chaque cercle.

La représentation en modèle 3D, qui constitue la dernière étape du modèle de Marr (1982), permet la reconnaissance d'objets. Un axe est tout d'abord défini au centre de

l'objet, puis un cylindre est déterminé autour de cet axe. Les objets peuvent alors être décrits par un ou plusieurs « cylindres généralisés ». Ainsi, lorsqu'un objet est composé de plusieurs cylindres, la représentation comportera plusieurs axes. Par exemple, un corps humain peut être décrit comme un ensemble de cylindres généralisés correspondant au tronc, à la tête, aux bras et aux jambes. Ces différents composants, représentés par des cylindres ont leurs propres axes et contribuent à l'axe global du corps humain. Un exemple est montré à la Figure 24. Le système de description de Marr (1982) a surtout été développé pour la reconnaissance d'objets chez des machines, et aucune validation expérimentale n'a été réalisée chez l'homme ou l'animal. Biederman (1987) a enrichi le modèle de Marr en apportant des données empiriques.

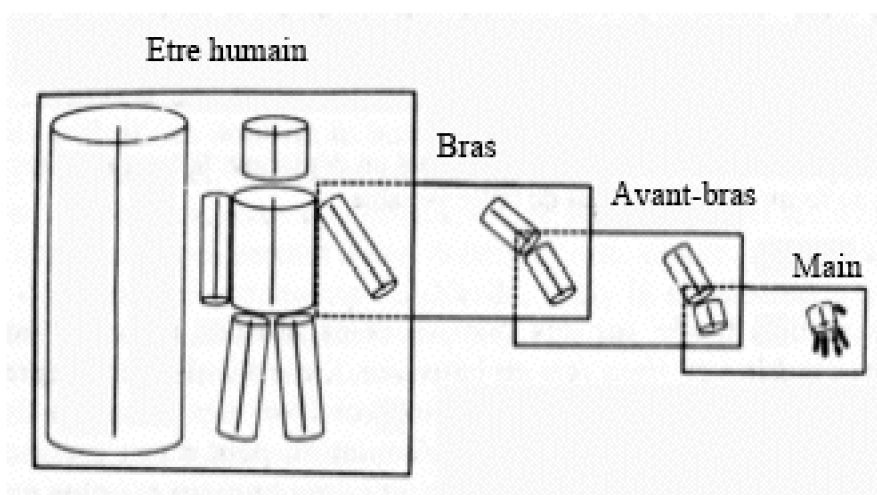


Figure 24 : Modèle hiérarchique du stade de la représentation en modèle 3D pour un être humain. Chaque rectangle montre l'axe principal de l'objet (à gauche), et les axes de ses composants (à droite).

Marr (1982) a formulé une hypothèse sur la localisation corticale des trois étapes de son modèle. L'aire V1 serait la région de l'ébauche primitive, tandis que l'ébauche en 2,5D et la représentation 3D seraient liées aux aires V4, MT et IT. Toutefois, le modèle de Marr (1982) n'explique pas comment se font les différentes transitions entre les trois étapes décrites (Pinker, 1984). Pour Palmer (1999), le processus de remplissage d'une surface n'est pas incompatible avec le modèle de Marr (1982), car ce processus pourrait survenir à une étape intermédiaire avant l'élaboration d'une représentation volumétrique complète.

2.4.2.L'approche physiologique et computationnelle de Grossberg

Grossberg et Mingolla (1985) ont proposé une théorie afin d'expliquer l'extraction des contours réels et illusoires, et les interactions qui existent entre les contours et la luminosité d'une surface. Dans l'introduction de leur article, Grossberg et Mingolla (1985) précisent que leur modèle tente d'expliquer le processus de remplissage dans le cadre de la tache aveugle, des scotomes, des images stabilisées ou encore des figures à contours illusoires. Par la suite, ce modèle a pu se généraliser au traitement de la surface des formes (Grossberg, 1994). Par ailleurs, la première version du modèle est limitée à une représentation monoculaire qui conduit à une perception en 2D. La présentation du

modèle de Grossberg décrite dans cette section se centre essentiellement sur des aspects physiologiques.

La théorie de Grossberg et Mingolla (1985) suppose l'existence de deux systèmes distincts dont le traitement s'effectue en parallèle. Ces deux systèmes sont appelés le « Boundary Contour System » (ou BCS) et le « Feature Contour System » (ou FCS). Le BCS génère les contours (réels ou illusoires), et le FCS induit un processus de remplissage. Pour Grossberg et Mingolla (1985), les contours (BCS) ne produisent pas un percept visible à l'œil humain, ils deviennent seulement perceptibles à partir du moment où ils restreignent le remplissage. Dans ce sens, tous les contours sont invisibles et la visibilité d'une forme est une propriété du remplissage de la surface (FCS).

Le BCS est construit à partir des champs récepteurs, et est donc sensible à l'orientation et au contraste mais pas à la direction¹⁴ du contraste. Ce premier niveau de traitement permet d'estimer la configuration globale des éléments d'une scène à travers deux interactions basées sur la position spatiale des champs récepteurs, soit des CR proches ou « short-range interaction », soit des CR plus éloignés ou « long-range interaction ». La première s'explique par l'activation des cellules par la présentation d'une double barre de type noir/blanc d'orientation fixe, qui va aboutir à inhiber les cellules qui codent la même orientation et qui sont proches. La seconde interaction se produit sur une distance plus importante. Cette interaction se traduit par une coopération entre les cellules qui répondent à la même orientation. Pour Grossberg et Mingolla (1985), ce dernier processus permettrait d'expliquer la formation des contours réels et illusoires (en particulier le carré de Kanisza).

Le FCS traite les contrastes ainsi que la direction des contrastes d'une forme, et l'orientation des contrastes n'est pas traitée par ce système. La sensibilité à la direction du contraste détermine la couleur qui est à l'intérieur ou à l'extérieur d'une forme. Le BCS active un processus de complétion de contours, et le FCS permet la propagation du mécanisme de remplissage. Grossberg et Mingolla (1985) ont émis l'hypothèse selon laquelle le mécanisme de remplissage se produisait au sein d'un syncytium de cellules. Par syncytium de cellules, Grossberg et Mingolla (1985) supposaient que les cellules étaient connectées entre elles dans des compartiments, ce qui permettait aux cellules contiguës de transmettre rapidement l'information entre les membranes. Le BCS a également une seconde fonction qui consiste à limiter la diffusion du remplissage aux contours de la forme. Ainsi, l'extraction des contours et le processus de remplissage permettent de finaliser un percept final.

Grossberg et Mingolla (1985) ont émis des prédictions sur la localisation physiologique de ces deux systèmes, et une mise à jour de ces hypothèses a ensuite été réalisée par Grossberg, Mingolla et Ross (1997) ainsi que Grossberg (2000). Les contours seraient élaborés dans une voie CGL-Blob- Inter bande, alors que la surface serait formée à partir de la voie : CGL-Interblob-Bande fine. Les projections, qui sont faites dans les régions blob et inter blob de V1, proviennent de la voie parvocellulaire du CGL. Par ailleurs, la voie magnocellulaire serait impliquée dans le traitement des contours en mouvement (Grossberg, 1994).

¹⁴ La direction du contraste est définie comme le changement de polarité du contraste : noir/blanc vs. blanc/noir.

Grossberg (1994) a par la suite étendu sa théorie à la vision 3D, c'est le modèle FACADE (« form and color and depth »). Pour obtenir une perception en 3D, Grossberg (1994) prend en compte la vision binoculaire. L'objectif est de mettre en lien les traitements BCS-FCS obtenus de façon monoculaire pour parvenir une représentation 3D. Dans son modèle, une première représentation est déterminée par le traitement BCS monoculaire, et par la suite, une interaction est réalisée entre les deux représentations BCS pour aboutir une séparation binoculaire de la scène. La représentation BCS binoculaire interagit avec le domaine de remplissage monoculaire (« filling in domain » ou FIDO) et est obtenue par les deux systèmes FCS. Par la suite, la représentation 3D du modèle FACADE est déterminée par l'interaction entre les deux domaines de remplissage et la représentation binoculaire du BCS. D'un point de vue physiologique, l'interaction finale se produirait dans l'aire V4. Cette région jouerait un rôle clé dans la séparation figure-fond, mais aussi dans la reconnaissance d'objets.

Des données empiriques (Dresp & Grossberg, 1997 ; 1999) viennent étayer le modèle portant sur les contours illusoires de Grossberg et Mingolla (1985). Toutefois, l'expérience de Mendola, Dale, Fischl, Liu et Tootell (1999) a démontré que des aires visuelles de haut niveau étaient fortement impliquées dans le traitement de contours illusoires. Cette étude ne remet pas en cause ce modèle, mais la relation entre les aires visuelles et les différentes étapes de traitement du modèle. Ainsi le modèle de Grossberg et Mingolla (1985) doit tenir compte de ces données, et notamment par rapport aux hypothèses physiologiques que Grossberg (2000) a émis en fonction du modèle FACADE. Dans le chapitre précédent, nous avons vu que l'aspect temporel était un facteur important dans l'intégration des contours (Beaudot, 2002 ; Hess, Beaudot & Mullen, 2001), or, le modèle de Grossberg (1994) mais aussi celui de Marr (1982) ne tiennent pas compte de cet élément.

2.5. Arguments comportementaux en faveur d'un traitement hiérarchique

L'objectif de cette partie est de présenter un ensemble d'études comportementales démontrant un traitement hiérarchique du contour et de la surface par l'intermédiaire du remplissage. La présentation de ces études permettra de confronter les données physiologiques qui ont été abordées précédemment mais également les modèles théoriques de Marr (1982) et de Grossberg et al. (Grossberg, 1994, 2000 ; Grossberg & Mingolla, 1985 ; Grossberg, Mingolla & Ross, 1997). Deux parties seront distinguées dans les études comportementales, tout d'abord une présentation des études portant sur les contours illusoires sera réalisée, puis une présentation des études relatives aux surfaces uniformes et des textures terminera cette partie. Les études physiologiques ont permis de montrer que le remplissage n'était pas déterminé de façon explicite pour le traitement de la surface, la question est donc de savoir si cette ambiguïté se retrouve dans les études comportementales.

2.5.1. Mise en évidence d'un traitement hiérarchique à l'aide de figures illusoires

Les processus impliqués dans le traitement des contours et des surfaces ont principalement été étudiés sur des figures à contour illusoire (pour une revue voir Leshner, 1995). L'intervention d'un processus de remplissage pour le traitement des figures à contours illusoires a été émise par Day et Joris (1980). On peut supposer que dans le cadre des contours illusoires, et conformément au modèle de Grossberg et Mingolla (1985), les bords illusoires permettraient une propagation de la luminance au même titre que les bords réels. Les expériences, présentées dans cette partie, testent directement l'hypothèse d'un traitement hiérarchique à l'aide des figures à contours illusoires.

Dresp et Bonnet (1991) ont déterminé le seuil de détection d'un point lumineux pour différentes positions : à l'intérieur, à l'extérieur, ou sur le contour d'une figure illusoire (le carré de Kanizsa). Les résultats ont montré que les seuils de détection étaient plus bas lorsque le point était situé à l'intérieur du carré de Kanizsa qu'à l'extérieur. De plus, le seuil diminuait lorsque la distance entre le point et le contour illusoire augmentait. Par ailleurs, lorsque le point était placé sur le contour même, le seuil était beaucoup plus faible par rapport à un seuil déterminé pour un point placé au centre de la figure illusoire. Des résultats identiques ont été observés lorsque la taille du carré de Kanizsa augmentait, et ce que le point soit situé à l'intérieur ou à l'extérieur (Dresp & Bonnet, 1993). Ces résultats montrent que le processus de remplissage est tardif par rapport au traitement du contour. Un modèle mathématique du processus de remplissage pour les contours à figure illusoire a été réalisé par Brigner (1980). Dans ce modèle, la luminosité de la surface est déterminée à partir des bords, par le produit scalaire de quatre vecteurs : les contours illusoires verticaux et horizontaux.

L'ensemble des études physiologiques (Lamme & Roelfsema, 2000 ; Lamme, Supér & Spekreijse, 1998 ; Macknik, Martinez-Conde & Haglund, 2000) et des modèles (Grossberg & Mingolla, 1985 ; Marr, 1982) supposent l'existence de deux systèmes impliqués dans la reconnaissance d'une forme visuelle : un système rapide traitant les contours et un système plus lent traitant la surface. Dans cette perspective Rogers-Ramachandran et Ramachandran (1998) ont utilisé une nouvelle classe de stimuli appelée les « contours fantômes ». Ce stimulus est composé d'un fond uniforme gris sur lequel est superposé des points blancs sur la partie supérieure de la figure et, des points noirs dans la partie inférieure. Puis, après un laps de temps très court, la luminance des points est inversée, les points blancs deviennent noirs et réciproquement. Les observateurs ont alors la perception d'un bord horizontal qui sépare les deux parties du stimulus, alors qu'aucun élément physique ne permet de distinguer l'une ou l'autre des deux parties. Cette observation a été confirmée dans une expérience où les sujets effectuaient une tâche de discrimination entre un bord horizontal ou vertical (les points blancs sont situés à droite puis à gauche, et inversement pour les points noirs). La moyenne pour l'ensemble des observateurs était alors de 99.4 % de discrimination correcte. Pour confirmer le fait que cette tâche permettait de juger la capacité des observateurs à discriminer les contours, Rogers-Ramachandran et Ramachandran (1998) ont superposé à leur stimulus une croix blanche qui séparait les points blancs des noirs. L'observateur déterminait les régions contenant des points identiques : verticale, horizontale ou en diagonale. Les performances de discrimination chutaient à 31 % de discrimination correcte. Dans cette condition, le contour fantôme est supprimé et la

réponse des observateurs est essentiellement basée sur la dimension de surface. Par ailleurs, la tâche était correctement réalisée à partir du moment où les observateurs ajustaient les temps de présentation des stimuli. Dans cette condition, la fréquence de présentation passait de 15 à 7 Hz. Ce dernier résultat indiquait que les caractéristiques de surface (les régions de point) étaient perçues uniquement pour des temps de présentation plus importants. Rogers-Ramachandran et Ramachandran (1998) ont proposé que ces deux temps de présentation différents reflètent l'implication de deux systèmes visuels distincts : un traitement rapide pour l'extraction des contours, et un système plus lent pour le traitement de la surface. D'autres expériences ont indiqué que les performances de discrimination des contours étaient identiques lorsque deux niveaux de gris étaient utilisés. De plus, les performances augmentaient quand les stimuli étaient présentés en vision périphérique ou dans une condition où les points étaient lissés. De même, les performances diminuaient quand les points étaient de couleurs différentes (rouge et vert). Au vue des résultats pour l'ensemble de ces conditions, les auteurs ont supposé que les propriétés sous-jacentes à la visibilité des contours, pour cette illusion, se rapprochaient des caractéristiques de la voie magnocellulaire. Ainsi, les deux systèmes que Rogers-Ramachandran et Ramachandran (1998) ont identifiés peuvent correspondre au système magnocellulaire et parvocellulaire.

L'existence d'un traitement hiérarchique a également été mis en évidence dans l'effet COCE par Davey, Maddess et Srinivan (1998). Une des hypothèses émises sur le mécanisme de cette illusion avance que le système visuel induit une diffusion de la luminosité à partir des contours entre chaque région par un mécanisme de remplissage latéral. L'objectif de leur expérience était de montrer que si un processus de remplissage existe dans cette illusion, alors ce mécanisme serait traité dans un espace fini (délimité par les contours), et par conséquent, des caractéristiques temporelles pourraient être déterminées. En d'autres termes, il s'agit de définir la vitesse du remplissage.

Un réseau COCE était présenté et les observateurs fixaient la région centrale. Par exemple, les barres présentées dans la Figure 25 apparaissent alternativement claires et sombres, notamment dans leur région centrale. Toutefois, le profil, indiqué par la courbe noire, montre que contrairement à ce que nous percevons, les régions centrales des barres, c'est-à-dire les surfaces, sont équiluminantes¹⁵. Si un mécanisme de remplissage est impliqué dans cette illusion, alors les variations de luminosité dépendent de la vitesse de remplissage, et de la distance entre le point fixé par l'observateur et le contour. Les fréquences temporelles et spatiales sont manipulées. Par conséquent, le contraste du bord est inversé continuellement, ce qui implique que les zones « sombres-claires » sont inversées en « claires-sombres ». Un changement rapide d'inversion du contraste produit une modification rapide de la luminosité de la surface. Toutefois, si la vitesse du changement de contraste augmente énormément, il y a une limite pour laquelle les variations de luminosité ne se produisent plus. En d'autres termes, quand l'observateur fixe l'une des bandes de ce réseau, il ne devrait plus percevoir de changement de luminosité. La fréquence temporelle qui induit cet effet est appelée le seuil limite.

¹⁵ Purves, Shimp et Lotto (1999) ont montré que la différence de luminance entre deux régions adjacentes dans l'effet COCE, était plus importante lorsque des stimuli tri-dimensionnels étaient utilisés.

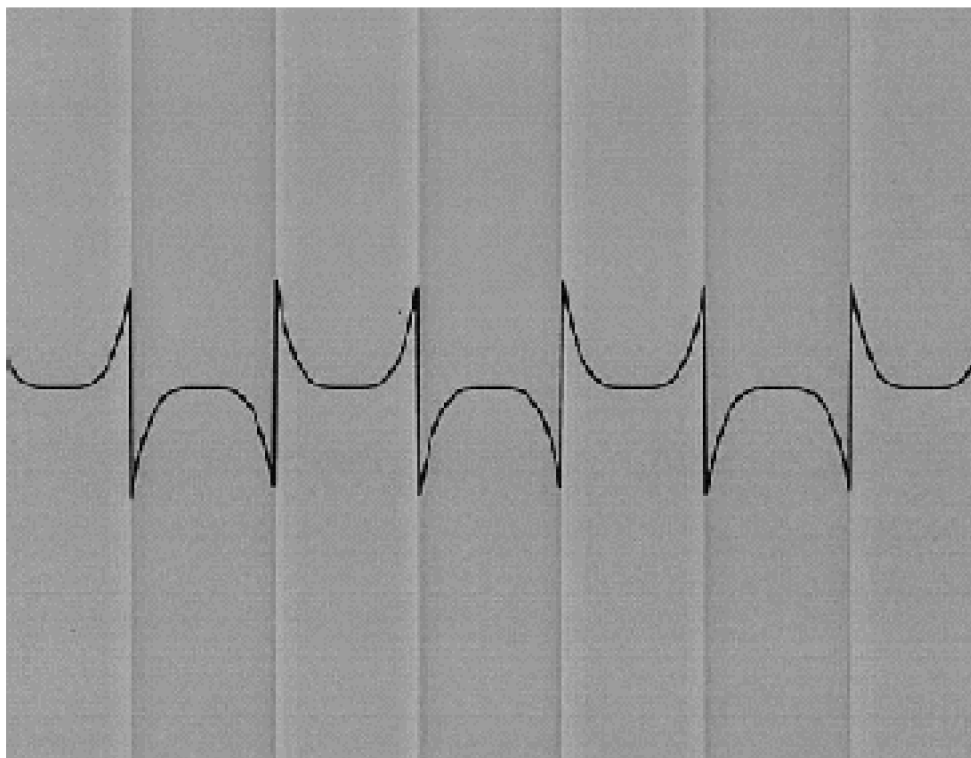


Figure 25 : Exemple d'un réseau COCE. Le profil (en ligne continue) du réseau a été ici superposé à la figure illusoire. On peut s'apercevoir que malgré les différences de luminosité entre les barres, les zones centrales des barres sont égales.

La tâche des observateurs était d'ajuster la luminance d'un réseau sinusoïdal ¹⁶ à celui d'un réseau COCE pour obtenir un niveau identique. Les fréquences temporelles des deux réseaux étaient synchronisées. Les résultats ont indiqué que plus l'espacement entre les bords est faible (les fréquences spatiales élevées) et plus le seuil limite des fréquences temporelles était important. De plus, lorsque l'espacement entre les bords est important (des fréquences spatiales basses), le seuil limite des fréquences temporelles est faible. Ce résultat implique une relation linéaire entre la fréquence temporelle et la fréquence spatiale (voir Figure 26). Un résultat identique a été trouvé dans une seconde expérience dans laquelle les auteurs ont présenté un réseau COCE avec un faible contraste. L'ensemble de ces résultats est cohérent avec l'hypothèse selon laquelle un mécanisme de remplissage est impliqué dans l'illusion COCE. En effet, plus la distance entre les contours du réseau est petit et plus le remplissage s'effectue rapidement (fréquences temporelles élevées), et inversement, lorsque la distance entre les contours est importante le remplissage est plus lent (fréquences temporelles basses). Les auteurs ont estimé que la propagation de l'information de surface se produisait pour une moyenne de 22.44 deg/sec (+/- 6.54 deg/sec selon les observateurs).

¹⁶ Dans le cas d'un réseau sinusoïdal, la différence de luminosité entre les bandes claires et sombres existent réellement.

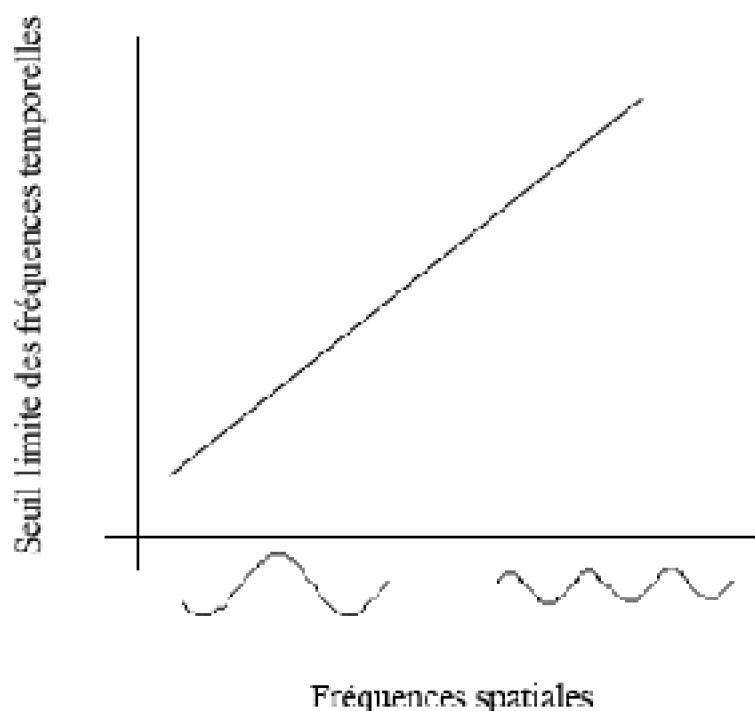


Figure 26 : Relation obtenue dans la première expérience de Davey, Madess et Srinivan (1998) entre le seuil des fréquences temporelles pour différentes fréquences spatiales.

Par ailleurs, l'induction de luminosité ne se limite pas à des conditions achromatiques. Pinna, Brelstaff et Spillman (2001) ont mis en évidence un processus de remplissage à l'aide d'une nouvelle illusion comportant de la couleur : l'effet « watercolor ». L'illusion est composée par de nombreuses variabilités au sein d'une ligne (ressemblant à une sinusoïde) et elle est présentée comme un boustrophédon, avec d'un côté une couleur pourpre et de l'autre une couleur orange. La surface confinée entre ces contours n'apparaît pas blanche, mais prend la teinte de la ligne colorée. De plus, cet effet augmente lorsque la fréquence spatiale de la sinusoïde augmente, et cet effet diminue lorsque l'espace inclus entre les variations de la sinusoïde augmente. Ce dernier résultat est un argument supplémentaire en faveur d'un mécanisme de propagation qui se produirait à partir des contours. Les auteurs ont également constaté qu'il n'y a pas de couleur particulière qui limite cet effet.

2.5.2. Mise en évidence d'un traitement hiérarchique à l'aide de figures naturelles

L'objectif de cette partie est de présenter une revue des données sur le traitement de remplissage en dehors des figures à contour illusoire. En effet, l'utilisation des figures à contours illusoires est critiquable pour deux aspects. La première critique provient des données obtenues en IRMf (Hirsch, DeLaPaz, Relkin, Victor, Kim, Li, Borden, Rubin & Shapley, 1995 ; Mendola, Dale, Fischl, Liu & Tootell, 1999), présentées dans la première partie. A partir des résultats trouvés dans ces études, on peut supposer que les figures à contours illusoires ne sont pas considérées de la même façon que les formes réelles,

puisqu'il existe une controverse sur les aires visuelles impliquées dans le traitement de ces figures. La seconde critique est de nature écologique, et pose la question suivante : comment la perception peut-elle s'expliquer à partir de figures illusoires, alors qu'il existe très peu d'illusions dans les scènes visuelles de la vie courante (Spillman & Werner, 1996) ? Pour répondre à cette question, l'ensemble des expériences, qui seront présentées au cours de cette partie, concerne des formes non illusoires avec une surface uniforme. De plus, la présentation de ces études nous permettra de faire le lien avec les études physiologiques qui utilise des stimuli identiques. Par ailleurs, cette partie est primordiale pour nos expériences, car nous utiliserons des stimuli uniformes dans le cadre du traitement contour/surface.

2.5.2.1. Cas d'une surface uniforme

Le traitement hiérarchique du contour et de la surface a été démontré dès les années 60 à l'aide du paradigme des images stabilisées. L'objectif premier de cette technique était d'éliminer les mouvements oculaires, ce qui permettait ainsi d'observer leur rôle en perception visuelle. Pour supprimer les mouvements, l'observateur porte une lentille de contact reliée à un projecteur permettant d'envoyer une image sur la rétine. Il est possible de stabiliser une image sur la rétine de telle manière que, dès que l'œil bouge, l'image se déplace avec lui. L'utilisation de cette méthode a permis de montrer que le système visuel construit des régions dans le champ visuel uniquement à partir de l'information de contour. Ce résultat a été observé par Krauskopf (1963). Dans cette expérience, un cercle vert était entouré par un anneau rouge, et la frontière entre le rouge et le vert était stabilisée sur la rétine. Pour les observateurs, le disque vert central disparaissait et était rempli par l'anneau rouge, ce qui créait la perception d'un seul disque rouge. Cette expérience indiquait que d'une part, les mouvements oculaires participaient à la détection du contraste entre deux régions, et d'autre part, l'information de contour détermine la perception d'une surface. En effet, à partir du moment où la frontière entre l'anneau et le disque est stabilisée, ce contour disparaît, et la perception de la surface également. Cette expérience démontrait l'importance de l'information de contour pour le traitement de la surface. En utilisant des réseaux sinusoïdaux, Blakemore, Muncy et Ridley (1971) ont mis en évidence que le phénomène des images stabilisées n'était pas purement rétinien, mais pouvait être également d'origine corticale.

Par ailleurs, l'importance de ce traitement hiérarchique est également démontré avec l'utilisation d'une figure classique : à savoir les bandes de Mach. Le stimulus se présente sous la forme d'un dégradé achromatique allant du blanc au noir. Dans ce cas, une bande fine supplémentaire est perçue de façon plus claire que la surface lumineuse (à gauche) ; tandis qu'une bande plus foncée équivalente est perçue pour la surface sombre (à droite). Ces deux effets sont observés sur les bords des deux surfaces. Un exemple est présenté à la Figure 27a. La Figure 27b montre le profil de luminance affichée, alors que la Figure 27c indique le profil de luminosité classiquement perçue. Ce phénomène est similaire à l'effet COCE, mais ici les contours ne sont pas physiquement rehaussés. Ce type de stimuli n'implique pas un processus de remplissage dans l'intégration de la surface, mais démontre l'importance du traitement des contours (et plus particulièrement les contrastes) dans la perception de surface.

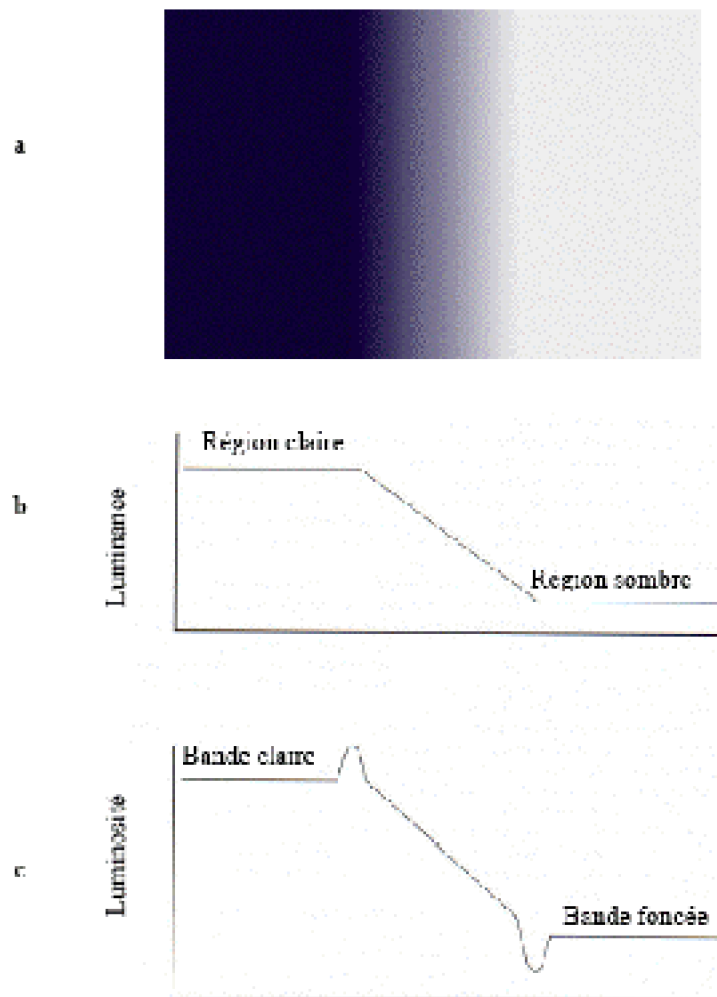


Figure 27 : Les bandes de Mach sont présentées à la figure 27a. La figure 27b présente le profil de luminance réel, et la figure 27c indique le profil de luminosité (l'exemple est issu de Schwartz, 1999, pp 194).

Toutefois, l'utilisation des images stabilisées permet certes, de démontrer l'existence d'un traitement hiérarchique, mais pas de façon binoculaire. De même, les bandes de Mach ne sont pas une démonstration de l'existence d'un traitement de remplissage. Par conséquent, l'hypothèse du remplissage a été testée à l'aide du paradigme de masquage. La première étude est celle de Stoper et Mansfield (1978) qui utilise un paradigme de métacontraste. Dans les expériences traditionnelles de métacontraste, un disque (la cible) est suivi par un anneau plus grand (le masque), et dans ce cas, l'anneau empêche de percevoir la cible. Une explication de ce phénomène est que le masque perturbe le traitement du contour de la cible ce qui bloque le traitement de sa surface. Dans l'expérience de Stoper et Mansfield (1978), pour empêcher que le masque n'interfère avec le contour de la cible, il était plus petit qu'elle. Les résultats ont montré que le contour de la cible était cette fois-ci perçu, mais les observateurs ne percevaient toujours

pas la surface. Le processus de remplissage, distinct du processus de codage du contour, a donc été sélectivement perturbé. Ces données sont cohérentes avec l'idée que le contour est traité par un système rapide précoce, alors qu'un système plus lent et plus tardif traite la surface.

L'étude de Paradiso et Nakayama (1991) va dans la continuité des expériences menées par Stoper et Mansfield (1978). L'objectif de leur étude était de savoir, plus précisément, si l'information de contour participait à la perception d'une région homogène. Leur raisonnement était le suivant : si le processus de remplissage implique une activité de propagation, on peut démontrer son existence en l'interrompant. En d'autres termes, si les bords interrompent le remplissage, que se produit-il quand de nouveaux bords sont introduits ? Est-ce que le processus de remplissage est affecté avant qu'il ne soit complet ?

Paradiso et Nakayama (1991) ont tout d'abord réalisé une observation préliminaire dans laquelle l'expérience de Stoper et Mansfield (1978) était répliquée, mais cette fois-ci avec différents masques. Les résultats étaient identiques (cf. Figure 28) à ceux trouvés par Stoper et Mansfield (1978). Ces observations impliquent que le remplissage d'une surface peut être interrompu par la présentation d'une nouvelle forme (le masque).

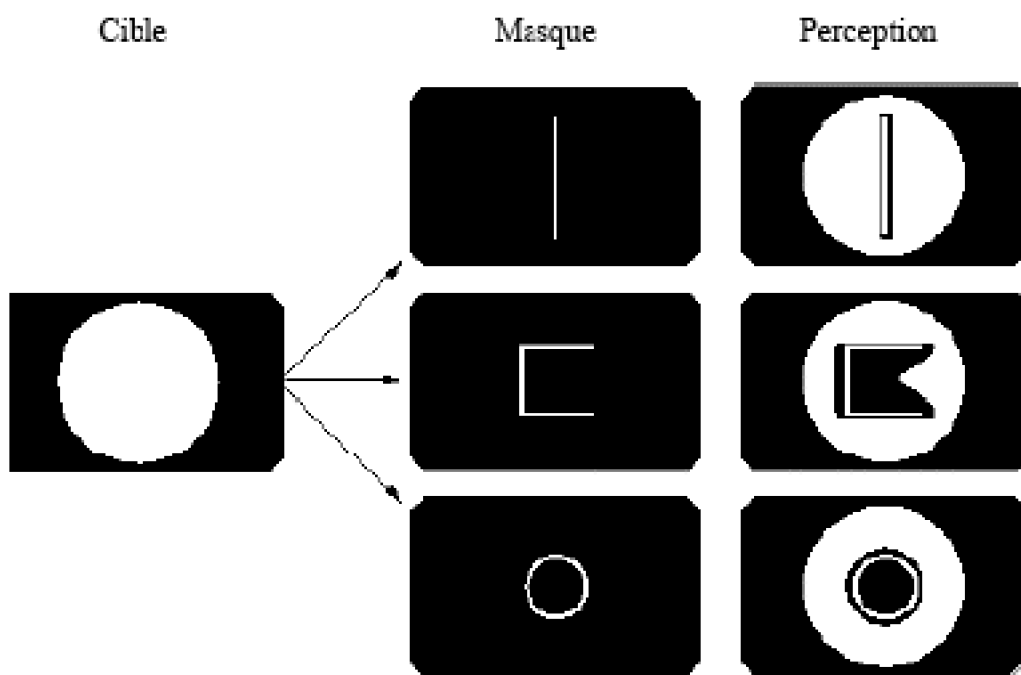


Figure 28 : Résultats des observations obtenues dans l'expérience de Paradiso et Nakayama (1991). La figure de gauche indique la cible utilisée. Les trois figures du centre montrent les différents masques utilisés, et enfin les figures de droite présentent ce que les observateurs percevaient à la suite de la présentation de la cible et du masque (l'exemple provient de l'article de Paradiso et Nakayama, 1991).

A la suite des observations préliminaires, une expérience a été réalisée dans laquelle une cible, à savoir un disque blanc sur fond noir, était masquée par un disque blanc plus petit. La taille de la cible variait de 2.3 à 3.4 deg, alors que la taille du masque était

constante (2 deg). La tâche de l'observateur était de juger, sur une palette de plusieurs niveaux de gris, la luminosité du centre de la cible. Deux principaux résultats ont été observés. Tout d'abord, les observateurs percevaient une diminution de la luminosité dans la zone de la cible se situant à l'intérieur du masque, mais également à proximité de la zone extérieure au masque. De plus, la suppression de la luminosité augmentait lorsque la taille de la cible augmentait. Ces résultats confirment que le remplissage est interrompu par les contours, mais il indique en outre que cet arrêt se produit graduellement à proximité du nouveau contour. Mais ceci est surtout vrai pour des délais de présentation entre le masque et la cible de -50 à +100 msec. Les auteurs ont par la suite estimé la vitesse de remplissage de la cible. Pour cela, ils ont affiché le délai entre la cible et le masque (msec) par rapport à la distance entre la cible et le masque (deg). Les résultats ont montré que le processus de remplissage se produisait pour un délai de 110 à 150 deg/sec.

Paradiso et Nakayama (1991) ont réalisé une seconde expérience où un disque segmenté était utilisé en tant que masque : la suppression variait de 0 (masque complet) à 360 degrés (pas de masque). Le processus de remplissage se produisait aussi lorsque le masque était segmenté. Ainsi, ces résultats, tout comme ceux observés dans l'expérience principale, reflètent l'intervention d'un processus de remplissage, qui commence dès les contours de la cible et qui continue jusqu'à ce que d'autres contours arrêtent ce phénomène. Ces résultats ne semblent donc pas être compatibles avec la théorie BCS/FCS de Grossberg et collaborateurs dans laquelle les contours arrêtent immédiatement le remplissage de la surface. Pour Arrington (1994), cette critique n'intègre pas les récents changements de la théorie BCS/FCS (Grossberg & Todorovic, 1988). Il montre par une simulation informatique que le modèle BCS/FCS permet de simuler les données des expériences de Paradiso et Nakayama (1991), pourvu que ces changements soient pris en compte. Cette étude conforte donc la thèse selon laquelle les contours n'arrêtent pas directement le remplissage de la surface, mais le font graduellement.

Le phénomène de remplissage a également été démontré en utilisant d'autres techniques que le paradigme de masquage. Dans une observation préliminaire, Paradiso et Hahn (1996) ont montré que lorsque la luminance d'un disque augmentait, la luminosité de ce disque n'était pas homogène. En effet, les observateurs percevaient soit une luminosité différente sur les contours du disque par rapport à la zone centrale, soit une zone sombre qui se remplissait vers l'intérieur. Les auteurs ont donc testé de façon qualitative ce phénomène en manipulant le délai entre chaque augmentation ou diminution de la luminance du signal. Le temps de présentation global du signal était toujours identique. Dans leurs expériences, les observateurs avaient pour consigne d'évaluer la force de leur sensation sur une échelle. Les résultats ont indiqué que pour des délais brefs (0.16 msec) ou longs (2 sec), aucune perception de remplissage n'était perçue. Mais la perception de ce phénomène était maximisée pour des durées intermédiaires (0.25 à 0.5 msec). De plus, la perception du remplissage était plus importante lorsque la luminance du signal diminuait. Des résultats identiques ont été observés quand le disque était placé autour de la tache aveugle. Ce dernier résultat suggère que la perception du remplissage est réalisée à partir de l'activité des neurones

du cortex visuel. Par la suite, Paradiso et Hahn (1996) ont manipulé le temps de présentation global du signal : le délai de présentation entre chaque stimulus était toujours identique et a été choisi pour maximiser le processus de remplissage. Les résultats ont montré que plus le temps de présentation était long, et moins les observateurs percevaient le phénomène du remplissage.

Pour Paradiso et Hahn (1996), le mécanisme de remplissage aurait un rôle fondamental dans la perception de la luminance. En effet, la luminance perçue au centre d'une région uniforme pourrait être le résultat d'une interaction latérale entre les neurones. Ainsi, le signal serait traité à partir du contraste entre deux bords, et ce contraste influencerait progressivement les zones plus éloignées du bord. Le traitement neuronal du remplissage déterminerait la luminosité. Le problème de cette interprétation est qu'elle n'explique pas pourquoi les observateurs percevaient une surface homogène pour les délais inter stimuli les plus brefs. Les auteurs supposent que ce phénomène pourrait exister durant cette période, mais il se produit trop rapidement pour être perçu. Ce qui expliquerait pourquoi nous ne percevons pas de surface non uniforme dans la vie courante. De même, pour les délais inter stimuli les plus longs, à chaque fois que la luminance augmente, le traitement de remplissage est effectué avant la prochaine augmentation de luminance. Ainsi, le remplissage d'une surface existe, mais ce processus est trop bref pour être perçu.

La luminosité d'une surface dépend de paramètres spatiaux, comme le contraste au bord d'une figure ou encore la taille de la figure (Schwartz, 1999). Par ailleurs, la luminosité d'une surface est également influencée par d'autres caractéristiques, et notamment les informations contextuelles. Par exemple, un carré gris sur un fond clair apparaît plus sombre qu'il ne l'est réellement. Cet effet est inversé lorsque le même carré est présenté sur un fond sombre. Pour tenir compte de l'ensemble de ces paramètres, Rossi et Paradiso (1996) ont utilisé un réseau sinusoïdal. Les fréquences temporelles et spatiales de ce réseau étaient manipulées. La tâche de l'observateur était d'ajuster la fréquence temporelle à la plus petite fréquence pour laquelle la luminosité de la bande centrale semble arrêter la variation de luminance entre les bandes claires et sombres du réseau. Les résultats ont indiqué que le seuil des fréquences temporelles diminuait lorsque la taille des bandes augmentait (les fréquences spatiales basses). Des résultats similaires ont été observés lorsque les observateurs ajustaient la luminosité d'un rectangle par rapport à la luminosité de la bande centrale du réseau. Ces résultats supposaient que l'induction de luminosité était plus lente pour des grandes tailles. Cette limite est en accord avec l'idée qu'un mécanisme de remplissage se produit depuis les bords jusqu'à l'intérieur de la figure.

De nombreuses études physiologiques (Komatsu, Murakami & Kinoshita, 1996 ; MacEvoy, Kim & Paradiso, 1998 ; Rossi & Paradiso, 1999 ; Rossi, Rittenhouse & Paradiso, 1996) ont cherché à déterminer l'existence d'un processus de remplissage d'une surface uniforme à partir de l'enregistrement des neurones chez le singe comme nous l'avons présenté en détail dans une partie précédente. L'une des premières questions que l'on pourrait se poser est de savoir si les singes sont sensibles aux mêmes phénomènes que l'observateur humain. Une étude comportementale a ainsi été réalisée afin de déterminer si le phénomène de remplissage était généralisable pour l'ensemble

des primates (Friedman, Zhou & von der Heydt, 1999). Les stimuli étaient composés d'un cercle et d'un anneau, et deux couleurs étaient utilisées alternativement (rouge et vert). Les bords du cercle étaient correctement définis ou lissés. Par ailleurs, le disque était statique ou en mouvement. La tâche consistait à appuyer sur une touche dès que le stimulus apparaissait comme uniforme. L'expérience était réalisée par des singes, mais également par des observateurs humains. Des essais contrôle étaient effectués pour lesquels la couleur du disque changeait régulièrement. Cette condition contrôle permettait de vérifier que les singes réalisaient correctement la tâche en ne répondant pas de façon systématique au bout d'un certain temps. Les premières réponses du singe ont été enregistrées 3-4 sec après la présentation du stimulus. Un plus grand nombre de réponses était donné lorsque le stimulus était composé d'un cercle statique et lissée. De plus, les singes percevaient le stimulus comme étant moins uniforme quand le disque était lissé sur les bords et en mouvement, ou statique et avec des contours bien définis. Très peu de réponses ont été données lorsque le cercle avait une frontière fixe et en mouvement. Des résultats similaires ont été obtenus chez l'observateur humain. Les résultats suggèrent donc que les singes perçoivent le processus de remplissage de la même façon que les humains.

L'ensemble de ces études ont montré la participation du processus de remplissage dans la perception de luminance. La propagation du signal neuronal a été ici mis en évidence en le stoppant (Paradiso & Nakayama, 1991 ; Stoper & Mansfield, 1978), ou en manipulant le temps de présentation (Paradiso & Hahn, 1996) ou la taille (Rossi & Paradiso, 1996) du signal. Le phénomène du remplissage est un processus rapide qui se produit à une vitesse de 110-150 deg/sec pour la propagation de la luminosité (Paradiso & Nakayama, 1991). Ce résultat est différent de la vitesse déterminée par Davey, Madess et Srinivan (1998) qui était de 22.44 deg/sec. Cette différence pourrait être liée à l'utilisation de figures illusoires (Davey, Madess & Srinivan, 1998) ou non (Paradiso & Nakayama, 1991).

Par ailleurs, une distinction contour/surface a également été réalisée pour l'activation des représentations d'objets. Biederman et Ju (1988) ont testé l'implication de la dimension de surface dans les représentations des objets. Dans une de leurs expériences, les observateurs devaient dénommer des objets constitués de dessins aux traits ou de photos en couleur des mêmes objets. Les résultats ne montraient pas de différence entre les temps de dénomination pour ces deux conditions, ce qui suggère que l'information de surface n'apporte pas d'information supplémentaire pour l'identification d'objet. Un résultat similaire a été trouvé par Ostergaard et Davidoff (1985) dans une tâche de catégorisation vivant/non vivant pour des dessins d'objet et des photos.

2.5.2.2. Cas de la texture

Dans la partie précédente, une revue de la littérature a été effectuée afin de montrer l'importance du processus de remplissage dans le cas de figure non illusoire. La question est ici de savoir si le mécanisme de remplissage s'applique à une surface constituée par une texture ? En effet, un mécanisme basé sur une propagation latérale ne peut pas expliquer le remplissage d'une texture puisque la propagation devrait s'arrêter pour chaque élément qui la compose. Caputo (1998) fait l'hypothèse que le mécanisme de

remplissage, impliqué pour la texture, se produirait à un niveau de traitement où la texture, composée de plusieurs éléments, est représentée comme une surface continue. Dans cette étude, Caputo (1998) a utilisé une procédure de masquage. La cible était une texture composée de petits traits blancs sur un fond noir, le pourtour était constitué d'une texture identique dont les traits étaient orientés différemment, et le masque était un carré dont le contour était blanc et l'intérieur noir. Un exemple des stimuli utilisés dans cette expérience est présenté à la Figure 29. Lors d'une observation préliminaire, une texture était tout d'abord affichée sur l'écran, elle était ensuite suivie d'un fond noir, et pour terminer un masque était présenté. Cet ordre de présentation rendait les éléments de la texture, situés à l'intérieur du masque, plus clairs.

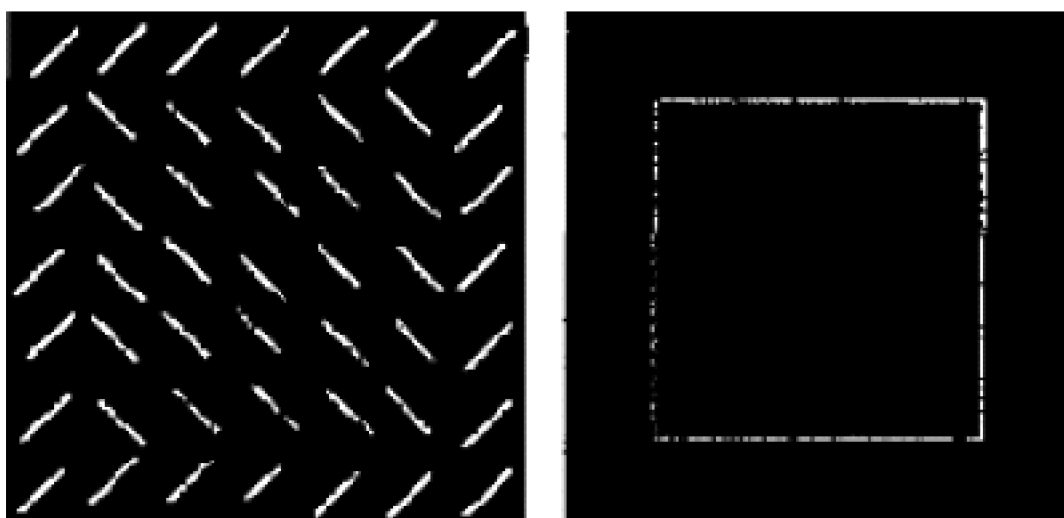


Figure 29 : La figure de gauche présente la texture et la figure de droite montre le masque utilisé dans l'expérience de Caputo (1998).

Par la suite, la procédure de leur expérience était en tout point identique excepté pour la tâche : les observateurs devaient ajuster la luminosité des traits, apparaissant à l'intérieur du masque, à la luminosité des traits de la cible. Les résultats ont montré que les observateurs avaient tendance à surestimer leurs ajustements avec des ISI (interval inter stimuli) courts (inférieur à 40 msec), tandis qu'avec des ISI longs, leurs ajustements étaient corrects. L'auteur fait l'hypothèse que deux étapes sous-tendent le remplissage de la texture. La première étape constitue la représentation d'une surface continue, dans laquelle seules les discontinuités des éléments composant la texture sont extraites. Puis dans une seconde étape, il y aurait une propagation de la luminosité des éléments qui compose la texture à partir de leurs propres bords.

Motoyoshi (1999) a répliqué les résultats de Caputo (1998) en utilisant une procédure totalement identique à celle de Paradiso et Nakayama (1991). La cible était une texture dont la forme globale était circulaire et était composée d'une texture de plusieurs traits noirs. Le masque était un cercle blanc avec un contour noir. Les résultats ont montré que la présentation du masque, successive à la cible, supprimait la perception de la texture de la cible à l'intérieur du masque. Ainsi, la suppression d'une aire par un masque ne se produit pas seulement pour un stimulus uniforme (Paradiso & Nakayama, 1991).

Les études comportementales, qui ont été décrites dans cette partie, présentent des résultats en faveur d'un traitement hiérarchique du contour et de la surface. De même, ces expériences démontrent l'existence d'un processus de remplissage dont le fonctionnement est de traiter une forme à partir des contours. Par ailleurs, une différence fondamentale existe entre les études physiologiques qui ne parviennent pas à démontrer directement l'existence du processus de remplissage dans le traitement de la surface, et les études comportementales qui vont dans le sens de cette hypothèse. Dans la thématique du traitement de contour et de la surface, notre approche sera novatrice puisque notre objectif est de déterminer les dimensions qui sont pertinentes pour l'observateur. De plus, nous envisageons d'utiliser une tâche de détection qui n'a jamais été réalisée jusqu'à présent dans les études comportementales utilisant une surface uniforme. En effet, les expériences comportementales, qui ont été présentées, emploient souvent des tâches de jugement qui favoriseraient plus l'utilisation de fonctions cognitives par les observateurs. De plus, lorsque nous avons abordé le paradigme des scotomes artificiels, nous avons vu qu'il existait un problème au niveau des temps de réponses. Cela nous a amené à distinguer un remplissage spécifique pour des conditions naturelles comme la tâche aveugle et pour ce paradigme. Nous pouvons donc faire l'hypothèse que le remplissage mis en évidence dans les études utilisant une surface uniforme ou une texture relève du même phénomène que celui des scotomes artificiels. Cette idée permettrait d'expliquer la dissociation entre le comportemental et le physiologique.

2.6. Implications des processus attentionnels

Les domaines de la vision et de l'attention sont souvent liés et pose la question de la précocité à laquelle la sélection attentionnelle s'opère. Les processus attentionnels sont souvent abordés comme un processus ascendant qui permet de faire ressortir une figure par rapport au fond. Cette partie présente comment la focalisation de l'attention modifie le remplissage et le traitement de la surface.

2.6.1. Evidences expérimentales de l'intervention des processus attentionnels dans le traitement du remplissage

L'objectif de cette partie est de montrer le rôle de l'attention dans le processus de remplissage. Brown et Thurmond (1993) ont déterminé si un facteur pré-attentionnel permettait la réalisation du processus de remplissage dans la tâche aveugle. Dans leur expérience, deux barres de couleurs différentes (rouge et vert) étaient présentées de chaque côté de la tâche aveugle. Par ailleurs, deux lignes étaient placées perpendiculairement aux barres colorées. La tâche des observateurs était d'indiquer ce qu'ils percevaient, et si la réponse était ambiguë, alors ils devaient préciser où les couleurs se rencontraient. Les observateurs percevaient les deux lignes comme une seule, qui interrompaient le remplissage des deux barres colorées. Ainsi, si la ligne était perçue de façon continue, le remplissage des barres était interrompu par la position que ces deux lignes occupaient dans la tâche aveugle. Dans une condition contrôle où les deux lignes n'étaient pas présentées, le remplissage ne s'effectuait pas. Par la suite, Brown et Thurmond (1993) ont testé les effets d'une couleur sur une autre dans le

processus de remplissage. Les stimuli étaient composés de quatre barres discontinues de couleurs différentes (rouge et vert), et étaient toujours présentés dans la tache aveugle. Les résultats ont indiqué que cette situation à choix forcé conduisait à un remplissage pour la couleur la plus saillante, à savoir le vert. Un résultat identique a été trouvé lorsque des stimuli achromatiques étaient utilisés : le remplissage de deux barres noires se produisait plus souvent par rapport à une condition où des barres grises étaient utilisées. Les données ne peuvent pas être attribuées à un traitement conscient, mais suggère l'implication d'un traitement pré-attentionnel dans la tache aveugle, influencé par des dispositions du système cognitif par rapport aux caractéristiques des stimuli.

En dehors de la tache aveugle, des effets attentionnels ont également été montrés dans le cadre des scotomes artificiels. Stürzel et Spillman (2001) ont déterminé si le temps de disparition d'un signal était corrélé avec sa saillance. La saillance de la cible par rapport au fond variait selon l'orientation, la forme et la « régularité ». Il n'y avait pas de frontière entre la cible et le fond, excepté pour une condition contrôle. Tout d'abord, la saillance pour la condition d'orientation était formé par des barres dont l'orientation changeait entre la cible et le fond. La saillance pour la condition forme était définie par le nombre d'angles que possédait la cible par rapport au fond. En effet, le fond était un cercle et la cible était représentée par six polygones dont le nombre d'angles augmentait. Le troisième type de stimuli était formé de points dont l'espacement était identique pour le fond, et différent pour la cible, ce que les auteurs ont appelé la condition de régularité. La cible était présentée à 8 deg d'excentricité d'un point de fixation. Les observateurs donnaient une estimation sur une échelle de 1 à 10 de la saillance perceptive de la cible. L'estimation des quatre cibles augmentait lorsque la saillance de la cible par rapport au fond était importante. Par la suite, les observateurs déterminaient le temps de disparition des cibles. Pour les quatre types de stimuli, les résultats ont indiqué une augmentation des temps de disparition des cibles, lorsque la saillance augmentait. La corrélation entre l'estimation de la saillance de la cible et le temps qu'elle mettait à disparaître était élevée. Pour Stürzel et Spillman (2001), le temps de disparition de la cible augmente systématiquement avec l'augmentation de la saillance perceptive.

Néanmoins, Lou (1999) a avancé l'hypothèse selon laquelle la disparition de la cible dans le cas des scotomes artificiels est le résultat d'un effet inhibiteur de l'attention volontaire sur le traitement sensoriel. Les stimuli, utilisés dans l'expérience, étaient les suivants : six disques, trois verts et trois oranges, étaient disposés en cercle de façon alternée autour d'un point de fixation sur un fond gris moyen. Les observateurs devaient diriger leur attention sur les trois disques d'une même couleur, les trois autres disques étant ignorés. Les disques disparaissaient plus fréquemment (dans 81.3 % des cas) lorsqu'ils étaient sous la focus attentionnel de l'observateur. De plus, la disparition des disques se produisait au bout de 10 sec. (moyenne = 11.35 sec) sur 1 à 2 cercles (moyenne = 1.46) selon les observateurs. De même que pour Brown et Thurmond (1993), la couleur verte disparaissait moins facilement. Par contre, les disques non sélectionnés par l'attention de l'observateur, demeuraient visible. La disparition sélective des cibles reflète un effet inhibiteur de l'attention sur le traitement sensoriel.

Les résultats de Lou (1999) ne paraissent pas logique. En effet, pour Spillman et DeWeerd (2003), l'attention top-down faciliterait la séparation figure-fond, ce qui

diminuerait le temps de disparition du signal dans le cas des scotomes artificiels. Les résultats expérimentaux de Stürzel et Spillman (2001) suggèrent que la saillance de la cible augmente la séparation figure-fond, ce qui a pour effet de ralentir le remplissage de la figure. L'effet de saillance est donc une caractéristique du signal qui permet d'augmenter l'attention de l'observateur. Au contraire, Lou (1999) montre que les cercles, sur lesquels l'attention de l'observateur est dirigée, disparaissent plus facilement que des cercles sur lesquels l'attention n'est pas dirigée. Dans ce cas, pourquoi l'attention top-down réduit le temps du remplissage alors qu'il devrait plutôt faciliter la séparation figure-fond, et par conséquent augmenter les temps de remplissage ?

Malgré cette contradiction, l'apport des travaux présentés dans cette partie est de montrer que les processus attentionnels peuvent interagir avec le traitement de remplissage naturel ou artificiel. Toutefois, la question du rôle de l'attention dans le processus de remplissage reste encore un véritable puzzle à reconstituer.

2.6.2. Importance de la surface dans les processus attentionnels

L'objectif des études, décrites dans la partie précédente, était de déterminer le rôle de l'attention dans le traitement du remplissage. Nakayama, He et Shimojo (1995) ont abordé les traitements attentionnels de façon différente. Nakayama et al. (1995) supposent qu'il ne peut y avoir d'attention sans considération préalable de la dimension de surface.

Récemment une étude neuropsychologique de Mattingley, Davis et Driver (1997) a apporté un argument en faveur de la précocité du traitement de la surface. Les auteurs ont étudié le cas d'un patient souffrant d'une extinction spatiale. L'extinction spatiale qui fait suite à une lésion pariétale est définie de la façon suivante : un patient est capable de détecter correctement un stimulus unique présenté dans le champ controlatéral à la lésion. Par contre, il ne détecte pas ce même stimulus quand un autre stimulus est présenté au même moment dans le champ ipsilatéral à la lésion. Ces patients présentent un désordre de l'attention sélective. Nakayama et al. (1995) suggèrent que le traitement de la surface intervient à un niveau pré-attentionnel, le traitement de la surface serait préservé chez ces patients. Mattingley et al. (1997) ont présenté, au patient, un carré de Kanisza uniquement dans le côté controlatéral ou les deux. Les résultats ont indiqué que le pourcentage de détection correcte était aussi élevé lorsque le carré était présenté dans le champ droit et gauche que dans le côté controlatéral à la lésion. La performance était affectée par différentes manipulations qui diminuait l'illusion en rendant le carré moins perceptible. Ce résultat indique que le traitement de la surface est réalisé avant que le traitement attentionnel soit effectué.

Les études physiologiques qui ont été présentées dans cette partie suggèrent que la dimension de surface est reconstruite, puisqu'elle n'est pas encodée directement par les cellules ganglionnaires de la rétine (Rossi & Paradiso, 1999). Ces données suggèrent un codage de la surface en deux étapes : un premier assez rapide et précoce puisque les études physiologiques ont permis de montrer que les neurones de l'aire V1 répondaient fortement à une surface uniforme, indépendamment du rôle que pourrait avoir les contours (Rossi & Paradiso, 1999). Dans une seconde étape, les bords d'une forme permettrait de rehausser l'activité des neurones pour le traitement de la surface (Hung,

Ramsden, Chen & Roe, 2001). De plus, l'hypothèse selon laquelle un processus de remplissage permettrait de traiter la dimension de surface n'a pas été démontrée physiologiquement dans le cortex strié.

A l'inverse de ces données, les études comportementales ont démontré expérimentalement un traitement de remplissage confortant l'hypothèse de Grossberg et al. (Grossberg, 1994, 2000 ; Grossberg & Mingolla, 1985 ; Grossberg, Mingolla, & Ross, 1997). Une distinction existe entre les données physiologiques et comportementales, et aucune réponse satisfaisante n'a pour le moment été donnée. Néanmoins, l'ensemble de ces études, physiologiques ou comportementales, sont en accord avec l'hypothèse d'un traitement hiérarchique du traitement de contour puis de la surface. L'objectif de notre thèse est de déterminer comment ces deux informations sont réparties dans un stimulus et comment elles permettent à l'observateur de prendre sa décision. Pour cela nous nous utiliserons le paradigme de l'image de classification qui sera présenté dans la prochaine partie.

3. Le bruit visuel : perturbation ou influence sur la prise de décision ?

Les deux précédentes parties théoriques ont permis de montrer les aspects théoriques du traitement de contour et de la surface. L'objectif de la thèse est de déterminer si la surface est recalculée ou si c'est une dimension indépendante du contour, extraite aussi précocément que lui. Pour répondre à cette question le paradigme de classification perceptive a été utilisé (Ahumada, 1996). Cette partie a pour but de présenter les différentes techniques utilisées en psychophysique, et d'expliquer le choix de ce nouveau paradigme.

La première partie présentera un paradigme classique dans lequel le bruit joue un rôle important en psychophysique : la théorie de la détection du signal (ou TDS). Dans une seconde partie, le paradigme de masquage sera décrit. Enfin, le paradigme de classification perceptive sera présenté ainsi que les expériences qui ont permis de valider cette technique. Nous avons choisi de présenter la TDS, car comme nous le verrons, ce paradigme est lié à la technique de l'image de classification par l'intermédiaire du critère de décision de l'observateur. De plus, nous nous sommes intéressés au paradigme de masquage car les idées sous-jacentes à cette technique semblent opposées à l'image de classification.

3.1. Le signal et le bruit : l'exemple de la théorie de la détection du signal

L'objectif de la psychophysique est de faire correspondre le domaine physique et le domaine sensoriel. Les premiers travaux en psychophysique ont développé la notion de seuil absolu, qui est la valeur minimale du stimulus qui entraîne une sensation juste

perceptible. La théorie des seuils suppose que la réponse du sujet dépend uniquement de la stimulation, il y a donc un lien constant entre la sensation et la réponse de l'observateur. L'élément important de cette relation se base sur le critère de réponse de l'observateur, c'est-à-dire le niveau de sensation à partir duquel l'observateur estime avoir perçu le stimulus. Par conséquent, il est important que l'observateur conserve un critère de réponse identique tout au long de l'expérience. Toutefois, les différences individuelles et aussi intraindividuelles observées conduisent à douter du bien fondé de cet apriori. On a donc cherché à évaluer la sensibilité indépendamment de la place du critère : c'est ce que permet la théorie de la détection du signal (TDS).

Pour bien comprendre la TDS il est nécessaire d'introduire la notion de signal et de bruit. Le signal est le stimulus qui est présenté à l'observateur, et le bruit représente l'ensemble des phénomènes non pertinents qui demeurent en l'absence du signal (MacMillan & Creelman, 1991 ; Schwartz, 1999). Ainsi, même si aucun signal n'est présenté, il y a toujours une activité physiologique spontanée (bruit interne) ou, pour d'autres situations, tout ce qui est présenté à l'observateur en plus du signal (bruit externe) (Croner, Purpura & Kaplan, 1993). Toutefois, les expériences qui appliquent la TDS utilisent rarement une variable totalement aléatoire, en effet, ce sont souvent des distracteurs et non pas un bruit comme nous venons de le définir.

La théorie de la détection du signal est un paradigme qui permet non seulement de caractériser et de quantifier la décision d'un observateur, mais également de déterminer sa sensibilité dans une tâche. Le paradigme de la détection du signal est une théorie probabiliste, elle suppose qu'un observateur donne une réponse relative à un critère interne. L'un des aspects de ce modèle est de déterminer une sensibilité (d') qui est indépendante de la place du critère de réponse de l'observateur. Par conséquent, ce modèle permet de faire des prédictions sur la nature des relations entre la performance à une tâche et le critère de l'observateur. L'utilisation de ces prédictions a été utilisée dans un large éventail de tâches psychophysiques.

Pour comprendre comment ces indices, sensibilité et critère sont dissociés, considérons une tâche simple de détection : il y a donc certains essais (la moitié) où le bruit est seulement présent (B), et les essais où le signal est superposé à un bruit (S+B). La tâche de l'observateur est d'indiquer si le signal était présent ou non. Le signal produit une réponse interne à l'observateur, et cette réponse interne sera utilisée pour prendre une décision. Un observateur idéal réaliserait la tâche de façon optimale, toutefois, l'observateur humain n'est pas idéal. En effet, l'une des principales différences consiste en l'existence de la notion de variabilité interne chez l'observateur humain qui n'existe pas pour l'observateur idéal (sauf si l'expérimentateur l'introduit). Si on présente à un observateur idéal plusieurs fois le même stimulus, il répondra exactement toujours de la même façon, il a donc des performances maximales (Geisler, 1989 ; Braje, Tjan & Legge, 1995). Contrairement à un observateur idéal, l'observateur humain a une variabilité interne ou bruit (Barlow, 1957). Cette variabilité, à travers les différents essais, a plusieurs origines : depuis les neurones jusqu'à des variations aléatoires dues à des informations parasites externes. Toutefois, un facteur aléatoire est souvent introduit dans les expériences utilisant un observateur idéal, mais même avec cette variabilité les performances sont systématiquement supérieures à celles de l'humain (Tjan, Braje,

Legge & Kersten, 1995). Une seconde distinction, entre l'observateur idéal et humain, repose sur leur habilité à déterminer l'information pertinente. Un observateur idéal utilise une réponse qui garantit une utilisation optimale de toutes les informations pertinentes pour une tâche donnée. A l'inverse, un observateur humain donnera des réponses non optimales. L'observateur idéal n'est donc pas un modèle des performances humaines, mais il permet d'en déterminer les limites (Geisler, 1989). Ainsi, un même signal ne produira pas la même réponse interne pour l'ensemble des présentations, mais au contraire une distribution de réponse à travers les essais. Dans le cas de la tâche de détection décrit plus haut, l'observateur humain et idéal auront deux distributions de réponse interne, une pour le bruit et une autre pour le signal et le bruit. Ces distributions représentent la variabilité « normale » des réponses. Ainsi, il existe une probabilité d'attribuer la réponse interne au bruit seul ou au signal et au bruit. Ces probabilités se répartissent selon deux distributions normales dont un exemple est donnée à la Figure 30.

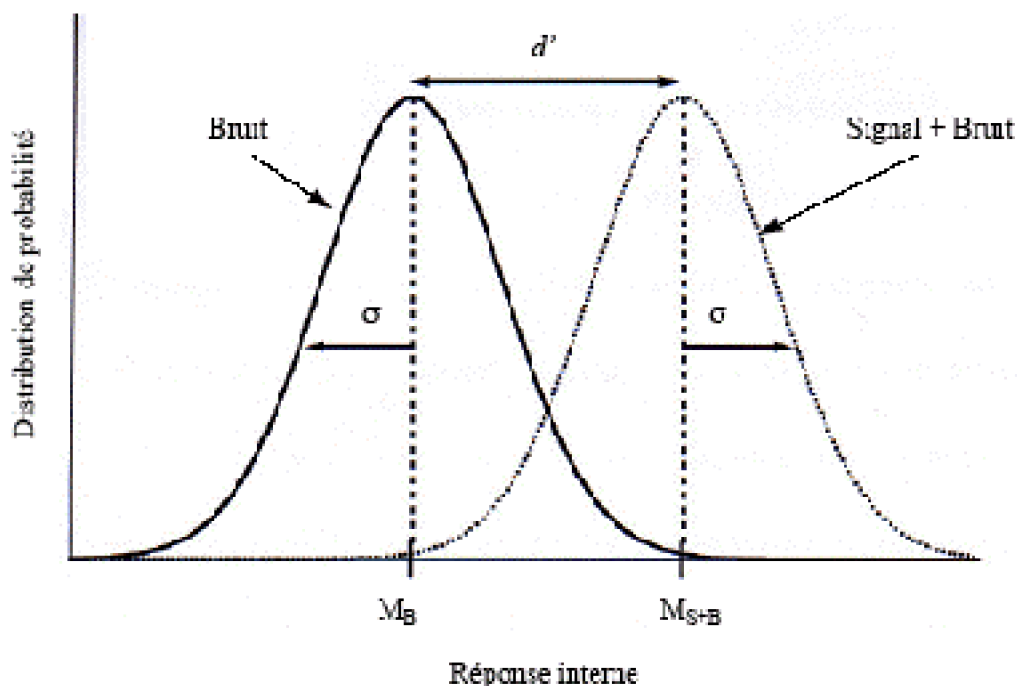


Figure 30 : Schématisation des distributions de réponse interne dans une tâche de détection. La distribution de gauche correspond au bruit (trait continu) et celle de droite au signal et au bruit (trait en pointillé), M_B et M_{S+B} sont les moyennes respectives des deux distributions. Les distributions supposent une variance similaire σ . La sensibilité de l'observateur (d') est égale à la différence entre les moyennes des deux fonctions normalisées par la déviation standard.

La Figure 30 indique deux distributions de réponse interne (hypothétique) réalisée par un observateur humain pour la tâche de détection qui a été décrite. Chaque distribution est centrée autour d'une moyenne de réponse M et a une variance σ . La théorie de la

détection du signal précise que les distributions utilisées sont de type gaussiennes (Green & Swets, 1966 ; MacMillan & Creelman, 1991). Dans ces conditions, la sensibilité (d') est définie comme la distance entre deux moyennes, normalisée par la variance de la distribution. Cette variabilité entre les moyennes correspond à un changement dans l'efficacité des réponses données. Par exemple, dans des conditions expérimentales où le signal est très difficile à distinguer du bruit, les deux distributions sont perceptivement semblables et sont voisines avec un recouvrement important. Au contraire, si le signal est très différent du bruit, les distributions seront très éloignées et leur recouvrement est négligeable. C'est le cas d'une tâche de détection facile où le signal se différencie facilement du bruit. Ainsi, la distance, calculée par l'indice d' , qui sépare les deux distributions donnent des indications sur les capacités perceptives des observateurs pour une tâche donnée. Une faible valeur de d' indique une situation où l'observateur distingue difficilement les essais B des essais S+B. D'un point de vue théorique cet indice est compris dans un intervalle entre $-\infty$ et $+\infty$. Toutefois, dans la plupart des expériences la valeur du d' varie entre -4 et 4 , cet indice est égal à 0 quand les réponses de l'observateurs sont données au hasard, une valeur positive élevée correspond à une situation dans laquelle l'observateur ne se trompe jamais, enfin une valeur négative indique que l'observateur donne systématiquement des réponses absentes alors que le signal est présent. L'indice d' se calcul à partir de la formule suivante d'après McMillan & Creelman (1991) :

$$d' = z(P_{DC}) - z(P_{FA})$$

dans cette équation, le d' est défini à partir du z qui est l'inverse d'une distribution normale, de plus, P_{DC} et P_{FA} sont les probabilité de détection correcte et de fausse alarme.

De plus, la TDS postule le fait que l'observateur se fixe un niveau particulier de sensation à partir duquel son jugement varie. En d'autres termes, toute sensation inférieure à ce niveau est attribué au bruit, et inversement toute sensation supérieure est attribuée au signal. Ce niveau correspond au critère de décision (ou c). La position de critère est placé le long du continuum de la réponse interne de la Figure 31. Par exemple, un observateur est prudent lorsque son critère est élevé, et il n'attribuera au signal que les quelques sensations dont l'intensité dépasse ce niveau. Le niveau du critère, pour cet observateur, se positionne plus sur la droite de la Figure 31. A l'inverse, un observateur aventureux prendra plus de risque pour prendre sa décision, dans ce cas, il se fixe un niveau bas et attribue au signal un nombre beaucoup plus important de sensation. Le niveau du critère est alors positionné sur la gauche de la Figure 31.

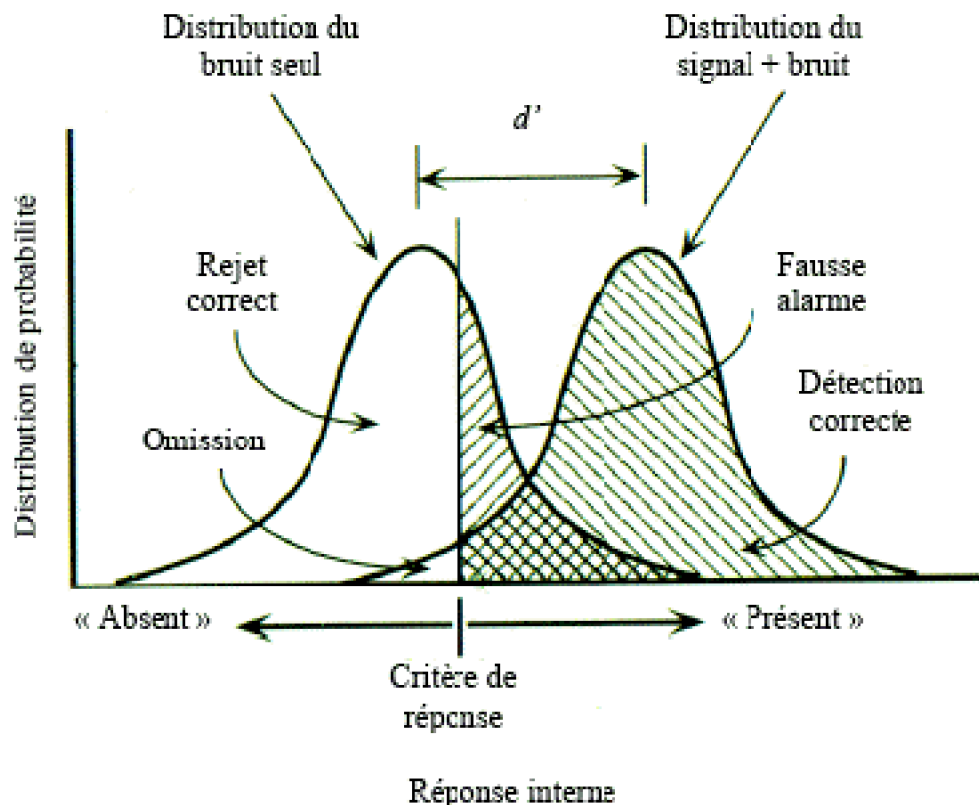


Figure 31 : Les quatre types de réponse possibles sont ici présentés en fonction de la distribution S+B ou B seul. Le critère de décision est représenté le long du continuum de la réponse interne, au-dessus de ce critère la réponse de l'observateur est de type « présent », et en dessous de ce critère la réponse est de type « absent ».

Le critère est alors défini par la fonction suivante (d'après McMillan & Creelman, 1991) :

$$c = -0,5 * [z(P_{DC}) + z(P_{FA})]$$

Quand le taux de fausse alarme et d'omission sont équivalents (ou le taux de détection correcte et de rejet correct), alors c est égal à 0. En d'autres termes, cela signifie que le critère de décision est positionné au point d'intersection des deux courbes.

3.2. Le phénomène de masquage

La situation expérimentale utilisée dans le paradigme de l'image de classification met en jeu deux stimuli : le signal et le bruit. Ces deux signaux sont présentés simultanément, mais ils sont distingués par l'expérimentateur, puisque la tâche implique que sujet tente

de dissocier le premier du second et puisque l'analyse se fonde uniquement sur l'influence de ce dernier. Cette situation est une variante de ce qui est classiquement appelé le masquage visuel utilisant une cible et un masque. Nous allons définir de façon générale ce phénomène avant de spécifier l'utilisation particulière qui en est faite dans le paradigme de l'image de classification.

3.2.1. Définition

Le masquage est avant tout un phénomène écologique qui se caractérise par la réduction de la visibilité d'un stimulus par la présentation simultanée ou dans un intervalle de temps suffisamment court, d'un autre stimulus. Par exemple, dans l'activité de lecture, une image est sans cesse remplacée par une autre sur la rétine. Des interactions d'inhibition et d'activation s'instaurent dans ces successions d'images. Le paradigme, qui est utilisé pour étudier les interactions entre deux stimuli, est appelée le paradigme de masquage visuel. Le paradigme de masquage consiste donc à perturber la perception d'une cible par la présentation d'un masque, qui est généralement une structure spatiale non significative (Delord, 1998, 2002).

3.2.2. Principales composantes du paradigme de masquage

Trois types de masquages sont généralement définis dans la littérature en fonction de l'ordre temporel entre la cible et le masque : le masque peut précéder la cible (masquage proactif), être présenté de façon simultanée ou lui succéder (masquage rétroactif). Dans la condition de masquage proactif, le masque est présenté avant la cible, par conséquent le SOA¹⁷ (" Stimulus Onset Asynchrony ") est négatif. Dans la condition de masquage simultané, le masque est présenté en même temps que la cible avec un SOA nul. Enfin, la condition de masquage rétroactif indique que le masque est présenté après la cible, le SOA est alors positif.

Trois variables interviennent dans le masquage visuel : le degré de similarité entre le masque et la cible, l'énergie des stimuli (selon la taille et le niveau de luminance) et le délai entre le masque et la cible (SOA). Pour expliquer les effets de l'ensemble de ces variables, deux mécanismes de masquage ont été distingués par les théories " tardives " de masquage : l'intégration (les propriétés du masque et de la cible sont intégrées en un même percept perturbant ainsi la perception de la cible) et l'interruption (le masque perturbant les traitements tardifs d'identification de la cible en les stoppant) qui n'intervient que pour le masquage rétroactif. Le mécanisme d'intégration dépend du rapport d'énergie entre la cible et le masque, et le mécanisme d'interruption dépend de l'intervalle qui existe entre la présentation de la cible et du masque (SOA). Pour l'intégration, le masquage est obtenu ou non selon que le masque ou la cible a le plus d'énergie. Par contre, pour le mécanisme d'interruption, la variable la plus importante est le délai entre le masque et la cible. Deux fonctions de masquages sont alors observées : la fonction est en U (le masquage est seulement tardif) lorsque les conditions expérimentales favorisent la cible sur le masque dans l'intégration, et elle est monotone lorsque l'inverse est vrai et que les processus précoces dominant (Turvey, 1973).

¹⁷ Le SOA est l'intervalle de temps séparant l'apparition d'une cible et l'apparition du masque.

Pour les aspects d'intégration, le raisonnement est le suivant : le masque va stimuler les mêmes canaux que ceux utilisés par la cible et va ainsi empêcher la réponse de ces canaux pour la cible. Si la présence du masque n'a pas d'effet sur la détection de la cible, les deux stimuli sont traités par des canaux indépendants. Si le masque a un effet, les deux stimuli excitent les mêmes canaux de traitement et le masque agit comme un piédestal sur lequel la cible doit être détectée. Le piédestal constitue une stimulation ajoutée à la stimulation engendrée par la cible. C'est l'hypothèse de la bande critique de masquage qui a été appliquée aux stimuli les plus simples, comme des réseaux de FS (revue : Breitmeyer, 1984), ou plus complexes comme des visages (Tieger et Ganz, 1976), des lettres (Solomon et Pelli, 1994), ou des dessins d'objets masqués par des masques " naturels " (i.e. non filtrés, comme des bruits ou des patterns ; Delord, 1998)

Par exemple, si la cible est composée de plusieurs fréquences spatiales (un dessin d'objet par exemple), la visibilité des informations qu'elle a en commun avec le masque va être sélectivement perturbée selon le contenu spectral du masque : lorsque la tâche implique l'extraction des composantes de FS basses des images (discrimination de forme globale), le masque le plus efficace sera celui dont l'énergie sur les composantes de FS basses est la plus importante ; par contre, si la tâche implique le traitement des contours précis (tâche d'identification), le plus efficace sera cette fois celui dont l'énergie est plus concentrée sur les FS moyennes et élevées (Delord, 1998).

3.2.3. Le modèle de masquage précoce de Breitmeyer

Nous présenterons dans cette partie une approche théorique qui se fonde seulement sur des mécanismes précoces pour expliquer le phénomène de masquage. Ce modèle a été proposé par Breitmeyer (1984) et a été complété récemment par Breitmeyer et Ogmen (2000) ; ces auteurs ont montré comment leur modèle précoce pouvait rendre compte de tous les effets de masquage, même ceux supposés comme plus tardifs..

L'hypothèse de départ suppose que deux canaux de traitement, avec des propriétés différentes, permettent d'analyser visuellement une image, ce sont les canaux toniques et phasiques (Legge, 1978). Ces deux canaux ont par la suite été intégrés dans une théorie du masquage visuel par Breitmeyer et Ganz (1976) : la théorie " transient-sustained ". Ces deux canaux de traitements se fondent sur l'existence en physiologie de cellules du même nom (Gouras, 1968). Gouras (1968) avait enregistré la réponse des cellules ganglionnaires pour différents niveaux de luminance, et avait obtenu deux types de réponse différentes (tonique et phasique). Toutefois, des études ultérieures ont établi que les cellules phasiques correspondaient aux cellules en parasol, qui se projettent sur la couche M du CGL, et les cellules toniques correspondaient aux cellules naines, qui se projettent dans la couche P du CGL (Kaplan, 2003).

Comme cela a été présenté au début, les canaux toniques (" sustained ") sont impliqués dans un traitement lent de l'intégration des caractéristiques de l'objet comme la couleur, les contours et les détails d'une figure. Les canaux phasiques (" transient ") sont impliqués dans une intégration rapide des éléments de l'objet, dans un traitement global mais aussi dans la localisation spatiale d'un stimulus. La présentation de tout stimulus engendre une activité du canal phasique puis du canal tonique. Breitmeyer (1984)

suppose que trois mécanismes de masquages sont nécessaires pour rendre compte des différentes perturbations engendrées par un masque : une interaction intercanal (inhibition) et deux interactions intracanal (inhibition et intégration). L'interaction la plus efficace est l'inhibition intercanal selon laquelle les canaux phasiques inhibent les canaux toniques (et inversement). Cette interaction s'effectue principalement lorsque le masque et la cible ne sont pas présentés simultanément à cause des différentes propriétés temporelles des canaux. Une autre interaction, la plus précoce, correspond à l'intégration intracanal qui est la compétition des activités de la cible et du masque pour les mêmes canaux de traitement (intégration des informations phasiques de la cible et du masque ou intégration des informations toniques des deux stimuli). Enfin, l'inhibition intracanal est le troisième mécanisme selon lequel l'activité tonique (ou phasique) de la cible est inhibée par l'activité tonique (ou phasique du masque), par exemple quand la cible active le centre d'un champ récepteur et le masque en active le pourtour, ce qui annule la réponse du canal.

Les interactions intracanal sont maximisées pour des SOA courts. Par contre, l'inhibition intercanal ne peut être activé uniquement lorsqu'un délai sépare les deux stimuli. Si le masque précède la cible, alors le masquage est faible car l'activité tonique du masque est inhibée par l'activité phasique de la cible. En revanche, si le masque succède à la cible alors l'activité tonique de la cible est inhibée par l'activité phasique du masque. Ces interactions permettent de rendre compte de l'ensemble des effets de masquage proactif et rétroactif (Breitmeyer (1984).

Suite aux avancées des connaissances dans le domaine physiologique, ce modèle a été revisité par Breitmeyer et Ogmen (2000). La différence essentielle réside dans le fait que cette approche tient compte de l'existence des connexions horizontales et " feedback " du système visuel. Ces modifications sont dérivées du modèle de Ogmen (1993). L'objectif du modèle de Ogmen (1993) concernait l'intégration des contours par le système visuel. Ainsi, l'intégration des contours serait réalisée par l'antagonisme centre/pourtour des cellules du canal tonique. Les contours seraient ensuite affinés par des connexions feedback qui modifieraient l'information en agissant sur le CGL ou le cortex visuel. Si un changement survient dans une scène visuelle, alors il est possible les variations ne soit pas prises en compte à cause de la persistance du " feedback ". Mais le système phasique (sensible au mouvement et aux variations soudaines dans le champ visuel) détecte ces modifications et réinitialise le système tonique par le biais de connexions latérales inhibitrices.

3.2.4. Rôle de l'attention dans le paradigme de masquage

Cette partie a pour objectif de présenter un modèle plus tardif de masquage qui remet en question l'intervention de processus précoces. L'implication de mécanisme attentionnel dans le masquage visuel a été étudié récemment à l'aide du paradigme RSVP (" Rapid Serial Visual presentation "). L'utilisation de ce paradigme a permis à Brehaut, Enns et Di Lollo (1999) de faire le lien entre les aspects temporels de l'attention visuelle et le paradigme de masquage. Une série de chiffres était présentée, parmi cette série deux lettres étaient insérées et la tâche de l'observateur était d'identifier les deux lettres. Les résultats indiquaient une identification importante de la première lettre mais pas de la

deuxième quand elle suivait la première de trop près. Pour Brehaut et al. (1999), ce résultat s'explique par la perturbation de la seconde lettre par la première. Cette perturbation était maximale pour un intervalle de 300 msec entre les deux lettres, et minimale à 100 ou 700 msec, une courbe en " U " est ainsi obtenue. Ainsi, la seconde cible n'a pas pu être traitée par l'observateur puisque ses ressources attentionnelles étaient portées sur la première cible. Ce phénomène est appelé la cécité attentionnelle (" attentional blink "). Si la cible est masquée par un masque simultanément, alors le déficit attentionnel caractérisé par la courbe en " U " ne se produit pas (Enns et Di Lollo, 2000). Dans ce cas, l'identification de la seconde cible est répartie de façon égale dans le temps, les performances ne sont plus influencées par l'intervalle temporel séparant les deux stimuli.

Enns et Di Lollo (2000) ont proposé une nouvelle situation de masquage qui est censée isoler les composantes attentionnelles du masque. Une matrice composée de 16 anneaux incomplets est présentée à l'observateur. La partie manquante de l'anneau a 4 positions différentes : à droite, à gauche, en haut ou en bas. L'un des anneaux (la cible) était entouré de 4 points, les autres anneaux étant alors des distracteurs. Ces 4 points restaient à l'écran après la présentation de la matrice. La tâche de l'observateur est de préciser la localisation manquante de l'anneau (à droite, à gauche, en bas ou en haut). Les résultats indiquaient que la persistance à l'écran des 4 points provoquait un effet de masquage. Par ailleurs, aucun effet de masquage n'était observé lorsque les distracteurs n'étaient pas présentés avec la cible. L'utilisation de ce nouveau paradigme indique que des conditions expérimentales bien particulières sont requises pour obtenir les effets de masquage. Ces résultats indiquent que pour obtenir ces effets de masquage l'attention de l'observateur doit être diffuse et non pas focalisée.

Enns et Di Lollo (2000) ont proposé un modèle pour expliquer ce type de phénomène. Dans ce modèle, deux niveaux sont activés : l'input (I) et le pattern (P). De plus, les informations du niveau pattern sont retranscrites au niveau de l'espace de travail (W), ce niveau permet une comparaison directe entre l'input et l'information du niveau pattern.

Le stimulus est reconnu au niveau pattern si les informations de l'input et de l'espace de travail sont identiques. Ainsi, quand la cible et le masque sont présentés simultanément, le niveau de l'input et du pattern sont transposés dans l'espace de travail. Dans l'expérience d'Enns et Di Lollo (2000) lorsque la cible disparaît, seuls les 4 points persistent. Par conséquent, une incompatibilité se produit entre l'information de l'espace de travail (masque et cible) et le niveau de l'input (masque seul). Le masque seul active l'ensemble du système, et du fait de sa persistance, cette information est la seule traitée. La cible n'a pas pu être traitée puisqu'elle a été remplacée par le masque, dans ce cas, seul le masque est perçu. Ce modèle est présenté dans la Figure 32.

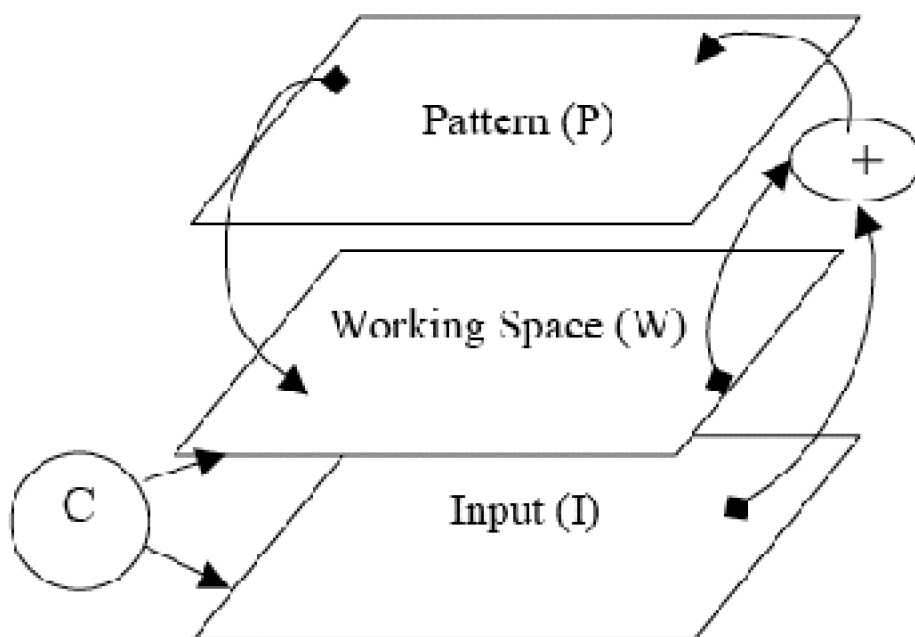


Figure 32 : Modèle de Enns et Di Lollo (2000). Les informations du niveau pattern (P) sont transmises dans un espace de travail (W), ce qui permet une comparaison (C) avec l'input (I). Les informations des niveaux I et W doivent être identiques pour être traitées.

Ce n'est que très récemment que les modèles sur le masquage (Breitmeyer & Ogmen, 2000 ; Enns & Di Lollo, 2000) ont inclus des processus feedback. D'un point de vue physiologique, il a été montré que l'activité neuronale de V1 est réduite dans un paradigme de masquage (Macknik & Livingstone, 1998). Une étude de Lamme, Zipser et Spekreijse (2002) a été réalisée afin de corroborer l'hypothèse de processus feedback dans le paradigme de masquage. Lamme et al. (2002) ont utilisé les mêmes stimuli que dans leurs études précédentes (Lamme, 1995 ; Lamme, Rodriguez-Rodriguez & Spekreijse, 1999 ; Zipser, Lamme & Schiller, 1996), mais ici une condition de masquage rétroactif est ajoutée dans laquelle le masque consiste en une texture uniforme à la suite de la présentation de la cible (un exemple est présentée à la Figure 18). Les résultats montrent que la réponse neuronale est supprimée pour un SOA correspondant à la suppression de la cible selon des données psychophysiques. Pour Lamme et al. (2002), la réponse neuronale de la cible dépend des processus feedback depuis les aires extra-striées (Hung, Ramsden, Chen, & Roe, 2001 ; Lamme & Roelfsema, 2000 ; Lamme, Supér & Spekreijse, 1998), par conséquent, il est tout à fait possible que le masquage interrompe les interactions entre V1 et les aires de plus haut niveau.

Nous avons présenté une revue de la littérature dans cette partie en comparant deux modèles de masquage : un modèle précoce (Breitmeyer et Ogmen, 2000) et un modèle basé sur une situation de masquage attentionnel qui isole les composantes les plus tardives du masquage (Enns & Di Lollo, 2000). Les expériences que nous avons présentées dans la partie expérimentale et celles de la thèse se positionnent à un niveau

précoce pour lequel le niveau d'attention du sujet n'est pas manipulé. Par conséquent, nous nous référerons principalement au masquage précoce fondé sur trois mécanismes inter et intracanaux décrit par Breitmeyer (1984) et par Breitmeyer et Ogmen (2000).

3.3. Le paradigme de classification perceptive

Dans la théorie de la détection du signal, la réponse d'un observateur, à une tâche, repose sur la décision obtenue par la corrélation entre le stimulus et un prototype, en tenant compte du bruit interne. L'objectif de la technique de l'image de classification est d'estimer le prototype utilisé par l'observateur en mesurant l'influence de chaque pixel du bruit externe.

3.3.1. Description du paradigme

Cette nouvelle méthode permet d'estimer le prototype d'un observateur en mesurant l'influence relative de chaque pixel d'un bruit externe sur les réponses de l'observateur. L'image qui montre l'influence de chaque pixel du bruit est appelé l'image de classification, elle permet ainsi de déterminer les éléments d'un stimulus qui sont utilisés pour prendre une décision dans une tâche donnée. L'avantage de cette nouvelle démarche est de pouvoir obtenir visuellement une représentation de ce prototype.

Cette technique a tout d'abord été utilisée dans le domaine de l'audition (Ahumada & Lovell, 1970 ; Ahumada, Marken & Sandusky, 1975 ; Sandusky & Ahumada, 1970), puis récemment appliquée à la vision (Ahumada, 1996 ; Beard & Ahumada, 1997 ; Gold, Murray, Bennett, & Sekuler, 2000 ; Thomas & Knoblauch, 1998 ; Watson, 1998 ; Watson & Rosenholtz, 1997). Le principe est le suivant. Au cours de l'expérience, on présente aléatoirement à l'observateur soit un bruit seul, soit un signal de faible intensité en même temps que le bruit. En vision, un bruit est une variation aléatoire des caractéristiques du signal. Un tirage aléatoire du bruit est effectué à chaque essai. Cette méthode est radicalement différente des techniques psychophysiques traditionnelles puisque les variables ne sont pas choisies au départ, mais sont au contraire totalement aléatoires pour chaque essai. Dans les expériences de vision spatiale, le bruit qui est utilisé est constitué de petit carré dont la luminance varie aléatoirement.

La tâche est par exemple de détecter le signal (présent/absent). De la même façon que pour la TDS, il existe quatre types de réponses possibles : une détection correcte (réponse correcte « présent »), une fausse alarme (réponse « présent » alors que seul le bruit était présenté), une omission (réponse « absent » alors que le signal était présent) ou un rejet correct (réponse correcte « absent »). Nous avons pris l'exemple d'une tâche de détection, mais ce paradigme peut aussi bien être utilisé dans une tâche de discrimination ou tout autre tâche ayant un nombre fini d'alternative.

Les caractéristiques du signal sont choisies pour le rendre difficilement visible, et il est présenté dans un bruit. Par conséquent, la probabilité de répondre « détecté » est plus importante lorsque les indices visuels présents dans le bruit correspondent à ceux utilisés par l'observateur pour traiter le signal.

Une fois que les données sont récoltées, nous allons déterminer comment la

distribution du bruit a influencé la réponse de l'observateur tout au long des essais. Ceci est réalisé par la corrélation entre les perturbations aléatoires et les réponses de l'observateur à chaque essai. Pour analyser les indices du bruit qui ont influencé l'observateur dans sa réponse, une première étape consiste à trier les différents bruits à chaque essai en fonction des quatre possibilités de réponses. La seconde étape est de moyenner les différents bruits pixel par pixel selon les quatre possibilités de réponses. Ainsi, les moyennes des bruits présentés aux essais ayant engendrés les deux types de réponse « présent » (détection correcte et fausse alarme) contiennent les indices visuels du bruit qui sont corrélés avec la stratégie de classification de l'observateur. Dans une dernière étape, les moyennes des bruits ayant entraîné la réponse « absent » (omission et rejet correct) donneront la partie qui est anti-corrélée. En combinant ces moyennes (les catégories anti-corrélées recevant un poids négatif), on obtient une image qui exprime l'influence de chaque pixel sur la réponse de l'observateur. En d'autres termes, ces pixels sont considérés comme des indices visuels utilisés par l'observateur pour traiter le signal. Un exemple du traitement des données de cette nouvelle technique est présentée à la Figure 33.

C'est l'image de classification perceptive, révélatrice de la façon dont l'observateur a pondéré ou classifié les informations du stimulus pour prendre sa décision. En effet, il s'agit d'une estimation des indices du signal qui sont utilisés par l'observateur pour prendre sa décision.

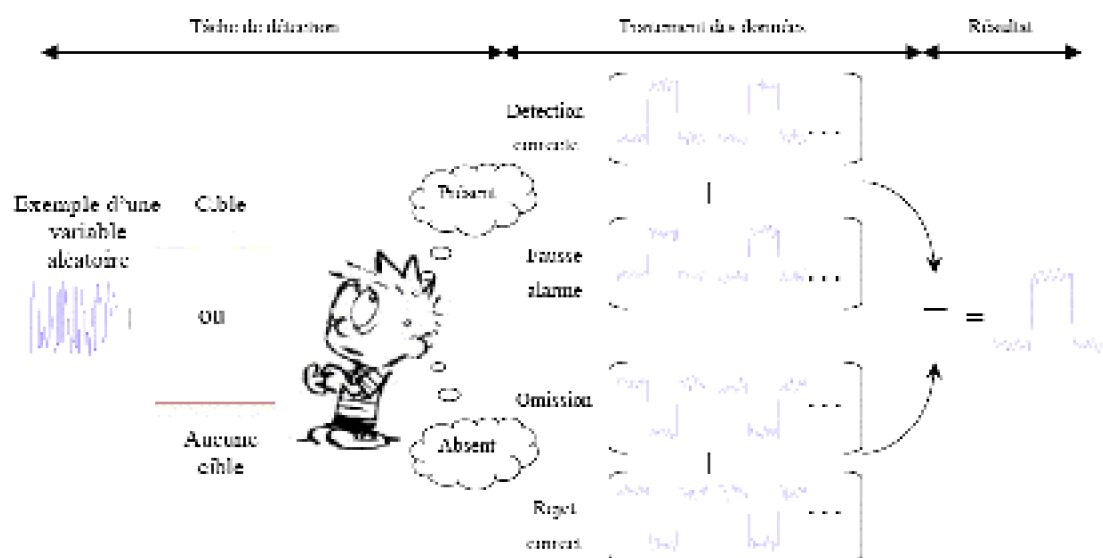


Figure 33 : Exemple du déroulement d'une expérience utilisant la technique de l'image de classification. La cible est ici un carré présent dans la moitié des essais, ce signal est alors superposé à une variation aléatoire. L'observateur doit alors détecter la présence ou non de la cible. Pendant le traitement, les bruits sont classés en fonction de la catégorie de réponse. L'image de classification est ensuite obtenue par la soustraction entre la moyenne des bruits ayant engendré une réponse de type présent et absent.

La question abordée avec ce paradigme est très nouvelle : l'intérêt ne se porte pas sur le signal, mais au contraire sur le bruit dans lequel le signal est présenté, puisque la réponse du sujet n'est pas seulement influencée par le signal lui-même mais est aussi déterminée par le bruit dans lequel il apparaît. Ce bruit, qui peut représenter les perturbations de l'environnement ou les perturbations du système lui-même, donne des indications sur les indices pertinents qui engendrent un percept. Pourvu qu'un nombre suffisant d'essais expérimentaux permette de faire émerger, des variations aléatoires de la structure du bruit, les indices ayant provoqués de façon systématique l'une ou l'autre réponse, on peut déterminer quelles dimensions du signal sont pertinentes pour la

réponse.

L'utilisation de ce paradigme est extrêmement pertinente puisqu'il nous permet de déterminer visuellement le prototype utilisé par l'observateur. Cette nouvelle démarche est donc liée à la TDS par l'intermédiaire du critère de décision. En effet, nous avons vu que la TDS permettait d'évaluer le critère que l'observateur se fixe tout au long des essais. L'image de classification que nous obtenons est en quelque sorte une estimation visuelle de la distribution spatiale ou temporelle du critère de l'observateur.

3.3.2. Arguments expérimentaux

La variable bruit n'est pas déterminée au début de l'expérience, mais elle est triée de façon totalement aléatoire à chaque essai. Cette partie présente des exemples expérimentaux qui ont pour objectif de montrer que l'on retrouve dans l'image de classification ce que l'observateur utilise du signal. Nous aborderons des expériences dans le domaine de la vision spatiale puis temporelle pour illustrer cette idée.

3.3.2.1. Expériences dans le domaine de la vision spatiale

La première expérience d'image de classification dans le domaine de la vision a été réalisée par Ahumada (1996). L'un des objectifs était de déterminer si la méthode utilisée dans le domaine de l'audition pouvait être appliquée dans le domaine visuel. Ahumada (1996) a conduit une expérience portant sur la tâche classique de l'acuité de Vernier. La tâche était de déterminer si deux barres sont alignées ou légèrement décalées l'une par rapport à l'autre. Les résultats indiquaient que la réponse de l'observateur se fondait sur la position relative des deux barres et non pas de la position absolue de la barre dont la position variait. Les résultats de cette expérience sont présentés à la Figure 34. Des résultats similaires étaient trouvés lorsque les deux barres ne se touchaient plus, mais étaient séparées par une distance de 10 min (Beard & Ahumada, 1998).

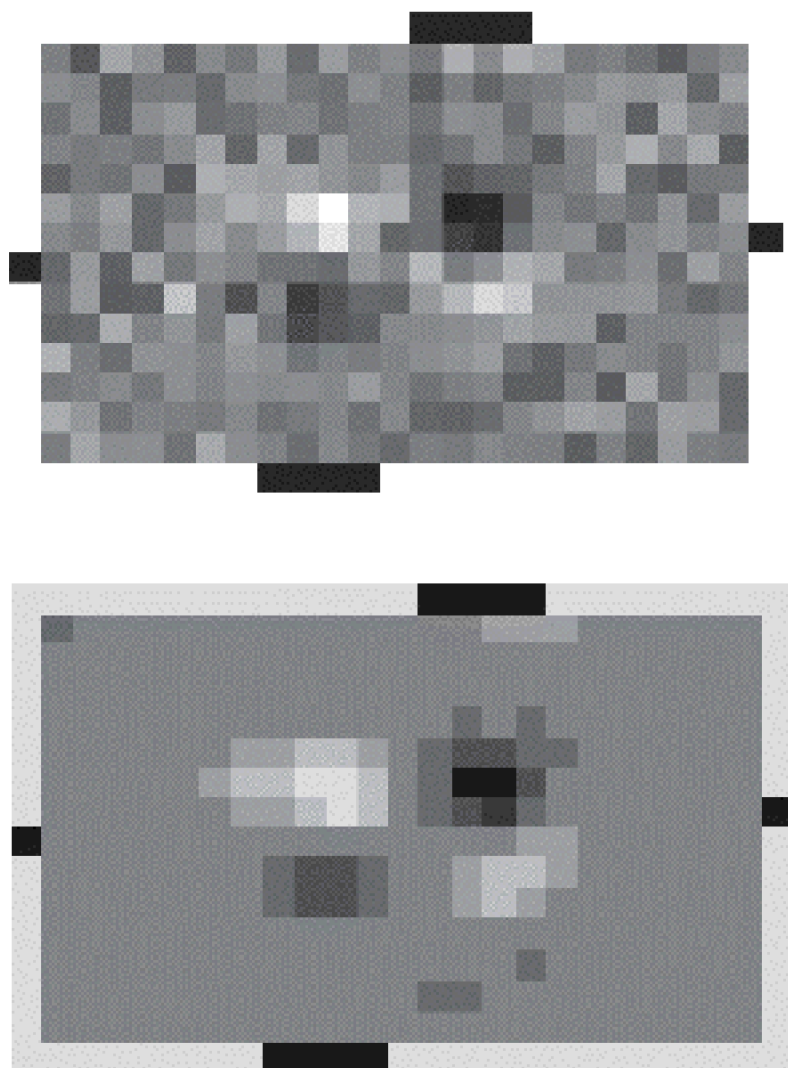


Figure 34 : Résultats de l'image de classification dans le cas de l'acuité de Vernier (Figure du haut), l'image est ensuite lissée (figure du bas). Dans la figure du bas, les pixels significativement différents de 0 sont colorés différemment par rapport au fond gris moyen. Les carrés noirs sur le côté de chaque image indiquent la position des barres, et les lignes (en bas et en haut de chaque image) montre la longueur des barres. Cet exemple est issu de Ahumada (1996), l'observateur est AJA qui a réalisé 1600 essais.

Eckstein, Shimozaki et Abbey (2002) ont appliqué la technique d'image de classification pour déterminer comment l'attention affecte le traitement de l'information dans le paradigme d'attention indicée de Posner. Eckstein et al. (2002) ont comparé le prototype utilisé par l'observateur humain pour des positions indicées et non indicées. La procédure était la suivante : un carré était tout d'abord présenté sur l'une des deux positions possibles pour les cibles ; les stimuli apparaissaient juste après et étaient constitués de deux blobs gaussiens, puis un bruit était présenté et le contraste de l'un des deux blobs gaussiens augmentait. L'observateur indiquaient si oui ou non cette augmentation était présente. Les résultats ne montraient pas de différence dans le

prototype utilisé par les observateurs pour les deux positions. Eckstein et al. (2002) ont montré une différence dans la taille des images de classification : les images étaient beaucoup plus grande pour une position indiquée que non indiquée. En d'autres termes, l'attention visuelle change le poids de l'information utilisée par l'observateur, mais ne change pas la qualité du traitement de l'information.

Abbey, Ekstein et Bochud (1999) ont étendu la technique d'Ahumada (1996) à une tâche de détection d'un patch gaussien. Les images de classification obtenues dans cette tâche montraient une zone claire centrale à l'emplacement du signal, mais également un faible pourtour antagoniste. Abbey et Eckstein (2000) ont répliqué cette expérience, mais cette fois, trois bruits différents étaient utilisés : un bruit blanc, un bruit passe-bas et un bruit passe-haut. Les images de classifications étaient identiques à l'expérience de Abbey et al. (1999), mais la zone claire centrale apparaît de façon plus petite dans le cas d'un bruit passe-bas.

Gold, Murray, Bennett et Sekuler (2000) ont cherché à déterminer si les contours illusoires étaient impliqués dans une tâche perceptive. Le carré de Kanisza était utilisé, et les observateurs effectuaient une tâche de discrimination entre deux conditions où les contours illusoires sont concaves ou convexes. Deux conditions contrôles étaient réalisées : l'une avec des contours réels et l'autre avec des « pac-men » retournés dans laquelle les contours illusoires n'apparaissaient plus. Les images de classification montraient que les observateurs utilisaient les contours verticaux pour réaliser la tâche, aussi bien pour des contours réels qu'illusoires. Une tâche identique a été effectuée par un observateur idéal qui utilisait uniquement l'information centrée sur les bords des « pac-men » pour répondre.

Les expériences présentées dans les paragraphes précédents utilisaient un bruit composé de petits carrés dont la luminance varie aléatoirement. Toutefois, certaines études (Levi & Klein, 2002 ; Neri, Parker & Blakemore, 1999) ont montré que la technique de l'image de classification était également applicable pour d'autres formes de variable aléatoire. L'image de classification est alors obtenue de la même façon que le calcul proposé par Ahumada (1996). L'étendue de la technique de l'image de classification se retrouve dans la littérature sous le nom de corrélation inverse (Levi & Klein, 2002 ; Neri & Heeger, 2002 ; Neri, Parker & Blakemore, 1999). Le principe général est toujours le même puisqu'il s'agit d'obtenir ce que l'observateur utilise du signal, mais ici la technique est étendue à d'autres possibilités de perturbation aléatoire.

Neri, Parker et Blackemore (1999) ont étudié les mécanismes impliqués dans la disparité binoculaire¹⁸. Dans l'expérience de Neri et al. (1999), les stimuli étaient constitués soit de points dont la position était aléatoire (noir et blanc) sur un fond gris avec une disparité binoculaire nulle, soit d'un carré composé également de points aléatoires qui apparaissait comme étant plus en avant par rapport au fond. Les observateurs détectaient la présence ou non de ce signal. Les images montraient un profil de réponse identique à celui d'une fonction en chapeau mexicain, c'est-à-dire, avec un centre et un pourtour antagoniste. Neri et al. (1999) ont ensuite réalisé une simulation informatique dans

¹⁸ La disparité binoculaire consiste en un déplacement d'une image issue d'un même environnement lorsque l'œil droit est ouvert et l'œil gauche fermé, puis l'inverse. Cet effet est dû à la distance qui existe entre les deux yeux.

laquelle deux filtres gaussiens optimisés (dont le profil est identique au champ récepteur de V1) étaient combinés pour effectuer la tâche de détection. L'un des deux filtres était centré sur l'image, et l'autre était décalé spatialement. Cette simulation prenait ainsi en compte la disparité binoculaire. Les images de classifications obtenues présentaient des profils de réponses identiques à celui de l'observateur humain.

Cette technique a également été utilisée pour comparer les prototypes utilisés par l'observateur entre une cible présentée en vision fovéale et parafovéale (Levi & Klein, 2002). Le signal et le bruit étaient caractérisés par des sinusoides. Une tâche de détection était tout d'abord réalisée, à chaque essai le signal était présenté avec l'un des quatre contrastes possibles (0%, 12%, 24% et 36%), les observateurs indiquaient s'il percevaient ou non le signal en précisant le degré de précision (de 1, pour un signal non visible, à 4, pour le signal le plus visible). Les images de classifications étaient proches du signal pour une position en fovéa et en parafovéa et ce quel que soit le niveau de contraste du signal. Dans une seconde expérience, les observateurs effectuaient une tâche de discrimination entre 5 positions possibles et cette fois le niveau de contraste était toujours le même (en fovéa et parafovéa). À l'inverse de la tâche de détection, les images de classification étaient différentes au centre et à la périphérie, en effet, la fréquence spatiale augmentait avec l'excentricité. En d'autres termes, les images de classifications étaient représentées par des sinusoides plus larges dans la parafovéa. Ces deux expériences montrent que les indices qui sont utilisés par l'observateur au cours de la tâche peuvent ressortir quel que soit le type de variable aléatoire.

3.3.2.2. Peut-on utiliser une variable qui varie aléatoirement dans le temps ?

En dehors des expériences en vision spatiale, notre objectif est de montrer qu'une variable aléatoire temporelle permet également de faire ressortir ce que l'observateur utilise du signal. Thomas et Knoblauch (1998) ont étudié la détection de la variation temporelle d'un patch lumineux qui variait dans le temps de façon gaussienne ($\sigma = 160$ msec ; durée = 640 msec) et qui était présenté dans un bruit dont la variation de luminance était aléatoire dans le temps. Les images de classifications étaient bien inscrites dans une courbe gaussienne, mais leur largeur (entre 160 et 220 msec) différait de celle de la stimulation réellement présentée, ce qui indiquait une incertitude temporelle des observateurs sur la largeur du signal. Par ailleurs, le maximum de la fonction gaussienne dans l'image de classification était systématiquement plus précoce (d'environ 100 msec) que le maximum de la gaussienne du signal. Un résultat similaire a été trouvé lorsque le signal gaussien était inconnu de l'observateur (Knoblauch & Yssaad-Fesselier, 2000). Ces expériences permettent de généraliser le principe des images de classification : nous pouvons déterminer un prototype à partir d'une variable aléatoire aussi bien spatiale que temporelle.

Neri et Heeger (2002) ont modifié la technique de l'image de classification afin d'utiliser une variable aléatoire spatiale le long d'un continuum temporel. L'objectif était de montrer l'intervention de mécanismes différents pour la détection et l'identification. Dans leur expérience une séquence de neuf bruits apparaissait les uns à la suite des autres (chaque image étant présentée durant 27 ms). Chaque bruit est composé de onze barres adjacentes. Sur la moitié des essais, une barre lumineuse (la cible) apparaissait au bout

de la cinquième séquence (au milieu), l'observateur indiquait alors si la barre était présente ou non. Une séquence expérimentale est présentée à la Figure 35.

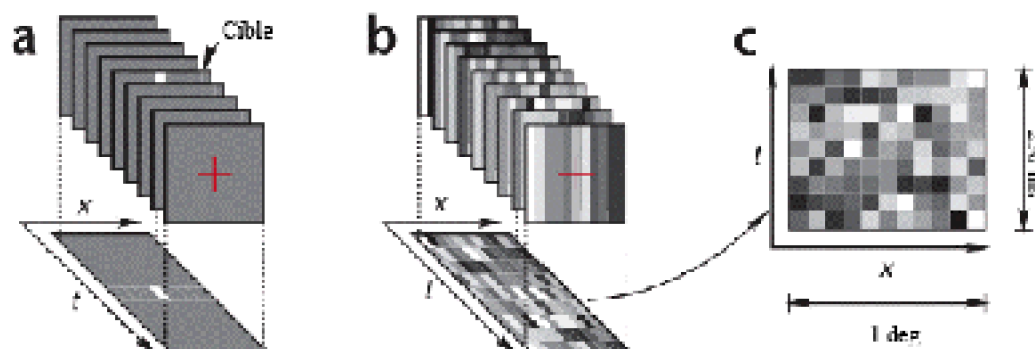


Figure 35 : La cible était une barre lumineuse présentée au milieu de la séquence (Figure de gauche). Le bruit était constitué de plusieurs barres dont la luminance était aléatoire (Figure du centre). La Figure de droite indique l'image qui était conservée à la fin de chaque essai. Pour chacune des figures x représente la taille et t le temps (l'illustration est issue de Neri & Heeger, 2002).

Les résultats étaient ensuite traités de la même façon que pour une image de classification, c'est-à-dire que les bruits étaient classés selon le type de réponse possible, et l'image de classification était obtenue par la soustraction entre la moyenne des bruits ayant engendrés le type de réponse signal présent et absent. Il est important de noter que la position temporelle de chaque bruit était conservée pendant le calcul de l'image de classification, ce qui permettait d'apprécier la stratégie de l'observateur au cours d'un essai. Les images indiquaient un pic positif à l'emplacement de la cible, mais également un pourtour négatif. Par ailleurs, une zone apparaissait dans l'image de classification qui précédait l'apparition du signal. Le calcul de la variance montrait un pic positif qui précédait de 100 ms la présentation de la cible. Cette expérience permettait de montrer que la détection de cible est basée sur un contraste entre le stimulus et la région adjacente. De plus, le résultat issu de la variance semblait indiquer l'existence d'une étape préalable pour la détection d'une cible. Dans une seconde expérience, Neri et Heeger (2002) ont testé directement l'hypothèse de deux étapes pour la détection et la discrimination. La cible apparaissait comme plus claire ou plus foncée, et à droite ou à gauche d'un point de fixation. Pour cette expérience, une séquence était composée de 22 barres (au lieu de onze dans la première) et la cible était placée au milieu des onze premières (ou dernière) barres. Les observateurs déterminaient le niveau de luminance (claire ou sombre) et la position (droite ou gauche) de la cible pour chaque essai. La distribution du bruit était calculée séparément pour la tâche de détection et d'identification. Pour la tâche de détection, la variance indiquait un pic positif qui précédait de 100 msec la présentation de la cible, mais rien n'apparaissait pour la moyenne de l'image. Pour la tâche d'identification, l'image de classification montrait un pic positif au moment où la cible apparaissait dans le temps, et qui dépendait du contraste avec la région du pourtour. Par ailleurs, aucun résultat n'était observé à partir du calcul de la variance. Ces résultats mettent en évidence l'existence de deux étapes de traitement : une étape de détection où

seule la position du contraste est traitée, puis d'une étape d'identification dans laquelle la position de l'image sélectionnée permet de déterminer plus précisément la cible (claire ou sombre).

Dans cette dernière partie théorique, nous avons présenté deux techniques utilisant la notion de bruit, mais aussi une technique classique : la masquage. Au premier abord ce dernier paradigme s'oppose d'un point de vue théorique au paradigme de classification perceptive. En effet, l'un suppose que le bruit influence positivement la décision de l'observateur et le second postule au contraire que le bruit perturbe la perception de la cible. La question que l'on peut alors se poser est de savoir si ces deux paradigmes sont aussi contradictoires qu'il ne le paraisse. Nous tenterons de répondre à cette question au cours de la conclusion générale.

En résumé ...

Le codage précoce d'une information, par l'antagonisme centre/pourtour des champs récepteurs, indique que l'une des premières informations qui est traitée se centre sur les bords d'une forme (Hubel & Wiesel, 1962, 1968). Notre première partie théorique a permis de montrer qu'il existait un consensus entre les études physiologiques et comportementales sur la façon dont le système visuel traite les contours.

La plupart des théories ont minimisé le rôle de la surface à un mécanisme secondaire : le remplissage. Dans la seconde partie théorique, nous avons vu que ce processus expliquait de nombreux phénomènes comme la tache aveugle ou les scotomes, les résultats des études physiologiques et psychologiques vont dans le même sens. Par contre, lorsqu'il s'agit d'intégrer le remplissage au traitement de la surface, l'implication de ce processus n'est pas aussi clair sur le plan physiologique. Toutefois, le traitement en deux étapes, basé sur les contours puis sur la surface, n'est pas remis en cause. Par conséquent, si les contours sont encodés par le système visuel précoce et que la surface est traitée ultérieurement, alors nous pouvons prédire que les bords d'une forme constituent un trait important utilisé par les observateurs pour la détection d'une forme. Nous avons voulu déterminer si cette hypothèse générale est valide dans une tâche simple de détection. La technique de l'image de classification va nous permettre de répondre directement à cette question. En effet, nous avons vu que ce paradigme permettait de déterminer les indices pertinents qui sont utilisés par l'observateur.

Chapitre II : Partie expérimentale.

1. Première série d'expériences : variation de la durée de présentation du stimulus

Les données expérimentales, présentées au cours la partie théorique, ont montré que l'information de contour est une dimension qui est codée précocement par le système visuel. L'organisation centre/pourtour des champs récepteurs des cellules ganglionnaires de la rétine est telle que seule l'information située sur les bords d'une forme visuelle est véhiculée. En effet, la conception classique des champs récepteurs ne permet pas d'intégrer la dimension de surface. Par conséquent, il est admis que l'intégration de la surface nécessite une étape de reconstruction à partir de l'information de contour. Pour de nombreux auteurs (Gerrits & vendrik, 1970 ; Grossberg & Mingolla, 1985 ; Spillman, 1997 ; Walls, 1954), le processus de remplissage permettrait de reconstruire la dimension de surface. Bien que le remplissage d'une forme n'ait pas été démontré de façon explicite dans les études physiologiques, l'existence de ce processus a été démontrée dans de nombreuses études comportementales (Caputo, 1998 ; Dresch & Bonnet, 1991 ; Paradiso & Nakayama, 1991 ; Stoper & Mansfield, 1978). Toutefois, cette contradiction entre les résultats physiologiques et comportementaux, concernant le processus de remplissage dans l'intégration d'une surface, ne remet pas en question l'existence d'un traitement

hiérarchique du contour et de la surface. Par conséquent, si les contours sont encodés par le système visuel et que la surface est ensuite traitée, alors nous nous attendons à ce que les contours constituent une dimension prépondérante pour la détection d'une surface uniforme. Ce type de tâche n'a jamais été utilisé pour mettre en évidence un traitement hiérarchique contour/surface, en dehors des études portant sur les contours illusoire. Mais, nous avons vu que les formes illusoire et non illusoire se distinguent par leur niveau de traitement (Mendola, Dale, Fischl, Liu & Tootell, 1999). Par ailleurs, les études comportementales utilisent souvent des tâches de jugement, et il est possible que cela affecte plus des fonctions d'ordre cognitive et non pas un traitement précoce. Ce qui permettrait d'expliquer la contradiction entre les expériences comportementales et physiologiques. En effet, cette distinction ne se retrouvait pas pour le traitement des contours puisque les expériences chez l'observateur humain utilisent souvent la détection. Cette tâche est ici rendue possible grâce au paradigme de l'image de classification. Ce paradigme va nous permettre d'analyser la représentation utilisée par l'observateur pour répondre, ce qui permettra de vérifier dans quelle mesure l'analyse des prototypes est compatibles avec l'importance accordée aux contours.

Nous savons très bien que manipuler le temps de présentation des stimuli influence les dimensions utilisées par l'observateur (Davey, Madess & Srinivan, 1998 ; Dresch & Bonnet, 1991, 1993 ; Paradiso & Hahn, 1996 ; Paradiso & Nakayama, 1991). Si la conception traditionnelle est vraie, nous nous attendons à ce que la répartition l'information de surface varie avec le temps de présentation, mais que l'information de contour soit toujours présente puisqu'il s'agit d'une information primordiale pour le système visuel.

1.1. Expérience princeps : durée de présentation intermédiaire (400 ms)

La première expérience était une étape princeps pour l'ensemble de notre recherche. L'objectif était de déterminer pour une durée de présentation intermédiaire, quelles informations (le contour ou le contour plus la surface) étaient pertinentes pour la détection d'une forme géométrique simple : un carré.

Pour l'ensemble des expériences que nous présentons dans cette thèse, deux étapes sont nécessaires. Une expérience de détermination de seuil est tout d'abord réalisée, puis une expérience dans laquelle l'image de classification proprement dite est calculée. La première expérience a servi à déterminer le niveau du bruit (un bruit supérieur au bruit interne du système visuel) et le niveau du signal (un niveau engendrant une performance imparfaite) qui seront utilisés dans l'expérience de détermination de l'image de classification.

1.2. Expérience préliminaire : détermination de seuil

Le seuil est la valeur de contraste du signal qui le rend juste perceptible, c'est-à-dire, la valeur qui engendre une performance intermédiaire de détection correcte. Le signal était

présenté simultanément à un bruit. Le contraste du bruit variait, ce qui augmentait la perturbation qu'il engendrait sur la perception du signal.

Nous nous attendions à ce que le seuil du signal augmente quand le contraste du bruit augmente, mais pas pour les plus faibles niveaux de bruit, car ces niveaux sont inférieurs au bruit interne du système visuel (Lu & Doshier, 1998). Cette expérience permettait ainsi de déterminer ce bruit interne, de manière à choisir un bruit expérimental supérieur à cette valeur.

Nous pouvons résumer la relation entre le contraste du bruit et le contraste seuil du signal de la façon suivante : 1) pour les plus faibles niveaux de contraste du bruit, le contraste seuil du signal n'augmente pas quand le contraste du bruit augmente, ce qui indique que le bruit est inférieur au bruit interne du système visuel lui-même, et 2) par contre, pour les valeurs supérieures de contraste du bruit, le contraste seuil du signal augmente quand le contraste du bruit augmente indiquant que le bruit est supérieur au bruit du système.

Les différents niveaux de contraste du bruit étaient présentés en blocs et à l'intérieur de chaque bloc, la méthode constante de présentation des niveaux de contraste du signal était utilisée pour déterminer le seuil de détection.

1.2.1. Méthode expérimentale

Participants :

Quatre observateurs, dont la vue était normale ou corrigée à la normale, ont participé à l'expérience (FD, SD, SH et VB). Les deux premiers étaient les expérimentateurs, le troisième et le quatrième observateur n'étaient pas prévenus des hypothèses.

Stimuli :

Les stimuli, qui étaient affichés sur un fond gris moyen (8 cd/m^2), étaient composés d'un signal et d'un bruit superposés l'un à l'autre : le signal était un carré de 3° d'angle visuel de côté, le bruit était une distribution uniforme de petits carrés (2×2 pixels soit 0.15×0.15 deg) de luminance aléatoire, recouvrant 10° d'angle visuel. Un exemple de bruit, de signal et la superposition des deux est présenté Figure 36. L'étalonnage de l'écran a permis le contrôle des luminances grâce à un photomètre : une « Look Up Table » (LUT), c'est-à-dire une échelle linéaire de luminance, comportant donc un intervalle identique entre tous les niveaux a été créé. Les niveaux de la LUT ont été choisis parmi les 255 niveaux RVB (rouge, vert, bleu) qui correspondent aux trois canaux de couleurs de l'écran

19 .

¹⁹ Lorsqu'une même valeur est envoyée aux trois canaux, un niveau de gris est obtenu.

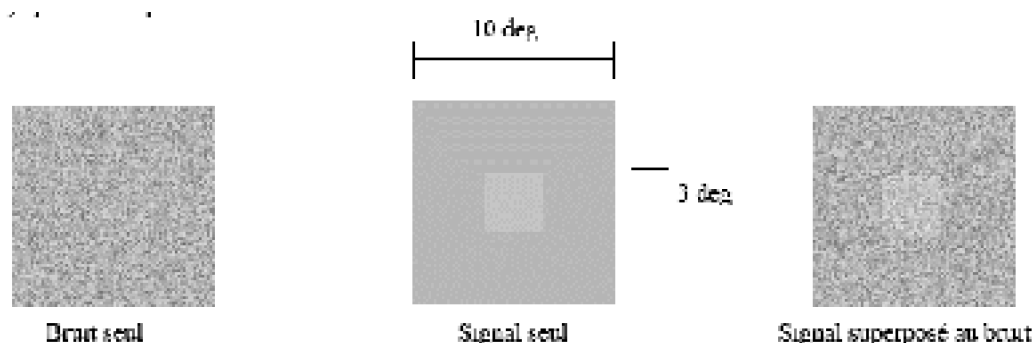


Figure 36 : Exemples de bruit, de signal seul et d'un signal superposé au bruit.

Le contraste du signal ΔL était spécifiée par le contraste Weberien : $\frac{\Delta L}{\bar{L}}$ où \bar{L} est la luminance moyenne du fond.

Le contraste de la distribution uniforme du bruit est déterminé par l'intervalle $\bar{L}(1 + m)$ et $\bar{L}(1 - m)$, sa variance a été définie de la façon suivante : $m^2 \bar{L}^2 / 3$, l'élément m est ici considéré comme la variation maximum du bruit à partir de la luminance moyenne. La répartition des différents pixels autour de la luminance du fond est présentée de façon schématique à la Figure 37. Nous obtenons ainsi une valeur de contraste compris entre 0 et 1. Quarze contrastes du bruit ont été utilisés dans l'expérience (0.5, 0.6, 0.7 et 0.8). Un exemple de ces différents niveaux de contraste est présenté à la Figure 37. les niveaux de luminance utilisés pour les petits carrés, formant le bruit, sont attribués aléatoirement dans un intervalle de la LUT autour de la luminance moyenne. Ainsi, lorsque le contraste du bruit augmente, les luminances attribuées aux pavés se répartissent sur un plus grand intervalle.

Figure 37 : La figure du haut précise de façon schématique la distribution aléatoire de chaque pixels (représenté par des flèches noires) compris dans un intervalle autour de la luminance moyenne. La figure du bas montre de façon plus concrète les différents niveaux de contrastes du bruit utilisés dans l'expérience de détermination de seuil (de gauche à droite : 0.5, 0.6, 0.7 et 0.8).

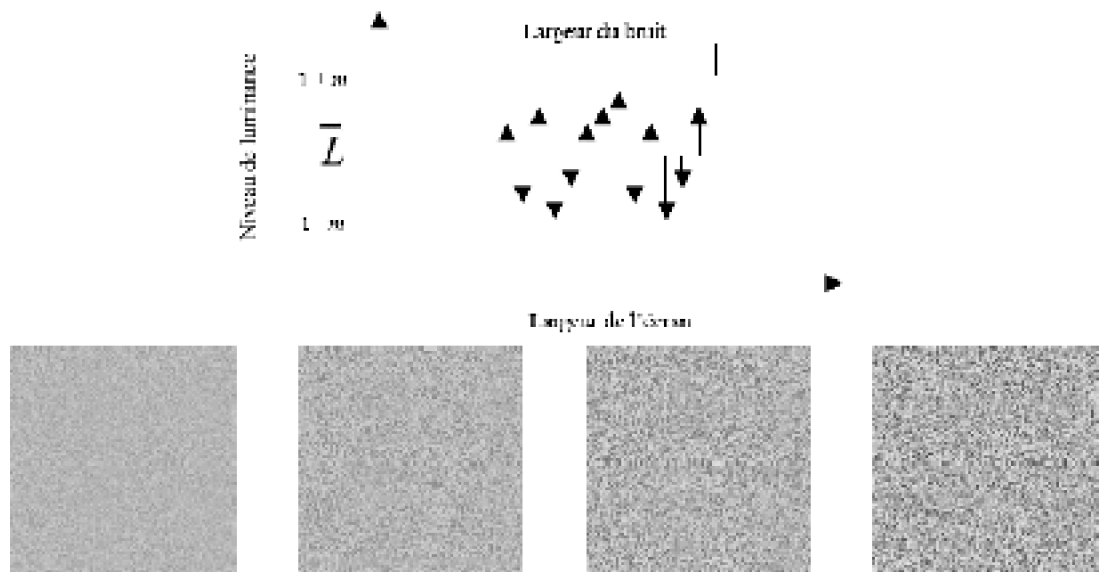


Figure 37 : La figure du haut précise de façon schématique la distribution aléatoire de chaque pixels (représenté par des flèches noires) compris dans un intervalle autour de la luminance moyenne. La figure du bas montre de façon plus concrète les différents niveaux de contrastes du bruit utilisés dans l'expérience de détermination de seuil (de gauche à droite : 0.5, 0.6, 0.7 et 0.8).

Le signal était, à l'inverse du bruit, composé d'une seule valeur de luminance : augmenter son contraste revient à augmenter sa luminance. Les six niveaux de luminance du signal utilisés dans l'expérience pour chaque niveau de bruit étaient choisis pour chaque observateur sur la base d'essais préliminaires et se répartissaient sur une échelle de $10^{0.1}$. Les valeurs de contraste du signal variaient ainsi de 0.045 à 0.1423 pour le bruit 0.5 et celui de 0.6 ; de 0.05 à 0.1581 pour le bruit 0.7 ; de 0.06 à 0.1714 pour le bruit 0.8 (Figure 38). Chaque sujet passait un bloc d'entraînement pour chaque contraste du bruit, sauf le sujet VB qui a répété 6 fois l'entraînement.

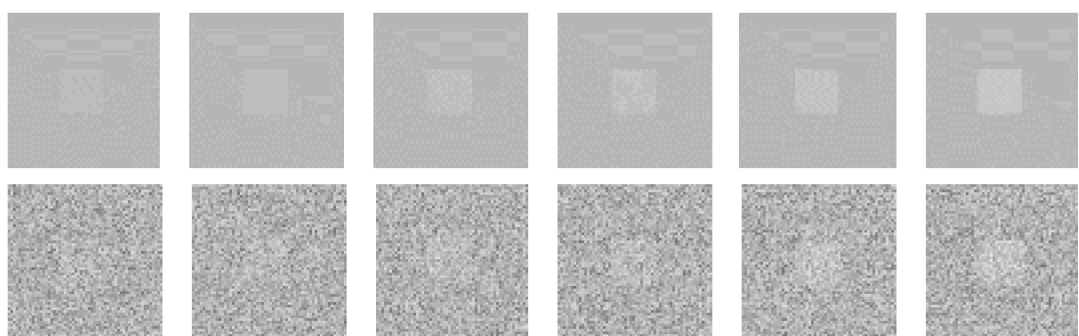


Figure 38 : Les différents contrastes du signal sont présentés ci-dessus (de gauche à droite), et les mêmes signaux présentés dans le bruit (contraste de 0.7) sont présentés au dessous (les contrastes du signal ont été augmentés par rapport aux valeurs réellement utilisées, pour des raisons d'impression).

· Matériel :

Un PC 233 Mhz équipé d'un écran SONY Multiscan SE II (20 pouces) et d'une carte graphique Matrox Millenium II (VGA), était utilisé. La programmation était réalisée à l'aide du logiciel Matlab 5.3 (MathWorks), équipé des deux boîtes à outils « signal processing » et « image processing » ainsi que de Matvis (Neurometrics Institute), un logiciel complémentaire pour les besoins de l'expérimentation. La pièce était faiblement éclairée. Les sujets étaient face à l'écran, une mentonnière permettait de placer l'observateur à 50 cm face à l'écran. La luminance de l'écran a été mesurée par un photomètre (Minolta, CS-100).

Procédure :

Le déroulement d'un essai était le suivant. Un point de fixation était présenté au centre de l'écran pendant une période de 200 ms, et, après un intervalle de 700 ms, le stimulus apparaissait au même endroit pendant 400 ms. L'observateur donnait alors sa réponse de détection (répondre si le signal était présent ou absent) sur une échelle de confiance en cinq points. Il devait évaluer son degré de certitude, selon une échelle de catégorie convenue à l'avance : « le carré est absent », « le carré est peut-être absent », « l'observateur ne sait pas », « le carré est peut-être présent » et « le carré est présent ». Les touches 1, 2, 3, 4, 5 du pavé numérique étaient utilisées pour enregistrer les réponses. Aucune consigne explicite de rapidité n'était formulée, seule la précision de la réponse était exigée. Le nouvel essai était lancé après la réponse. La Figure 39 illustre la séquence temporelle d'un essai.

Avant de débiter les expériences, une présentation du signal seul était réalisée, afin de définir pour les observateurs la forme qu'ils devaient détecter. Cette présentation n'était faite qu'une seule fois.

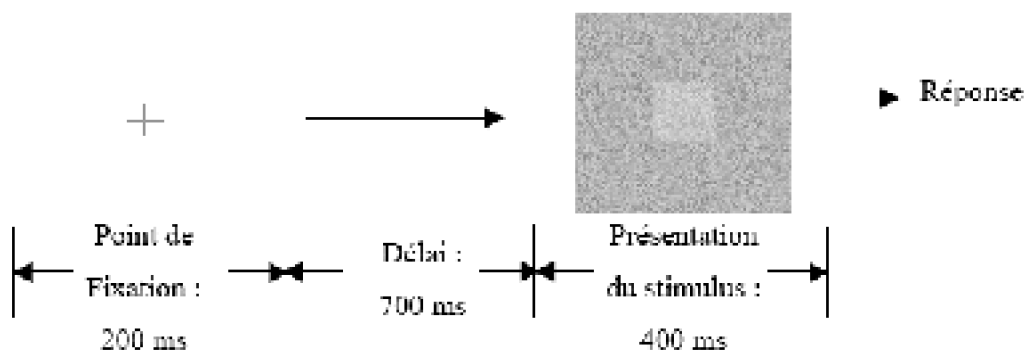


Figure 39 : Diagramme de la séquence d'un essai de l'expérience de détermination de seuil.

L'expérience comportait 16 blocs expérimentaux pour chaque observateur, soit quatre blocs expérimentaux pour chacun des quatre contrastes du bruit. Chaque bloc se composait de 196 essais, soit 28 répétitions des 6 niveaux de contraste du signal plus la condition « signal absent ». Les niveaux de contraste du signal étaient présentés selon un tirage aléatoire sans remise, effectué tous les 7 essais. Deux pauses étaient proposées à l'intérieur d'un bloc. Le signal carré était présenté aux observateurs au début de l'expérience et n'était plus représenté par la suite.

1.2.2. Résultats

Fonctions psychométriques :

Pour chaque contraste du bruit, un seuil de contraste du signal a été déterminé (voir Tableau I). Pour cela, la probabilité de réponse « présent » pour chaque contraste du signal est déterminée de la façon suivante : le nombre total de réponse de type 4 et 5 est divisé par le nombre total de réponse. Ensuite, le seuil, qui est la valeur de luminance du signal permettant de le détecter, est obtenu par l'ajustement de la fonction Weibull aux données de chaque observateur. L'ajustement a été réalisé par un programme sous le langage Matlab qui utilisait la fonction *Fmins* permettant de minimiser les erreurs entre les données et la fonction. La fonction Weibull a la forme suivante (Watson, 1979 ; Watson & Pelli, 1983) :

$$P(c) = 1 - (1 - \gamma)e^{-(c/\alpha)^\beta}$$

Ici, $P(c)$ est la probabilité de bonne réponse en fonction du contraste c , le paramètre α correspond au niveau le plus bas de l'asymptote, β est la paramètre seuil correspondant à $P(c)$, et γ détermine la pente de la fonction psychométrique. Les fonctions sont déterminées à partir du pourcentage de détection correcte et du nombre de fausse alarme pour un niveau de contraste nul (le carré est absent).

La Figure 40 montre les fonctions psychométriques des sujets FD, SH, SD et VB. On s'attendait à ce que la pente de ces fonctions soit identique, et que le seuil de détection augmente lorsque le contraste du bruit augmente, excepté pour les plus petits niveaux du bruit où les seuils de détection doivent être équivalents. Conformément à cette hypothèse, le Tableau I montre que les pentes variaient très peu pour un même sujet. De plus, le seuil de détection augmente lorsque le contraste du bruit augmente à partir du contraste 0.6.

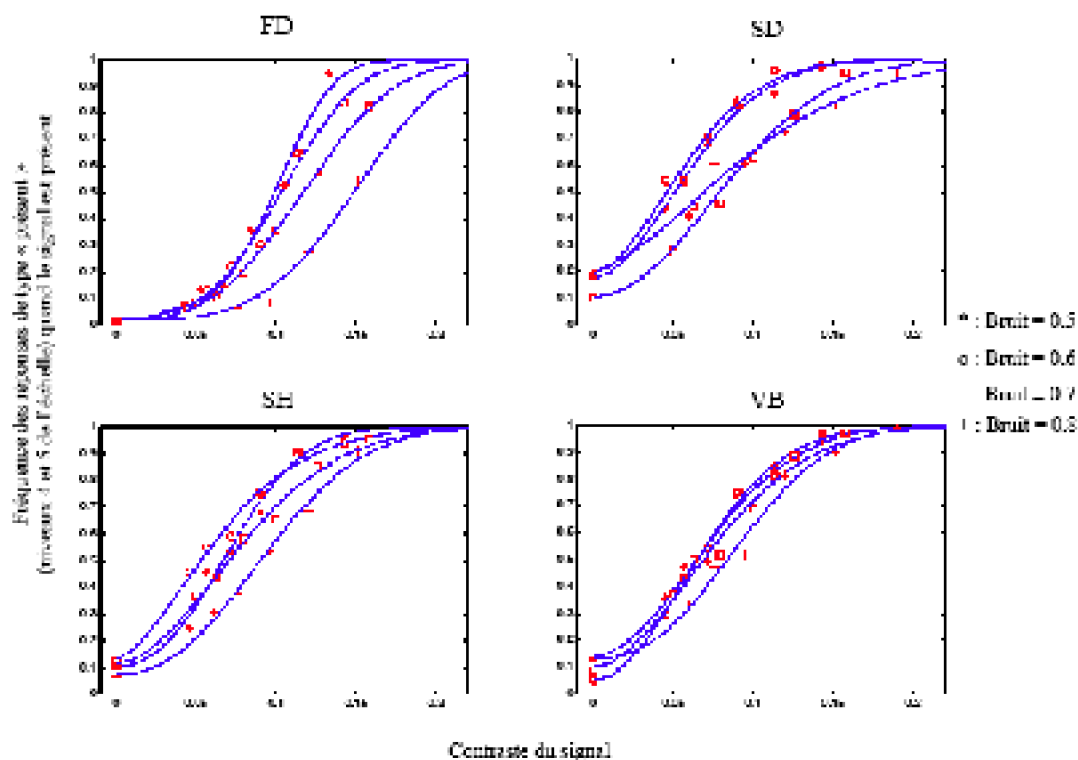


Figure 40 : Expérience préliminaire 1 :Fonctions psychométriques des quatre observateurs pour les quatre contrastes du bruit : 0.5, 0.6, 0.7 et 0.8.

Tableau I : Valeurs des seuils du contraste du signal, des pentes et des probabilités correspondant à chaque seuil pour chaque contraste du bruit des observateurs FD, SD, SH et VB.

	Contraste du bruit	0.5	0.6	0.7	0.8
FD	Seuil	0.1181	0.1168	0.1327	0.1643
	Pente	3.98	3.25	3.02	3.85
	$P(c)$	0.6342	0.6430	0.6393	0.64
SD	Seuil	0.0742	0.07	0.1033	0.1123
	Pente	1.83	1.75	2.03	1.59
	$P(c)$	0.6965	0.7035	0.6709	0.7074
SH	Seuil	0.0844	0.0769	0.0953	0.1110
	Pente	2.34	1.59	1.84	2.29
	$P(c)$	0.6706	0.6801	0.6740	0.6588
VB	Seuil	0.0881	0.0818	0.0916	0.1069
	Pente	1.96	1.96	1.85	2.41
	$P(c)$	0.6843	0.6511	0.6696	0.6805

Les seuils de détection sont également présentés pour chaque contraste du bruit et pour chaque observateur dans la Figure 41. On constate que la performance augmentait entre les contrastes du bruit de 0.6 et 0.7, mais aussi 0.7 et 0.8. Par ailleurs, les seuils de détection avaient tendance à diminuer entre les contrastes du bruit de 0.5 et 0.6 pour trois observateurs. L'augmentation du contraste du bruit engendrait une augmentation du contraste seuil du signal à partir du niveau de bruit de 0.6 pour les quatre observateurs. Au-delà de cette valeur, il semble que le bruit expérimental soit supérieur au bruit interne du système. On observe également que trois observateurs avaient des performances très proches alors que la sensibilité de l'observateur FD était inférieure (ses seuils étaient plus élevés). Des analyses supplémentaires ont été réalisées afin de déterminer si un biais de réponse ne pouvait pas expliquer cette différence de performance entre les observateurs. Pour cela, la sensibilité de chaque observateur a été calculée à partir de l'indice d' de la TDS.

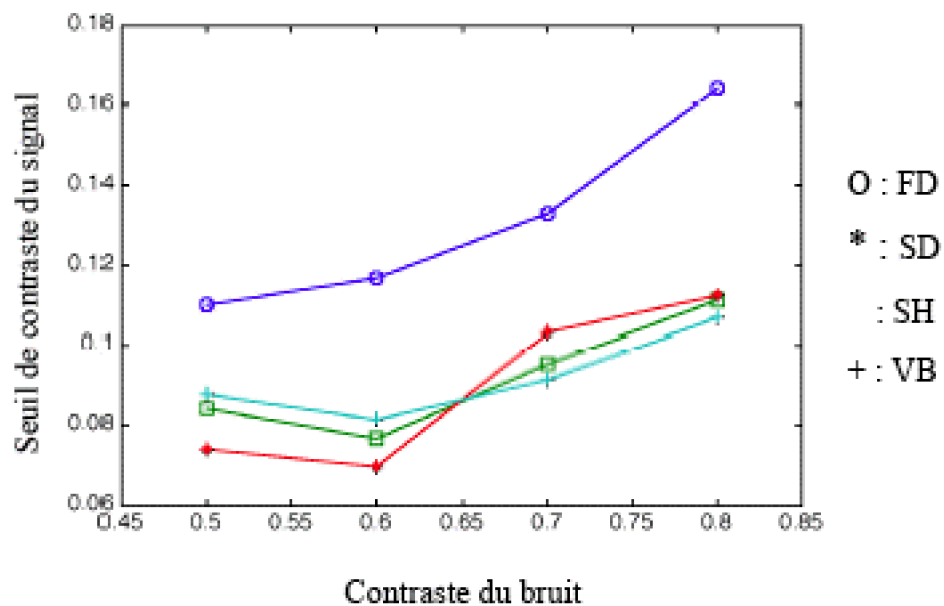


Figure 41 : Expérience préliminaire 1 : Relation entre les seuils de contraste du signal et le contraste du bruit des quatre observateurs.

- Sensibilité : l'indice d' .

La proportion de réponse pour chaque niveau de certitude était relevée, ce qui a permis de calculer la sensibilité de l'observateur au moyen de l'indice d' à partir des courbes ROC. La formule que nous avons utilisée pour déterminer cet indice est présentée dans la partie théorique.

On le préfère au pourcentage de réponse correcte, car c'est une mesure de performance indépendante de la place du critère de réponse du sujet. Pour chaque niveau de contraste du signal, un indice d' a été calculé, ces résultats sont présentés à la Figure 42.

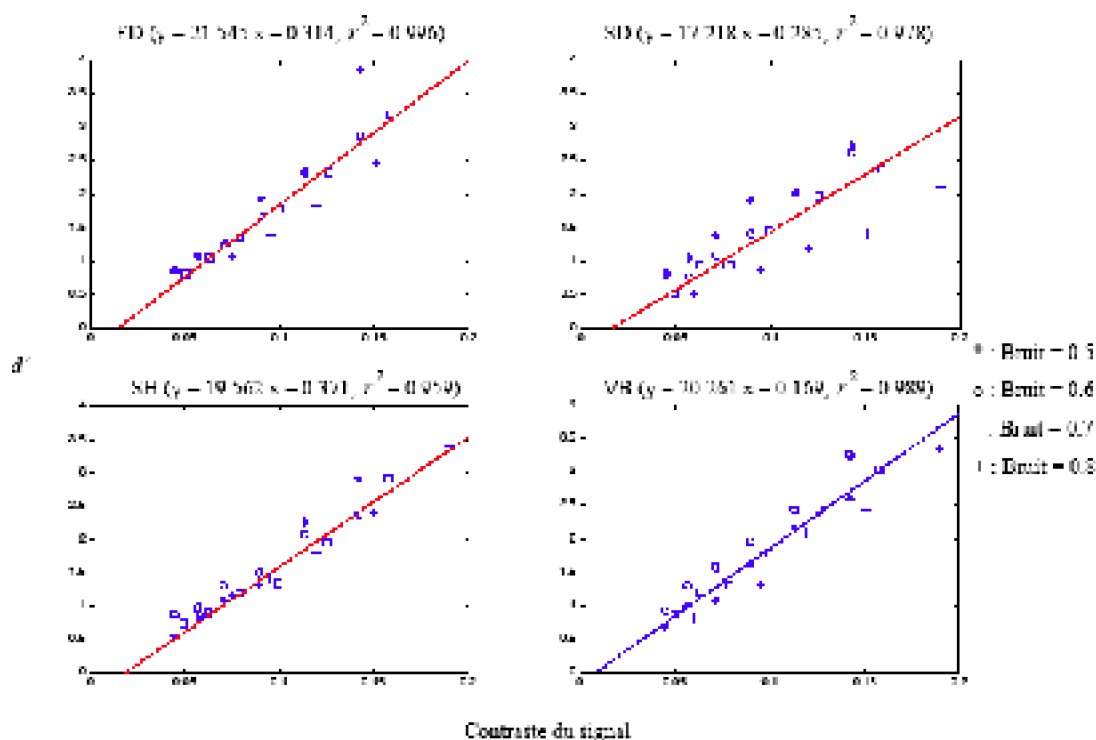


Figure 42 : Expérience préliminaire 1 : Effet du contraste du bruit (dans le corps du graphique) en fonction du contraste du signal (en abscisse), sur la sensibilité de la

détection (en ordonnée) pour les quatre sujets sur les quatre graphiques. La droite de régression de la condition de contraste du bruit 0.7 est affichée.

Pour estimer les seuils de détection individuels, nous avons choisi d'utiliser un seuil équivalent à un $d' = 1$ (avec le même raisonnement que précédemment pour un pourcentage de réponse correcte). Ce niveau correspond à une performance plus faible qu'un seuil de détection correcte, et il est classiquement utilisé pour estimer un seuil dans le cadre de la TDS (MacMillan & Creelman, 1991). De la même façon que pour le pourcentage de réponse correcte, deux étapes sont nécessaires pour déterminer les seuils de contraste du signal. La première étape consistait à calculer les droites de régression pour chaque observateur. Puis, dans une seconde étape, le seuil de contraste du signal était déterminé par interpolation pour un d' égal à 1. L'ensemble des contrastes seuil du signal pour un d' égal à 1 sont présentés dans le Tableau II et dans la Figure 43.

La comparaison des performances des 4 observateurs (voir Figure 43 et le Tableau II) montrait en outre que les seuils de détection de FD et VB étaient globalement inférieurs à ceux de SD et SH. Lorsqu'un indicateur de performance non biaisé est utilisé pour calculer le seuil, l'observateur FD a une sensibilité meilleure que SD et SH alors que les résultats précédents, contaminés par le biais, amenaient à croire l'inverse. De façon similaire à l'analyse portant sur le pourcentage de détection correcte, le contraste du signal augmentait lorsque le contraste du bruit augmentait, mais ceci n'est pas vrai pour les plus bas niveaux de contraste du bruit, c'est-à-dire les contrastes de 0.5 et 0.6 (voir la Figure 43, ainsi que la Figure 44 pour un résumé des données).

Tableau II : Valeurs des seuils du contraste du signal correspondant à un d' de 1 pour chaque contraste du bruit (0.5, 0.6, 0.7 et 0.8) et pour chaque observateur.

	0.5	0.6	0.7	0.8
FD	0.061	0.055	0.061	0.065
SD	0.053	0.063	0.074	0.099
SH	0.066	0.067	0.07	0.071
VB	0.06	0.047	0.058	0.071

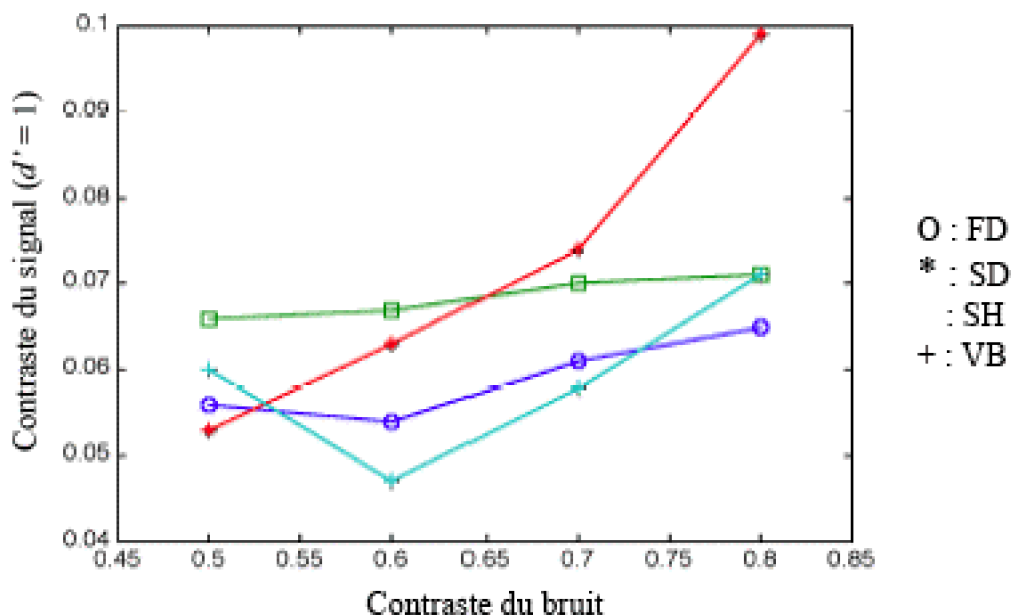


Figure 43 : Expérience préliminaire 1 : Relation entre le contraste du signal correspondant à un indice d' de 1 et le contraste du bruit pour chaque observateur.

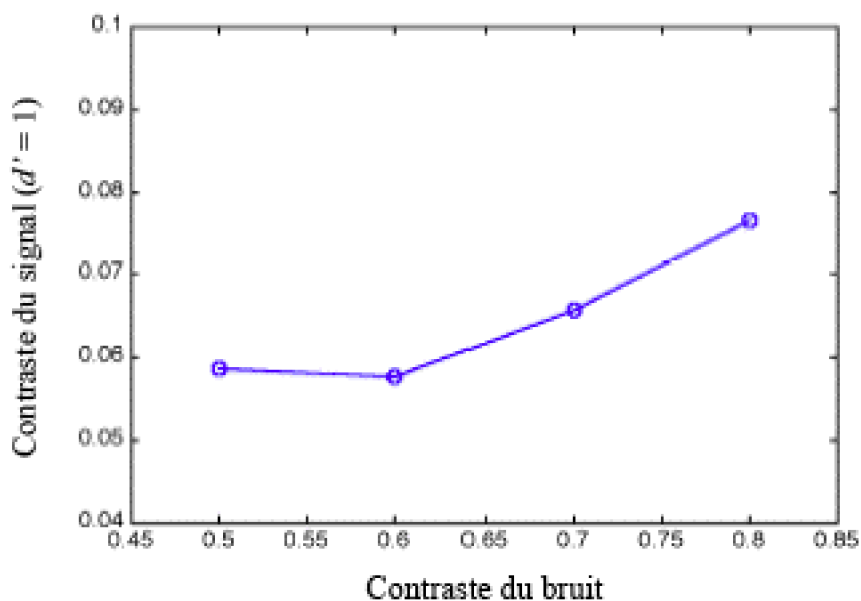


Figure 44 : Expérience préliminaire 1 : Relation entre la moyenne des contrastes du signal correspondant à un indice d' de 1 et le contraste du bruit.

Une analyse des courbes ROC (« Receiver Operating Characteristic curve ») a été effectuée. Ces courbes sont affichées dans l'Annexe 1. Une courbe ROC est représentée pour chaque contraste du signal et pour chaque niveau de contraste du bruit en fonction des cinq types de réponses possibles. La représentation graphique des courbes ROC nous permettait de vérifier si la variance entre le signal et le signal+bruit était similaire. Les résultats indiquaient que plus le contraste du signal augmentait, et plus le signal se

distinguaient facilement du bruit. Ce qui se traduit par une courbe de plus en plus creuse avec l'augmentation du contraste du signal. En d'autres termes, l'indice d' était de plus en plus important. Par ailleurs, pour un même niveau de contraste du bruit, les courbes étaient proches pour l'ensemble des observateurs.

Pour l'expérience de détermination de l'image de classification, nous avons préféré utiliser le calcul basé sur le d' , à la place du pourcentage de détection correcte. Deux raisons nous ont amené à ce choix : la première concerne l'existence du biais que nous avons mis en évidence dans la comparaison entre le pourcentage de détection correcte et l'indice d' , la seconde se rapporte au fait qu'un $d' = 1$ correspond à un pourcentage de détection correcte inférieure à celui du seuil de la fonction Weibull (MacMillan & Creelman, 1991). En effet, un nombre élevé d'erreurs est nécessaire pour maximiser l'influence du bruit et pour obtenir suffisamment d'observations dans l'expérience suivante.

Cette expérience avait pour objectif de choisir un niveau de bruit supérieur à celui du bruit interne. La Figure 43 indique les seuils de contraste du signal sont stables pour la moitié des observateurs entre les contrastes du bruit de 0.5 et 0.6 et la Figure 44 montre que en moyenne les seuils de contraste du signal sont équivalents pour ces deux niveaux de bruit. Par contre, les seuils augmentaient pour chaque observateur à partir du contraste du bruit de 0.7. C'est à partir de ce niveau que le contraste bruit expérimental était supérieur au bruit interne des observateurs.

Nous n'avons pas utilisé de situation contrôle dans laquelle il n'y aurait pas de bruit pour une raison d'ordre technique, en effet nous n'avons pas la possibilité de diminuer suffisamment le contraste du signal (en raison du nombre restreint de niveaux différents dans la LUT) pour obtenir un nombre minimum de réponse de type fausse alarme. Ce qui nous aurait permis de déterminer que notre bruit était effectivement perturbant.

Nous avons choisi pour l'expérience de détermination de l'image de classification un contraste du bruit de 0.7. De plus, les niveaux individuels de contraste du signal qui était associés à ce niveau de contraste du bruit ont été utilisés dans l'expérience principale. Les droites de régression, correspondant au niveau de bruit de 0.7, sont présentées à la Figure 42, et le seuil de contraste du signal de chaque observateur pour ce niveau de bruit est indiqué dans le Tableau II.

1.3. Détermination de l'image de classification

1.3.1. Introduction

Le but de l'expérience est de connaître quelle information visuelle est la plus pertinente dans la tâche de détection : le contour, ou la surface conformément au modèle proposé par Grossberg et al. (Grossberg, 1994 ; Grossberg, 2000 ; Grossberg & Mingolla, 1985 ; Grossberg & Todorovics, 1988 ; Grossberg, Mingolla & Ross, 1997). De plus, l'existence de ce traitement hiérarchique, mais aussi du processus de remplissage, a été mise en évidence par de nombreuses études qui utilisaient une surface uniforme (Krauskopf, 1963 ; Paradiso & Hahn, 1996 ; Paradiso & Nakayama, 1991 ; Rossi & Paradiso ; 1996 ; Stoper

& Mansfield, 1979). Toutefois, notre approche se distingue de l'ensemble de ces expériences par le fait que nous utilisons une tâche de détection. L'utilisation de cette tâche est rendue possible grâce à l'utilisation du paradigme de classification perceptive. Cette technique suppose que les zones du signal qui émergent dans l'image de classification donnent des indications sur la nature des informations qui sont utilisées par l'observateur dans sa prise de décision.

La tâche de détection est une tâche extrêmement simple, théoriquement, la réponse (présent/absent) peut être déclenchée dès qu'une information visuelle du signal est encodée. Dans une perspective traditionnelle, la rétine ne contient pas de cellule sensible à la luminosité, seulement au contraste. Par conséquent, la seule information véhiculée au cours du traitement est le contour des objets, et la surface est reconstruite de façon indépendante à ce qui est présenté réellement. Dans ce cas, un contraste positif devrait émerger au niveau des contours du carré dans l'image de classification sans la surface, la zone centrale ne devrait pas être différente du fond. Cette alternative est représentée dans la Figure 45a.

Si il existe des cellules sensibles à la luminosité, alors la surface pourrait très bien participer au traitement même faiblement. Nous avons effectivement vu que des neurones situés dans le CGL et V1 étaient stimulés par une surface uniforme (Rossi & Paradiso, 1999 ; MacEvoy, Kim & Paradiso, 1998 ; Rossi, Rittenhouse & Paradiso, 1998). Toutefois, il est possible que dans ces expériences la stimulation des cellules soit la conséquence de la propagation de l'information à partir d'un contraste. Dans ce cas, l'information principale reste le contour avec en plus une participation de la surface. Dans l'image de classification, l'information de contour serait présente et la luminance de la partie centrale se situerait à niveau intermédiaire entre les contours et le fond. Cette hypothèse est représentée à la Figure 45b.

Par ailleurs, nous pouvons nous attendre à obtenir dans l'image de classification un résultat qui soit conforme au signal que nous avons présenté durant l'expérience. Ceci suppose que l'observateur ajuste sa représentation exactement au signal, il réaliserait alors la tâche de manière optimale (alternative 45c).

Enfin, il est également possible d'obtenir un résultat alternatif par rapport aux deux premières hypothèses qui font référence à un traitement hiérarchique. Si l'information de surface est prévalente à celle du contour alors l'image de classification l'intègre. En effet, nous pourrions obtenir une image dont la réponse de l'observateur serait centrée uniquement sur la zone centrale du signal (Figure 45d).

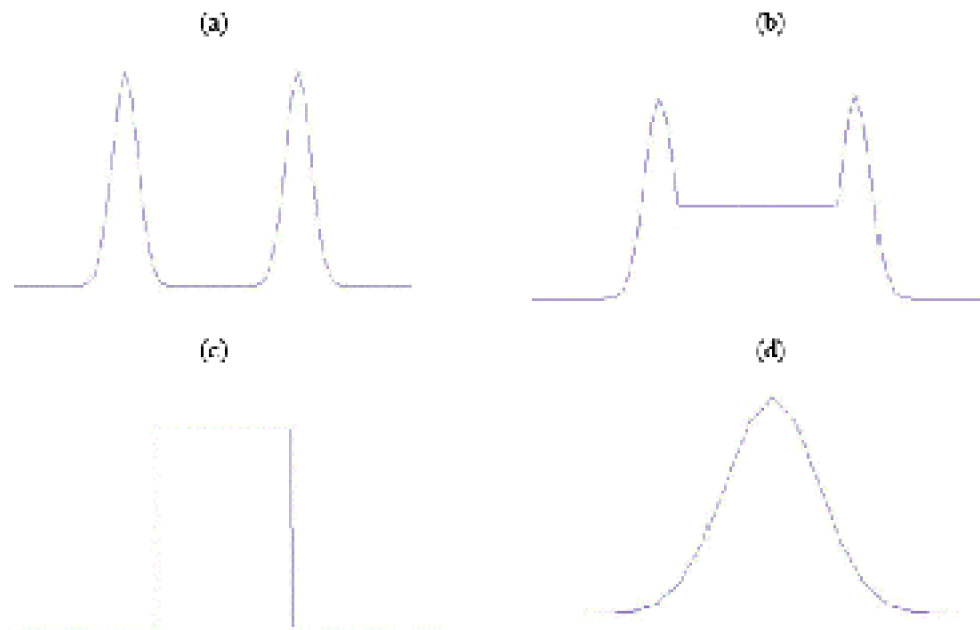


Figure 45 : Hypothèses concernant l'image de classification lorsque le signal est un carré.

La Figure 46a représente l'hypothèse d'une influence du contour seul, la Figure 46b montre l'hypothèse pour laquelle les deux dimensions seraient présentes, la Figure 46c illustre l'hypothèse dans laquelle le signal que nous avons présenté ressort dans nos résultats, enfin la Figure 46d indique l'image que nous devrions obtenir si la surface émerge dans l'image de classification. Les images sont affichées de façon unidimensionnelle mais elles sont généralisables pour des images bi-dimensionnelles.

Comme cela est mentionné dans la partie théorique, l'image de classification est calculée de la façon suivante : les bruits tirés aléatoirement à chaque essai sont triés en fonction des réponses : détection correcte (DC), rejet correct (RC), fausse alarme (FA) et omission (OM). Une moyenne pondérée des bruits est alors calculée pour chacune de ces réponses (Ahumada, 1996) :

$$\text{Image de classification} = (\overline{DC} + \overline{FA}) - (\overline{OM} + \overline{RC})$$

Si un nombre suffisant d'essais est pris en compte, les indices émergent significativement dans l'image. Cette analyse se fonde sur l'a priori selon lequel dans le stimulus composite qui est envoyé à chaque essai au système visuel (le signal plus le bruit particulier à cet essai), la réponse est non seulement influencée par le signal lui-même, mais également par la structure du bruit, surtout dans des conditions de visibilité faible du signal, comme ici. La structure du bruit est aléatoire, elle est tirée à chaque essai, mais pour un essai donné elle est fixe. Ainsi, la moyenne des bruits ayant engendrés la réponse « signal présent » contient les indices qui aide le système à donner cette réponse, et la moyenne des bruits ayant engendrés la réponse « signal absent » contient les indices anti-corrélés au signal.

1.3.2. Méthode expérimentale

- Participants.

Les observateurs sont les mêmes que dans l'expérience de détermination de seuil (FD, SD, SH et VB).

- Stimuli.

Les stimuli étaient composés, comme dans la première expérience, du signal superposé à un bruit, ou d'un bruit seul. Mais le contraste du bruit (0.7 pour tous les observateurs) et le contraste du signal (0.06 pour FD et VB, et 0.07 pour SH et SD) étaient fixes.

- Procédure.

Le déroulement d'un essai était identique à l'expérience précédente (un point de fixation pendant 200 ms, un intervalle de 700 ms, puis le stimulus pendant 400 ms) mais les alternatives de réponse étaient simplifiées. Le sujet devait répondre à la présence/absence du signal en s'aidant des touches 1 (présent) et 0 (absent) du pavé numérique.

Les sujets passaient 30 blocs de 200 essais, à l'exception de FD qui a passé 50 blocs. Sur ces 200 essais, le signal était présent/absent dans la moitié des cas. Un tirage aléatoire, tous les 8 essais (parmi lesquels 4 signaux étaient présents), était nécessaire pour éviter un trop grand nombre de réponses identiques à la suite. Chaque bloc comprenait donc 100 répétitions pour chacune des deux conditions expérimentales, ce qui fournit 3000 répétitions par condition sur l'ensemble de l'expérience (5000 répétitions pour l'observateur FD). Les observateurs étaient également prévenus que le signal n'était présent que dans la moitié des essais et que ce dernier était tiré de façon aléatoire.

1.3.3. Résultats

- La sensibilité et le critère de décision

La sensibilité de la détection du signal a été appréciée par l'indice d' . Les valeurs de contraste du signal avaient été choisies sur la base des résultats de l'expérience

précédente pour que la performance soit faible ($d' \approx 1$). Un indice global d' (sur l'ensemble des blocs) a été calculé pour vérifier cette affirmation. L'indice d' était également calculé tous les 200 essais pour tester un éventuel effet d'apprentissage. Les résultats montraient que la sensibilité globale était proche de 1 (FD = 1.15, SD = 1.03, SH = 1.28 et VB = 1.31). Toutefois, il est important de noter que l'indice d' ne variait pas au cours de l'expérience, comme le montre la Figure 39. Ceci indique que les observateurs étaient à leur plateau de performance, pour ces conditions expérimentales, ce qui est conforme à l'apprentissage important déjà acquis par ces observateurs qui ont passé un nombre élevé d'essais. Nous avons également calculé un critère de réponse pour chaque observateur à partir de la formule qui est indiquée dans la partie théorique. Un calcul global du critère indique que celui-ci est proche de 0 (FD = -0.18, SD = -0.05, SH = -0.07 et VB = 0.15). De plus, un critère a été calculé tous les 200 essais, le résultat est affiché à la Figure 46. Les résultats indiquaient que la place du critère de décision est stable tout au long de l'expérience.

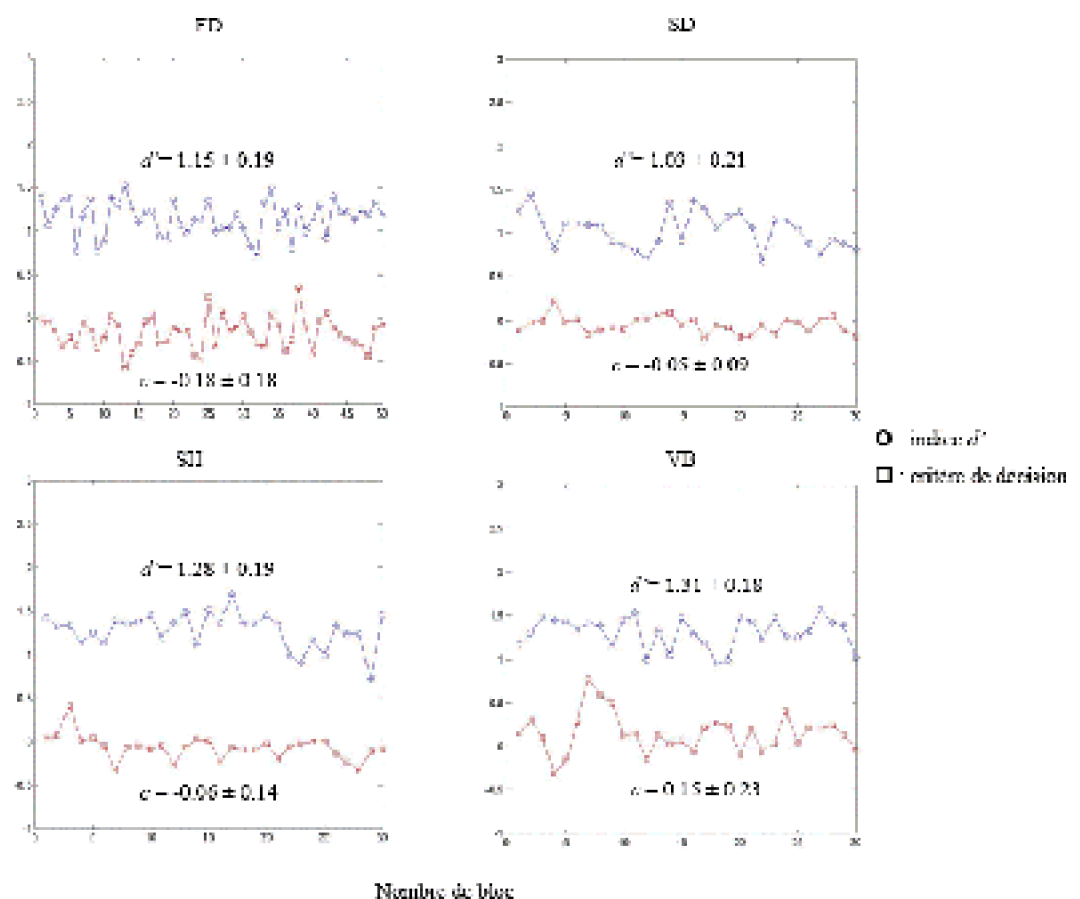


Figure 46 : Expérience 1 : Résultats de l'indice d' et du critère de décision calculés tous les

200 essais pour chaque observateur. L'indice d' et l'écart-type sont précisé pour chaque distribution.

L'analyse de l'image de classification s'est déroulée en quatre étapes. Une moyenne des bruits selon les quatre réponses possibles a tout d'abord été calculée, puis l'image de classification a été dérivée de ces moyennes, elle a ensuite subi un lissage et un traitement statistique, puis le contraste des pixels significatifs a été exagéré pour qu'on puisse visualiser les zones de l'image qui participait significativement à la réponse de l'observateur.

Moyenne des bruits

La Figure 47 présente la moyenne des bruits pour chaque type de réponse : détection correcte (DC), rejet correct (RC), fausse alarme (FA), et omission (OM) pour les observateurs FD, SD, SH et VB. Une zone centrale claire apparaissait pour les deux moyennes de bruit correspondant aux réponses « signal présent » (FA et DC) alors qu'une zone sombre est observée à cet endroit dans les moyennes des bruits correspondant à la réponse « signal absent » (OM et RC). Cette zone de l'image correspondait à l'emplacement du signal.

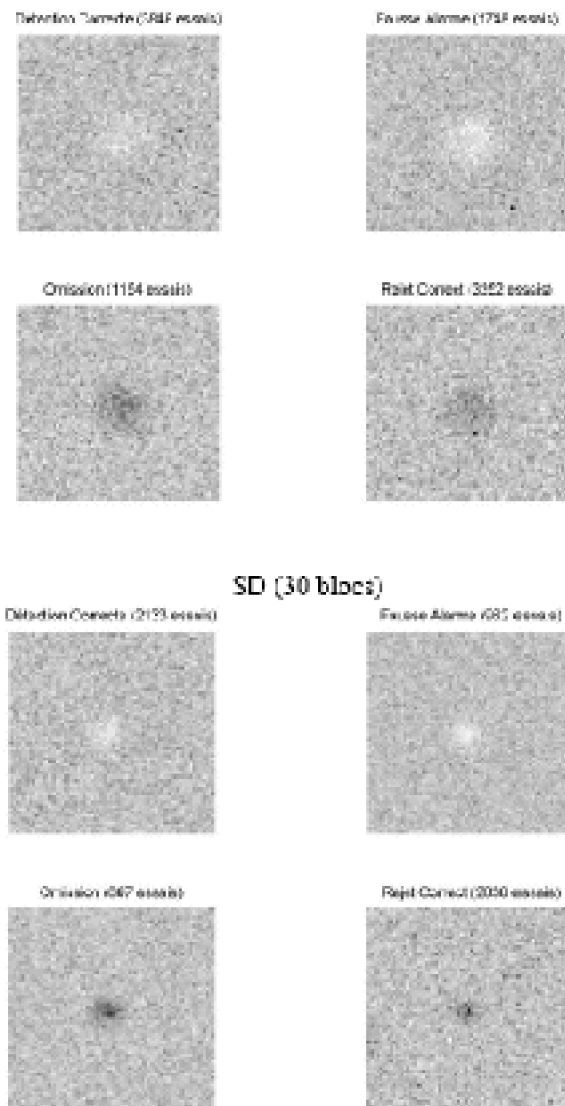
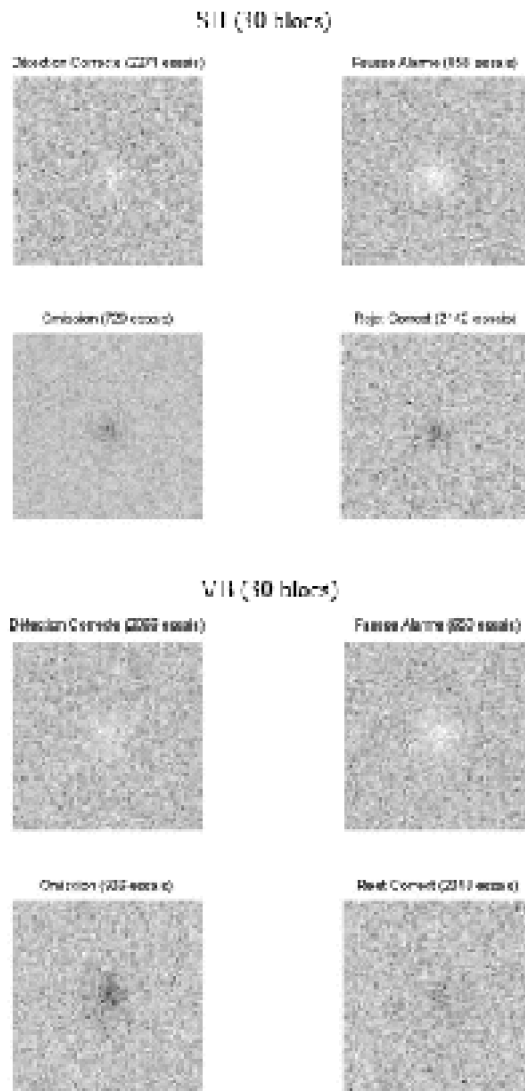


Figure 47 : Expérience 1 : Moyenne des bruits pour chaque type de réponse (DC, FA, RC et OM).



- Image de classification

Les images de classification dérivées de ces moyennes de bruit, pour les quatre observateurs, sont présentées dans la Figure 48. Une zone claire centrale est observée à l'emplacement du carré. Les résultats sont similaires pour les quatre observateurs.

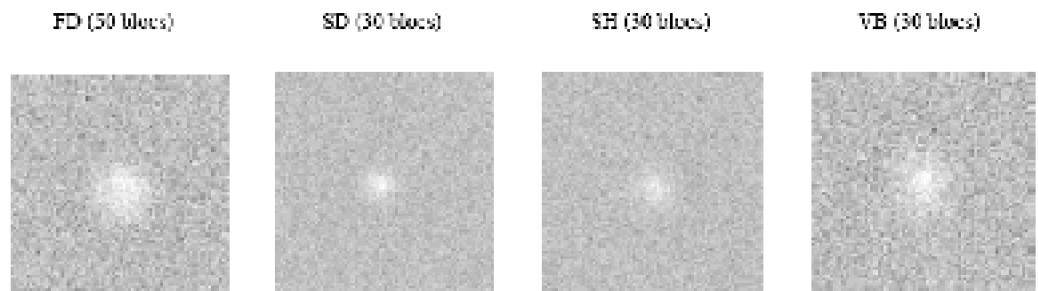


Figure 48 : Expérience 1 : Image de classification de chaque observateur pour un signal de

3 deg et un temps de présentation de 400 ms.

Lissage et test statistique

Le lissage de l'image de classification a été effectué à partir de la moyenne pondérée des pixels contigus. Ce lissage correspond à un filtrage triangulaire (9*9 pixels) de l'image et sert à dégager les éléments importants des détails. Les images de classification lissées ont été soumises à un test statistique. Pour chaque pixel, un test de significativité a été réalisé à partir d'un intervalle de confiance théorique ($p < .05$), cet intervalle était déterminé à partir de l'erreur standard de la distribution uniforme du bruit. La Figure 49 représente le résultat après le test statistique effectué sur l'image de classification lissée, pour les observateurs FD, SD, SH et VB. Dans ces images, les pixels représentant une moyenne qui différait significativement de la luminance du fond ont été préservés alors que les autres ont été ramenés à la luminance du fond. Pour les quatre observateurs, une zone claire apparaissait à l'emplacement du carré et un faible contraste négatif l'entourait (difficilement visible sur la version imprimée). Le test statistique valide ainsi le résultat observé de façon descriptive dans l'image de classification.

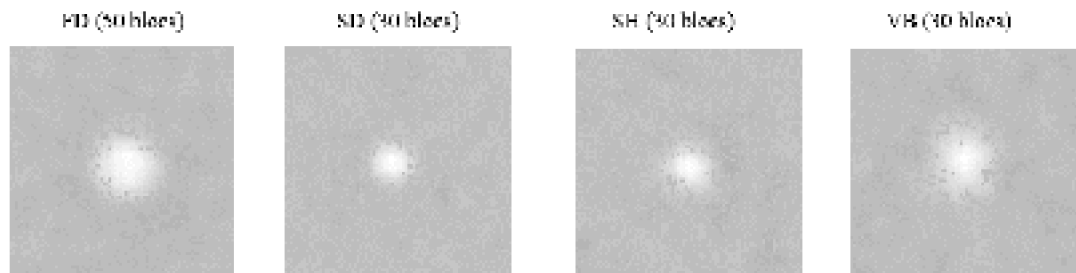


Figure 49 : Expérience 1 : Test statistique effectué sur l'image de classification de chaque observateur pour un signal de 3 deg et un temps de présentation de 400 ms.

z-score

De la même façon que pour le test statistique, l'image de classification est tout d'abord lissée par un filtre triangulaire de 9*9. Un test a ensuite été réalisé à partir de la division entre chaque pixel de l'image de classification lissée et l'erreur standard. Les images étaient identiques de celles obtenues pour l'image statistique de la Figure 50, mais ici les niveaux de gris représentent l'erreur standard.

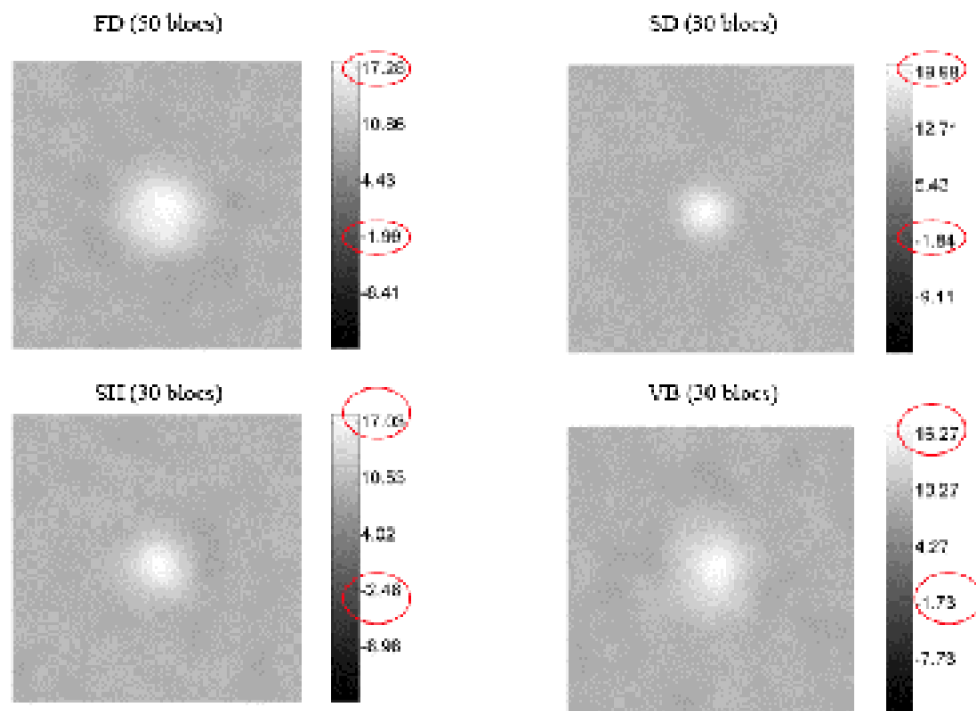


Figure 50 : Un test « z-score » a été réalisé pour chacune des images de classification. La barre situé sur la droite des images permet d'estimer le niveau de luminance qui correspond au valeur du test. De plus, les valeurs maximales et minimales sont entourées pour chaque observateurs.

- Pixels significatifs dans l'image de classification.

A la suite du test statistique, et afin de fournir une appréciation globale des zones qui contribuaient significativement à la réponse, l'amplitude des moyennes des pixels a été seuillée. Ainsi, les moyennes significativement plus grandes que la luminance du fond sont représentées par une luminance maximum en blanc (contraste positif) et les pixels significativement plus petits que la luminance moyenne sont représentés en noir (contraste négatif). Les pixels non significatifs apparaissaient en gris (voir Figure 51).

Le seuillage des pixels significatifs permettait de visualiser les régions, différent significativement du fond, qui passaient inaperçues dans l'image lissée puisque leur valeur était très proche de celle du fond (i.e. l'amplitude de leur contribution est faible).

Cette manipulation a permis de mettre en évidence un résultat qui était jusque-là passé inaperçu et complètement inattendu : des zones aux contraste positif et négatif émergeaient sur les bords extrêmes de l'image de classification. Ceci signifie que la réponse « signal présent » était favorisée quand l'extrême pourtour du masque comportait des pixels cohérents (plus clairs et plus sombres) sur une région assez large. Ces zones, qui sont très éloignées du signal, contribuaient donc significativement à la réponse de détection. Ainsi, il semble que les observateurs ne se limitaient pas à la zone contenant le stimulus, mais basaient également leur réponse sur des détails beaucoup plus éloignés. Toutefois, ce résultat est à relativiser puisque comme le montre l'amplitude des valeurs moyennes, les zones de l'extrême pourtour participent beaucoup plus faiblement que les deux zones centrales.

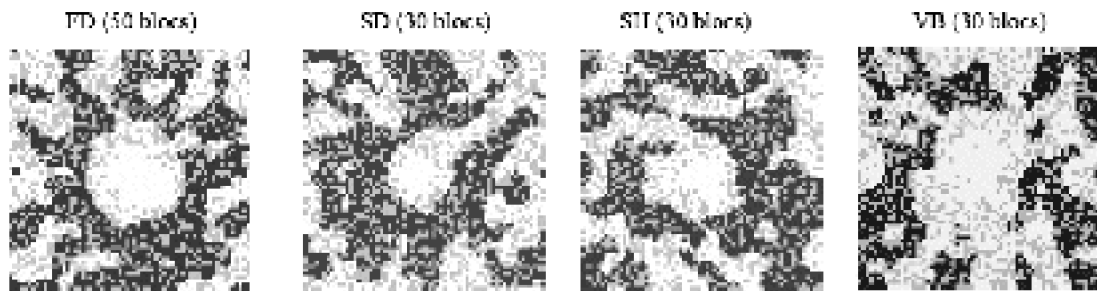


Figure 51 : Expérience 1 : Pixels significatifs qui émergent dans l'image de classification de chaque observateur pour un signal de 3 deg et une durée de présentation de 400 ms.

Estimation de la taille par ajustement de la différence de deux gaussiennes (« DoG »).

L'objectif de cette analyse était de caractériser la zone centrale claire et le pourtour obtenu dans l'image de classification. Les résultats issus de l'image de classification (zone claire au centre et zone sombre au pourtour), indiquaient en effet que ces données sont proches d'une fonction en « chapeau mexicain » ou fonction « DoG » (Difference of Gaussians, voir un exemple Figure 52). La fonction « DoG » était obtenue par la soustraction de deux fonctions gaussiennes, l'une pour le centre (c) et l'autre pour le pourtour (p). La Figure 53 illustre ces deux fonctions. Une fonction gaussienne se définit par deux valeurs : la valeur crête (l'amplitude : A_c , A_p) et l'écart type (le sigma : σ_c , σ_p). La fonction DoG se présente sous la forme suivante :

$$DoG_{2D}(x, y) = A_c e^{-\pi \left(\frac{(x-x_c)^2 + (y-y_c)^2}{\sigma_c^2} \right)} - A_p e^{-\pi \left(\frac{(x-x_p)^2 + (y-y_p)^2}{\sigma_p^2} \right)}$$

Dans cette équation, x_0 et y_0 sont les coordonnées centrales de la fonction gaussienne, et x, y sont les coordonnées spatiales de l'image. L'analyse avait pour but d'ajuster individuellement la fonction « DoG » aux données des observateurs. L'ajustement de cette fonction permettait d'obtenir les valeurs de la taille et de l'amplitude de la zone centrale et du pourtour. Le résultat de cet ajustement, pour chaque sujet, est montré à la Figure 54. Les valeurs du σ et de l'amplitude pour la zone centrale et le

pourtour sont présentées dans le Tableau III. On observe que la zone centrale correspondait à une taille comprise entre 2.3 et 4.3 deg, et le pourtour à une taille comprise entre 7 et 10 deg. Le calcul du rapport taille pourtour / taille centre montre que le pourtour avait une taille de 2 à 4 fois plus grande selon les observateurs que la zone centrale. L'amplitude de la zone centrale variait de 3 à 5 et l'amplitude du pourtour de 0.01 à 0.15 et le rapport entre l'amplitude centrale et l'amplitude du pourtour indiquait que le pourtour représentait entre 3 et 8% selon les observateurs de la luminance de la zone centrale.

L'ajustement de la fonction « DoG » permet ainsi de d'apprécier quantitativement la participation du traitement de la surface du signal. L'ajustement a permis également de déterminer la position de la zone centrale par rapport au fond. Les résultats, qui sont présentés dans les deux dernières colonnes du Tableau III, montraient un faible biais dans l'estimation (inférieure à 13 min) de la position du signal.

L'ensemble de ces résultats montrent qu'il y a quelques différences individuelles notamment sur le rapport de taille. Ces différences pourraient indiquer que les observateurs utilisaient des prototypes qui leur sont propres pour donner leur réponse. Il est également possible que cela soit dû à des mouvements oculaires ou des erreurs de fixation. Mais le résultat majeur, solidement observé pour les 4 observateurs est que la zone correspondant à la surface du signal domine l'image.

Tableau III : Représentation des valeurs d'amplitude (colonnes 2 et 4) et du sigma (colonnes 1 et 3) pour la zone du pourtour et la zone centrale, le rapport entre ces deux zones est également présenté (colonnes 5 et 6), ainsi que la position (colonnes 7 et 8), de la zone centrale pour les quatre observateurs.

Tableau III : Représentation des valeurs d'amplitude (colonnes 2 et 4) et du sigma (colonnes 1 et 3) pour la zone du pourtour et la zone centrale, le rapport entre ces deux zones est également présenté (colonnes 5 et 6), ainsi que la position (colonnes 7 et 8), de la zone centrale pour les quatre observateurs.							x_0 (min)	Y_0 (min)
	Zone centrale		Pourtour		Rapport pourtour / zone centrale			
Sujet	Sigma (deg)	Amplitude	Sigma (deg)	Amplitude	Sigma	Amplitude		
FD	4.31	3.52	10.55	0.028	2.45	0.08	12.3	12.9
SD	2.31	5.05	9.76	0.15	4.23	0.03	-13.5	6.3
SH	2.85	3.72	12.18	0.1	4.26	0.03	4.9	15.5
VB	4.05	2.84	7.02	0.01	1.75	0.04	4.6	-3.5

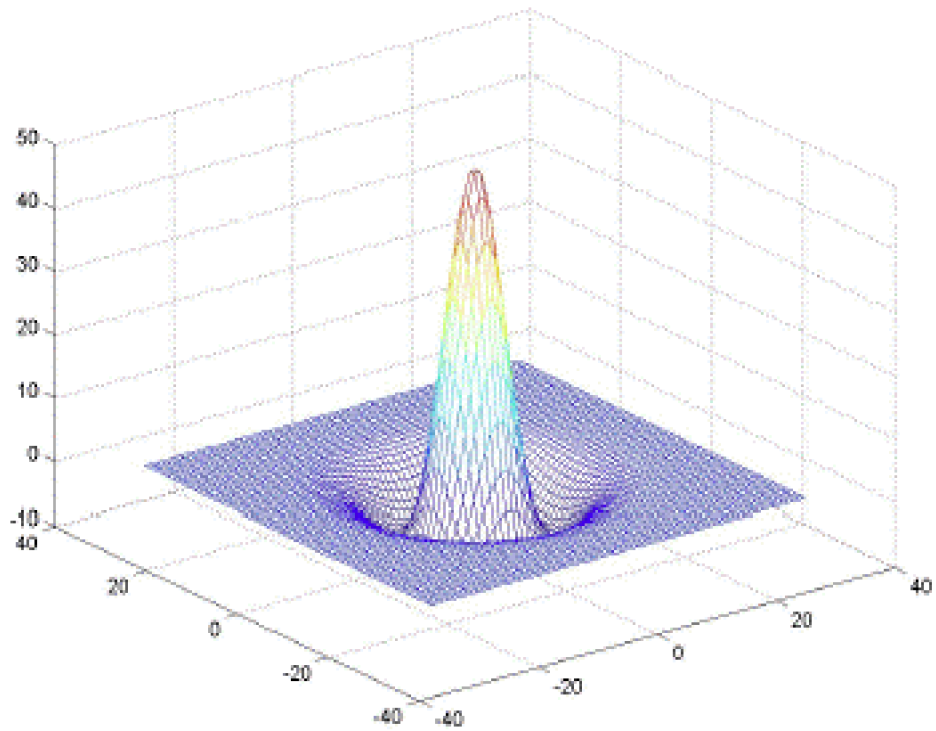


Figure 52 : Représentation d'une fonction « DoG » qui ajuste les résultats obtenus dans l'image de classification.

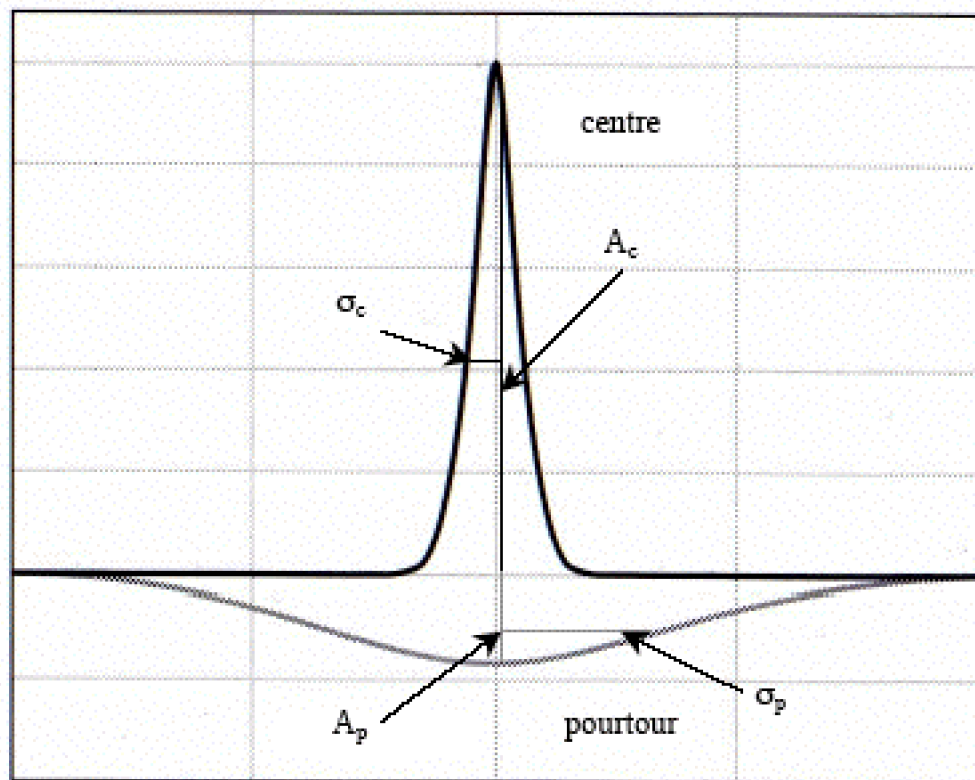


Figure 53 : La fonction gaussienne est représentée par un trait foncé, elle correspond à la zone centrale, et la fonction gaussienne qui est représentée par un trait gris clair correspond à la zone du pourtour, la soustraction entre ces deux fonctions permet d'obtenir une fonction DoG. L'amplitude (A_p , A_c) et le sigma (σ_c , σ_p) sont indiqués pour chacune des deux fonctions.

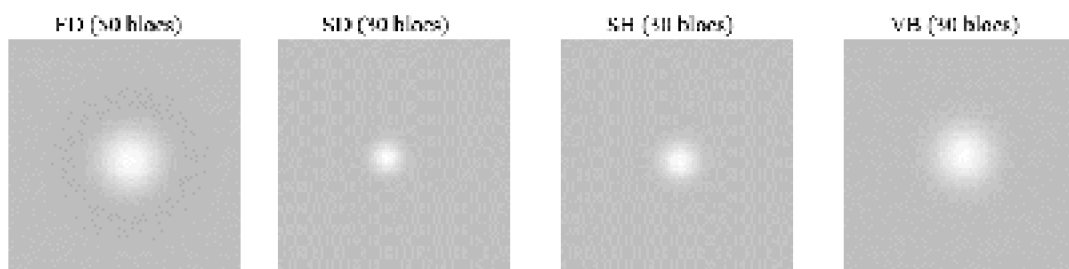


Figure 54 : Expérience 1 : Ajustement de la fonction « DoG » aux résultats obtenus dans l'image de classification pour chaque observateur.

- Image de classification liée à la présence ou l'absence du signal.

Les images de classification indiquaient la présence d'une zone centrale claire de forme plus circulaire que carré, ce qui laisse penser que l'information basée sur les bords du carré n'était pas présente, et ceci est vrai pour les quatre observateurs. Mais, il est tout à fait possible que ce flou soit le résultat d'une incertitude spatiale sur la localisation du signal, incertitude qui, si elle existe, serait plus importante lorsque le signal est absent. En

effet, lorsque le signal est présent, même faiblement, le bruit permet de renforcer les indices utilisés par l'observateur quand ils sont alignés avec le signal, alors que, quand le signal est absent, un tel alignement n'est pas possible et l'observateur estimerait la position du signal de façon plus aléatoire. Ainsi, la moyenne des bruits pour lesquels le signal était absent devrait comporter moins d'information sur les contours que celle correspondant aux bruits présentés avec un signal.

Nous avons effectué des analyses complémentaires en distinguant les images de classifications pour lesquels le signal était présent et absent, de la façon suivante :

$$\text{Image de classification}_{\text{Signal Présent}} = \overline{DC} - \overline{OM}$$

$$\text{Image de classification}_{\text{Signal Absent}} = \overline{FA} - \overline{RC}$$

avec la moyenne des bruits correspondant aux détections correctes (DC), omissions (OM), fausses alarmes (FA) et rejets corrects (RC).

Les résultats sont présentés à la Figure 55. L'inspection de la figure montre que le profil des images diffère très peu que le signal soit présent ou absent.

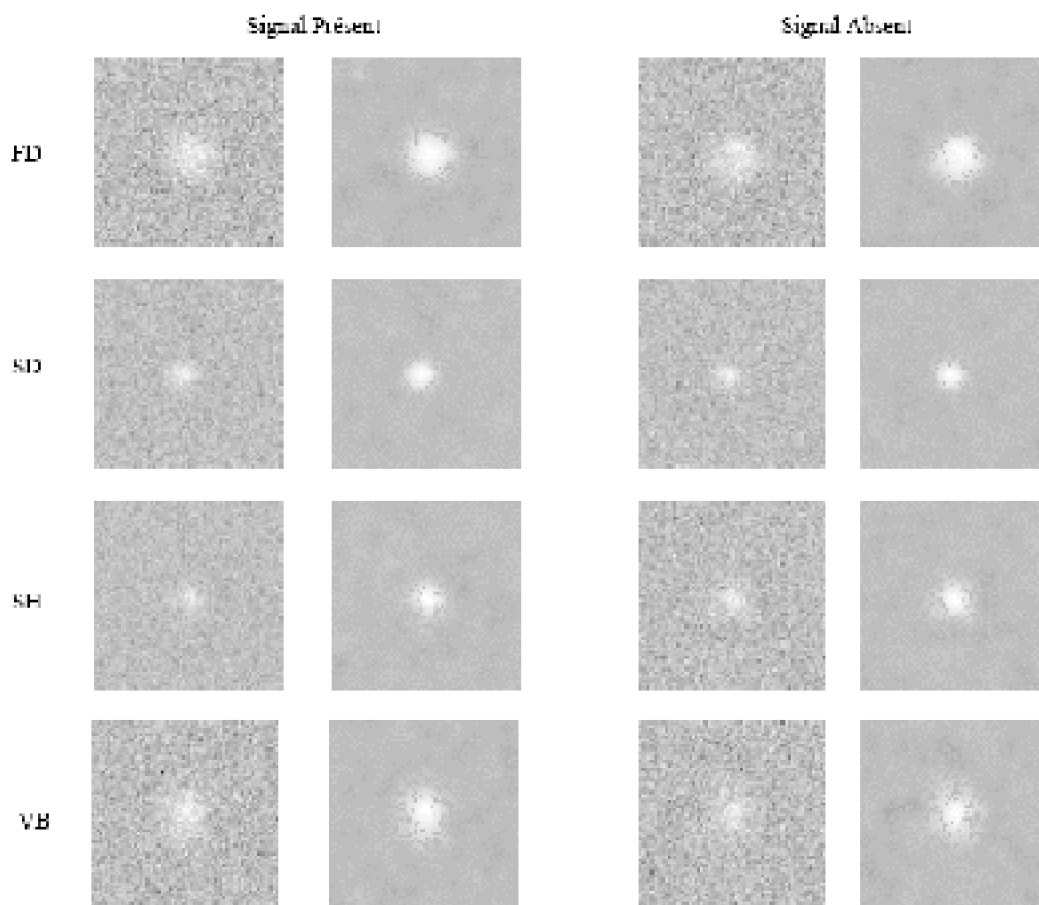


Figure 55 : Expérience 1 : Résultats des images de classification lorsque le signal est présent ou absent, des images lissées et du test statistique appliquées aux images de classification pour l'ensemble des observateurs (FD, SD, SH et VB).

Pour quantifier l'incertitude spatiale, un ajustement de la fonction DoG a ensuite été réalisé pour chacune de ces images. Les résultats sont montrés dans le Tableau IV. Les ajustements pour les deux conditions confondues (signal présent et absent) indiquaient que la taille de la zone centrale était comprise entre 2.16 et 4.47 deg, et le pourtour entre 7.54 et 17.08 deg. Le pourtour participait faiblement à la détection puisqu'il représentait entre 1.5 et 9% de l'amplitude de la zone centrale. De même que pour l'ajustement de l'image de classification globale, l'estimation de la position de la zone centrale présentait un faible biais (inférieur à 15 min).

Si l'incertitude spatiale est plus importante lorsque le signal est absent, alors nous nous attendons à ce que la zone centrale (représentée par σ_c) soit plus large pour les images de type signal-absent par rapport aux images de type signal-présent. En effet, les observateurs donneraient leur réponse de détection sur une zone plus importante en

l'absence d'indice perceptif lié au signal. De plus, l'incertitude spatiale est également déterminée par l'emplacement de la zone centrale, mais les coordonnées par rapport au centre de l'image, représentées aux colonnes 7 et 8 du Tableau IV, variaient très peu. Les paramètres de l'ajustement variaient approximativement de 30% entre les images où le signal est présent ou absent (colonne 5 du Tableau IV), toutefois, cette variation se faisait dans un sens ou dans l'autre suivant l'observateur. Si la réponse aux contours est plus importante dans les images de classification pour les essais où le signal est présent, alors le ratio de l'amplitude A_p / A_c devrait être plus élevé dans cette condition par rapport à la condition où le signal est absent. Toutefois, les résultats n'allaient pas dans le sens de cette hypothèse (colonne 4 du Tableau IV). Les réponses basées sur le pourtour de l'observateur VB étaient très faibles, et nous n'avons pas pu estimer les paramètres pour la condition où le signal est absent.

Par ailleurs, pour chaque observateur, les paramètres étaient proches de ceux observés lorsque l'image de classification était calculée de façon globale.

Tableau IV : Estimation des paramètres de l'ajustement DoG à l'image de classification pour les essais où le signal est présent et absent. Les colonnes 1 et 2 présentent les valeurs d'amplitude, les colonnes 4 et 5 indiquent la taille du sigma en degrés, les ratios centre/pourtour sont indiqués aux colonnes 3 et 6. La position du centre est présentée aux colonnes 7 et 8.

Observateur		A_c	A_p	A_p / A_c	σ_c (deg)	σ_p (deg)	σ_c / σ_p	x_0 (min)	y_0 (min)
FD	Présent	1.98	0.29	0.015	4.47	8.98	2.01	13.8	9.6
	Absent	1.76	0.08	0.05	4.07	11.87	2.92	13.8	22.2
SD	Présent	2.52	0.09	0.04	2.39	8.52	3.56	-15.1	5.6
	Absent	2.59	0.07	0.03	2.16	10.71	4.96	-11.9	7.3
SH	Présent	2.03	0.04	0.02	2.61	12.86	4.93	7.1	13.0
	Absent	1.81	0.16	0.09	3.24	7.54	2.33	2.2	18.1
VB	Présent	1.47	0.01	0.01	4.01	17.08	4.26	3.9	-4.4
	Absent	1.37	-	-		-	-	5.4	0.0

1.3.4. Discussion

L'objectif de la première expérience était de déterminer les éléments utilisés par l'observateur au cours d'une tâche de détection grâce à la technique de l'image de classification perceptive. Cette expérience a mis en évidence deux résultats. Le premier résultat montrait une zone centrale claire pour les quatre observateurs. Bien que le signal soit un carré, les images de classification présentaient une zone centrale arrondie. Ce résultat est conforme à notre hypothèse pour laquelle l'information de surface serait la seule utilisée par l'observateur. Il est donc tout à fait possible d'interpréter nos données de la façon suivante : la surface est une information précoce et très importante pour le traitement visuel.

La partie théorique suggère que cette dernière interprétation est contradictoire avec l'hypothèse dominante de la littérature et conduit à réviser la conception classique du traitement de la surface par un processus tardif. En effet, la notion de champ récepteur

classique développée dans le modèle de Hubel et Wiesel (1962, 1968) laissait à penser que les bords d'un stimulus stimulaient fortement les neurones du système visuel. La surface serait ensuite reconstruite à partir de l'information de bord. Mais, la présence de la surface dans l'image de classification ne peut pas être une reconstruction puisqu'elle implique que, à chaque essai, les cellules qui codaient les bords ou la luminance de la surface dans la zone centrale véhiculaient une information qui correspondait à ce que le système visuel attendait du signal.

Par ailleurs, nos résultats ne démontrent pas l'implication du processus de remplissage pour le traitement de la surface. En effet, le processus de remplissage est déterminé, par définition, à partir de l'information des bords, et notre expérience indique que les observateurs utilisent principalement le centre du stimulus pour répondre, sans que l'information de contour n'apparaissent dans les images. Nos résultats ne remettent pas en cause l'existence du processus de remplissage qui a été démontré dans de nombreux phénomènes comme la tache aveugle, les scotomes ou les images stabilisées (voir partie théorique). Toutefois, notre expérience indique que le processus de remplissage n'est pas une condition nécessaire pour traiter l'information de surface.

Ce premier résultat était identique pour les quatre observateurs, mais l'ajustement de la fonction DoG a permis de montrer que la taille de la zone centrale était différente entre les observateurs. Ces différences indiqueraient que les observateurs utilisent différents prototypes pour la tâche de détection ou qu'ils n'arrivent pas à localiser correctement la position du signal. Cette dernière explication est probablement due aux mouvements oculaires ou à des erreurs de fixation.

Le deuxième résultat est la zone sombre entourant cette zone claire. Ceci indique que la réponse « signal présent » est privilégiée lorsque le pourtour immédiat de la zone où le signal était attendu était sombre. L'ajustement de la fonction DoG a permis de montrer que le pourtour était 2-4 fois plus large que la zone centrale et contribuait faiblement à la détection (3-8 % de l'amplitude de la zone centrale). Ainsi, la région qui entoure le centre contribuait, de façon complémentaire, à la détection du carré. Il apparaît donc que le système visuel traite une information relative, plus qu'une information absolue. Tout se passe comme si la décision, concernant la détection, reposait sur la différence de luminance entre la région du centre et la région qui l'entoure. Cette faible influence du pourtour a déjà été démontrée dans des expériences utilisant l'image de classification pour la détection d'une fonction Gaussienne (Abbey & Eckstein, 2002 ; Solomon, 2002). La faible participation d'un pourtour sombre pourrait révéler une tendance à intégrer l'information de contour du stimulus dans la détection. En effet, la présence du pourtour sombre reflète un contraste entre deux régions qui est par définition un contour.

De plus, la taille importante du pourtour indiquerait que l'étape de décision où l'observateur compare le prototype et le signal ne serait pas d'ordre pré-cortical, puisque cela implique une intégration sur plusieurs degrés d'angle visuel. Le traitement cortical nécessite une intégration de l'information de surface dans une étape préalable. De nombreuses études (MacEvoy, Kim & Paradiso, 1998 ; Rossi & Paradiso, 1999 ; Rossi, Rittenhouse & Paradiso, 1996) indiquent que les neurones de V1 sont sensibles à la luminosité d'une surface, mais l'origine de cette réponse est incertaine. En effet, les neurones sont-ils activés à partir des bords de la stimulation ou à partir des

caractéristiques de la surface. Cette question n'a pas encore trouvé de réponse satisfaisante à l'heure actuelle. Toutefois, si les neurones sont activés par les contours du signal, un mécanisme de remplissage permettrait de traiter l'information de surface par l'intermédiaire des connexions horizontales long-range et indépendamment de la surface réelle. Ce mécanisme ne permet pas d'expliquer notre résultat principal puisque l'information sur laquelle l'observateur base sa réponse est essentiellement située au centre du carré. Ce résultat pourrait être expliquée par le fait que les neurones corticaux répondent à une surface uniforme, mais pas à la dimension de contours qui induit un processus de remplissage.

Les résultats qui ont été présentés ne remettent pas en cause le fait que l'information de contour soit essentielle dans d'autres contextes, ou encore que le mécanisme de remplissage est réalisé par le système visuel. En effet, le processus de remplissage a été démontré de façon comportementale (Paradiso & Hahn, 1996 ; Paradiso & Nakayama, 1991 ; Ramachandran & Gregory, 1991) et physiologique (DeWeerd, Gattass, Desimone & Ungerleider, 1995). De même, les contours sont importants dans certains contextes comme les contours illusoires (Gold, Murray, Bennett & Sekuler, 2000) ou encore pour une tâche de discrimination (Neri & Heeger, 2002). Notre résultat montre par contre que l'information de surface dans le stimulus est aussi transmise à travers le système visuel, puisque cette dimension est utilisée par l'observateur pour donner sa réponse de détection. Ainsi, un mécanisme « top-down » comme le remplissage n'est pas nécessaire pour expliquer comment la surface est perçue à l'intérieur des contours (processus « bottom-up »), puisque la surface est codée par les neurones (MacEvoy, Kim & Paradiso, 1998 Rossi & Paradiso, 1999 ; Rossi, Rittenhouse & Paradiso, 1996).

Par ailleurs, les images, pour lesquelles les pixels significatifs sont exagérés, indiquaient la présence de zones significatives dans l'extrême pourtour de l'image. Toutefois, ce traitement est assez grossier puisque ces pixels sont d'amplitude trop faibles pour être visibles dans l'image statistique. Néanmoins, il est tout à fait possible que les observateurs prennent en considération le champ visuel dans son entier pour donner leur réponse de détection.

1.4. Diminution du temps de présentation

La deuxième série d'expérience cherchait à déterminer si l'information de surface restait une information pertinente pour une durée de présentation brève dans une tâche de détection. Si l'information de surface apparaît à nouveau, la prépondérance de cette dimension pour le traitement visuel sera démontrée. De plus, la manipulation de ce facteur nous permettra de déterminer si la répartition spatiale de l'information de surface est constante ou si elle varie en fonction du temps de présentation. Cette hypothèse sera mise à l'épreuve grâce aux résultats de l'ajustement DoG concernant la taille et l'amplitude de la zone centrale.

1.5. Expérience de détermination de seuil pour une durée de 40 ms

1.5.1. Introduction

De façon identique à la première série expérimentale, une expérience de détermination de seuil a été réalisée afin de déterminer les conditions expérimentales à utiliser dans l'expérience de détermination de l'image de classification, avec une nouvelle durée de présentation du stimulus. Toutefois, ici, un seul niveau de bruit était étudié, celui ayant pour contraste 0.7. Puisque la diminution de la durée d'un stimulus entraîne une diminution de son énergie, nous nous attendions à ce que les seuils de détection observés pour une durée de 40 ms soient plus élevés (c'est-à-dire à ce que la sensibilité soit plus faible) que ceux observés pour les mêmes observateurs où la durée de présentation était de 400 ms.

1.5.2. Méthode expérimentale

- Participants.

Deux observateurs (FD et SH) ont participé à l'expérience, le dernier n'était pas au courant des hypothèses.

- Stimuli.

La configuration spatiale des stimuli était similaire à celle de la première série expérimentale : un carré plein qui servait de signal et dont le contraste variait en 7 niveaux (un niveau 0 pour la condition « signal absent » et 6 niveaux de contraste qui se répartissaient sur une échelle de $10^{0.1}$ à partir de la valeur 0.05 pour FD et 0.09 pour SH) était superposé à un bruit dont le contraste était fixe (0.7). Des essais préliminaires étaient nécessaires pour déterminer à tâtons les valeurs de contraste du signal à utiliser pour chaque observateur. Nous ne présenterons ici que les blocs d'essais correspondant aux niveaux de contraste du signal qui ont été finalement choisis pour chaque observateur.

- Procédure.

La seule différence de procédure par rapport à celle de l'expérience 1 était la durée de présentation du signal qui était diminuée : 40 ms contre 400 ms auparavant.

1.5.3. Résultats

La performance, exprimée en terme de sensibilité par l'indice d' , est présentée à la Figure 56. De la même façon que pour l'expérience princeps, une droite de régression a été calculée afin de réaliser une interpolation, à partir des résultats, pour obtenir la valeur de contraste du signal correspondant à un $d' = 1$. Les seuils de contraste du signal étaient plus élevés (respectivement 0.13 et 0.14 pour FD et SH) que ceux observés, pour les mêmes observateurs, dans la première expérience, pour laquelle la durée de présentation était de 400 ms (0.07 pour FD et SH). Conformément à notre hypothèse, la sensibilité diminuait quand la durée de présentation du stimulus diminuait.

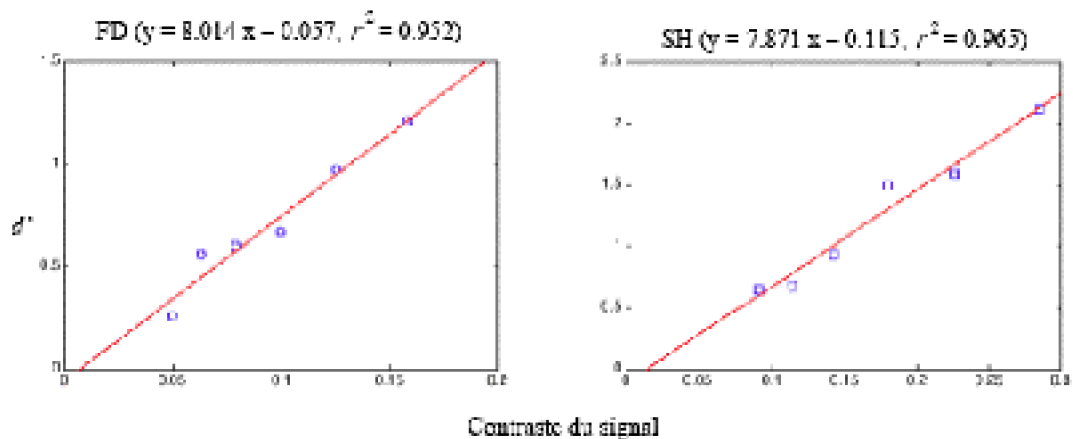


Figure 56 : Expérience préliminaire 2 : Droite de régression entre l'indice d' et le contraste du signal des observateurs FD et SH pour un contraste du bruit de 0.7 et un temps de présentation de 40 ms.

Une analyse des courbes ROC a également été réalisée. Les résultats sont indiqués dans l'Annexe 2. Pour chaque niveau de contraste du signal, il y a très peu de différence entre les deux observateurs. Ces courbes indiquent que le signal se distinguait de plus en plus facilement du bruit avec l'augmentation du contraste. De plus, les courbes ROC étaient beaucoup moins creusées que celles de l'expérience précédente. Les performances étaient de façon générale plus faibles que dans la première expérience de seuil.

Il est possible que FD et SH aient eu plus de difficultés pour distinguer le signal. Pourtant, les niveaux de contraste choisis pour le signal dans cette seconde expérience avaient été augmentés pour compenser la diminution de durée et pour que la performance soit équivalente dans les deux expériences.

1.6. Détermination de l'image de classification pour une durée de 40 ms

1.6.1. Introduction

Cette expérience a pour but de déterminer l'image de classification pour une durée brève de présentation du stimulus.

Notre première expérience a permis de montrer un résultat inattendu par rapport au modèle classique concernant le traitement contour/surface. En effet, nos résultats montraient une prédominance de l'information de surface. Il est bien connu que le facteur temporel influence le traitement de l'information visuelle. Si cette dimension continue d'être utilisée dans le prototype pour cette durée brève, alors, cela témoignera de sa prépondérance dans les traitements précoces.

Deux hypothèses sont possibles : 1) les observateurs utilisent un prototype qui est toujours le même quelles que soit les conditions expérimentales, ou 2) le poids de la surface diminue avec le temps de présentation. Pour la première hypothèse, nous

devrions obtenir des images dont les paramètres de l'ajustement DoG (taille et amplitude) sont similaires à ceux de la première expérience pour un même observateur. Par contre, dans la seconde hypothèse, nous supposons que la taille et l'amplitude de la zone centrale diminueraient par rapport aux données obtenues dans la première expérience, de même l'amplitude du pourtour augmenterait.

1.6.2. Méthode expérimentale

Les stimuli étaient les mêmes que dans l'expérience précédente : soit le bruit seul présenté avec un contraste de 0.7, soit le signal présenté à son niveau de contraste seuil (0.13 pour FD et 0.14 Pour SH) superposé à ce bruit. La luminance du fond ne variait pas (8 cd/m²). La seule différence de procédure par rapport l'expérience précédente était la durée de présentation du stimulus : 40 ms au lieu de 400 ms. La tâche du sujet était identique : répondre si le signal était présent ou absent. De plus, il y avait 100 répétitions par condition pour chaque bloc expérimental (soit 200 essais par blocs). FD a passé 30 blocs et SH 25 blocs.

1.6.3. Résultats

- L'indice d' et le critère de décision.

La Figure 57 montre l'indice d' calculé tout les 200 essais ne s'améliorait pas pour chaque observateur : on constate que les performances étaient stables autour de 1, ce qui indique que l'apprentissage était au plateau. La sensibilité globale (calculée sur l'ensemble des essais) était de 1.41 pour FD et SH. De plus, la place du critère de décision est stable tout au long des essais et le calcul global de celui-ci est proche de 0 (FD = -0.17 et SH = -0.18).

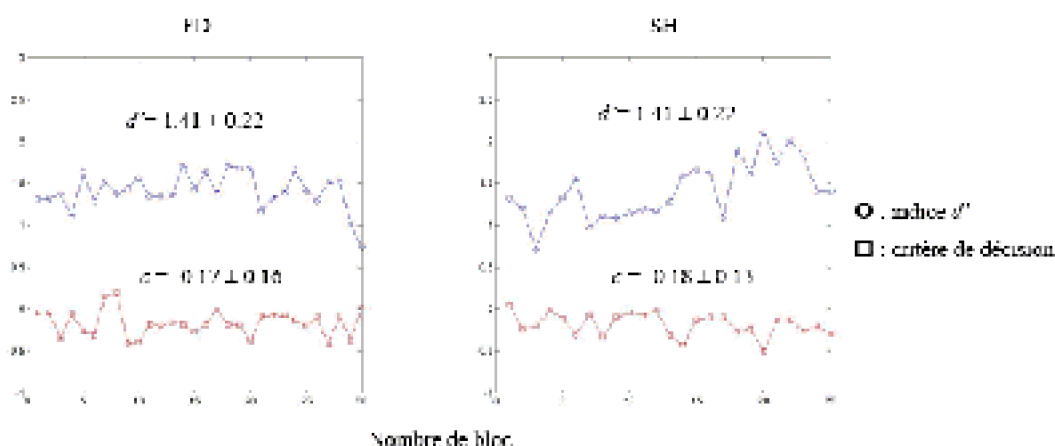


Figure 57 : Expérience 2 : Résultats de l'indice d' et du critère de décision au cours des blocs expérimentaux obtenue par FD et SH pour une durée de présentation de 40 ms.

- Image de classification.

Les images de classifications sont montrées dans la Figure 58. D'une manière globale, on

observe que les images étaient relativement proches de celles observées précédemment pour une durée plus longue : une zone claire émergeait au centre, à l'emplacement du signal, et un pourtour sombre l'entourait.

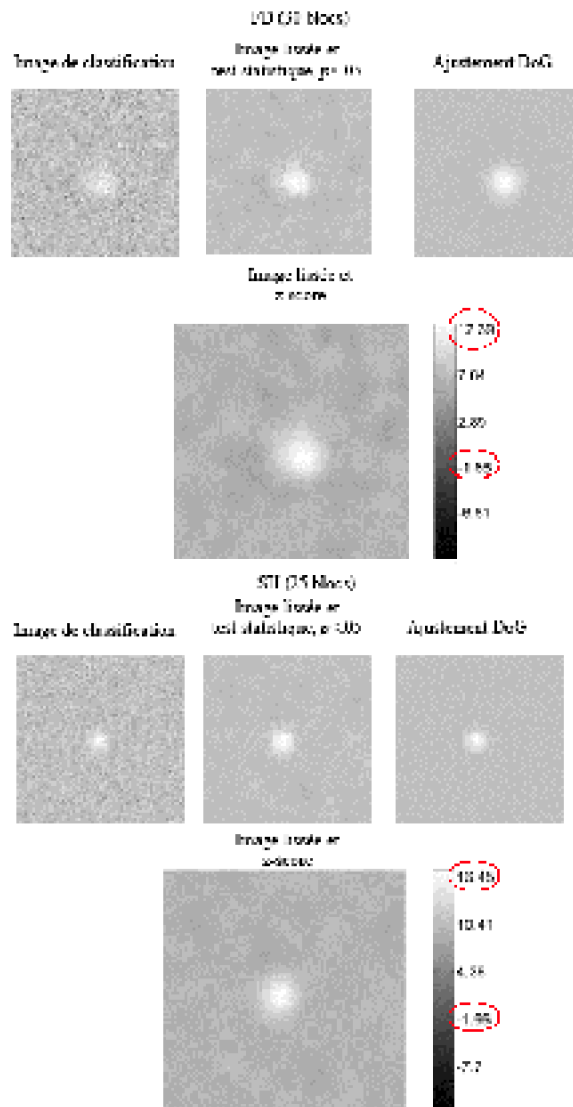


Figure 58 : Expérience 2 : Résultats de l'image de classification, du test statistique et du z-score effectué sur l'image de classification mais aussi de l'ajustement DoG des observateurs FD et SH pour une durée de présentation de 40 ms. La bande qui est présentée à la droite de l'image permet d'estimer la valeur de chaque pixel. Par ailleurs, les valeurs minimales et maximales pour chaque matrice sont entourées.

Les ajustements DoG qui sont présentés dans ces figures permettent d'estimer la taille et l'amplitude de la zone centrale et du pourtour. Les résultats sont présentés dans le Tableau V. On observe que la taille était modifiée par la diminution de la durée. En effet, la zone centrale se centrait sur un sigma de 3.23 et 1.99 deg, et le pourtour sur un sigma de 6.08 et 2.01 deg, respectivement pour FD et SH. Le pourtour représentait ainsi environ 1 à 2 fois la surface centrale. Par contre, les valeurs d'amplitude n'allaient pas dans le même sens pour les deux observateurs. En effet, l'amplitude du pourtour

représentait 10 % de l'amplitude du pourtour pour FD (contre 8% dans l'expérience précédente), et pour SH, l'amplitude du pourtour représente 1% de l'amplitude de la zone centrale (contre 3% dans l'expérience précédente). Nous n'avons pas mesuré s'il existait des différences significatives entre les ajustements DoG de toutes les expériences. Néanmoins, nous pouvons observer que FD avait besoin de plus de contraste entre la zone centrale et le pourtour immédiat pour donner une réponse de type présent par rapport à la première expérience, mais l'observateur SH manifestait la tendance inverse. L'ajustement de la fonction a également permis d'estimer la position du signal. Comme dans l'expérience princeps, les résultats montraient toujours un léger biais de quelques dizaines de minutes (compris entre -11 et 35 min).

Tableau V : Valeur de l'amplitude (colonne 2 et 4) et du sigma (colonne 1 et 3) pour la zone du pourtour et la zone centrale, rapport existant entre ces deux zones (colonne 5 et 6), et position du signal (colonne 7 et 8) des observateurs FD et SH pour les deux durées de présentation.

Tableau V : Valeur de l'amplitude (colonne 2 et 4) et du sigma (colonne 1 et 3) pour la zone du pourtour et la zone centrale, rapport existant entre ces deux zones (colonne 5 et 6), et position du signal (colonne 7 et 8) des observateurs FD et SH pour les deux durées de présentation.							x_0 (min)	y_0 (min)
	Zone centrale		Pourtour		Rapport pourtour / zone centrale			
Sujet	Sigma (deg)	Amplitude	Sigma (deg)	Amplitude	Sigma	Amplitude		
FD (400 ms)	4.31	3.52	10.55	0.28	2.45	0.008	12.38	12.94
FD (40 ms)	3.23	2.83	6.08	0.29	1.88	0.1	25.69	35.34
SH (400 ms)	2.85	3.72	12.18	0.1	4.26	0.03	4.88	15.47
SH (40 ms)	1.99	4.51	2.01	0.06	1.01	0.01	-11.16	8.34

Nous avons calculé de la même façon que pour l'expérience princeps l'incertitude spatiale. En d'autres termes, la présence du signal renforcerait les indices perceptifs portant sur les contours, nous nous attendons à ce que la taille du sigma soit plus importante lorsque le signal est absent. Les images de classification, spécifique à la présence du signal ou non, sont présentées à la Figure 59. Les images de classifications étaient les mêmes pour les deux observateurs et pour les deux conditions (signal présent et absent). L'ajustement DoG a ensuite été déterminé, et les résultats sont présentés dans le Tableau VI. Les paramètres estimés de la taille de la valeur centrale étaient légèrement supérieurs lorsque le signal était présent pour les deux observateurs. L'hypothèse d'une incertitude spatiale ne permet pas d'expliquer les résultats obtenus pour l'image de classification globale. De plus, si la présence du signal renforce la présence des contours, alors le ratio A_c / A_p serait plus important pour les essais où le signal est présent. Néanmoins, les résultats indiquaient que le pourtour participait faiblement à la détection, de plus ces valeurs n'étaient pas stables pour les deux observateurs. La dimension de contour n'est donc pas renforcée pour les essais de type

signal présent.

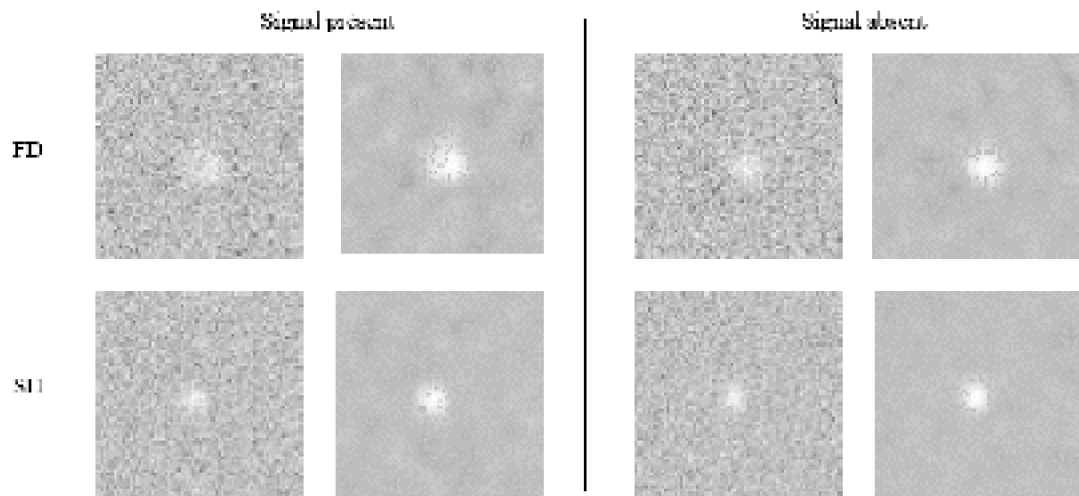


Figure 59 : Expérience 2 : Images de classification calculées lorsque le signal est présent (à gauche) et absent (à droite) pour les deux observateurs FD et SH. Les images de classification brutes mais également les images lissées et le test statistique sont présentées pour chaque condition.

Tableau VI : Valeurs de l'amplitude (colonne 1 et 2), du sigma (colonne 3 et 4) pour la zone du pourtour et la zone centrale, rapport entre ces deux zones (colonne 5 et 6), et position du signal (colonne 7 et 8) des observateurs FD et SH pour une durée de présentation de 40 msec en distinguant les réponses où le signal est présent et absent.

Observateur		A_c	A_p	A_p / A_c	σ_c (deg)	σ_p (deg)	σ_c / σ_p	x_0 (min)	y_0 (min)
FD	Présent	1.29	0.05	0.04	3.21	6.59	2.05	22.27	35.19
	Absent	1.40	0.1	0.07	2.9	6.33	2.18	28.66	36.2
SH	Présent	2.18	0.06	0.03	2.21	8.39	3.79	-15.21	12.43
	Absent	2.37	0.001	0.002	1.8	1.9	1.05	-7.99	4.83

1.6.4. Discussion

Cette deuxième expérience témoigne de la robustesse des effets précédemment démontrés, qui s'observent indépendamment de la durée de présentation du stimulus. La zone centrale et le faible pourtour continue d'être observés quand la durée de présentation diminue. Ces deux informations sont donc disponibles très rapidement. Il apparaît ainsi que l'information de surface est une dimension importante des traitements les plus précoces, nécessaires à la détection d'une forme.

Ceci nous amène à croire que même si le résultat principal reste inchangé par rapport à la première expérience, il est tout à fait possible que les observateurs utilisent une stratégie différente pour répondre. S'il existe une différence entre les deux expériences, alors elle doit principalement se manifester par des variations dans les paramètres des fonctions DoG pour chaque observateur.

Nous avons présenté au dessus qu'il n'y avait pas de différence d'amplitude entre l'expérience princeps et l'expérience deux. Toutefois, il y avait des différences de taille. Les tailles de la zone centrale et du pourtour étaient moins importantes que celles de la première expérience. Ceci indique que la zone du champ visuel utilisée pour répondre diminuait. Ce changement dans la taille de la zone centrale et du pourtour pourrait indiquer l'application d'une stratégie différente de la part des observateurs qui ont tendance à focaliser leur attention au centre de l'image pour donner leur réponse de détection. En d'autres termes, diminuer les temps de présentation reviendrait à restreindre le nombre d'indices que les observateurs utilisent. De plus, il est possible que les observateurs aient une stratégie dans laquelle le contraste entre la zone centrale et le pourtour serait plus élevé par rapport à la première expérience. Ce qui permettrait de faciliter leur réponse de détection. Dans ce cas, le ratio A_C/A_D devait être plus important dans la seconde expérience. Mais, les résultats de ce rapport ne vont pas dans le même sens pour les deux observateurs. Ceci montre que les observateurs n'ont pas ajusté exactement le prototype utilisé pendant la première expérience à la seconde expérience, en effet, dans ce cas le rapport concernant l'amplitude aurait dû augmenter.

Notre seconde expérience a ainsi permis de montrer, en plus de l'importance accordée à la surface, que les observateurs n'utilisent pas de prototype fixe pour donner leur réponse de détection. Nous n'avons pas eu la possibilité de diminuer le temps de présentation du stimulus en dessous d'une durée de 40 msec car l'utilisation du logiciel Matvis ne le permet pas. Mais, il est probable que diminuer le temps de présentation ne changerait pas notre principal résultat.

1.7. Conclusion

Notre principal résultat indique que l'observateur humain donne sa réponse de détection sur la luminance des pixels de la région centrale, et non pas sur les contours du signal. L'importance de la surface pour la détection est surprenant puisque d'après les théories dominantes cette information est reconstruite à partir des contours, or cette dernière dimension n'apparaît pas clairement dans nos images. Toutefois, les résultats obtenus ne signifient pas que l'information de contour ne joue pas un rôle déterminant dans la perception. En effet, l'importance des bords d'une forme a été démontré dès les années 60 par les travaux de Hubel et Wiesel (1962, 1968). Ces études indiquaient que les champs récepteurs de la rétine et de V1 sont sensibles aux contours. Notre résultat démontre toutefois que l'information de surface est transmise à travers le système visuel puisque cette information est utilisée par l'observateur durant la détection. Par ailleurs, le mécanisme de remplissage ne semble pas pertinent pour expliquer l'utilisation de la surface. En effet, ce processus est généralement déterminé à partir de l'information de contour, or, cette dimension ne se trouve pas dans les images de classification. De plus, à l'inverse du mécanisme de propagation qui intervient durant le remplissage, la réponse des observateurs est principalement basée au centre du signal. Néanmoins, cette conclusion ne rejette pas l'existence du processus de remplissage qui a été démontré non seulement d'un point de vue physiologique (DeWeerd, Gattass, Desimone & Ungerleider, 1995) mais aussi comportementale (Paradiso & Nakayama, 1991 ; Ramachandran &

Gregory, 1991). Notre principal résultat suggère que l'information de surface peut être utilisée sans qu'un processus de remplissage ne soit réalisé. Ceci est conforme aux études physiologiques récentes qui montre une sensibilité des neurones du cortex strié dont le champ récepteur est situé sur une surface uniforme. Toutefois, l'origine de cette sensibilité est encore incertaine. De plus, quel que soit le temps de présentation des stimuli, les images de classification montrent une prépondérance de l'information de surface dans la réponse des observateurs pour la détection. Néanmoins, l'ajustement de la fonction DoG suggère la possibilité que la taille de la zone centrale diminuait avec la diminution de la durée de présentation du stimulus. En effet, nous avons pu observer que les observateurs avaient tendance à restreindre la zone du champ visuel utilisée pour répondre.

Notre second résultat montrait qu'un faible pourtour sombre entourait la zone centrale claire, ce qui indiquerait que les observateurs traitent une information relative et non pas absolue. La présence de ce pourtour négatif est également présent lorsque la présentation du signal diminue à 40 msec. Par ailleurs, le rapport de l'amplitude centre/pourtour, de l'ajustement DoG, variait très peu entre les deux expériences pour un même observateur. Ce résultat tendrait à indiquer que la contraste utilisé pour détecter un signal est identique quelle que soit la condition de présentation.

2. Seconde série d'expériences : variation de la taille du signal et du bruit

L'objectif de cette deuxième série d'expérience était vérifier si les résultats obtenus dans les expériences précédentes étaient généralisables quelle que soit la taille du signal, mais aussi du bruit, utilisé. En d'autres termes, cette deuxième série d'expériences cherchait à déterminer s'il existe une limite spatiale aux mécanismes de codage de la surface.

La perception visuelle est sous tendue par de nombreux canaux qui sont sensibles aux facteurs spatiaux et temporels, dans ce cas les résultats seraient influencés par la variation de ces facteurs. Dans notre première expérience, nos résultats étaient très peu influencés par la durée de présentation. Nous nous sommes ensuite intéressés aux facteurs spatiaux qui sont susceptibles d'influencer la réponse de détection des observateurs.

2.1. Expérience de détermination de seuil pour deux nouvelles tailles du signal

2.1.1. Introduction

Une expérience de détermination de seuil a été effectuée afin de déterminer les conditions de contraste du signal à utiliser dans l'expérience de détermination de l'image

de classification pour deux nouvelles tailles du signal (1 et 6 degrés d'angle visuel). La diminution de la taille du signal entraîne une diminution de l'énergie, puisque l'énergie d'un stimulus est proportionnelle à sa luminance, sa durée et sa surface. Par conséquent, le seuil de contraste du signal devrait donc être plus élevé que celui observé dans l'expérience princeps. De même, l'augmentation de la taille du signal entraîne une augmentation son énergie. Nous faisons l'hypothèse que le seuil de contraste du signal devrait diminuer par rapport à ce qui été observé (pour les mêmes observateurs) dans l'expérience où le signal était de 3 deg.

2.1.2. Méthode expérimentale

- Participants.

Deux observateurs (FD et VB) ont passé l'expérience pour la condition dans laquelle la taille du signal était de 1deg. L'un des deux observateurs n'était pas au courant des hypothèses (VB). Par ailleurs, les observateurs FD et SD ont réalisé l'expérience pour une taille du signal de 6deg.

- Stimuli.

Les stimuli étaient identiques à ceux utilisés dans la première série d'expérience. La seule différence par rapport aux expériences précédentes était la taille du carré qui était plus petite ou plus grande : 1 et 6 degrés d'angle visuel contre 3 degrés auparavant. Les niveaux de contraste du signal se répartissaient sur une échelle logarithmique ($10^{0.1}$) en 7 niveaux. Pour une taille de 1 deg, les valeurs étaient supérieures à celles de la première expérience : de 0.09 à 0.2846 pour FD et 0.07 à 0.2214 pour VB. De plus, lorsque la taille du signal était de 6 deg, le contraste du signal variait de 0.045 à 0.1423 pour FD, et de 0.05 à 0.1581 pour SD. Un seul niveau de contraste du bruit était utilisé (0.7). Les valeurs de contraste du signal étaient déterminés par des essais préliminaires qui ne sont pas présentés.

- Procédure.

Une présentation du signal à détecter est tout d'abord réalisée, sans présentation du bruit, pour permettre aux observateurs de connaître le signal à détecter. Le signal était montré une seul fois, avant le commencement des premiers essais, et la présentation n'était pas répétée. La procédure était en tout point identique aux précédentes expériences de détermination de seuil.

2.1.3. Résultats

La Figure 60 présente les résultats de FD et VB pour une taille du signal de 1 deg. Le contraste du signal interpolé pour un d' de 1 de FD était ici de 0.17, celui de VB était de 0.12. Le seuil est donc plus élevé que le signal plus large utilisé dans la première d'expérience, pour laquelle le seuil de contraste du signal était de 0.06 pour les deux observateurs. Conformément à nos hypothèses, la sensibilité diminuait lorsque la taille du signal diminuait.

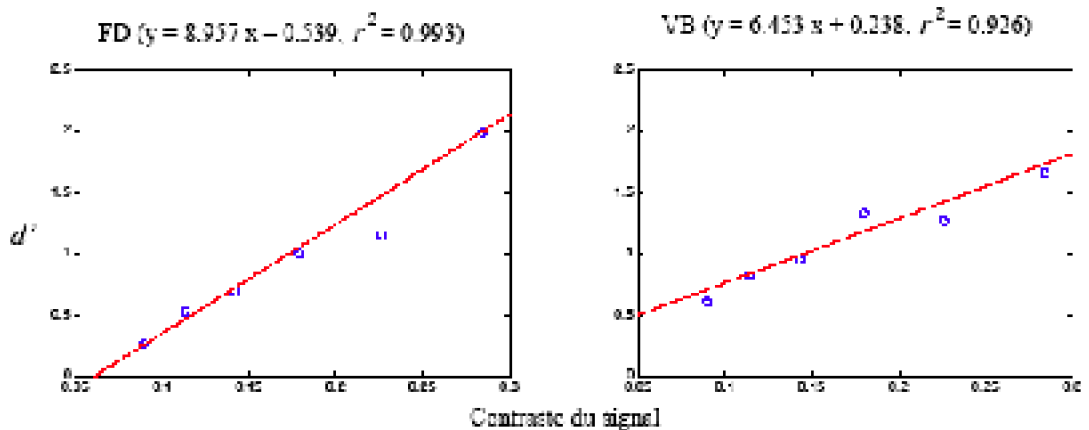


Figure 60 : Expérience préliminaire 3 : Droite de régression entre l'indice d' et le contraste du signal des observateurs FD et VB pour une taille du signal de 1° .

Les courbes ROC ont été affichées pour les six niveaux de contraste du signal spécifique à chaque observateur. Les résultats sont présentés dans l'Annexe 3. Les résultats indiquent que les courbes étaient de plus en plus creusées pour les derniers niveaux de contraste. Ce résultat indiquait que le signal était de plus en plus visible par rapport au bruit. Si l'on compare les courbes avec celles obtenues dans l'expérience où la taille du carré était de 3 deg, alors nous pouvons remarquer que le signal se distinguait moins facilement du bruit lorsque la taille était de 1 deg.

Les résultats de FD et SD sont présentés dans la Figure 61 pour une taille du signal de 6 deg. Le seuil de contraste du signal de FD était de 0.06, celui de SD de 0.08. Les seuils étaient proches de ceux utilisés dans l'expérience princeps, excepté celui de SD qui augmentait très légèrement (0.07 dans la première expérience). Ainsi, contrairement à notre hypothèse, l'augmentation de la taille du signal n'améliore pas la sensibilité des observateurs. Il est donc possible que les mécanismes mise en jeu soit identique pour une taille au-delà de 3 deg.

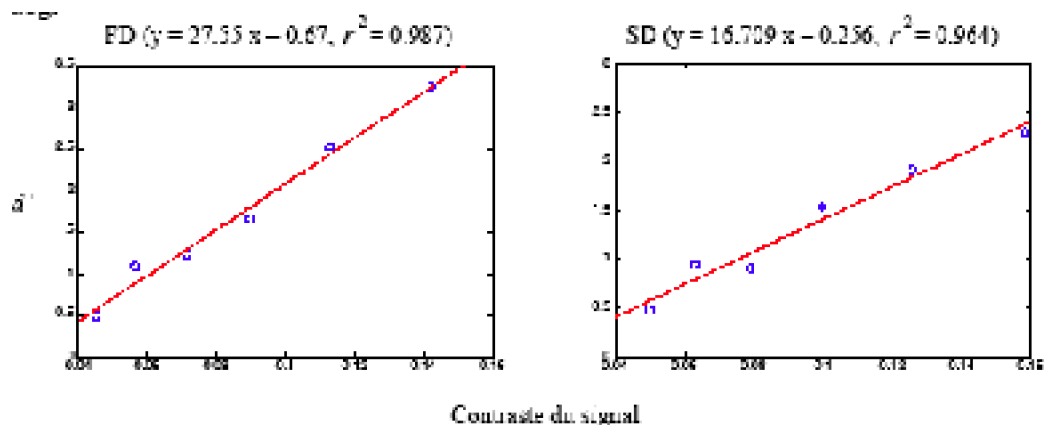


Figure 61 : Expérience préliminaire 3 : Droite de régression entre l'indice d' et le contraste du signal des observateurs FD et SD pour une taille du signal de 6° .

Les courbes ROC sont affichées pour FD et SD dans l'Annexe 4. Les résultats sont identiques pour les deux observateurs, et montrent que signal se distingue plus facilement du bruit. Aucune variation n'est constatée par rapport aux précédentes expériences.

2.2. Détermination de l'image de classification pour différentes variations de taille du signal

2.2.1. Introduction

L'objectif de cette expérience est de montrer qu'une différence se manifeste au niveau du traitement des contours entre la fovéa et la périphérie, alors la dimension de contour serait plus facilement utilisée par les observateurs en vision centrale par rapport à une présentation des contours dans la périphérie. En effet, il y a une perte d'information dans le champ visuel périphérique, mais pas en vision centrale.

La première série d'expérience indiquait que l'information de surface était principalement utilisée par les observateurs pour donner leur réponse de détection. La taille du signal était de 3 deg, et il était possible que les contours de ce signal (à +/- 1.5 deg) soient positionnés dans la périphérie du champ visuel de l'observateur. Par conséquent, les observateurs auraient tendance à ne pas se servir de cette dimension dans leur réponse. L'objectif de cette expérience est de montrer que la dimension de contour peut être utilisée si les contours ne sont pas positionnés dans le champ périphérique. En effet, il est bien connu que le centre de la rétine (la fovéa) est sensible au détail des stimuli. En revanche, lorsqu'un même stimulus est présenté dans la périphérie rétinienne, une partie plus ou moins importante pour la reconnaissance de ce stimulus est perdue (DeWeerd, Desimone & Ungerleider, 1998 ; Levi & Klein, 1985, 2002 ; Rovamo & Virsu, 1979). Pour cela, nous avons réduit la taille du signal à 1 deg. Nous nous plaçons donc dans une condition où la fovéa est principalement stimulée, puisque le diamètre de cette région est de 2 deg (Schwartz, 1999). Le signal est à proximité du point de fixation, mais aussi au sein de la fovéa. Nous avons également utilisé une condition dans laquelle la taille du signal est plus importante (6 deg). Dans ce cas, les contours du signal sont situés dans la périphérie du champ visuel et la dimension de surface devrait être principalement utilisée par les observateurs. Pour cette condition, nous nous attendons donc à ce que les images de classification soient identiques aux expériences précédentes.

Des hypothèses identiques de la première expérience peuvent être formulées : 1) les contours sont la seule information utilisée par l'observateur (Figure 45a) ou 2) il est possible que la surface apparaisse dans l'image de classification mais que son niveau de luminance est plus faible que celui des bords (Figure 45b).

En outre, si l'observateur base sa réponse sur un emplacement précis correspondant au signal en adaptant sa fenêtre attentionnelle, dans ce cas, la taille de la zone centrale devrait diminuer ou augmenter (selon le signal présenté) par rapport à la taille estimée dans la première expérience. Au contraire, si les observateurs utilisent un prototype fixe indépendamment de la taille du signal, alors la taille de la zone centrale devrait être proche de celle déterminée pour l'expérience princeps.

2.2.2. Méthode expérimentale

- Participants.

Les observateurs sont les mêmes que dans l'expérience de détermination de seuil (FD, SD et VB). Ainsi, pour la condition dans laquelle le signal est de 6 deg, aucun observateur « naïf » n'a réalisé l'expérience.

- Stimuli.

Les stimuli étaient composés soit d'un signal (un carré de 1 ou de 6 degrés d'angle visuel de côté) superposé à un bruit de 10 degrés d'angle visuel de côté, soit d'un bruit présenté seul. Le signal était présenté au niveau de contraste seuil pour chaque observateur. Pour une taille de 1 deg, les valeurs de contraste étaient de 0.17 pour FD et 0.12 pour VB, et pour une taille du signal de 6 deg ces valeurs étaient de 0.06 pour FD et 0.08 pour SD. Les stimuli étaient donc identiques aux expériences antérieures, exceptée la taille du signal.

- Procédure.

La tâche des observateurs était la même que précédemment : détecter le signal. Le temps de présentation était identique à la première expérience (400 ms). Le nombre de répétitions a été diminué, puisque FD a passé 30 blocs, SD et VB 25 blocs, mais un même nombre d'essais était demandé par bloc.

2.2.3. Résultats

- L'indice d' et le critère de décision.

La sensibilité de chaque sujet a tout d'abord été calculée pour chaque bloc. Les résultats sont présentés à la Figure 61. Les résultats indiquaient que les performances de FD, SD et de VB étaient stables et proches de 1. Dans la condition « signal 1 deg », l'indice d' calculé pour l'ensemble des essais était de 1.35 pour FD et de 0.87 pour VB. De même, lorsque la taille du signal était de 6 deg, les résultats indiquaient que la sensibilité globale des observateurs était également proche de 1 (FD = 1.45 et SD = 1.52). Par ailleurs, le calcul du critère de décision indique que celui ne bouge pas tout au long de l'expérience, et ce, quelle que soit la condition expérimentale. De plus, nous avons également déterminé un critère pour l'ensemble des essais, le résultat indique qu'il est proche de 0 pour un signal de 1 deg (FD = -0.19 et VB = 0.18), mais aussi pour un signal de 6 deg (-0.14 = FD et -0.23 = SD).

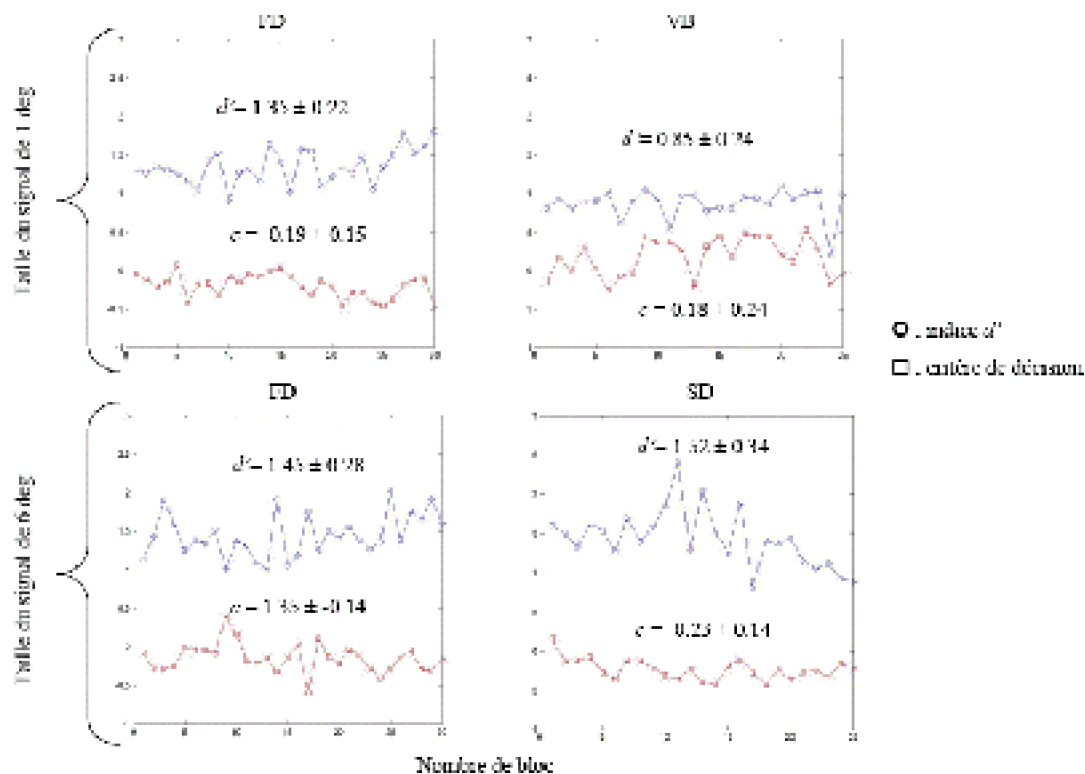


Figure 61 : Expérience 3 : Résultats de l'indice d' et du critère de décision au cours des blocs expérimentaux obtenue par FD, SD et VB pour une taille du signal de 1 (figures du haut) et 6 (figures du bas) degrés d'angle visuel.

- Image de classification.

Les résultats des images de classification sont présentés dans la Figure 62. On observe que les images obtenues par les deux observateurs étaient relativement identiques à celles observées pendant la première série d'expérience, et ce, quelle que soit la taille du signal. En effet, une zone centrale claire émergeait à l'emplacement du signal et un faible pourtour sombre l'entourait.

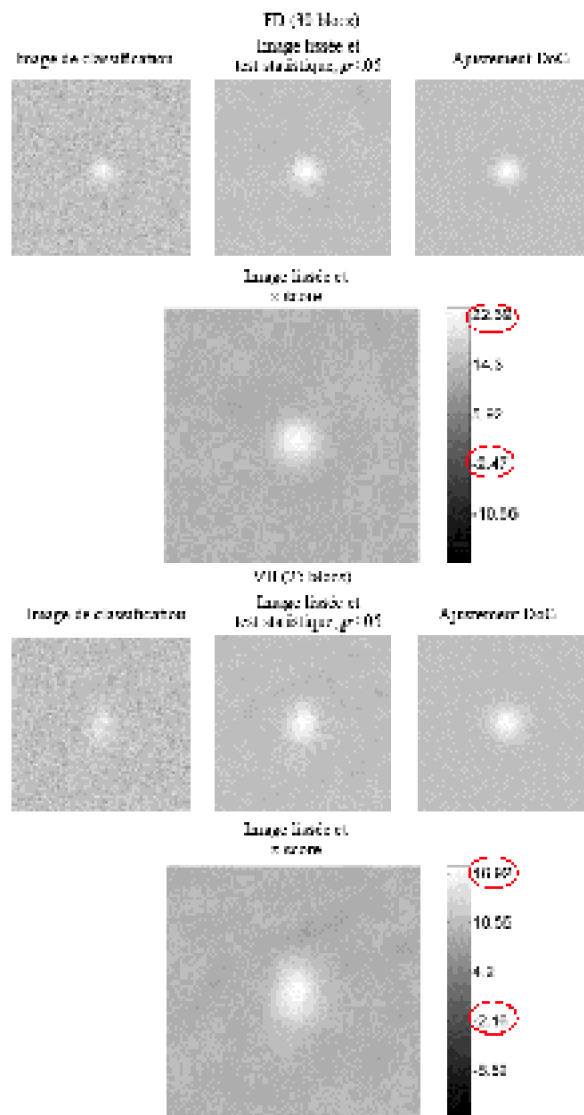


Figure 62 : Expérience 3 : Résultats de l'image de classification, du test statistique, du z-score et de l'ajustement « DoG » de FD et VB pour une taille du signal de 1 deg. Les valeurs du test sont affichées à droite de l'image, et les valeurs maximales et minimales des images sont entourées.

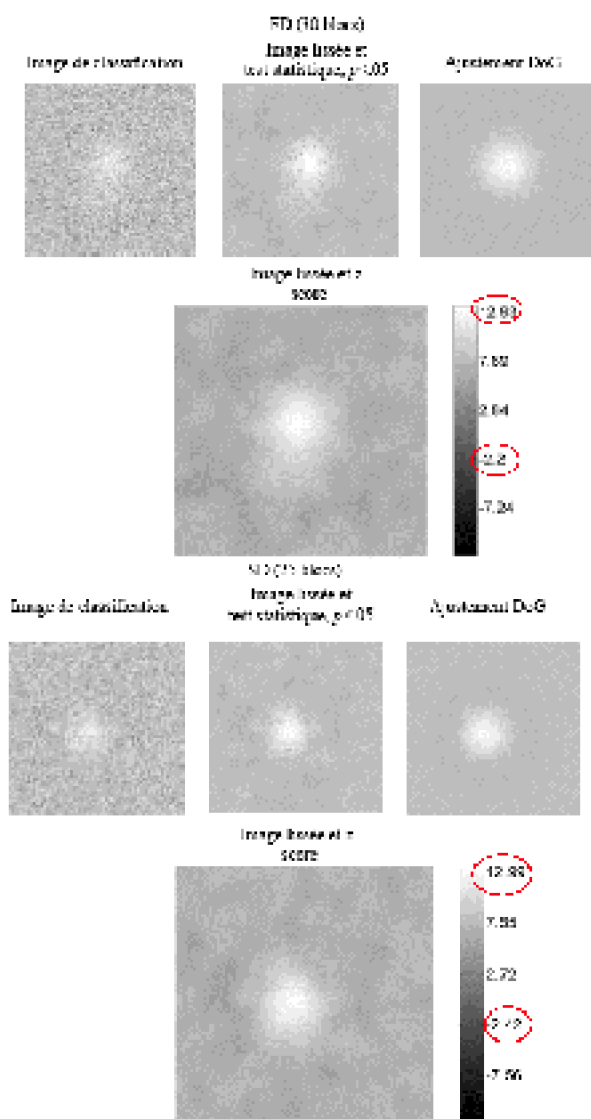


Figure 63 : Expérience 3 : Résultats de l'image de classification, du test statistique et normalisé, mais aussi de l'ajustement « DoG » des observateurs FD et SD pour une taille du signal de 6 deg. Les valeurs entourées dans la barre de droite correspondent aux minima et aux maxima de l'image.

Les résultats de l'ajustement DoG, pour une taille du signal de 1 deg, sont présentés dans le Tableau VII et ont permis d'estimer la participation de la zone centrale et du pourtour. La taille du centre était plus petite que la taille obtenue par les mêmes observateurs dans l'expérience princeps. Par contre, la taille de la zone centrale était 2-3 fois plus grande que la taille réelle du signal, ce qui indique que les deux observateurs ont surestimé le signal qui devait être pris en compte. De plus, le rapport entre le pourtour et le centre montrait que les réponses des observateurs reposaient sur un pourtour environ 4 fois plus important que la taille de la zone centrale. Ce rapport était beaucoup plus important que celui utilisé dans la première expérience. La réponse des observateurs était basée sur la taille d'un pourtour plus grand lorsque la taille du signal diminuait. Par contre, le rapport centre/pourtour montrait également que l'amplitude du pourtour diminuait avec

la taille du signal, puisque ce dernier représentait entre 1 et 3% de l'amplitude de la zone centrale (contre 4 et 8 % dans l'expérience princeps).

Tableau VII: Valeur de l'amplitude et du sigma pour la zone du pourtour et la zone centrale ainsi que le rapport entre ces deux zones pour les tailles du signal de 1 et 3 degrés d'angle visuel des observateurs FD et VB.

Tableau VII: Valeur de l'amplitude et du sigma pour la zone du pourtour et la zone centrale ainsi que le rapport entre ces deux zones pour les tailles du signal de 1 et 3 degrés d'angle visuel des observateurs FD et VB.							x_0 (min)	y_0 (min)
	Zone centrale		Pourtour		Rapport pourtour / zone centrale			
Sujet	Sigma (deg)	Amplitude	Sigma (deg)	Amplitude	Sigma	Amplitude		
FD (3 deg)	4.31	3.52	10.55	0.28	2.45	0.008	12.38	12.94
FD (1 deg)	2.3	5.75	8.9	0.02	3.87	0.003	11.89	12.15
VB (3 deg)	4.05	2.84	7.02	0.01	1.75	0.04	4.6	-3.46
VB (1 deg)	3.13	3.41	13.7	0.05	4.39	0.01	5.15	0.33

De même, le tableau VIII présente les résultats des observateurs FD et SD pour la condition dans laquelle le signal est de 6 deg . On observe que la taille de la zone centrale était modifiée par la variation de la taille du signal : en effet, lorsque la taille du signal augmentait (de 3 à 6 degrés) la taille de la zone centrale augmentait (de 4.31 et 2.31 à 4.8 et 4.01 respectivement pour FD et SD). Par contre, la taille de la zone centrale est inférieure à la taille réelle du signal. De même, la taille du pourtour augmentait avec la taille du signal (de 10.55 et 9.76 à 18.3 et 10.08 respectivement pour FD et SD). Néanmoins, la taille du pourtour était 2 à 4 fois plus importante que celle de la zone centrale, le rapport est approximativement identique à celui de la première expérience. Les valeurs d'amplitude allaient dans le même sens pour les deux observateurs. En effet, l'amplitude de la zone centrale diminuait lorsque la taille du signal augmentait (de 3.52 et 5.05 à 2.25 et 2.88 respectivement pour FD et SD), et l'amplitude du pourtour augmentait avec la taille du signal (de 0.028 et 0.15 à 0.07 et 0.28 respectivement pour FD et SD). Le rapport centre/pourtour montrait que le pourtour participait toujours aussi faiblement à la détection (entre 3 et 8% de la zone centrale). Ce qui indique que les observateurs n'utilisent pas un contraste plus important ou plus faible pour détecter le signal.

Tableau VIII : Valeur de l'amplitude et du sigma pour la zone du pourtour et la zone centrale ainsi que le rapport entre ces deux zones pour les tailles du signal de 3 et 6 degrés d'angle visuel de FD et SD.

Tableau VIII : Valeur de l'amplitude et du sigma pour la zone du pourtour et la zone centrale ainsi que le rapport entre ces deux zones pour les tailles du signal de 3 et 6 degrés d'angle visuel de FD et SD.							x_0 (min)	y_0 (min)
	Zone centrale		Pourtour		Rapport pourtour / zone centrale			
Sujet	Sigma (deg)	Amplitude	Sigma (deg)	Amplitude	Sigma	Amplitude		

FD (3 deg)	4.31	3.52	10.55	0.28	2.45	0.008	12.38	12.94
FD (6 deg)	4.8	2.25	18.3	0.07	3.81	0.03	4.56	13.31
SD (3 deg)	2.31	5.05	9.76	0.15	4.23	0.03	-13.5	6.3
SD (6 deg)	4.01	2.88	10.08	0.26	2.51	0.09	-28.40	19.77

De la même façon que pour l'expérience princeps, les images de classification ont été déterminées pour les essais où le signal est présent et absent. La Figure 64 présente les résultats obtenus, et ces derniers sont similaires aux calculs de l'image de classification globale pour les deux conditions. L'information de contour n'apparaît pas directement dans nos images. Par ailleurs, les résultats de l'ajustement DoG (Tableau IX pour un signal de 1 deg et Tableau X pour un signal de 6 deg) indiquent que les paramètres estimés, pour un même observateur, sont proches des images de classifications globales quel que soit la présence ou non du signal. De même, ces valeurs sont similaires à celles déterminées pour l'image de classification globale. Les résultats de cet ajustement ont permis de montrer que l'étendue de la zone centrale était non seulement similaire lorsque le signal est présent ou absent, pour un même observateur. Ce résultat ne permet donc pas de confirmer l'hypothèse de l'incertitude spatiale. Toutefois, pour un signal de 1 deg, les résultats indiquaient que l'amplitude du pourtour étaient plus faible pour les essais où le signal est absent. En effet, le ratio de l'amplitude indique que le pourtour participait faiblement à la détection (de 0.1 à 5 % de l'amplitude du centre) pour les essais où le signal est présent. Néanmoins, ce ratio est encore plus faible pour les essais où le signal est absent. Ces résultats permettent de valider l'hypothèse selon laquelle la dimension de contour serait plus importante lorsque le signal est présent. L'estimation des coordonnées centrales indiquait un biais relativement faible est identique avec ceux du calcul précédent quel que soit la taille du signal.

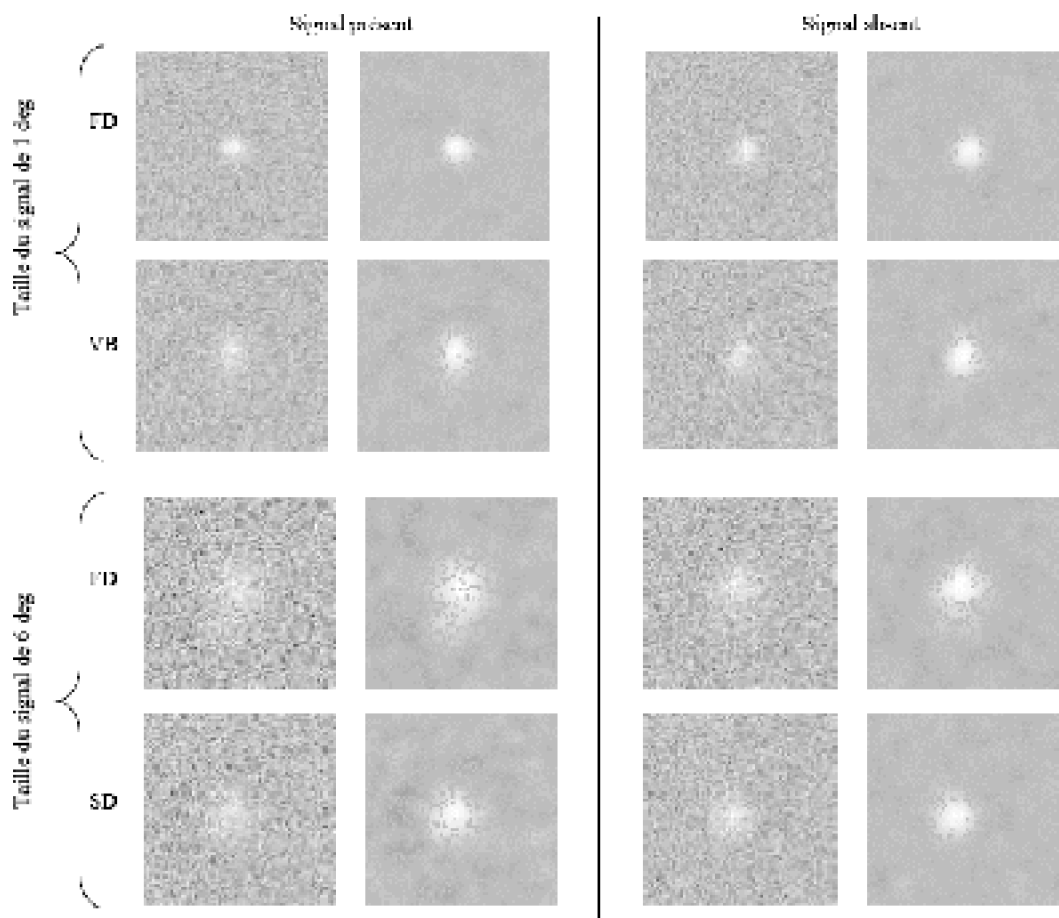


Figure 64 : Expérience 3 : Résultats des images de classification pour les essais où le signal était présent (à gauche) ou absent (à droite) pour les deux conditions (signal de 1 ou 6 deg).

Tableau IX : Résultats déterminés par l'ajustement DoG pour les conditions où le signal est présent ou absent des observateurs FD et VB lorsque la taille du signal est de 1 deg. Le tableau indique les valeurs d'amplitude et du sigma de la zone du pourtour (colonnes 2 et 5) et la zone centrale (colonnes 1 et 4), pour le rapport entre ces deux zones (colonnes 3 et 6), et enfin les positions des zones centrales sont indiquées aux colonnes 7 et 8.

Observateur		A_c	A_p	A_p/A_c	σ_c (deg)	σ_p (deg)	σ_c/σ_p	x_0 (min)	y_0 (min)
FD	Présent	3.41	0.17	0.05	2.2	6.28	2.85	7.54	8.42
	Absent	2.58	0.06	0.02	2.37	16.45	6.94	16.86	16.92
VB	Tableau X : L'estimation des paramètres de l'amplitude et du sigma sont présentées pour la zone centrale et le pourtour pour un signal de 6 deg. Les deux ratios sont indiqués aux colonnes 3 et 6. De plus, la position du point centrale de la fonction DoG est précisée dans les colonnes 7 et 8. L'ensemble de ces valeurs ont été déterminé pour les essais où le signal est présent ou absent et ce de façon spécifique à chaque observateur (FD et SD).								
	Présent	1.71	0.03	0.02	3.12	17.82	5.71	6.18	5.43
	Absent	1.75	0.003	0.001	3.04	12.57	4.13	3.86	5.54

Tableau X : L'estimation des paramètres de l'amplitude et du sigma sont présentées pour la zone centrale et le pourtour pour un signal de 6 deg. Les deux ratios sont indiqués aux colonnes 3 et 6. De plus, la position du point centrale de la fonction DoG est précisée dans les colonnes 7 et 8. L'ensemble de ces valeurs ont été déterminé pour les essais où le signal est présent ou absent et ce de façon spécifique à chaque observateur (FD et SD).

Observateur		A_c	A_p	A_p/A_c	σ_c (deg)	σ_p (deg)	σ_c/σ_p	x_0 (min)	y_0 (min)
FD	Présent	1.01	0.001	0.0004	4.97	7.58	1.53	-4.68	-11.44
	Absent	1.28	0.07	0.05	4.49	11.79	2.63	-4.68	-15.70
SD	Présent	1.35	0.03	0.02	5.69	7.02	1.23	-26.93	20.55
	Absent	1.49	0.05	0.03	3.68	13.6	3.7	-27.1	20.46

2.2.4. Discussion

Les résultats obtenus dans cette expérience permettent de conforter les résultats de la première série d'expériences. Les images de classification pour une tâche de détection où la taille du signal était réduite à 1 deg ou augmentait à 6 deg, indiquaient une zone centrale claire entourée d'un faible pourtour sombre.

Pour expliquer les résultats de la première série d'expérience, nous émettons l'hypothèse que l'absence de la dimension de contour dans la réponse donnée par les observateurs était la conséquence d'un traitement du champ visuel périphérique. En effet, il était possible que l'absence de contour soit due au fait que les contours du carré tombent dans la périphérie et non pas dans la fovéa. Cette hypothèse nous a conduit à utiliser des conditions expérimentales telles que la taille du signal soit de 1 deg et que le signal soit placé au centre de l'écran. Néanmoins, la sensibilité de la fovéa ne semble pas permettre pas d'intégrer plus facilement l'information de contour puisque cette information ne ressort pas dans les images de classification. Cette expérience démontre donc la pertinence du traitement de la surface pour la détection.

L'hypothèse alternative était d'obtenir un résultat identique à ce que nous avons dans les expériences précédentes. En effet, les aspects spatiaux et temporels d'un signal influencent souvent les traitements qui sont effectués par les observateurs. Par

conséquent, il était possible que si la manipulation du temps de présentation n'influait pas les réponses données par les observateurs, alors la variation de taille du signal ne changerait pas les indices utilisés pour répondre. Cette hypothèse est ici vérifiée pour la diminution et l'augmentation de la taille du signal.

Dans le cadre de cette dernière hypothèse, nous supposons que les observateurs devraient tenir compte de la taille du signal qui est présentée. En effet, la deuxième expérience a permis de montrer que les observateurs n'utilisaient pas un prototype fixe pour répondre, mais que ce dernier changeait en fonction du temps de présentation. Les résultats observés, à partir de l'ajustement de la fonction DoG semblent indiquer que la fenêtre attentionnelle des observateurs tenait compte des différentes variations de taille du signal. En effet, la taille de la zone centrale était moins importante par rapport à la première expérience pour un signal de 1 deg, de façon réciproque la zone centrale était plus importante lorsque le signal était de 6 deg.

Toutefois, l'analyse a également permis de montrer que la zone centrale utilisée par les observateurs n'était pas exactement ajustée à la taille du signal présenté.

En effet, pour un signal de 1 deg, les observateurs avaient tendance à surestimer la taille de la surface. La taille du pourtour est approximativement quatre fois plus importante que la taille de la zone centrale. Ce rapport est plus élevé par rapport à la première expérience pour les mêmes observateurs. L'ajustement DoG fait également ressortir le fait que l'amplitude du pourtour ne représente plus que 1 à 3 % de l'amplitude de la zone centrale. Ce dernier résultat semble indiquer que les observateurs utilisaient moins le pourtour de la zone centrale pour donner leur réponse de détection. La fovéa qui a une plus grande sensibilité permettrait ainsi aux observateurs d'attribuer une réponse plus importante à la surface sans avoir besoin du pourtour pour répondre. Il est donc possible que la présence du pourtour faciliterait la réponse des observateurs lorsque le signal n'est pas présenté en vision centrale.

De même, pour un signal de 6 deg la taille de la zone centrale est inférieure à la taille réelle du carré. En d'autres termes, les observateurs ont tendance à sous estimer la surface du signal. A la fin du paragraphe précédent, nous avons formulé l'hypothèse que le pourtour aiderait plus facilement les observateurs à détecter le signal lorsque ce dernier est grand. Cette hypothèse est confirmée ici par le rapport de l'amplitude entre la zone centrale et le pourtour. En effet, ce rapport est d'autant plus important que la taille du signal est large. Ce résultat indique que les observateurs utilisent le pourtour sombre lorsque les contours du signal sont placés dans la périphérie du champ visuel. Néanmoins, ce résultat est à relativiser puisque nous n'avons pas effectué d'analyse permettant d'affirmer si il y a une différence significative entre les ajustements DoG pour chaque condition.

2.3. Comparaison des trois tailles du signal

Les résultats ont montré que quelle que soit la taille du signal (1, 3 et 6 deg) deux résultats apparaissent : une zone centrale claire et un pourtour sombre. Cette partie a pour objectif de relier l'ensemble des résultats obtenus pour les trois différentes tailles du

signal. Ce qui nous permettra de déterminer les relations qui existent entre la taille réelle et la taille prise en compte par les observateurs à partir des résultats de l'ajustement de la fonction DoG. Toutefois, les idées qui seront émises dans cette partie ne relèvent que de l'observation.

2.3.1. Résultats

Un récapitulatif, de l'ajustement de la fonction DoG sur l'ensemble de nos expériences, est indiqué dans le Tableau XI. Trois résultats importants émergent sur l'ensemble des données. Tout d'abord, la taille de la zone centrale n'était pas exactement ajustée aux différentes tailles du signal. De plus, le pourtour représente 2 à 4 fois la taille de la zone centrale, et participe faiblement à la détection de 1 à 9 % de la luminance de la zone centrale.

De façon plus spécifique, l'objectif était de déterminer l'évolution des images de classification selon la taille du signal. Pour cela, les données de l'observateur FD ont été comparées sur les trois différentes tailles du signal, ainsi que les résultats de VB à 1 et 3 degrés et de SD à 3 et 6 degrés.

Pour un signal de 1 degré, la taille prise en compte par les observateurs était plus importante (2.3 et 3.13 degrés) que la taille réelle du signal. De plus, pour un signal étendu (6 degrés), la taille de la zone centrale était plus petite (4.8 et 4.01 degrés) que la taille réelle du signal. Par contre, pour une taille intermédiaire du signal (3 degrés), la taille du centre variait entre 2 et 4 degrés. Ainsi, plus la taille du signal devenait importante, plus la taille de la zone centrale augmentait. De plus, l'amplitude de la zone centrale diminuait avec l'augmentation de la taille du signal pour un même observateur. La taille du pourtour augmentait en fonction de la taille du signal, excepté pour l'observateur VB dont la taille du pourtour diminuait. L'amplitude du pourtour avait tendance à augmenter avec la taille du signal, excepté pour VB. Par ailleurs, le rapport du sigma entre le pourtour et la zone centrale était approximativement identique pour les trois tailles du signal. Par contre, le rapport de l'amplitude entre ces deux zones augmentait lorsque la taille du carré augmentait. Ce dernier résultat indique la possibilité que plus la taille du signal est importante et plus le contraste entre la zone centrale et le pourtour est élevé.

Tableau XI : Valeur de l'amplitude et du sigma pour la zone du pourtour et la zone centrale ainsi que le rapport entre ces deux zones pour les tailles du signal de 1, 3 et 6 degrés d'angle visuel de chaque observateur.

Sujet	Zone centrale		Pourtour		Rapport pourtour / zone centrale	
	Sigma (deg)	Amplitude	Sigma (deg)	Amplitude	Sigma	Amplitude
FD (1 deg)	2.3	5.75	8.9	0.02	3.87	0.003
VB (1 deg)	3.13	3.41	13.7	0.05	4.39	0.01
FD (3 deg)	4.31	3.52	10.55	0.03	2.45	0.008
SD (3 deg)	2.31	5.05	9.76	0.15	4.23	0.03
SH (3 deg)	2.85	3.72	12.18	0.1	4.26	0.03
VB (3 deg)	4.05	2.84	7.02	0.01	1.75	0.04
FD (6 deg)	4.8	2.25	18.3	0.07	3.81	0.03
SD (6 deg)	4.01	2.88	10.08	0.26	2.51	0.09

2.3.2. Discussion

Les expériences de la variation de la taille du signal ont permis mettre en évidence trois principaux résultats. Le premier résultat montre que, quelles que soient les variations de la taille du signal, une zone claire centrale entourée par une zone sombre continue d'être observé. Ceci permet de généraliser les résultats de la première série d'expérience, puisque la manipulation des facteurs temporels et spatiaux ne modifie pas la pertinence de l'information de surface dans la détection.

Le second résultat provient de l'observation des différents ajustements DoG : les tailles de la zone centrale et du pourtour immédiat variaient en fonction de la taille du signal. Ce résultat conforte notre hypothèse selon laquelle les observateurs ajustent leur fenêtre attentionnelle à la taille du signal. Néanmoins, la taille de la zone centrale ne correspond pas exactement à la taille du signal qui est présentée. En effet, pour un signal de 6°, la taille de la zone centrale est plus petite et, pour une taille du signal de 1°, elle est plus grande. Par contre, la taille du pourtour est approximativement et toujours 2 à 4 fois plus grande que la taille de la zone centrale. Il se peut ainsi que FD et SD tiennent moins compte de la surface du signal quand celle-ci est plus grande que la taille des champs récepteurs. Réciproquement, FD et VB ont tendance à exagérer l'importance de la surface pour un signal plus petit que la taille des champs récepteur, puisque celui-ci active l'ensemble des neurones.

Le troisième résultat est obtenu par l'ajustement de la fonction DoG et concerne la participation de la luminance du pourtour immédiat. En effet, le rapport d'amplitude entre la zone centrale et le pourtour augmente avec la taille du signal. Ce résultat indique la réponse des observateurs nécessite une plus grande participation du pourtour lorsque le carré s'agrandit. Les observateurs ont tendance à donner leur réponse de détection quand les pixels autour de la zone centrale sont de plus en plus foncés. Ce résultat pourrait être interprété comme une stratégie de l'observateur pour pallier une perte d'information issue du système visuel périphérique.

2.4. Expérience de détermination de seuil pour une taille du signal de 3 degrés d'angle visuel et une taille du bruit de 30 degrés d'angle

visuel

2.4.1. Introduction

Une expérience de détermination de seuil a tout d'abord été réalisée afin de définir les conditions spécifiques à utiliser pour l'expérience de détermination de l'image de classification. Nous nous attendions à ce que quel que soit le champ du bruit, celui-ci n'influence pas le seuil de détection.

2.4.2. Méthode expérimentale

- Participants.

Un seul observateur (FD) a passé l'expérience.

- Stimuli.

La seule différence qui existait avec la première expérience était l'augmentation du champ du bruit : 30 degrés d'angle visuel au lieu de 10 degrés. Les éléments du bruit sont toujours identiques (2*2 pixels). La luminance du signal se répartissait sur 7 niveaux depuis une échelle logarithmique ($10^{0.1}$) à partir de la valeur 0.045. Comme dans les expériences précédentes sauf la première, une seule valeur de contraste du bruit a été utilisée (0.7).

- Procédure.

La procédure était en tout point identique aux précédentes expériences de détermination de seuil.

2.4.3. Résultats

Les résultats de FD sont présentés dans la Figure 65. Le seuil de contraste du signal pour un $d' = 1$ était de 0.06. Conformément à notre hypothèse, la valeur de contraste du signal était semblable à celle de la première expérience pour laquelle la taille du bruit était de 10 degrés. L'augmentation de la taille du bruit n'entraîne pas une perturbation ou une facilitation de la détection du signal.

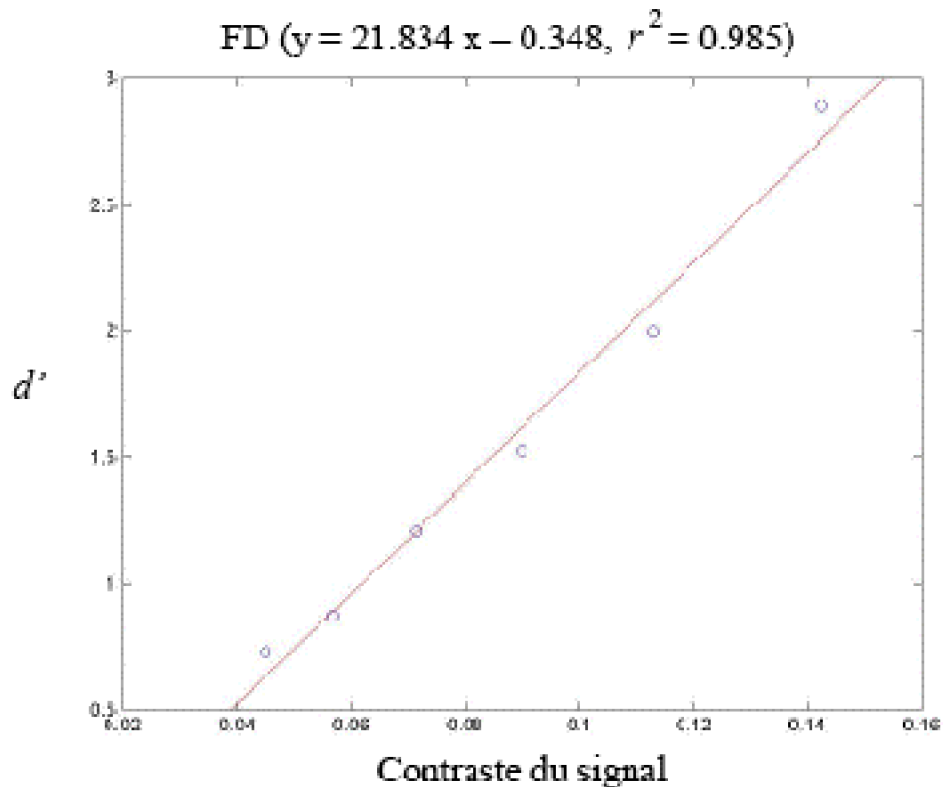


Figure 65 : Expérience préliminaire 5 : Droite de régression entre l'indice d' et le contraste du signal du sujet FD pour une taille du signal de 3° et une taille du bruit de 30° .

Par la suite, une représentation des courbes ROC a été réalisée, ces dernières sont présentées dans l'Annexe 5. Les figures présentent des résultats identiques à ceux observés dans les expériences précédentes. Les courbes étaient de plus en plus creusées, ce qui signifie que le signal apparaissait distinctement au fur et à mesure de l'augmentation du contraste du signal. Par ailleurs, si l'on compare ces courbes avec celles du même observateur mais lorsque la taille du bruit était de 10° , alors nous pouvons constater que les courbes avaient une apparence similaire en fonction du niveau de contraste.

2.5. Expérience de détermination de l'image de classification pour une taille du signal de 3 degrés et une taille du bruit de 30 degrés

2.5.1. Introduction

L'objectif de cette expérience est de généraliser les résultats que nous avons obtenus dans les expériences précédentes pour une autre propriété spatiale que la taille du signal. Nous nous sommes intéressés à la variation du champ du bruit. En effet, la distance qui sépare les bords du signal et ceux du bruit a très bien pu influencer notre principal résultat.

Les deux précédentes expériences ont montré que l'augmentation ou la diminution de

la taille du signal ne changeait pas notre principal résultat. L'augmentation de la taille du bruit ne devrait pas changer l'importance accordée à la surface pour la détection. Ainsi, nous nous attendons à ce que l'image de classification soit proche de celles obtenues dans nos expériences précédentes.

Par ailleurs, cette expérience permettra de déterminer l'influence de la taille du bruit sur les paramètres estimés par l'ajustement DoG. En effet, DeBonet et Zaidi (1997) ont réalisé une expérience dans laquelle le contraste entre le fond et le signal était identique, mais la taille du fond augmentait. Les résultats indiquaient que la luminosité était plus importante lorsque la taille du fond augmentait. Si nous appliquons ce raisonnement à notre expérience, le ratio A_c/A_p augmenterait lorsque le champ du bruit est important, et ce pour une taille du signal identique. En d'autres termes, nous nous attendons donc à ce que l'amplitude de la zone centrale soit plus élevée par rapport à l'expérience où le bruit était de 10 deg.

Au cours de notre expérience princeps, une exagération des pixels significatifs a été effectuée. Les résultats mettaient en évidence la participation des informations de l'extrême pourtour. Il est donc possible que ces informations aient pu être prises en compte si la taille du bruit de 10 deg correspondait à une zone du champ visuel utilisé par les observateurs. Par conséquent, si l'observateur utilise le champ visuel dans son entier ou une partie du champ visuel centré autour du signal, alors l'augmentation de la taille du bruit engendrerait une influence locale de l'extrême pourtour. Par contre, si l'observateur tient compte du champ visuel dans son entier, dans ce cas, les informations issues de l'extrême pourtour seraient toujours présentes même pour un bruit extrêmement grand.

2.5.2. Méthode expérimentale

Cette expérience a été effectuée par un seul observateur (FD). La méthode expérimentale était en tout point semblable à celle des expériences antérieures, seule la taille du bruit variait (30 degrés au lieu de 10 degrés). FD a réalisé 25 blocs expérimentaux. Chaque bloc expérimental comportait 200 essais, et le signal était présent uniquement dans la moitié des essais.

2.5.3. Résultats

Le résultat de la sensibilité de FD est présenté dans la Figure 66. L'effet d'apprentissage est vérifié de la même façon que pour les expériences préalables grâce à l'indice d' , et la place du critère de décision a également été calculée. Ces derniers ne montraient aucune variation pendant l'expérience. De plus, le calcul de l'indice d' pour l'ensemble des essais (1.15) se rapproche de 1, et le critère tend vers 0 (-0.32). L'image de classification a ensuite été déterminée (Figure 67). Les résultats sont similaires à ceux trouvés dans les précédentes expériences : une zone claire centrale est entourée d'une zone sombre. De même, l'exagération des pixels significatifs indiquaient une participation des pixels de l'extrême pourtour.

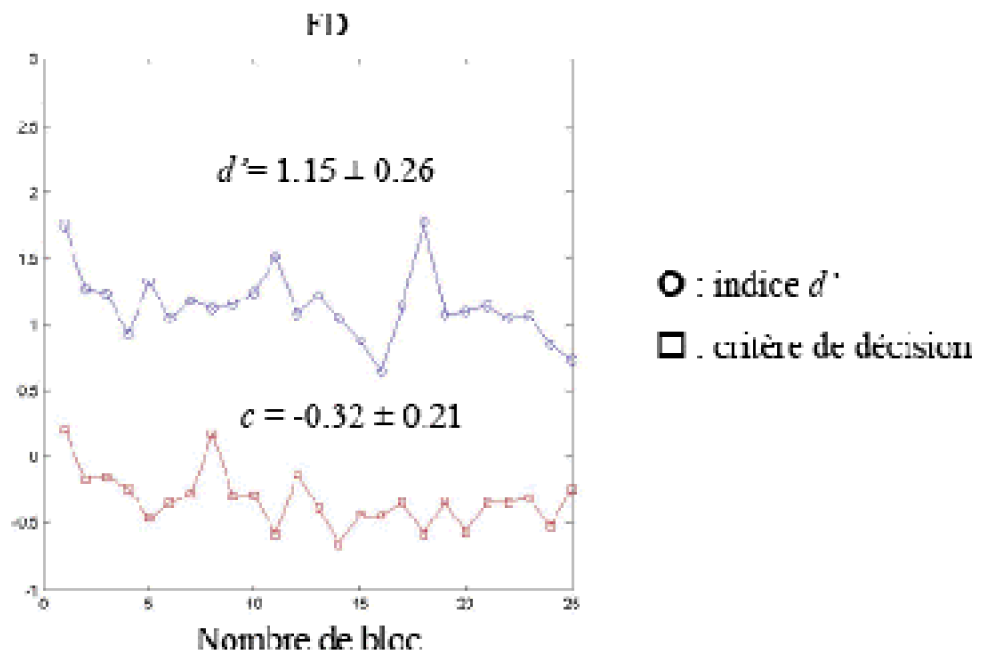


Figure 66 : Expérience 5 : Résultats de l'indice d' et du critère de décision au cours des 25 blocs expérimentaux obtenue par FD pour une taille du signal de 3° et une taille du bruit de 30° d'angle visuel.

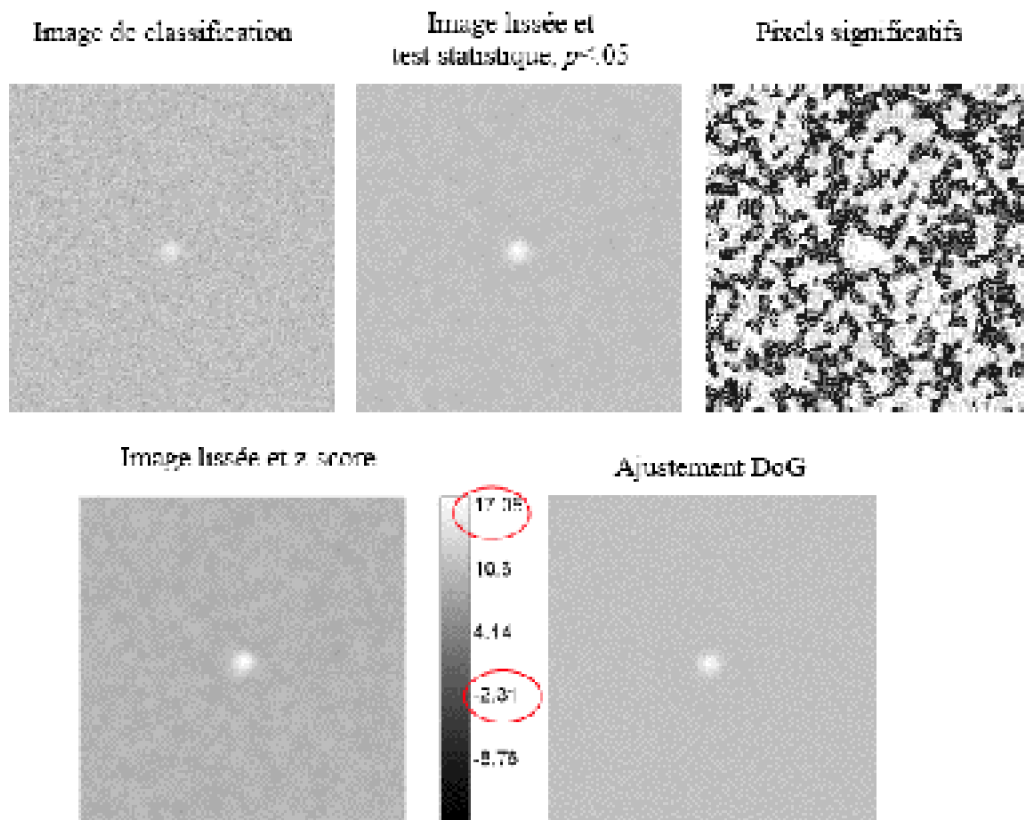


Figure 67 : Expérience 4 : Résultats de l'image de classification, du test statistique et normalisé, des pixels significatifs et de l'ajustement « DoG » de FD pour une taille du signal de 3° et une taille du bruit de 30°. La bande à la droite de l'image z-score permet de faire correspondre les valeurs du test au niveau de luminosité de l'image, et les chiffres entourés représentent les valeurs minimales et maximales obtenues pour FD.

Par la suite, l'ajustement « DoG » (cf. Tableau XII) a permis de montrer que cette zone sombre contribuait faiblement à la détection (3 %) et que la taille de la zone centrale (3.13°) était proche de la taille du signal (3 deg). De plus, des zones claires et sombres continuaient d'émerger sur l'ensemble de l'image de classification.

Nous avons ensuite comparé la condition où le bruit était de 10 deg et celle où le bruit était de 30 deg. Ce rapprochement a été réalisé entre les deux conditions puisque l'observateur FD a participé aux deux expériences, cette comparaison est à relativiser puisqu'aucun test statistique n'a été réalisé entre les différents ajustements de la fonction DoG. La taille de la zone centrale diminuait lorsque la taille du bruit augmentait, mais la taille du pourtour était identique. Par ailleurs, l'amplitude du pourtour était plus importante

pour un bruit de 30 deg, par contre l'amplitude de la zone centrale était stable entre les deux conditions. Par conséquent, le ratio A_c/A_p indiquait que la zone du pourtour participait plus à la détection du signal. De même, la position centrale de la fonction DoG était soumise à un faible biais qui était toutefois plus important pour l'axe des y_0 .

Tableau XII : Valeur de l'amplitude et du sigma pour la zone centrale (colonnes 1 et 2) et le pourtour (colonnes 3 et 4), ainsi que le rapport entre ces deux zones (colonnes 5 et 6) pour une taille du signal de 3 degrés d'angle visuel et un bruit de 30 degrés de l'observateur FD. La position centrale de l'ajustement est indiquée aux colonnes 7 et 8. Un rappel des estimations obtenues pour le même observateur dans l'expérience princeps est également présenté.

Tableau XII : Valeur de l'amplitude et du sigma pour la zone centrale (colonnes 1 et 2) et le pourtour (colonnes 3 et 4), ainsi que le rapport entre ces deux zones (colonnes 5 et 6) pour une taille du signal de 3 degrés d'angle visuel et un bruit de 30 degrés de l'observateur FD. La position centrale de l'ajustement est indiquée aux colonnes 7 et 8. Un rappel des estimations obtenues pour le même observateur dans l'expérience princeps est également présenté.							x_0 (min)	y_0 (min)
	Zone centrale		Pourtour		Rapport pourtour / zone centrale			
Sujet	Sigma (deg)	Amplitude	Sigma (deg)	Amplitude	Sigma	Amplitude		
FD (3 deg)	4.31	3.52	10.55	0.03	2.45	0.008	12.38	12.94
FD (30 deg)	3.13	3.59	11.17	0.09	3.57	0.03	4.43	26.41

Nous avons effectué une dernière analyse afin de déterminer si les contours du signal apparaissaient dans les images de classification où le signal est présent ou absent. Les images étaient similaires à celle observée lorsque ces deux conditions ne sont pas distinguées (cf Figure 68). Par ailleurs, les résultats de l'ajustement sont présentés dans le Tableau XIII. Les paramètres étaient similaires à ceux de l'image de classification globale, et aucune différence notable ne ressortait entre les deux conditions. Ces résultats ne validaient pas l'hypothèse de l'existence d'une incertitude spatiale qui renforcerait la présence des contours pour les essais où le signal est présent.

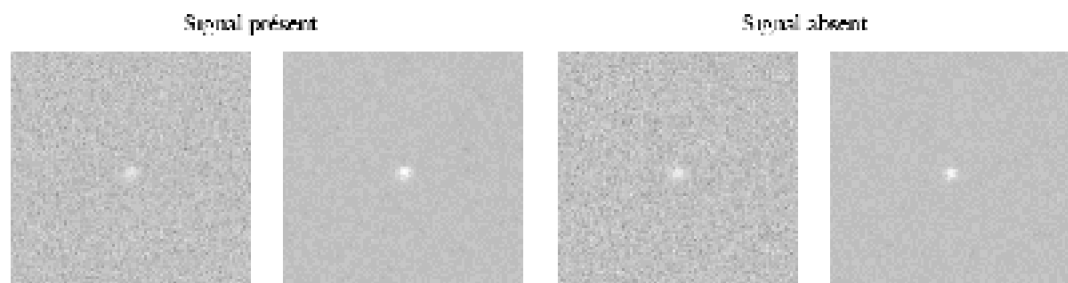


Figure 68 : Expérience 4 : Résultats des images de classification de l'observateur FD pour une condition expérimentale où la taille du bruit est de 30 deg. Les images de classification brutes et les images lissées avec le test statistique sont présentées pour les essais où le

signal est présent (à gauche) et absent (à droite).

Tableau XIII : Le tableau présente les résultats de l'ajustement de la fonction DoG de l'observateur FD pour la condition expérimentale où le bruit est de 30 deg, en distinguant les essais où le signal est présent et absent. Les deux premières colonnes indiquent les valeurs d'amplitude du centre et du pourtour, le ratio est indiqué à la colonne 3. Les colonnes 4 et 5 présente la taille (en deg) de ces deux paramètres mais aussi le ratio à la colonne 6. La position centrale de la fonction ajustée est présentée aux colonnes 7 et 8.

Observateur		A_c	A_p	A_p / A_c	σ_c (deg)	σ_p (deg)	σ_c / σ_p	x_0 (min)	y_0 (min)
FD	Présent	1.87	0.05	0.03	3.15	13.6	4.32	2.73	24.59
	Absent	1.74	0.06	0.03	3.05	10.31	3.38	6.52	26.63

2.5.4. Discussion

Cette expérience a permis de montrer que la participation de la surface pour la détection était stable. Par la suite, nous avons émis deux hypothèses concernant l'ajustement de la fonction DoG. Cette hypothèse est confirmée (cf colonne 6 du Tableau XI). Une seconde hypothèse concernait le fait que l'amplitude de la zone centrale serait plus élevée que lorsque le masque est seulement de 10 deg, conformément aux données de DeBonet et Zaidi (1997). Toutefois, nos données indiquaient que l'amplitude de la zone centrale était identique quelle que soit la taille du bruit. Cette différence pourrait s'expliquer par le fait que dans l'expérience de DeBonet et Zaidi (1997), le fond n'était pas constitué d'un bruit.

Un résultat non attendu a également été mis en évidence par l'ajustement DoG. En effet, la taille de la zone centrale diminuait lorsque la taille du bruit augmentait. Néanmoins, ce résultat est à relativiser puisque la taille de la zone centrale était correctement ajustée à la taille réelle du signal. Deux possibilités peuvent expliquer ce résultat : soit l'augmentation de la taille du bruit perturbait la perception du signal, soit FD était maintenant habitué à la taille du signal (après 31000 essais). Cette dernière explication a été retenue. En effet, l'expérience de détermination de seuil a montré que FD avait un contraste seuil du signal identique à celui de la première expérience. Par ailleurs, si le bruit facilitait ou perturbait la détection du signal, nous aurions pu nous attendre à ce que le contraste du signal diminue ou augmente par rapport à l'expérience princeps, ce qui n'est pas le cas ici.

Nos images ont également montré une participation des zones de l'extrême pourtour pour la détection. Toutefois, il est possible que le lissage triangulaire consiste en quelque sorte à moyenniser localement un petit ensemble de pixels proches et introduit peut-être ainsi des zones de cohérences. Cette hypothèse pourrait expliquer l'émergence de ces pixels dans l'extrême pourtour.

2.6. Conclusion

Notre première série d'expérience a permis de montrer le rôle déterminant de la dimension de surface pour différentes durées de présentation du stimulus dans une tâche

de détection. Par ailleurs, nous savons que le codage d'un signal est déterminé par de multiples canaux, mais est également influencé pas les facteurs spatiaux et temporels. Par conséquent, cette seconde série d'expériences avait pour objectif de déterminer si la dimension de surface était aussi importante pour la détection lorsque la taille du signal et du bruit variait. En effet, l'une de nos hypothèses était que la surface était principalement utilisée car l'information de contour n'était pas disponible pour l'observateur à cause d'une perte possible due à un traitement périphérique. Dans ce cas, cette dernière dimension aurait dû émerger dans l'image de classification lorsque la taille du signal était de 1 deg, par contre, pour une taille de 6 deg, la surface serait une information prépondérante. Les images de classification indiquaient, au contraire, une zone centrale claire à l'emplacement du signal, ce qui montrait la robustesse de la dimension de surface quelle que soit la taille du signal et du bruit. De même, un faible pourtour sombre entourait la zone centrale claire. Par conséquent, la manipulation des facteurs spatiaux et temporels ont permis de généraliser la prépondérance de la surface pour la détection.

D'un point de vue théorique ce résultat paraît surprenant. En effet, les études physiologiques classiques (Hubel & Wiesel, 1962, 1968) ont montré que l'une des premières informations, qui était encodée par le système visuel, reposait sur la dimension de contour des objets. De plus, de nombreuses études psychophysiques (Dresp & Bonnet, 1991, 1993 ; Paradiso & Nakayama, 1991) ont démontré l'importance du processus de remplissage pour le traitement de la surface. Néanmoins, les études physiologiques n'ont pas mis en évidence, de façon explicite, le traitement de ce processus dans le traitement d'une surface uniforme (Lamme, Rodriguez-Rodriguez & Spekreijse, 1999 ; Rossi & Paradiso, 1996, 1999). Par ailleurs, cette dimension est traitée dès le CGL (Rossi & Paradiso, 1999). Toutefois, dans ces expériences il est tout à fait possible que l'activation de ces neurones soit la conséquence de l'activité d'autres cellules à partir d'un contraste. Mais, notre expérience permet de montrer que le système visuel ne traite pas la surface à partir d'un processus de remplissage.

Nos expériences contribuent également à montrer la précocité du traitement d'une surface uniforme pour la détection. Les études comportementales, mettant en évidence un traitement hiérarchique contour/surface, utilisent rarement une tâche de détection. A notre connaissance, les expériences de Dresp et Bonnet (1991, 1993) ont été les seules à démontrer ce traitement hiérarchique pour la détection. Toutefois, les auteurs utilisaient la figure de Kanisza, et de nombreuses études en IRMf ont montré que le système visuel ne réalisait pas un traitement identique entre des formes illusoire et non illusoire (Hirsch, DeLaPaz, Relkin, Victor, Kim, Li, Borden, Rubin & Shapley, 1995 ; Mendola, Dale, Fischl, Liu & Tootell, 1999).

L'estimation des paramètres de l'ajustement de la fonction DoG à nos données a permis de montrer que le « prototype » utilisé par l'observateur n'était pas fixe quelle que soit les conditions expérimentales. Au contraire, les résultats ont permis de montrer que les observateurs surestimaient la taille du signal lorsque celui-ci était petit, et à l'inverse, ils sous-estimaient la taille du signal lorsque ce dernier était plus grand. Ainsi, les observateurs ajustent leur fenêtre attentionnelle à la taille du signal.

3. Troisième série d'expériences : changement de la consigne de détection

Les expériences qui ont été présentées précédemment ont permis de montrer l'importance de la dimension de surface dans une tâche de détection. Ce résultat était indépendant des variations temporelles et de taille du signal. Une critique peut être formulée concernant les deux premières séries d'expériences. En effet, Spillman et Kurtenbach (1992) ont émis l'hypothèse selon laquelle les contours ne sont pas perçus nettement, lorsque ces derniers sont placés dans un fond bruité. Ainsi, les contours d'une forme ne sont pas utilisés par l'observateur car le bruit visuel « ronge », en quelque sorte, les contours du carré. C'est une information qui n'est pas disponible pour l'observateur dans la tâche de détection, et par conséquent, il est normal que cette dimension ne se retrouve pas dans l'image de classification.

Cette critique suppose donc que l'information de contour ne serait pas disponible pour l'observateur. L'objectif de cette expérience était de déterminer que si les contours sont perçus, alors cette dimension peut être utilisée par l'observateur pour répondre. Pour cela, la consigne a été manipulée : l'observateur devait répondre si le signal était présent uniquement lorsqu'il percevait les contours. Nous utilisons dans cette expérience une situation qui n'est pas du tout une stratégie naturelle pour les observateurs.

3.1. Expérience de détermination de seuil

3.1.1. Introduction

De la même façon que pour l'ensemble de nos expériences, un seuil de contraste du signal a été déterminé pour chaque observateur. Ce seuil sera utilisé dans l'expérience principale. Nous faisons l'hypothèse que le seuil de contraste du signal soit plus élevé que pour l'expérience princeps puisque les observateurs doivent tenir compte d'une consigne spécifique qui n'est pas du tout économique pour eux.

3.1.2. Méthode expérimentale

- Participants.

Cette expérience a été effectuée par deux observateurs (FD et SG).

- Stimuli.

Les stimuli, affichés sur un fond gris moyen (8 cd/m^2), étaient composés d'un signal et d'un bruit superposés : le signal était un carré de $3 \times 3^\circ$ d'angle visuel, le bruit était une distribution de luminance aléatoire de $10 \times 10^\circ$ d'angle visuel. Un seul contraste du bruit

(0.7) était utilisé. Le contraste du signal était distribué sur une échelle de $10^{0.1}$, le contraste du signal variait à partir de la valeur de 0.06 jusqu'à 0.1897 pour l'observateur FD et de 0.09 à 0.2846 pour SG. Le contraste du signal était ensuite déterminé pour un d' égal à 1.

Procédure.

La procédure était identique aux expériences de détermination de seuil que nous avons réalisé pour l'ensemble des tâches de détection. Une description détaillée est donnée dans la première série d'expérience. La seule différence de procédure concernait la consigne. En effet, l'observateur donnait sa réponse de détection seulement lorsque les contours du carré étaient perçus.

3.1.3. Résultats

Les résultats de l'expérience préliminaire sont présentés à la Figure 69. Les performances, représentées par l'indice d' , sont affichées pour chaque contraste du signal, et une droite de régression a été calculée afin de déterminer le contraste du signal qui était égal à un d' égal à 1. Conformément à notre hypothèse, le seuil du contraste du signal (0.08) était plus élevé que celui utilisé par le même observateur dans l'expérience princeps (0.06). De plus, le seuil de SG (0.12) était supérieur à celui des observateurs naïfs (entre 0.06 et 0.07) de la première expérience.

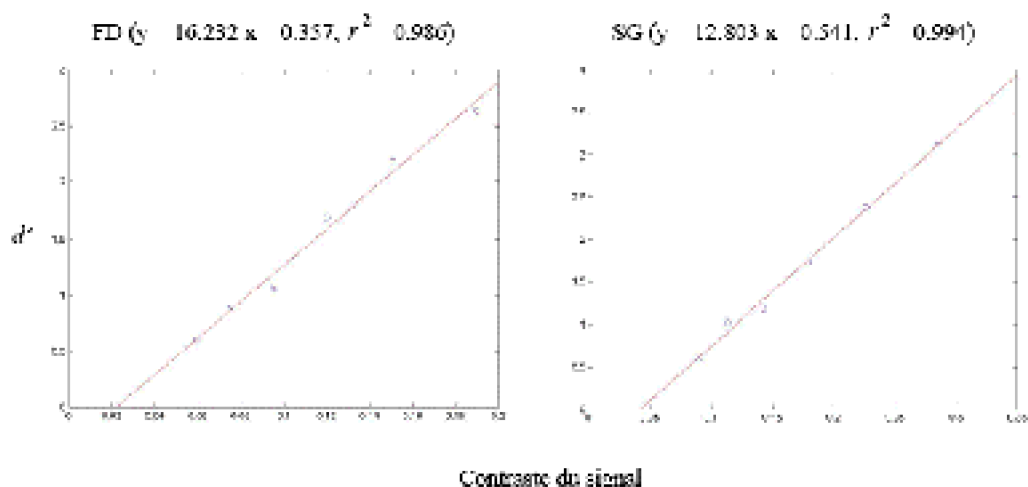


Figure 69 : Expérience préliminaire 5 : Droite de régression entre l'indice d' et le contraste du signal de l'observateurs FD et SG pour l'expérience de changement de consigne.

Les courbes ROC ont ensuite été affichées, les résultats sont indiqués dans l'Annexe 6 et sont proches pour les deux observateurs. De façon similaire aux précédentes expériences, les courbes étaient de plus en plus creusées en direction du coin supérieur gauche indiquant que l'indice d' augmente avec l'augmentation du contraste du signal.

3.2. Expérience de détermination de l'image de classification pour la condition du changement de consigne

3.2.1. Introduction

L'objectif de cette expérience était de déterminer si les contours d'une forme pouvaient quand même être utilisés par l'observateur dans une tâche de détection. Pour cela, nous avons obligé l'observateur à donner sa réponse de détection à partir des contours de la forme visuelle. La manipulation de la consigne obligeait ainsi l'observateur à ne pas utiliser la stratégie utilisée au cours des expériences précédentes qui consistait à donner une réponse de détection basée sur la surface. Ainsi, cette expérience nous permettra de répondre à une critique selon laquelle les contours ne sont pas disponibles lorsqu'ils sont présentés dans un bruit. Si l'observateur est capable d'utiliser les contours pour répondre alors cette dimension devrait émerger dans l'image de classification. De plus, il est possible que l'observateur utilise en plus l'information de surface pour donner sa réponse de détection. Dans ce cas, les dimensions de contour et de surface devraient apparaître dans l'image de classification. Ces deux hypothèses ont été présentées par la Figure 46b.

3.2.2. Méthode expérimentale

- Participants.

Les observateurs de l'expérience préliminaire (FD et SG) ont effectué l'expérience de détermination de l'image de classification. Seul FD était au courant des hypothèses formulées.

- Stimuli.

Un contraste seuil du signal, individuellement déterminé (0.08 pour FD et 0.12 pour SG), était utilisé. Ce contraste du signal était superposé à un bruit uniforme semblable à celui de l'expérience princeps. Les tailles du bruit et du carré étaient également identiques.

- Procédure.

La procédure était similaire à l'expérience princeps, excepté le nombre d'essai effectué par l'observateur. En effet, FD passait 25 blocs et SG n'a réalisé que 11 blocs. Les blocs étaient composés de 200 essais, le signal était présent dans la moitié des cas.

3.2.3. Résultats

Les résultats de l'observateur FD et SG sont présentés à la Figure 70. L'image de classification montrait que les contours du carré apparaissaient visiblement. De plus, une zone sombre se situait à l'intérieur et à l'extérieur de la zone claire. Ainsi, la réponse de l'observateur était facilitée lorsque le pourtour immédiat et le centre du carré étaient sombres. Ce résultat confirme notre hypothèse selon laquelle les contours du signal sont utilisés par l'observateur au cours de la tâche. Ainsi, notre expérience permet de montrer que l'observateur peut utiliser les contours du signal pour répondre mais pas de façon automatique.

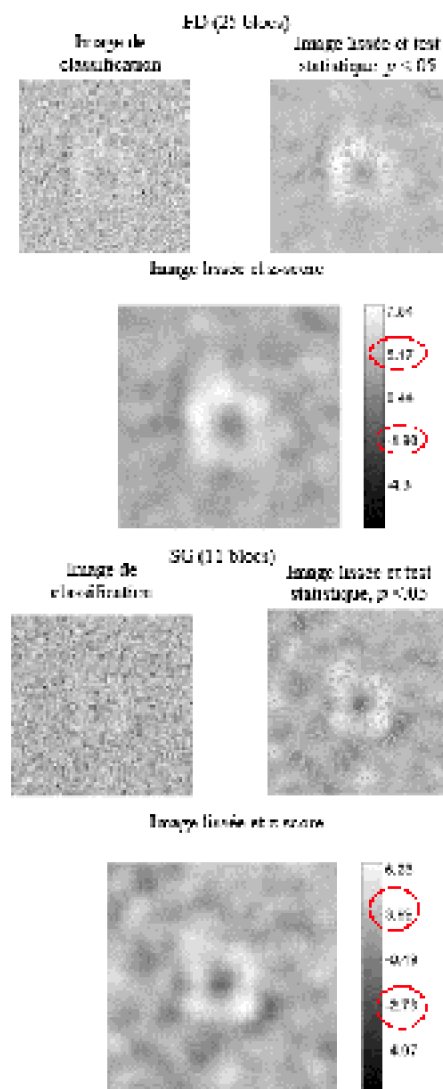


Figure 70 : Expérience 5 : Présentation de l'image de classification des observateurs FD et SG, du test statistique et normalisé effectués sur l'image de classification. Les valeurs minimales et maximales du z-score sont entourées sur la barre qui est présentée à la droite de l'image.

Cette expérience utilisait une situation artificielle dans laquelle les conditions de présentation étaient différentes de celles des expériences précédentes, puisque les niveaux luminance étaient supérieurs. Malgré cette augmentation de la luminance du signal, la sensibilité globale de FD au cours de l'expérience (exprimé par l'indice d') était inférieure à celle des précédentes expériences (0.74). A l'inverse, le d' calculé sur l'ensemble des essais pour SG était légèrement supérieur à 1 (1.35).

Nous avons calculé l'indice d' pour chaque blocs expérimentaux. Les résultats sont présentés à la Figure 71. Les deux observateurs avaient des résultats différents, SG avait une stratégie conforme aux conditions déterminées par l'expérience préliminaire, contrairement à FD. En effet, l'indice d' , calculé tous les 200 essais, semblait indiquer que l'observateur FD a changé de critère de décision entre le début et la fin de l'expérience.

En effet, au début de l'expérience, l'indice d' était supérieur à 1. Puis, à la fin de l'expérience, l'indice d' était proche de 0. Ce changement se manifestait principalement à partir du 13^{ème} bloc. Cette variation de l'indice d' au cours de l'expérience ne s'observait pas pour SG. Par ailleurs, le critère de décision variait constamment pendant l'expérience pour FD, ce qui n'était pas le cas pour SG. Par conséquent, les images de classifications ont été recalculées avant et après le 13^{ème} bloc pour FD. Les images sont présentées à la Figure 74. Les résultats montrent que les images après le 13^{ème} bloc étaient similaires à l'image de classification calculée sur l'ensemble des essais. Par contre, l'image de classification, calculée pour les 13 premiers blocs, indiquait que la réponse de l'observateur était basée sur la surface du carré. Il semble également que les contours sont également utilisés par l'observateur.

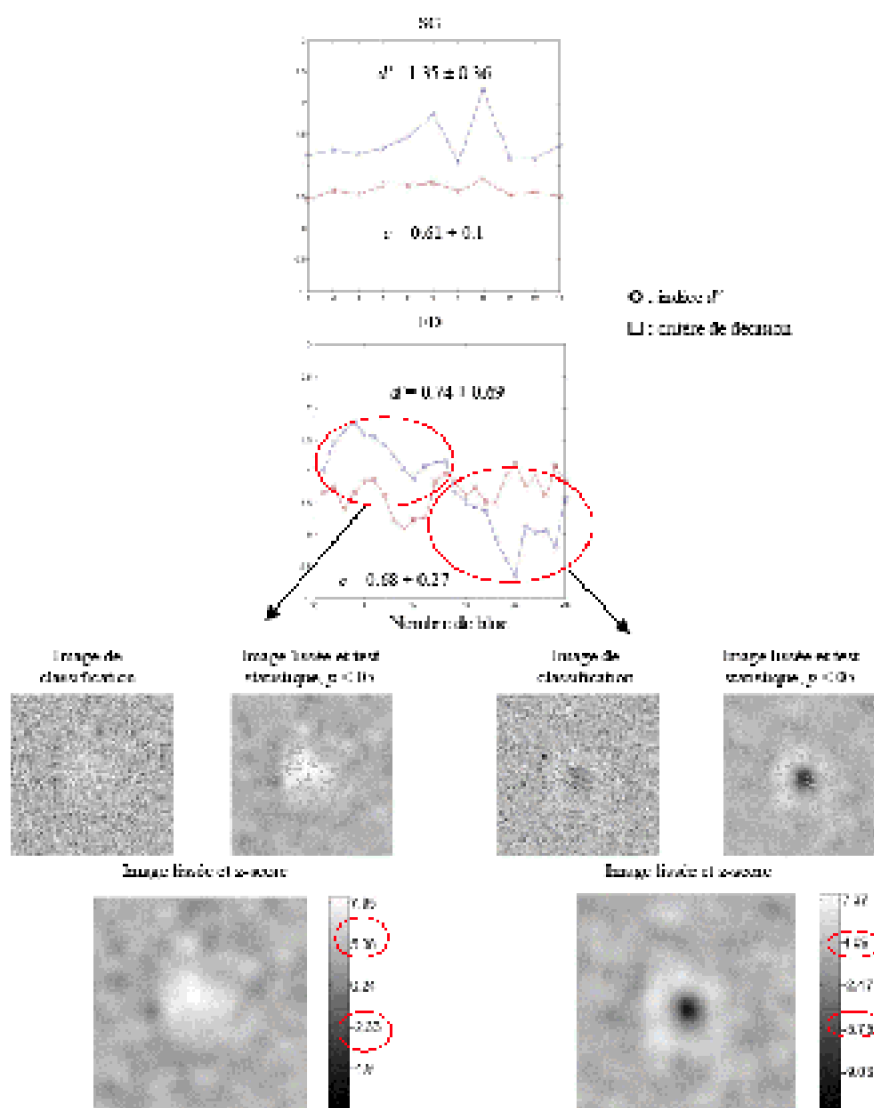


Figure 71 : Expérience 5 : L'indice d' et le critère de décision sont calculés tous les 200 essais pour FD et SG (graphiques du haut). Les Images de classification de l'observateur FD réalisées avant et après le 13^{ème} bloc, le test statistique et le test « z-score » effectués sur l'image de classification, sont présentées en bas. Les valeurs entourées sur la bande

représente les valeurs maximales et minimales des matrices.

Le calcul du critère de décision indique que celui-ci est stable pour l'observateur SG tout au long de l'expérience, par contre le critère de FD varie à travers les essais (Figure 74). Le critère calculé sur l'ensemble des essais était de 0.61 pour SG et de 0.68 pour FD.

Le fait que l'on obtienne un d' de 0 indiquait certainement que FD avait donné autant de réponses de type fausse alarme que de réponses de type détection correcte, c'est-à-dire qu'il ne voyait pas le signal pour les derniers blocs. La fréquence des réponses de type fausse alarme et détection correcte a été relevée et présentée à la Figure 72. Les résultats indiquaient qu'à partir du 11^{ème} bloc le nombre de détection correcte chute brutalement et à partir du 18^{ème} bloc, le taux de cette dernière catégorie de réponse est équivalente à celle des fausses alarmes. L'observateur utilisait ici une stratégie non-optimale qui ne lui permet pas de réussir la tâche.

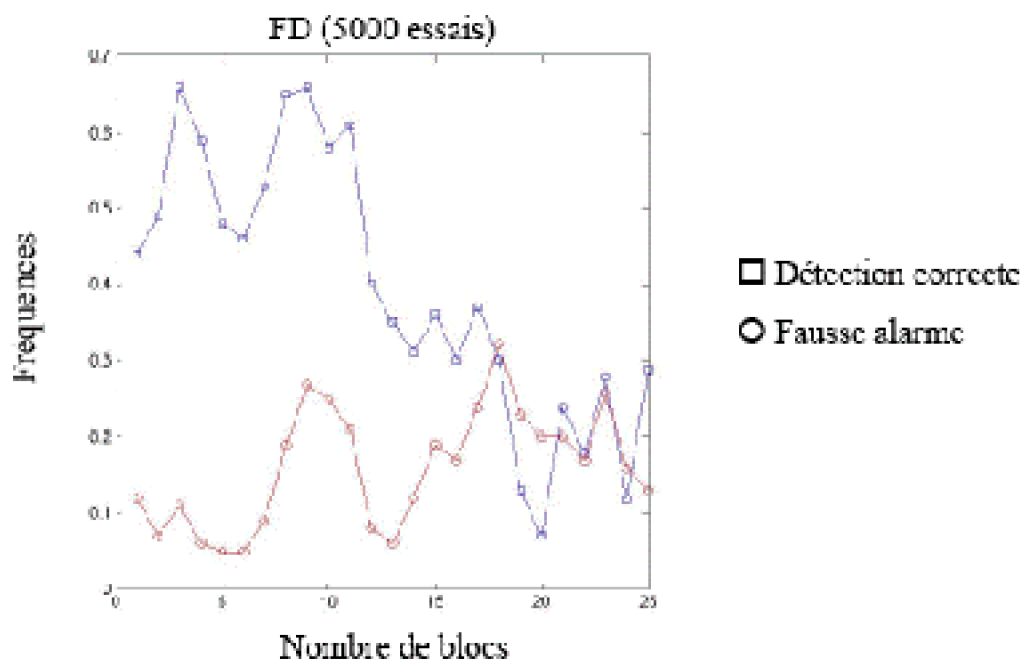


Figure 72 : Expérience 5 : Résultats des fréquences correspondant aux réponses de types détection correcte et fausse alarme pour chaque blocs.

Nous avons également réalisé un calcul de l'image de classification spécifique aux essais où le signal est présent et absent. Ce calcul a déjà été effectué pour les expériences précédentes et ne montrait que de très faibles variations entre les deux conditions. Les résultats sont présentés à la Figure 73. Les images montrent que la surface apparaissait faiblement pour les essais de type signal présent et négativement pour les essais de type signal absent. Ce résultat indique que l'observateur FD réalise mieux la tâche lorsque le signal est absent, de plus la surface est traitée automatiquement lorsqu'elle est présente réellement. La séparation des deux conditions a permis également de montrer un second résultat. En effet, les contours semblaient être centrés autour du signal réel pour les essais de type signal présent. A l'inverse, pour les essais de type signal absent, cette dimension semblait s'étendre au-delà du signal réel. Nous avons également réalisé des analyses concernant les images de classification où le signal est

présent ou absent, mais en distinguant cette fois les treize premiers blocs et les douze derniers blocs. Les images que nous avons obtenues étaient proches de la Figure 76, ces résultats ne sont pas présentés dans le présent manuscrit. Pour SG, les bords du carré émergent dans les deux conditions, toutefois, les images de type signal présent sont plus proches de l'image de classification présentée à la Figure 70.

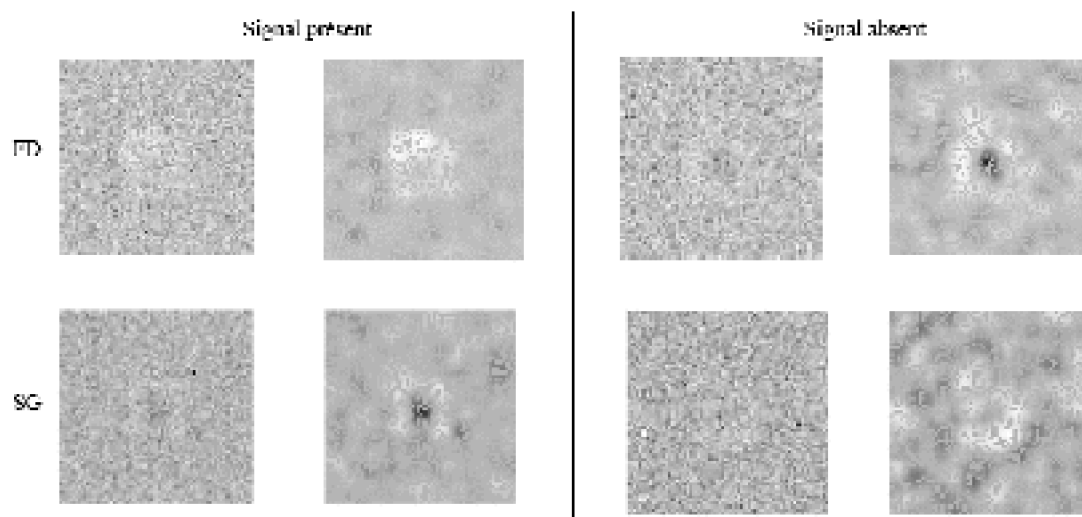


Figure 73 : Expérience 5 : Résultats des images de classification mais aussi de l'image lissée et du test statistique de FD et SG pour l'expérience portant sur le changement de la consigne. Les essais où le signal est présent ou absent sont distingués.

3.2.4. Discussion

Cette expérience permet de répondre à la critique formulée par Spillman et Kurtenbach (1992). En effet, pour ces auteurs, les contours d'une forme, placés dans un fond bruité, n'apparaissent pas nettement. Bien que les auteurs n'y font pas référence, cette idée est confortée par l'existence de l'inhibition intra-canal dans le modèle de Breitmeyer (1984) (voir partie théorique). Ceci aurait pu expliquer l'absence de contour dans les expériences précédentes. Or, cette nouvelle expérience permet de rejeter cette hypothèse. En effet, les résultats indiquaient que les contours sont disponibles, puisque l'observateur peut utiliser spécifiquement cette information pour répondre. Par contre, cette dimension n'est pas traitée de façon automatique et est coûteuse. Ainsi, les contours sont utilisés si l'attention de l'observateur est dirigée vers ces derniers comme critère de réponse.

Les résultats de nos deux premières séries d'expériences ont montré que la surface était une dimension prépondérante. Le résultat que nous obtenons dans cette dernière expérience, un peu paradoxalement, est un argument fort en faveur de l'importance de la dimension de surface. En effet, bien que la dimension de contour puisse être utilisée par les observateurs, elle n'est pas donnée de façon automatique.

Nos premières expériences n'ont pas permis de montrer l'importance des bords pour la détection. Toutefois, nous pouvons supposer que dans la détection ce traitement ne s'applique pas, et il se pourrait très bien qu'un traitement de type « coarse-to-fine » (ou du global au local) permette d'expliquer nos résultats (Ginsburg, 1986 ; Navon, 1977). Cette

hypothèse peut se résumer de la façon suivante : il est certainement plus efficace pour le système visuel d'encoder une description globale de l'image (donc imprécise) avant de faire une description détaillée de l'image. Cette idée est sous-tendue par l'existence des fréquences spatiales basses (« coarse ») et des fréquences spatiales élevées (« fine ») (DeValois & DeValois, 1988). Par ailleurs, cette hypothèse a également un corrélat physiologique puisque nous avons vu, au cours de la partie théorique, que les voies M et P répondaient de façon spécifique à ces deux canaux de fréquences spatiales respectivement. Or, la voie M est activée 20 msec avant la voie P. Ce traitement neuronal de l'information permettrait de valider l'hypothèse « coarse-to-fine ». Ainsi, nous avons pu mettre en évidence un avantage de la surface qui serait un traitement global de la forme, et notre dernière expérience a permis de montrer que les bords sont utilisés si l'attention de l'observateur se focalisait sur cette dimension, ce qui pourrait être interprété comme un traitement plus local.

4.Variation de la tâche : discrimination de luminance

Les séries d'expériences précédentes ont permis de montrer que la dimension de surface est prépondérante pour la détection. Cette nouvelle partie aborde la question des mécanismes neuronaux qui pourrait véhiculer l'information de surface, sans remplissage. Il faut que dès la rétine, des cellules répondent spécifiquement à cette dimension. Dans la rétine, le traitement de l'information visuelle est divisé en deux principales voies : M et P. Ces deux systèmes ont des propriétés distinctes et complémentaires. En effet, les cellules M répondent à des fréquences spatiales basses et les cellules de la voie P à des fréquences spatiales élevées. La voie P permet ainsi de transmettre une information basée sur les détails de l'image comme les contours (Kaplan, 2003). Notre objectif était de pouvoir différencier ces deux voies d'un point de vue comportemental.

Au cours de la partie théorique, les travaux de Pokorny et Smith (1997) ont été présentés (p. 41). L'objectif de leur étude était de distinguer les voies M et P de façon comportementale, ce qui a été rendu possible à partir des données physiologiques de Kaplan et Shapley (1986). Ces deux voies se distinguent par les réponses neuronales pour plusieurs niveaux de contrastes. Ceci a conduit Pokorny et Smith (1997) à utiliser trois paradigmes pour différencier la réponse des cellules M et P chez l'observateur humain. Nous nous limiterons à l'utilisation de la condition « steady-pedestal » qui permettait de faire émerger sélectivement les réponses issues de la voie M. L'originalité de cette nouvelle série d'expériences est d'utiliser la condition « steady-pedestal » avec la technique de l'image de classification. Toutefois, nous simplifierons ici la procédure expérimentale afin de permettre une utilisation des deux techniques simultanément, mais aussi pour comparer les résultats obtenus avec ceux des expériences précédentes. En effet, nous réduisons le nombre de stimuli à un seul au lieu de quatre dans les expériences de Pokorny et Smith (1997).

Cette partie se divise en deux expériences, tout d'abord une expérience de seuil qui permettra de déterminer les conditions expérimentales à utiliser pour l'expérience

principale où nous utilisons la technique de l'image de classification.

4.1. Expérience de détermination de seuil

4.1.1. Introduction

L'objectif est de déterminer deux seuils de contraste du signal : pour un contraste du signal qui est inférieur au signal de référence et pour un contraste du signal supérieur au signal de référence. Ces deux seuils seront par la suite utilisés dans l'expérience de détermination de l'image de classification.

4.1.2. Méthode expérimentale

- Participants

Deux observateurs (FD et NO) ont participé à l'expérience. Leur vue était normale ou corrigée à la normale.

- Stimuli

Les stimuli étaient composés d'un signal (un carré de 3 degrés d'angle visuel de côté) superposé à un bruit uniforme (10 degrés d'angle visuel de côté) composé de petit carrés (2*2 pixels) dont la luminance variait aléatoirement. Un seul niveau de contraste du bruit a été utilisé (0.7). Le signal était composé de 6 niveaux de luminance qui se répartissaient selon deux échelles logarithmiques ($10^{0.038}$ et $10^{-0.038}$ pour FD, ainsi que $10^{0.1}$ et $10^{-0.4}$ pour NO) à partir d'une valeur fixe de luminance d'un premier carré qui servait de pédestal pour la discrimination (0.2 pour FD et 0.1 pour NO). Ce pédestal était présenté sur un fond gris moyen au début de l'essai. Puis, la luminance du signal superposé au bruit était supérieure ou inférieure au pedestal. Ensuite, le premier carré de référence était affiché de nouveau. Les échelles logarithmiques ainsi que la valeur de luminance du carré présenté seul étaient déterminées par des essais préliminaires.

- Procédure

La séquence d'un essai se déroulait de la façon suivante : un carré seul était présenté au centre de l'écran pendant 400 ms, il était immédiatement suivi du stimulus pendant 100 ms, puis un carré seul apparaissait à nouveau pendant 400 ms. Les observateurs donnaient leur réponse après la présentation du dernier carré, un nouvel essai était alors lancé. La tâche des observateurs était de dire si la luminance du signal (superposé au bruit) augmentait ou diminuait par rapport au signal de référence (le carré qui est présenté au début et à la fin de l'expérience). Le carré de référence était présenté une seconde fois afin de faciliter la tâche. La réponse des observateurs était donnée par les touches 0 (plus foncé) et 1 (plus clair) du pavé numérique. La séquence d'un essai est présentée dans la Figure 74.

Les observateurs passaient 4 blocs expérimentaux. Chaque bloc comportait 180

essais, les 6 niveaux de contraste du signal étaient ainsi répétés 30 fois. Les niveaux de luminance du signal étaient tirés aléatoirement tous les 6 essais. Deux pauses étaient proposées aux observateurs pendant l'expérience.

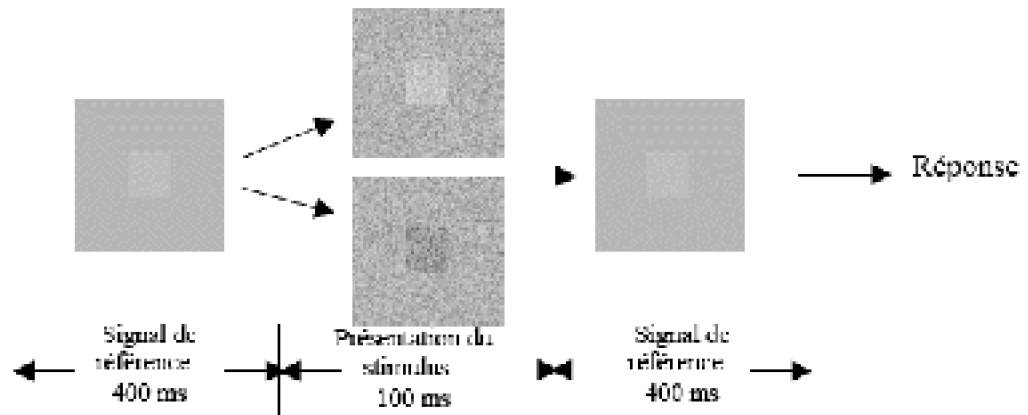


Figure 74 : Diagramme de la séquence d'un essai de l'expérience de détermination de seuil pour la tâche de discrimination. Les contrastes du signal ont été rehaussés pour des raisons d'impression.

4.1.3. Résultats

Dans cette expérience, un seuil a été déterminé pour un contraste du signal supérieur au signal de référence, puis un autre seuil a été calculé pour un contraste du signal inférieur au signal de référence. La méthode et les valeurs de référence utilisées pour calculer ces deux seuils étaient issues de MacMillan et Creelman (1991).

La méthode utilisée était la suivante : un pourcentage de réponse « plus clair » pour les six contrastes du signal a été calculé. Ces pourcentages sont ensuite transformés en z afin de déterminer le point d'égalité subjectif (ou PSE : « Point of Subjective Equality »). Une droite de régression est ensuite ajustée à ces données, et le PSE est défini pour une valeur de $z = 0$. Le calcul du PSE permet de déterminer si celui-ci est identique au contraste du signal de référence. S'il y avait une différence entre le signal de référence et le PSE, c'est ce dernier qui serait employé comme nouveau signal de référence dans l'expérience de détermination de l'image de classification. Par la suite, un d' , pour chaque contraste du signal, est déterminé à partir de la soustraction entre les z des deux premiers stimuli, puis des deux suivant, etc ... Les d' cumulés sont ensuite calculés pour connaître le niveau de sensibilité du PSE ainsi que les deux contrastes du signal à utiliser dans l'expérience de détermination de l'image de classification. Une droite de régression est ensuite ajustée afin de connaître la valeur de d' correspondant au PSE. Les deux niveaux de contrastes du signal sont ensuite définis à partir de la sensibilité du PSE en rajoutant ou en enlevant 0.5. Les deux valeurs obtenues permettent de juste distinguer une augmentation ou une diminution du contraste du signal par rapport au PSE (communément appelé « just noticeable difference » ou JND).

Les résultats ont montré que le PSE calculé pour un $z = 0$ était de 0.2017 pour FD et de 0.1063 pour NO. Le PSE correspondait à un d' de 1.45 pour FD et de 1.97 pour NO.

Le signal plus clair à utiliser, pour l'expérience de détermination de l'image de classification, correspond à un d' de 1.95 pour FD et de 2.47 pour NO, et est égal à un contraste du signal de 0.2173 et de 0.1240 respectivement. De même, le signal plus foncé à utiliser correspond à un d' de 0.95 pour FD et de 1.47 pour NO, et le contraste du signal utilisé pour FD était de 0.1859 et de 0.072 pour NO. Les résultats du z sont indiqués dans la Figure 75, et ceux du d' cumulé dans la Figure 76.

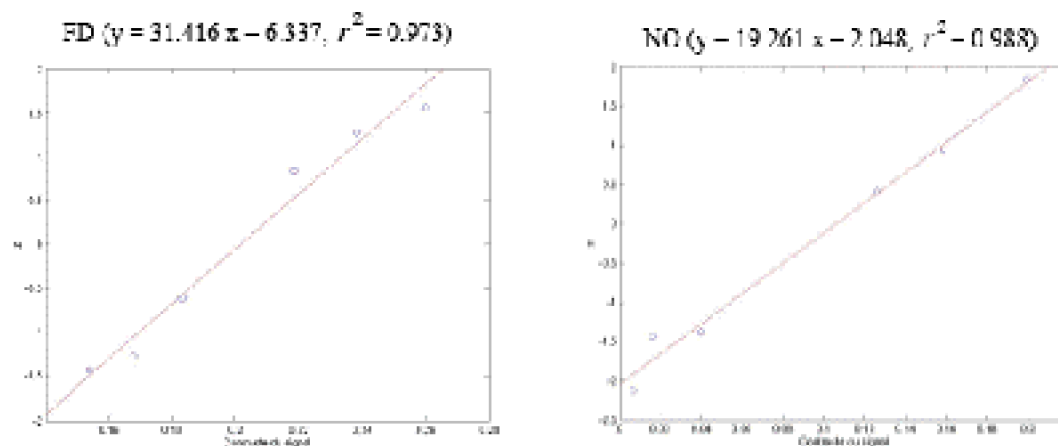


Figure 75 : Expérience préliminaire 6 : Droite de régression entre les valeurs du z et chaque contraste du signal pour les observateurs FD et NO.

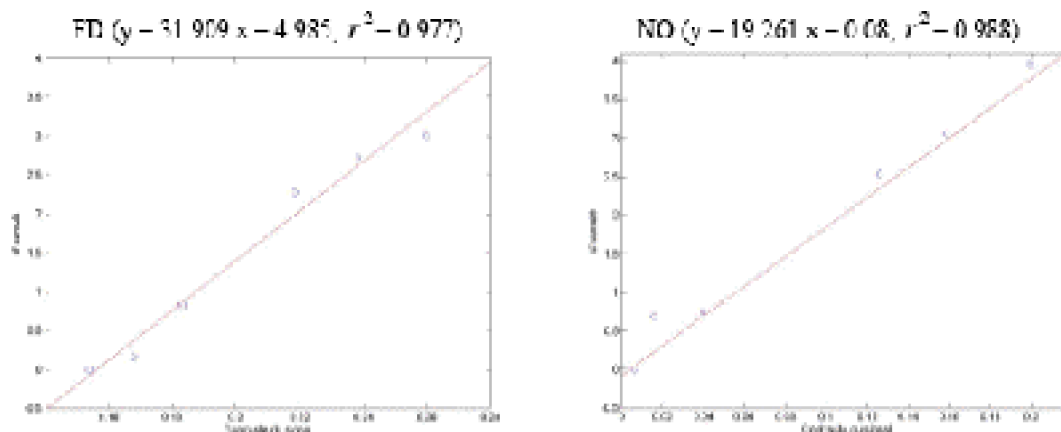


Figure 76 : Expérience préliminaire 6 : Droite de régression entre les valeurs de d' cumulé et chaque contraste du signal pour les observateurs FD et NO.

4.2. Expérience de détermination de l'image de classification pour une tâche de discrimination

4.2.1. Introduction

L'objectif de cette expérience était de déterminer la participation de la voie M dans le traitement de la surface, d'un point de vue comportemental.

Comme nous l'avons présenté, les travaux de Smith et Pokorny (1997) ont permis

d'établir deux paradigmes qui dissocient de façon comportementale les voies M et P. La distinction est rendue possible à partir de la différence de sensibilité au contraste des deux voies. Dans le premier paradigme (« pulse paradigm »), quatre carrés apparaissaient brièvement sur un fond gris moyen. Dans le second paradigme (« steady-pedestal paradigm »), les mêmes stimuli étaient présentés mais cette fois de façon continue. Dans les deux conditions, la luminance de l'un des carrés variait, la tâche de l'observateur était de discriminer le carré dont la luminance était différente des trois autres. Le seuil de luminance (le plus petit changement de luminance perçu) était relevé pour chaque observateur et pour différents niveaux de luminance de base des carrés. A partir de ces résultats une comparaison était ensuite réalisée avec les données neurophysiologiques de Kaplan et Shapley (1986) qui différençaient la réponse des cellules M et P en fonction du gain de contraste (cf. Figure 5). Ainsi, Pokorny et Smith (1997) ont montré que la condition « pulse paradigm » était spécifique à la voie P, tandis que la condition « steady-pedestal paradigm » était sensible à la voie M.

La technique de Smith et Pokorny (1997) a été adaptée pour être utilisée avec la technique de l'image de classification. Dans notre expérience, la procédure a été simplifiée puisque nous utilisons une tâche de discrimination simple de luminance. La tâche de l'observateur était de déterminer si la luminance du carré était plus claire ou plus foncée par rapport à un carré de référence. La simplification de la procédure nous permettra de comparer ces résultats avec ceux obtenus dans les expériences précédentes.

Nous proposons que la surface soit véhiculée plus particulièrement au système M car cette voie répond aux fréquences spatiales basses, et que la surface constitue une information globale. Cette information devrait donc être présente dans l'image de classification dans cette nouvelle tâche de discrimination de luminance.

Par ailleurs, Neri et Heeger (2002) ont montré une distinction entre les tâches de détection et de discrimination en utilisant le paradigme de l'image de classification perceptive. Leur expérience a permis de montrer l'existence de deux étapes liées spécifiquement à la détection et à l'identification. Pour la détection, les images basées sur le calcul de la variance indiquaient un pic positif 100 ms avant l'apparition du signal. Par ailleurs, les images de classification calculées pour la tâche de discrimination montraient une zone claire à l'emplacement de la cible. Cette différence entre les deux types de tâches a également été démontrée par de nombreux auteurs dans un paradigme de masquage (Nothdurft, 2002 ; Ziebell & Nothdurft, 1999). Il est donc tout à fait possible que des stratégies différentes soient mises en place pour une tâche de discrimination. Dans ce cas, si la tâche de discrimination permet une appréciation plus fine des composantes du signal, alors les contours de la cible devraient émerger dans notre image de classification.

L'image de classification est calculée de la façon suivante : les bruits tirés aléatoirement à chaque essai sont triés en fonction des réponses, les quatre types de réponses possibles sont présentés dans le Tableau XIV.

Tableau XIV : Représentation des 4 types de réponses possibles dans une tâche de discrimination. « 0 » correspond à un stimulus ou une réponse plus claire, et « 1 » à un stimulus ou une réponse plus foncée.

		Réponse	
		Plus clair	plus foncé
Stimulus	Plus clair	S0R0	S0R1
	Plus foncé	S1R0	S1R1

Une moyenne pondérée des bruits est alors calculée pour chacune de ces réponses, et l'image de classification est obtenue de la façon suivante (Ahumada, 1996) :

$$\text{Image de classification} = (\overline{S0R1} - \overline{S0R0}) + (\overline{S1R1} - \overline{S1R0})$$

4.2.2. Méthode expérimentale

- Participants.

Après avoir participé à l'expérience de détermination de seuil, FD et NO ont réalisé l'expérience de détermination de l'image de classification. Leur vue était normale ou corrigée à la normale.

- Stimuli.

Les stimuli étaient les mêmes que dans l'expérience précédente, mais ici un seul contraste du signal, supérieur au signal de référence, était utilisé, et un seul contraste du signal, inférieur au signal de référence, était employé. Ces deux contrastes du signal, spécifiques à chaque observateur, étaient déterminés à partir des résultats de l'expérience de détermination de seuil. Les contrastes du signal de référence étaient les suivants : 0.20 pour FD et 0.10 pour NO. De plus, un seul contraste du signal, supérieur au signal de référence, était utilisé (0.22 pour FD et 0.13 pour NO), et un seul contraste du signal, inférieur au signal de référence, était employé (0.19 pour FD et 0.08 pour NO).

- Procédure.

Une procédure proche à celle de l'expérience de détermination de seuil était utilisée. Les observateurs ont réalisé chacun 25 blocs de 200 essais.

4.2.3. Résultats

- La sensibilité.

La Figure 77 montre l'indice d' calculé tous les 200 essais. Les résultats montraient que l'indice d' était relativement bas tout au long de l'expérience pour l'observateur FD. L'indice d' calculé sur l'ensemble des essais indiquait que la sensibilité de FD était de 0.4919 et de 0.8659 pour NO. Le critère de décision a également été déterminé pour chaque bloc, cet indice est stable pendant la durée de l'expérience pour les deux

observateurs. Le calcul du critère sur l'ensemble de l'expérience était de -0.04 pour FD et -0.21 pour NO.

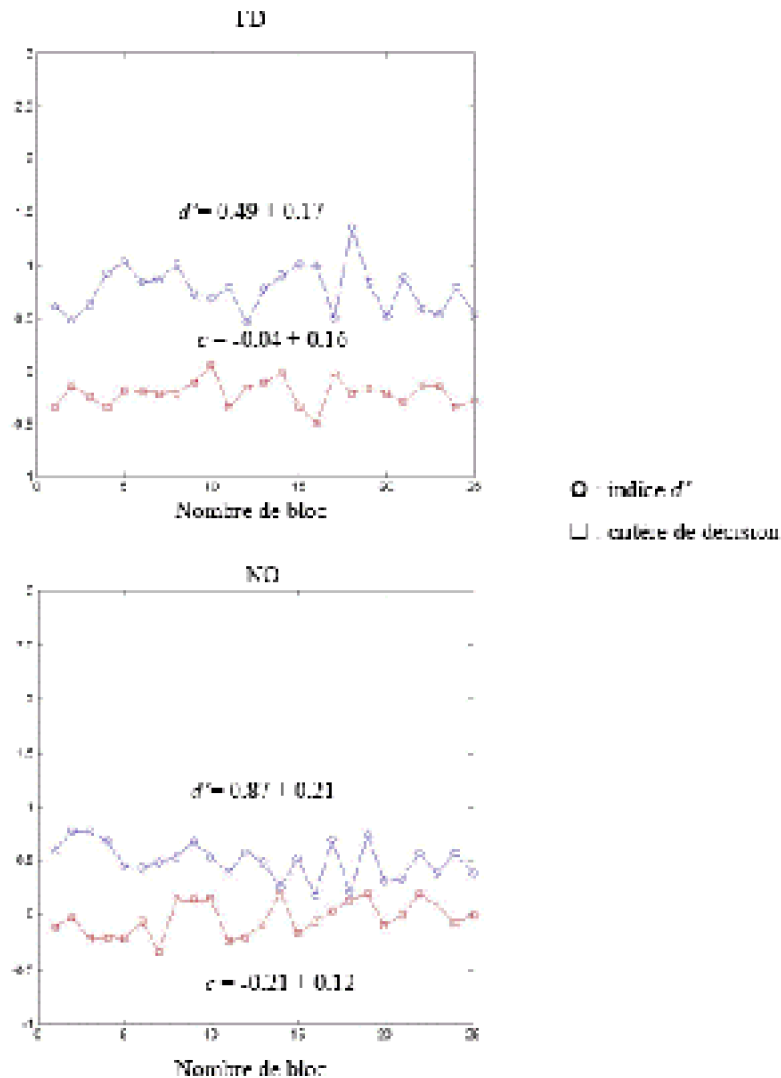


Figure 77 : Expérience 6 : Résultats de l'indice d' et du critère de décision calculé tous les 200 essais des observateurs FD et NO.

- Image de classification.

La Figure 78 indique les résultats obtenus par les observateurs FD et NO. Les images de classifications montrent deux principaux résultats : une zone centrale claire entourée par un pourtour sombre. Ces résultats sont identiques à ceux obtenus dans les expériences précédentes qui utilisaient une tâche de détection.

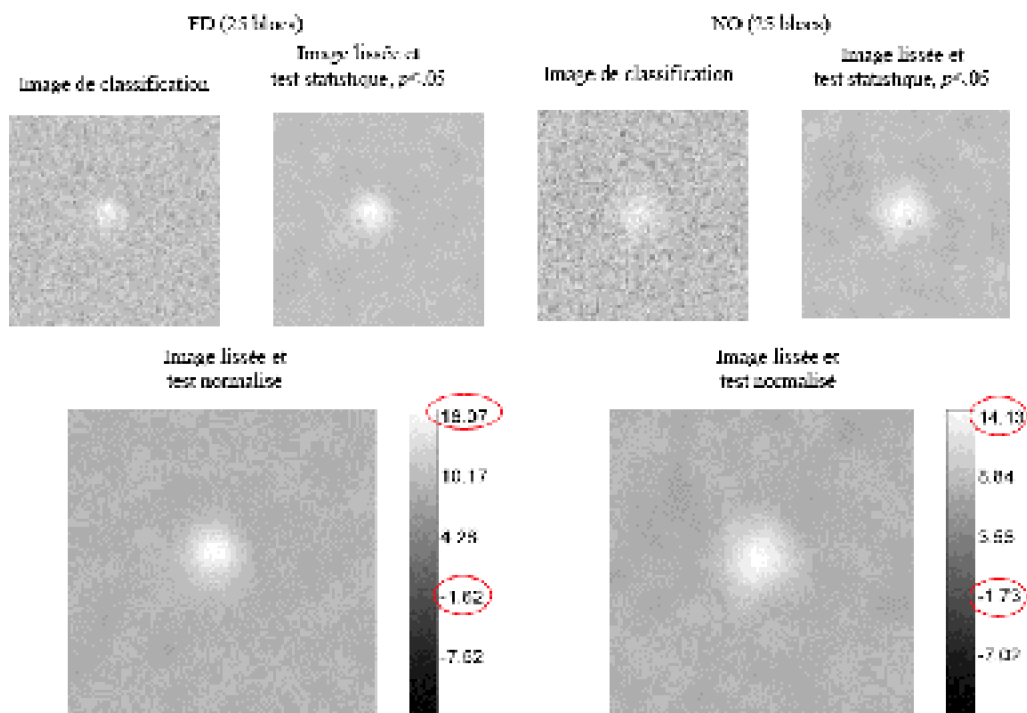


Figure 78 : Expérience 6 : Résultats de l'image de classification, du lissage et du test statistique effectué sur l'image de classification des observateurs FD (25 blocs) et NO (25 blocs) pour la tâche de discrimination. Pour le z-score, une bande est présentée sur la droite de l'image qui permet de faire correspondre un niveau de gris avec la valeur du pixel. De plus, les valeurs minimales et maximales ont été entourées pour chaque observateur.

Nous avons réalisé une estimation des paramètres de la fonction DoG, ajustée à nos données. Les résultats sont présentés dans le Tableau XV. Ces estimations indiquaient que la taille de la zone centrale était correctement ajustée à la taille du signal (colonne 1). De plus, le pourtour participait faiblement à la détection : de 0.2 à 3 % de la taille de la zone centrale (colonne 6). Par ailleurs, la taille du pourtour était approximativement 4 fois plus importante que la taille de la zone centrale (colonne 5). Les résultats étaient proches pour les deux observateurs excepté la participation de l'amplitude de l'observateur FD qui était extrêmement faible. Les paramètres de position indiquaient un faible biais pour la position de la zone centrale.

Tableau XV : Présentation des résultats des ajustements de la fonction DoG aux données des observateurs FD et NO. Les tailles de la zone centrale et du pourtour sont indiquées aux colonnes 1 et 3 respectivement. De même, les valeurs d’amplitude sont montrés aux colonnes 2 et 4 pour ces deux dimensions. Les ratios de la taille et de l’amplitude sont indiqués aux colonnes 5 et 6. Les colonnes 7 et 8 présentent une estimation de la position de la zone centrale.

Tableau XV : Présentation des résultats des ajustements de la fonction DoG aux données des observateurs FD et NO. Les tailles de la zone centrale et du pourtour sont indiquées aux colonnes 1 et 3 respectivement. De même, les valeurs d’amplitude sont montrés aux colonnes 2 et 4 pour ces deux dimensions. Les ratios de la taille et de l’amplitude sont indiqués aux colonnes 5 et 6. Les colonnes 7 et 8 présentent une estimation de la position de la zone centrale.							x_0 (min)	y_0 (min)
	Zone centrale		Pourtour		Rapport pourtour / zone centrale			
Sujet	Sigma (deg)	Amplitude	Sigma (deg)	Amplitude	Sigma	Amplitude		
FD	2.42	3.79	9.50	0.01	3.93	0.002	-15.98	-21.13
NO	3.1	2.76	11.68	0.09	3.77	0.03	-7.5	-6.93

4.2.4. Discussion

L’objectif de cette expérience était de déterminer de façon comportementale l’implication des voies précoces dans le traitement de contour et de la surface d’un signal. Pour cela, nous avons réalisé une tâche de discrimination, proche du paradigme de Pokorny et Smith (1997), mettant en jeu une participation de la voie M. Les résultats ont permis de montrer la présence d’une zone claire à l’emplacement du signal. Ainsi, cette expérience a permis de valider l’hypothèse selon laquelle si la voie M est responsable du traitement de la surface. Toutefois, il est possible que la modification du « steady-pedestal paradigm » implique un gain de contraste différent de celui mis en évidence dans l’étude Pokorny et Smith (1997). Il est donc nécessaire de réaliser une expérience a posteriori afin de vérifier la participation de la voie M dans notre expérience. De plus, une seconde série d’expériences complémentaires pourrait être réalisée où l’objectif serait de déterminer la participation de la voie P dans le traitement des dimensions de contour et de surface. Pour cela, l’utilisation du « pulse paradigm » nous paraît le plus approprié. Si les cellules de type P intègre l’information de contour, alors cette dimension devrait apparaître dans les images de classification.

5. Réponse de détection par un observateur modèle

5.1. Introduction

Les résultats expérimentaux qui ont été présentés au cours de la thèse indiquaient l'importance du traitement de la surface pour l'observateur. Nous voulons maintenant déterminer comment cette dimension pouvait être encodée par le système visuel.

La notion de champ récepteur classique de Hubel et Wiesel (1962, 1968) indique que la première information qui est encodée par le système visuel est centrée sur les contrastes d'une forme. Ainsi, de nombreux auteurs (Grossberg & Mingolla, 1985 ; Hubel, 1994 ; Marr, 1982) ont conclu que les contours étaient intégrés avant l'information de surface. La Figure 79 présente une convolution entre un champ récepteur et un carré, le résultat obtenu consiste en une image dans laquelle seule le contour du carré est présent et non pas sa surface.

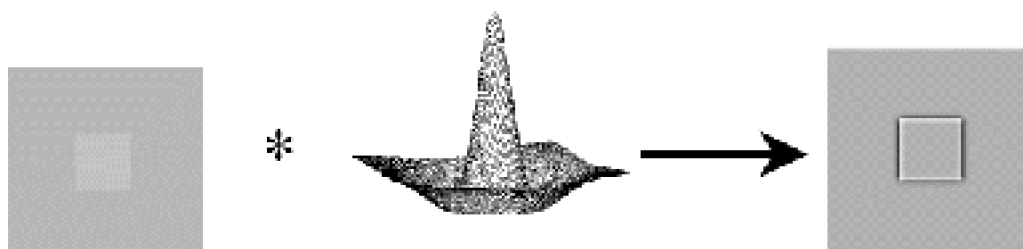


Figure 79 : Résultats de la convolution entre un carré et un champ récepteur, le résultat obtenu est une information centrée sur les bords du signal. Le champ récepteur est ici réalisé à partir d'une fonction DoG.

Les champs récepteurs des cellules ganglionnaires de la rétine permettent l'intégration d'un signal, il nous a semblé pertinent de modéliser cette intégration. Pour déterminer comment une cellule avec un antagonisme centre/pourtour (donc codant un contraste) pouvait aussi véhiculer l'information de surface. La forme des champs récepteurs de la rétine est identique, et elle est proche d'une fonction DoG (« Difference of Gaussians ») (Rodieck, 1998). Le profil d'un champ récepteur comporte une région centrale et un pourtour antagoniste. Notre modélisation a pris en compte ces deux caractéristiques des champs récepteurs. Ces propriétés sont essentiellement rattachées aux voies M et P. Tout d'abord, il est bien établi que la taille du centre et du pourtour est plus importante pour la voie M que la voie P (Lee, Kremers & Yeh, 1998). Toutefois, la simulation ne présente pas les tailles exactes de ces deux voies qui ne sont pas encore connues. Une seconde caractéristique de ces deux voies concerne le déséquilibre centre/pourtour des champs récepteur. En effet, certains auteurs (Croner & Kaplan, 1995 ; Dacey, Packer, Diller, Brainard, Peterson & Lee, 2000 ; Lee, Kremers & Yeh, 1998) ont montré que dans les voies M et P, l'amplitude du pourtour inhibiteur représente 50 à 80 % de la zone excitatrice centrale. Nous nous sommes donc placé dans des conditions expérimentales similaires. L'ensemble de ces paramètres va nous permettre de spécifier la forme et la taille des champs récepteurs des canaux visuels qui sont cohérents avec l'image réelle obtenue, c'est-à-dire qui sont susceptibles d'être impliqués dans le codage de l'information de surface.

Le principe de la simulation est le même que pour un observateur humain : on examinait à chaque essai si le bruit possédait suffisamment d'éléments pour conclure que le signal était présent ou absent (les quatre types de réponse : omission, fausse alarme,

rejet correct et détection correcte, sont toujours valables pour la simulation). De cette manière, des images de classification, proches de la voie M et de la voie P, étaient obtenues.

5.2. Expériences préliminaires

5.2.1. Introduction

Des essais préliminaires ont été effectués pour déterminer le contraste du signal à utiliser dans l'expérience de détermination de l'image de classification. L'objectif de cette expérience préliminaire était de déterminer un niveau de contraste du signal équivalent à un $d' = 1$. Ce niveau de contraste du signal a été estimé pour chacune des conditions expérimentales, à savoir les facteurs de taille, mais aussi de l'équilibre, ou non, du centre et du pourtour des champs récepteurs.

La modélisation a été réalisée sur une image unidimensionnelle dans un souci de simplification, mais les résultats peuvent être applicables et généralisables pour une image bi-dimensionnelle.

5.2.2. Méthode expérimentale

Stimuli :

Le signal est défini comme un vecteur de 1200 éléments où les 360 éléments du milieu sont fixés pour une valeur de 1 et les autres à 0. Les valeurs de 1 représentent ainsi le signal carré. De plus, le bruit est construit à partir de 60 éléments tirés aléatoirement, et fait partie d'une distribution uniforme. Les éléments sont répétés 20 fois successivement pour afficher un vecteur de la même longueur que le signal. Un stimulus est constitué soit du bruit tout seul soit du bruit et du signal. Si la taille du stimulus est de 10 deg, alors chaque élément correspond à une taille de 0.5 min (ou 120 éléments/deg). De même, les 60 éléments du bruit ont chacun une taille de 10 min, et le signal est de 3 deg.

Le profil des champs récepteurs est proche d'une fonction DoG (Roedieck, 1998), c'est pour cette raison que nous avons utilisé cette fonction pour notre modélisation. La fonction se présente sous une forme unidimensionnelle de la façon suivante :

$$DoG_{1D}(x) = A_c e^{-\frac{0.5(x-c)^2}{\sigma_c^2}} - A_p e^{-\frac{0.5(x-p)^2}{\sigma_p^2}}$$

Dans cette fonction, A représente l'amplitude et σ la taille de chaque fonction gaussienne spécifique au centre (c) et au pourtour (p). La distinction entre les deux systèmes se fait essentiellement sur la taille, nous avons manipulé l'écart type (σ) afin d'obtenir des petites ($\sigma_c = 1$ et $\sigma_p = 2$) ou des grandes tailles de champ récepteur ($\sigma_c = 10$ et $\sigma_p = 20$). De plus, nous avons choisi de manipuler l'équilibre du centre et du pourtour par le biais des paramètres d'amplitude de la fonction DoG. Lorsque le centre et le pourtour étaient équilibrés : le pourtour inhibiteur était identique à l'excitation de la zone

centrale. Tandis que, lorsque le centre et le pourtour étaient déséquilibrés : l'inhibition du pourtour était plus faible que l'excitation du centre. Pour cette dernière situation, deux conditions ont été effectuées, le pourtour inhibiteur pouvait représenter 80 % ou 50 % de l'excitation du centre. La réponse de l'observateur est donc testée sur six filtres. Par ailleurs, les écarts-types du centre sont déterminés pour obtenir deux tailles équivalentes à 1 et 10 min. De plus, le ratio du pourtour/centre était de 2, ainsi les tailles du pourtour étaient respectivement de 2 et 20 min.

Le contraste du bruit était identique à celui utilisé lors de nos expériences avec l'observateur humain (0.7). En revanche, le contraste du signal a été recalculé pour obtenir un niveau de performance qui correspondait à un indice $d' = 1$ (comme dans les expériences avec des observateurs humains).

· Procédure :

Le principe de la simulation était le suivant. Une première étape était de déterminer le prototype neuronal du carré, pour cela une convolution entre le carré est réalisée avec une fonction DoG. Puis, on filtre le carré présenté dans un bruit avec la fonction DoG, le bruit était tiré de façon aléatoire pour chaque essai. Cette dernière opération permettait d'obtenir une image neuronale du stimulus. Les paramètres des deux fonctions DoG étaient identiques pour le prototype neuronal et l'image neuronale. Les essais étaient partagés entre des stimuli contenant uniquement un bruit uniforme, et ceux où un carré est ajouté. Puis, un produit scalaire était effectué, à chaque essai, entre le prototype neuronal du carré et l'image neuronale du stimulus. Le produit scalaire nous permettait d'obtenir une réponse de détection à partir d'un critère de décision, si celui-ci était supérieur à 0 la réponse était « signal présent » et lorsque le critère était inférieur ou égal à 0 la réponse était « signal absent ». En effet, la réponse de l'observateur modèle est basée sur la similarité entre le prototype neuronal et l'image neuronale. La Figure 80 présente le schéma de réponse de l'observateur modèle.

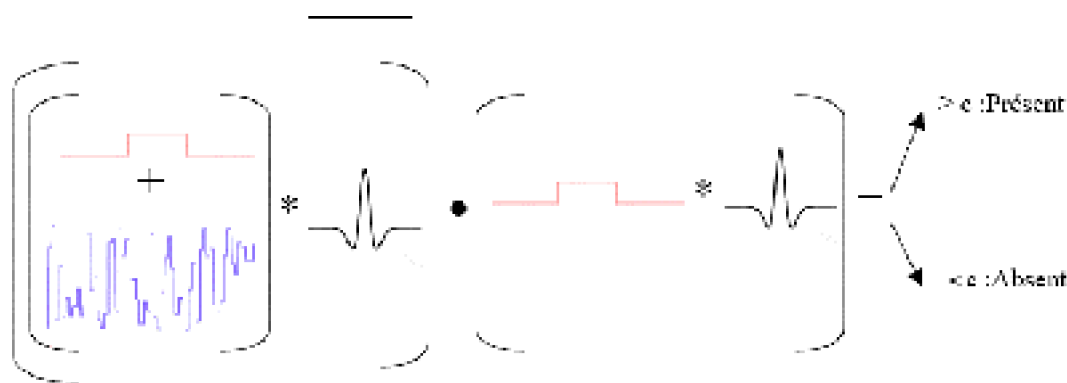


Figure 80 : Diagramme de la séquence d'un essai préliminaire de la simulation informatique. La partie droite de la figure indique l'opération réalisée pour obtenir un prototype neuronal, le résultat est toujours le même quels que soient les essais. La partie gauche de la figure présente l'opération qui est effectuée à chaque essai, le carré est présent uniquement dans la moitié des essais. Par la suite, la comparaison entre ces deux opérations (par l'intermédiaire d'un produit scalaire représenté par le point) permet

d'obtenir une réponse de type présent ou absent en fonction d'un critère supérieur ou inférieur à 0.

5.2.3. Résultats

L'ensemble des résultats sont présentés dans l'annexe 7 et les contrastes utilisés sont indiqués dans le Tableau XVI. Les résultats indiquaient que quelle que soit la condition expérimentale, les contrastes du signal qui correspondaient à un indice $d' = 1$ étaient supérieurs à ceux utilisés avec l'observateur humain. Ainsi, pour un centre et un pourtour équilibré, les contrastes du signal étaient 6 fois plus importants que chez l'observateur humain. Par contre, lorsque le centre et le pourtour de la fonction DoG étaient déséquilibrés, les contrastes du signal étaient 1,5 fois plus importants que les contrastes employés au cours de notre expérience princeps. Ces résultats étaient indépendants de la taille de la fonction DoG. Lorsque le centre et le pourtour étaient déséquilibrés, les résultats obtenus se rapprochaient des performances utilisées avec l'observateur humain. La différence de sensibilité au contraste peut être attribuée au fait que le système visuel humain somme les différents canaux de traitements, alors qu'au cours de notre simulation un seul canal de traitement a été modélisé.

Tableau XVI : Représentation des contrastes du signal correspondant à un indice $d' = 1$ pour chaque condition expérimentale.

Conditions		Contraste du signal
Centre/Pourtour équilibré	Petite taille	0.4
	Grande Taille	0.35
Centre/Pourtour déséquilibré ($A_p = 0.8 * A_c$)	Petite taille	0.1
	Grande Taille	0.1
Centre/Pourtour déséquilibré ($A_p = 0.5 * A_c$)	Petite taille	0.1
	Grande Taille	0.1

5.3. Détermination de l'image de classification de l'observateur modèle

5.3.1. Introduction

L'observateur modèle a été développé afin de déterminer la façon dont les filtres précoces du système visuel influencent les informations obtenues dans les images de classification. Pour répondre à cette question, les réponses de l'observateur modèle sont centrées sur les différents profils des champs récepteurs. Les images de classification seront ensuite calculées, ce qui permettra une comparaison avec les données de l'observateur humain. D'un point de vue physiologique, le profil des champs récepteurs est proche d'une fonction DoG (Rodieck, 1998). Par ailleurs, les champs récepteurs sont déterminés par

deux caractéristiques, à savoir leur taille et l'équilibre ou non de l'antagonisme centre/pourtour. Ainsi, la manipulation des paramètres de la fonction nous permettra de tester directement les facteurs qui caractérisent les champs récepteurs.

Les voies M et P se différencient par leur taille, en effet, les cellules M sont plus grandes que les cellules P. Cette différence anatomique conduit ces deux populations de cellules à répondre aux fréquences spatiales basses (voie M) ou élevées (voie P). Par conséquent, les petites tailles de champs récepteurs sont sensibles aux détails de l'image comme les contours, et les grandes tailles de champs récepteurs sont sensibles aux informations globales comme la surface. Nous faisons ainsi l'hypothèse que les images de classification comporte uniquement l'information de contour pour des petites tailles de champs récepteurs. Par contre, l'information de surface serait intégrée par grands champs récepteurs. Toutefois, les fréquences spatiales basses véhiculent également une information de contour qui est traité de façon floue, contrairement aux fréquences spatiales élevées où la transition est nette. Ainsi, dans cette condition, des contours devraient également être plus large dans l'image de classification. Les tailles que nous avons utilisées sont équivalentes à celles des voies M et P situées à plus ou moins 1.5 deg de la fovéa (Croner & Kaplan, 1995 ; Lee, Kremers & Yeh, 1998).

Nous avons également manipulé un second facteur à savoir l'équilibre centre/pourtour des champs récepteurs. Bien que classiquement, les profils des champs récepteurs sont souvent présentés de façon équilibrée (Marr, 1982), de nombreux auteurs (Croner & Kaplan, 1995 ; Lee, Kremers & Yeh, 1998) ont montré que l'amplitude du pourtour inhibiteur des cellules de la voie M et P est souvent plus faible que l'amplitude de la zone excitatrice. Ainsi, les données physiologiques ont déterminé que l'inhibition du pourtour représente de 50 à 80 % de l'amplitude de la zone centrale. L'utilisation de ce facteur est de tenir compte des contraintes physiologiques. Nous nous attendions à ce que si l'excitation de la zone centrale est plus importante, alors la réponse de la surface devrait être plus importante. Mais ceci n'est pas vrai pour les petites tailles de champs récepteurs, puisque ces derniers répondent préférentiellement à des fréquences spatiales élevées, par conséquent, l'image de classification devrait comporter une information essentiellement basée sur les contours.

De la même manière que dans les expériences préliminaires, le modèle est réalisé de façon unidimensionnelle afin de faciliter la vitesse de calcul. Mais les résultats pourraient être généralisés à une image bidimensionnelle.

5.3.2. Méthode expérimentale

- Stimuli.

Les stimuli étaient identiques à ceux utilisés durant les essais préliminaires. De même, le contraste du bruit ne variait pas. Les contrastes du signal utilisés, pour chaque condition, correspondaient aux niveaux déterminés par les essais préliminaires.

- Procédure.

La procédure était identique à celle utilisée au cours des essais préliminaires, mais ici un

seul niveau de contraste du signal était utilisé. Les bruits étaient ensuite classés en fonction des 4 types de réponses possibles, et l'image de classification était déterminée de la même manière que pour l'observateur humain. La méthode de calcul de l'image de classification est précisée dans la première série d'expérience.

5.3.3. Résultats

L'ensemble des résultats de la simulation sont présentés à la Figure 81. Les images de gauche présentent les résultats d'un filtre DoG de petit diamètre, alors que les images de droite correspondent aux résultats d'un filtre de grande taille. La première ligne représente les images obtenues lorsque le centre et le pourtour sont équilibrés. Pour cette condition, les images indiquaient une prédominance de la dimension de contour pour la détection, cette information était d'autant plus large que la taille du filtre était grande. Par ailleurs, l'information de surface ne participait pas à la détection. Les images de la deuxième ligne présentent les images obtenues lorsque le pourtour inhibiteur représente 80 % de l'excitation de la zone centrale. Les résultats indiquaient que l'information de surface était utilisée par l'observateur modèle. Ces images montraient également une participation des contours pour la détection. De la même façon que pour un filtre dont l'antagonisme est équilibré, la dimension de contour s'élargissait pour une taille plus importante du champ récepteur. De plus, la participation l'information de surface était plus importante pour un filtre de petite taille, puisque les contours étaient plus importants pour un filtre de grande taille. La troisième ligne indiquait les résultats quand le pourtour inhibiteur était réduit à 50 % de la zone centrale excitatrice. Les images montraient une participation importante de la surface, et ce que quelle que soit la taille du filtre DoG, aucune participation de l'information de contour n'était notée.

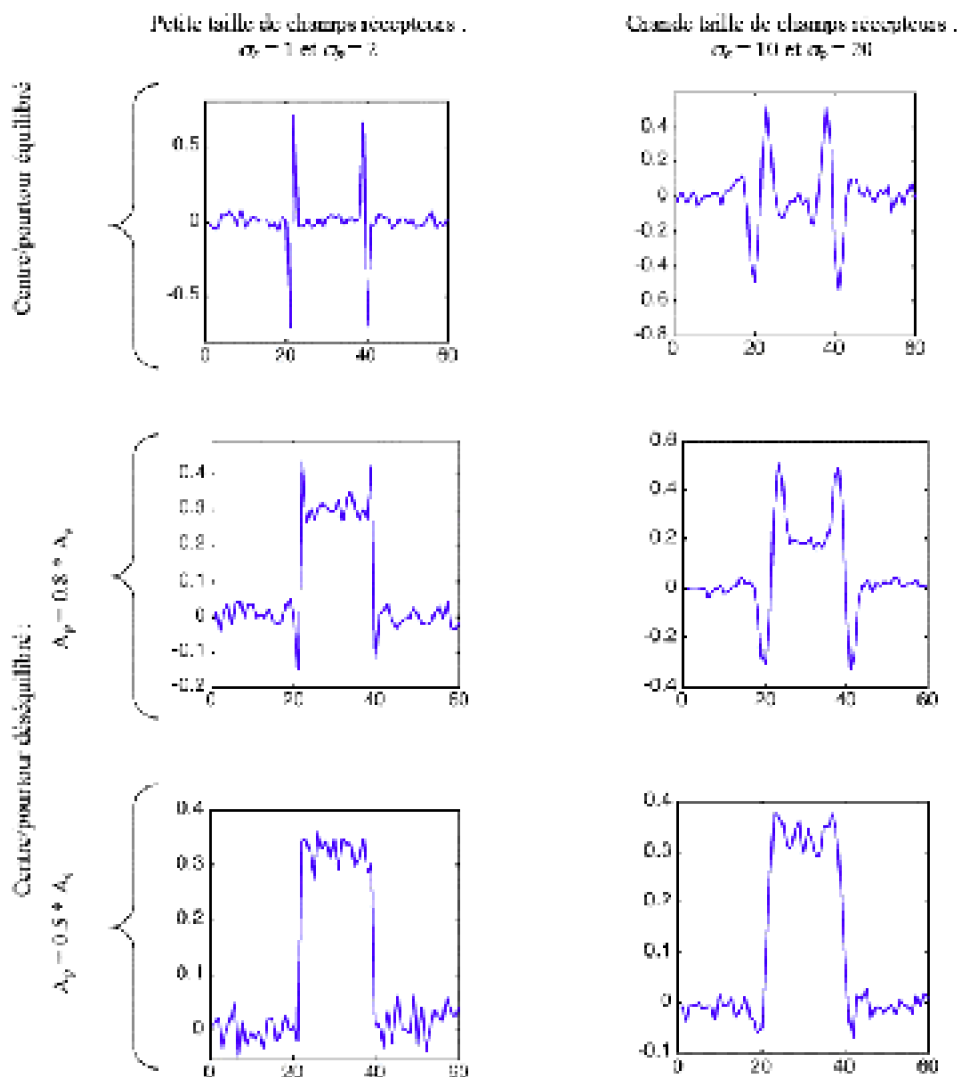


Figure 81 : Expérience 7 : La première ligne montre les résultats obtenus pour un centre/pourtour équilibré. Les deux dernières lignes indiquent les résultats pour un centre/pourtour déséquilibré : dans la seconde ligne l'inhibition du pourtour représente 80% de l'excitation du centre, et pour la troisième ligne, l'inhibition du pourtour représente, cette fois, 50% de la zone excitatrice centrale. Chacune de ces lignes présente l'image de classification brute (à gauche) et l'image de classification après lissage et test statistique (à droite) pour différentes tailles de champs récepteurs. Les petites tailles des champs récepteurs sont présentées à gauche ($\sigma_x = 1$) et les grands tailles des champs récepteurs ($\sigma_x = 10$) sont à droite. On observe que le champ récepteur équilibré (le produit taille x amplitude de la région centrale est égale à celui du pourtour) préserve l'information de surface, alors que les champs récepteurs déséquilibrés permettent le codage de la surface du signal.

5.3.4. Discussion

Les résultats indiquaient que les champs récepteurs présentant un déséquilibre du centre

et du pourtour intégrait l'information de surface, tandis que les champs récepteurs équilibrés véhiculaient principalement l'information de contour. Par contre, contrairement à notre hypothèse, la taille du champ récepteur n'avait aucune influence sur la distinction entre les dimensions de contour et de surface. Toutefois, l'effet de l'équilibre centre/pourtour était beaucoup plus marqué pour les petites tailles des champs récepteurs que pour les tailles plus importantes. Contrairement à l'observateur humain, les images montrent une réponse importante des contours, et aucune influence de l'information de surface lorsque les champs récepteurs sont équilibrés. Par contre, plus l'antagonisme centre/pourtour est déséquilibré et plus l'information de surface est utilisée par l'observateur modèle. Ce dernier résultat est proche des images que nous avons obtenues pendant les expériences comportementales chez l'observateur humain. Ainsi, l'observateur modèle permet de montrer que si l'information de surface n'était pas codée par les canaux visuels précoces, alors l'observateur humain ne pourrait pas utiliser cette information pour la détection.

Les résultats de cette simulation indiquent que la surface est transmise par les canaux visuels précoces quel que soit la taille des champs récepteurs, car seul le déséquilibre entre le centre et le pourtour est important. Des études physiologiques (Croner & Kaplan, 1995 ; Lee, Kremers & Yeh, 1998) ont permis de montrer que le déséquilibre centre/pourtour était équivalent à des valeurs de 50 à 80 %, ce qui est consistant avec les conditions requises pour obtenir une réponse de la surface chez les observateurs humains.

De plus, une distinction est faite entre les petites et les grandes tailles des champs récepteurs. En effet, plus la taille des champs récepteurs est grande, et plus le traitement des contours est étendu spatialement. Ceci est conforme à l'hypothèse selon laquelle les voies M et P répondent aux fréquences spatiales basses et élevées. En effet, comme le montre la Figure 4, les fréquences spatiales élevées véhiculent précisément les détails d'une forme comme les contours, alors que les fréquences spatiales basses transmettent une information plus large. Le fait que la voie M répond préférentiellement à des fréquences spatiales basses signifie que la transition entre deux régions se fait de façon plus étendue, à l'inverse, la voie P répond à des fréquences spatiales élevées ce qui montre que la transition entre deux régions se fait de manière plus localisée. Nos résultats montrent que cette différence est conservée.

Ce dernier résultat est d'autant plus vrai lorsque l'équilibre centre/pourtour est équivalent. Pour cette condition, il est normal que nous obtenions des images proches des résultats que Marr (1982) obtenait durant l'étape du « zero-crossing ». Par ailleurs, Marr (1982) ne pouvait pas manipuler l'équilibre du centre et du pourtour à partir du Laplacien. En effet, cette fonction était obtenue à partir de la dérivée seconde d'une fonction gaussienne qui ne permet de manipuler uniquement que deux paramètres, le résultat permet d'obtenir une fonction en chapeau mexicain équilibré. A l'inverse, le modèle DoG permet de manipuler quatre paramètres ce qui explique pour quoi Marr ne pouvait pas obtenir une participation de la surface dans la première étape de son modèle.

Les résultats de cette modélisation sont d'autant plus importants qu'ils permettent de répondre en partie à la problématique que nous avons posée dans l'introduction. En effet, il est souvent difficile de faire le lien entre codage et perception. Mais ici, nous mettons en

évidence que les indices qui sont perçus par l'observateur humain pour la détection d'une surface uniforme se retrouve dans les dimensions qui sont encodées à travers les champs récepteurs.

Conclusion et Perspectives

Notre objectif était de modéliser les stratégies perceptives que le système visuel met en œuvre pour traiter une information. Plus précisément, nous nous sommes intéressés à l'un des premiers aspects proposés dans la littérature, et qui concerne la distinction entre le traitement du contour et de la surface d'une forme. De nombreux auteurs (Gerrits & Vendrik, 1970 ; Grossberg & Mingolla, 1985 ; Marr, 1982 ; Walls, 1954) ont proposé que ces deux dimensions étaient traitées de façon hiérarchique. En effet, notre première partie théorique a permis de montrer que l'antagonisme centre/pourtour des champs récepteurs permettait d'extraire une information centrée sur un contraste aux bords d'une forme (Hubel & Wiesel, 1962, 1968). Ainsi, il a été suggéré que la dimension de contour est l'une des premières information encodée par le système visuel. Ainsi, notre perception de la surface ne serait pas issue directement de la lumière, mais nécessiterait au contraire une étape de reconstruction. Il est classiquement admis qu'un processus de remplissage permet de traiter l'information de surface à partir des contours de la forme. Par conséquent, si les contours sont encodés par le système visuel et si la surface est reconstruite dans un deuxième temps, alors les bords d'une forme constituent une dimension primordiale pour détecter un signal. En étudiant l'information utilisée par l'observateur, c'est-à-dire sa stratégie de détection, nous pouvons ainsi déterminer la nature du codage des dimensions de contour et de surface. Nous avons tenté de répondre à cette question en utilisant le paradigme de l'image de classification (Ahumada, 1996). Dans cette technique, un bruit est ajouté au signal à chaque essai, et ce bruit est corrélé avec la réponse de l'observateur, ce qui permet de définir les indices perceptifs qui sont pris en considération au cours de la tâche. Le résultat final est une image du

« prototype » du signal utilisé par l'observateur.

Pertinence de la surface pour la détection.

Notre première série d'expérience avait pour but de mettre en évidence ce traitement hiérarchique entre la dimension de contour et de surface dans une tâche de détection. Pour cela, le temps de présentation du stimulus a été manipulé. En effet, nous faisons l'hypothèse que pour un temps de présentation long, alors la dimension de contour et de surface devrait émerger dans l'image de classification. Par contre, pour des temps de présentations brefs, seule la dimension de contour devrait ressortir. Nous supposons ainsi que pour des durées brèves l'observateur utilise principalement une information basée sur les premiers niveaux de traitement. Toutefois, les images de classification indiquaient une zone centrale claire à l'emplacement du signal. Ce résultat est surprenant puisqu'il indique que les observateurs utilisaient principalement une information basée sur la surface et non pas sur les contours, et ce quel que soit le temps de présentation des stimuli (400 ou 40 msec). De plus, un faible pourtour sombre entourait cette zone centrale claire. Une estimation de l'ajustement d'une fonction DoG à nos données a permis de montrer que ce pourtour était plus large que la zone centrale et participait faiblement à la décision. Toutefois, il est possible que les contours du signal soient situés dans le champ visuel périphérique, et il est classiquement admis que le système périphérique dégrade l'information perçue (Rovamo & Virsu, 1979). Par conséquent, la dimension de contour ne pouvait pas être utilisée par les observateurs puisque cette information est restreinte. Pour vérifier cette hypothèse, dans une seconde série d'expériences, nous avons manipulé la taille du signal pour que les contours soient placés dans la fovéa (1 deg) ou au contraire dans la périphérie (6 deg). Néanmoins, ces deux expériences ont permis de montrer la robustesse de la dimension de surface pour la détection.

Le résultat important de ces deux études est que l'observateur humain base sa réponse de détection sur la luminance des pixels de la région centrale et non pas sur les contours. La surface est une information transmise par le système visuel puisqu'elle est utilisée par l'observateur pour la détection. Par ailleurs, nos résultats ne peuvent pas être expliqués par un mécanisme de remplissage puisque le centre du signal contribue directement à la détection. Mais ici, nous montrons que ce processus n'est pas nécessaire pour expliquer comment une surface uniforme est perçue. Toutefois, nos expériences ne remettent pas en cause l'importance des contours et du processus de remplissage dans d'autres contextes. En effet, le remplissage a été démontré non seulement physiologiquement (DeWeerd et al., 1995 ; Komatsu et al., 1996) mais également perceptivement (Paradiso & Hahn, 1996 ; Paradiso & Nakayama, 1991 ; Ramachandran & Gregory, 1991). De plus, de nombreuses études ont montré l'importance des contours pour les figures illusoires (Gold et al., 2000), mais aussi pour une tâche de discrimination (Neri & Heeger, 2002). Nos données ont également mis en évidence la participation d'un pourtour sombre, ce qui pourrait indiquer une tendance de la part des observateurs à intégrer les contours dans leur réponse, mais cette participation

est extrêmement faible, même si on la retrouve systématiquement dans nos expériences.

Par la suite, nous avons voulu savoir si les contours du signal pouvaient quand même être utilisés par l'observateur humain. Pour cela, une tâche de détection, identique aux expériences précédentes, a été réalisée. Mais cette fois, la réponse de type signal présent devait être donnée uniquement si les contours étaient perçus. Nos résultats ont permis de montrer que cette information ressortait dans les images de classification. Ainsi, les bords sont utilisés lorsque l'attention des observateurs est centrée sur cette dimension. En d'autres termes, les contours peuvent être utilisés mais pas de façon automatique, ce qui permet de renforcer l'importance de la surface pour la détection.

Rôle des voies précoces pour le traitement de la surface.

Une tâche de discrimination a ensuite été réalisée afin de déterminer la participation des voies précoces dans le traitement de la surface. Nous nous sommes ici intéressés plus particulièrement à la distinction entre les voies M et P. Ces deux systèmes ont été mis en évidence de façon comportementale à l'aide de deux paradigme par Pokorny et Smith (1997). Nous avons ici utilisé une simplification de la condition « steady-pedestal » qui permet de mettre en évidence la sensibilité au contraste de la voie M. Par conséquent, l'information de surface devrait être plus importante dans le traitement d'une forme. Conformément à notre hypothèse, cette dimension était principalement utilisée par les observateurs.

Les expériences chez l'observateur humain permettaient de montrer la distinction qui existe entre la perception (ici la surface) et le codage neuronal effectué par l'antagonisme centre/pourtour des champs récepteurs (basé essentiellement sur les bords). Nous avons voulu déterminer si la dimension de surface pouvait être codée par les champs récepteurs. Pour cela, une simulation informatique a été effectuée. Les réponses de l'observateur modèle sont basées à partir du profil des champs récepteurs qui est identique à une fonction DoG. Deux paramètres étaient manipulés : la taille et l'équilibre centre/pourtour des champs récepteurs. L'image de classification était ensuite calculée à partir de la réponse de l'observateur modèle. Nos résultats indiquaient que pour un champ récepteur équilibré, la réponse était principalement centrée sur les contours du signal quelle que soit la taille des champs récepteurs. Par contre, la dimension de surface était principalement utilisée lorsque l'antagonisme centre/pourtour était déséquilibré. Cet effet était d'autant plus robuste quand le déséquilibre était important. Par ailleurs, la taille des champs récepteurs n'était pas un facteur déterminant. Ainsi, les images étaient proches de celles obtenues avec les observateurs humains. Les résultats de l'observateur modèle sont importants puisqu'ils montrent que la surface est codée par les champs récepteurs quand l'antagonisme centre/pourtour est déséquilibré. Par conséquent, notre modélisation permet d'apporter des éléments de réponse par rapport à la problématique que nous avons développée entre le codage et la perception au cours de notre

introduction.

Par ailleurs, un mécanisme top-down comme le remplissage ne permet pas d'expliquer notre perception de la surface à partir d'une information bottom-up, c'est-à-dire les contours. Nous ne remettons pas en cause l'existence du phénomène de remplissage, mais d'après les résultats obtenus avec notre observateur modèle, il semble que ce processus ne soit pas pertinent pour expliquer le traitement de la surface dans des conditions naturelles. En effet, ce mécanisme a été mis en évidence dans des conditions naturelles bien précises mais également artificielles. Toutefois, les études physiologiques, qui utilisent une surface uniforme (Lamme, Rodriguez-Rodriguez & Spekreijse, 1999 ; Rossi & Paradiso, 1999 ; Rossi, Rittenhouse & Paradiso, 1996), ne montrent pas directement l'existence de ce mécanisme. Ces expériences indiquent que l'aire V1 est fortement impliquée dans le traitement de la surface, ainsi que les contours, mais la réponse des neurones au centre ou à proximité des contours est similaire. Ce résultat va à l'encontre de la définition même du mécanisme de remplissage.

Existe-t-il une contradiction entre la technique de l'image de classification et le paradigme de masquage ?

L'image de classification peut paraître paradoxale et notamment lorsque nous prenons en compte le paradigme de masquage. En effet, cette dernière technique suppose que le masque perturbe la perception d'un signal, et au contraire, dans le paradigme d'Ahumada (1996, 2002), le bruit influence positivement la réponse de l'observateur. Ce paradoxe est d'autant plus étonnant que les effets de masquage sont maximaux lorsqu'il y a une superposition du signal et du masque, ce qui correspond aux conditions expérimentales que nous avons utilisées. Nous pensons toutefois que ces deux paradigmes ne sont pas incompatibles.

Dans une expérience de masquage, on mesure la façon dont le bruit influence les réponses de l'observateur. Lorsque le masque se superpose au signal, sa perturbation est d'autant plus importante et le critère de décision est tel que l'observateur donne plus de réponses de type absent. Mais, si l'on augmente l'énergie du signal ou si l'on diminue l'énergie du bruit, la probabilité d'obtenir des réponses de type signal présent est plus élevée. Dans ce cas, rendre le bruit plus ou moins perturbant influence le critère de décision. Il n'y a donc pas de contradiction entre ces deux paradigmes puisque dans le masquage le bruit est manipulé pour faire en sorte qu'il empêche le traitement d'un signal et dans l'image de classification le bruit est manipulé pour ne pas détériorer entièrement la perception du signal.

Dans une expérience qui utilise la technique de l'image de classification, le niveau du bruit et du signal est fixé, ainsi le critère et l'indice d' sont stables. Ce paradigme étudie la distribution du bruit qui conduit l'observateur à prendre sa décision.

Perspectives de recherches.

Trois grandes thématiques seront développées concernant nos perspectives de recherches. La première porte sur l'extension de nos travaux sur la distinction contour/surface chez l'observateur humain et plus précisément dans le cadre de la discrimination. Nous pensons dans un premier temps adapter la condition « pulse paradigm » afin de déterminer la participation de la voie P dans le traitement d'une forme simple. Pour cela, nous pensons effectuer une tâche de discrimination de luminance entre deux carrés positionnés côte à côte et dans un même intervalle temporel. Nous supposons que si la technique de Pokorny et Smith (1997) met en évidence une réponse de la voie P, alors les contours du signal devraient principalement être utilisés pour répondre. En dehors de la distinction M/P, Neri & Heeger (2002) ont mis en évidence des différences de stratégie sur la façon de traiter l'information dans une tâche de détection et de discrimination. La tâche de discrimination permet une appréciation plus fine des composantes du signal. Ainsi, nous pensons réaliser une discrimination entre deux formes géométriques. En effet, si ces dernières se différencient essentiellement par la configuration des contours, alors cette information devrait ressortir dans l'image de classification.

La seconde perspective de recherche est de poursuivre nos travaux sur la problématique de la distinction des voies M/P. L'étude d'un patient atteint d'un déficit de la voie M sera réalisée, en collaboration avec Muriel Boucart (CHR, Lille). Un déficit du canal M a été postulé chez ce patient sur la base des signaux électriques recueillis à la surface du scalp (méthode des potentiels évoqués) montrant un dysfonctionnement des régions cérébrales sous-jacentes lorsque la stimulation visuelle implique spécifiquement la voie M. Nous pensons dans un premier temps utiliser les paradigmes de Pokorny et Smith (1997) qui permettent d'isoler les réponses au contraste des voies M et P. Par la suite, nous étudierons les conséquences de la détérioration du système M sur le percept du patient dans notre tâche princeps de détection. Cette dernière expérience nous permettra de préciser la participation de la voie P dans notre tâche princeps de détection. Dans la continuité de la distinction des voies M/P, une hypothèse a été formulée à propos d'un déficit de la voie M pour la dyslexie de surface (Schwartz, 1999 ; Stuart, McAnally & Castles, 2001). Il nous semble pertinent d'appliquer les deux paradigmes de Pokorny et Smith (1997) à ce déficit.

La dernière perspective de recherche a pour but de déterminer l'image de classification d'un signal positionné autour de la tache aveugle. Nous nous attendons à ce que seule l'information centrée sur les bords du signal émerge dans l'image de classification, et que, la partie qui est positionnée à l'intérieur de la tache aveugle ait un niveau de luminance identique à celle du fond puisqu'aucune information n'est transmise. Cette étude sera ensuite étendue à des patients rétinopathes²⁰ traités par photocoagulation. En effet, à la suite de ce traitement, de nombreux scotomes apparaissent, la dernière partie de ce projet sera réalisée en collaboration avec John

Werner (Université de Californie, Davis).

²⁰ Le diabète se caractérise par une hyperglycémie, et, lorsque cette hyperglycémie devient chronique, elle entraîne des troubles vasculaires. Ces complications peuvent toucher les yeux où les vaisseaux sont de petite taille : on parle alors de rétinopathie. Cette pathologie peut provoquer un décollement de la rétine, et conduire à la cécité si elle n'est pas traitée. Elle survient généralement 10 à 15 ans après le début du diabète. Le principal traitement qui existe à l'heure actuelle est la photocoagulation. Le principe repose sur la destruction des zones lésées, par des impacts au laser, la rétine périphérique étant la seule zone touchée dans cette pathologie.

Bibliographie

- Abbey, C.K., Eckstein, M.P (2000). Estimates of human-observer templates for a simple detection task in correlated noise. *Proceedings of SPIE*, 3981, 70-77.
- Abbey, C.K., Eckstein, M.P (2002). Classification image analysis : estimation and statistical inference for two-alternative forced choice experiments. *Journal of Vision*, 2, 66-78.
- Abbey, C.K., Eckstein, M.P., Bochud F.O. (1999). Estimation of human observer templates in two alternative forced-choice experiments. *Proceedings of SPIE*, 3663.
- Ahumada, A. (1996) Perceptual classification images from vernier acuity masked by noise. *Perception (Suppl.)*, 25, 18.
- Ahumada, A.J., & Beard, B.L. (1998). Response classification images in vernier acuity. *Investigative Ophthalmology and Visual Science*, 39, 1109.
- Ahumada, A. J. & Lovell, J. (1970). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, 49, 1751-1756.
- Ahumada, A. J., Marken, R. & Sandusky, A. (1975). Time and frequency analyses of auditory signal detection. *Journal of the Optical Society of America*, 57, 385-390.
- Alexander, K.R., Pokorny, J. Smith, V.C., Fishman, G.A., & Barnes, C.S. (2001). Contrast discrimination deficits in retinis pigmentosa are greater for stimuli that favor the magnocellular pathway. *Vision Research*, 41, 671-683.
- Allison, J.D., Melzer, P., Ding, Y., Bonds, A.B ;, & Casagrande, V.A. (2000). Differential

- contribution of magnocellular and parvocellular pathways to the contrast response of neurons in bushy primary visual cortex (V1). *Visual Neuroscience*, 17, 71-76.
- Angelucci, A., Levitt, J.B., Walton, E.J.S., Hupé, J.M., Bullier, J., & Lund, J.S. (2002). Circuit of local and global signal integration in primary visual cortex. *Journal of Neuroscience*, 22, 8633-8646.
- Anstis, S. (1996). Adaptation to peripheral flicker. *Vision Research*, 36, 3479-3485.
- Altmann, C.F., Bühlhoff, H.H., & Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology*, 13, 342-349.
- Barlow, H.B. (1956). Retinal noise and absolute threshold. *Journal of Optical Society of America*, 46, 634-639.
- Barlow, H.B. (1958). Temporal and spatial summation in human vision at different background intensities. *Journal of Physiology*, 141, 337-350.
- Beard, B.L. & Ahumada, A.J. (1997). Relevant image features for Vernier acuity. *Investigative Ophthalmology and Visual Sciences (Suppl.)*, 38, p. 118.
- Beard, B.L. & Ahumada, A.J. (1998). Response classification images in vernier acuity. *Investigative Ophthalmology and Visual Sciences (Suppl.)*, 39, p. 118.
- Beaudot, W.H.A. (2002). Role of onset asynchrony in contour integration. *Vision Research*, 42, 1-9.
- Bex, P.J., Simmers, A. J., & Dakin, S.C. (2001). Snakes and ladders : the role of temporal modulation in visual contour integration. *Vision Research*, 41, 3775-3782.
- Biederman, I. (1987). Recognition-by-component : A theory of human image understanding. *Psychological Review*, 94, 115-147.
- Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, 14, 143-177.
- Blakemore, C., Mansey, J.P.J., & Ridley, R.M. (1971). Perceptual fading of a stabilized image cortical image. *Nature*, 233, 204-205.
- Boucart, M. (1996). *La reconnaissance des objets*. Presses Universitaires de Grenoble.
- Boucart, M., Delord, S., & Giersch, A. (1994). The computation of contour information in complex objects. *Perception*, 23, 399-409.
- Boucart, M., Ferrand, L., & Grainger, J. (1995). Three-dimensional features facilitate object recognition. *Visual Cognition*, 2, 451-478.
- Boynton, G., Demb, J.B., Glover, G.H., & Heeger, D.J. (1999). Neuronal basis of contrast discrimination. *Vision Research*, 39, 257-269.
- Braje, W.L., Tjan, B.S., & Legge, G. (1995). Human efficiency for recognizing and detecting low-pass filtered objects. *Vision Research*, 35, 2955-2986.
- Brigner, W.L. (1982). Mathematical model for the filling-in of illusory contour. *Perceptual & Motor Skills*, 54, 815-820.
- Brehaut, J.C., Enns, J.T., & Di Lollo, V. (1999). Visual masking plays two roles in the attentional blink. *Perception & Psychophysique*, 61, 1436-1448.
- Breitmeyer, B.G. (1984). *Visual masking : an interactive approach*. Oxford : Clarendon Press.

- Breitmeyer, B.G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. *Psychological Review*, 83, 1-36.
- Breitmeyer, B.G., & Ogmen, H. (2000). Recent models and findings in visual backward masking : a comparison, review, and update. *Perception and Psychophysics*, 62, 1572-1595.
- Brown, R.J., & Thurmond, J.B. (1993). Preattentive and cognitive effects on perceptual completion at the blind spot. *Perception and Psychophysics*, 53, 200-209.
- Bullier, J. (1998). Architecture fonctionnelle du système visuel. In Boucart, M., Hénaff, M.A., & Belin, C. (Eds.), *Vision : aspects perceptifs et cognitifs*. Marseille : Solal.
- Bullier, J. (2001). Integrated model of visual processing. *Brain Research Review*, 36, 96-107.
- Bullier, J ; (2002). Neural basis of vision. In H. Pashler (Ed.), *Stevens' handbook of experimental psychology (3rd edition) Volume 1 : sensation and perception*. New-York : Wiley and Sons.
- Campbell, F.W., & Robson, J.G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 197, 551-566.
- Caputo, G ; (1998). Texture brightness filling-in. *Vision Research*, 38, 841-851.
- Chavane, F., Monier, C., Bringuier, V., Baudot, P., Borg-Graham, L., Lorenceau, J., & Frégnac, Y. (2000). The visual cortical association field : a gestalt concept or a psychological entity ? *Journal of Physiology (Paris)*, 94, 333-342.
- Chino, Y.M., Kaas, J.H., Smith III, E.L., Langston, A.L., & Cheng, H. (1992). Rapid reorganisation of cortical maps in adult cats following restricted deafferentation in retina. *Vision Research*, 32, 789-796.
- Churchland, P.S., & Ramachandran, V.S. (1996). Filling-in : Why Dennett is wrong. In K. Atkins (Ed.) *Perception*. New-York : Oxford University Press.
- Croner, L.J., & Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Research*, 35, 7-24.
- Croner, L.J., Purpura, K., & Kaplan, E. (1993). Response variability in retinal ganglion cells of primates. *Proceedings of the National Academy of Science of the United States of America*, 90, 8128-8130.
- Dacey, D.M., & Petersen, M.R. (1992). Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proceedings of the National Academy of Science of the United States of America*, 89, 9666-9670.
- Dacey, D.M., Packer, O.S., Diller, S., Brainard, D., Peterson, B., & Lee, B.B. (2000). Center surround receptive field structure of cone bipolar cells in primate retina. *Vision Research*, 40, 1801-1811.
- Davey, M.P., Madess, T., & Srinivasan, M.V. (1998). The spatiotemporal properties of the Craik-O'Brien-Cornsweet effect are consistent with fill-in. *Vision Research*, 38, 2037-2046.
- Day, R.H., & Jory, M.K. (1980). A note on a second stage in the formation of illusory contours. *Perception and Psychophysics*, 27, 89-91.

- Delord, S. (1998). Which mask is the most efficient : A pattern or a noise ? It depends on the task. *Visual Cognition*, 5, 313-338.
- Delord, S. (2002). Le masquage. In *Dictionnaire des Sciences Cognitives* (G. Tiberghien, Ed.). Paris : Armand Colin.
- Dennett, D.C. (1991). *Consciousness explained*. Boston : Little Brown.
- DeAngelis, G.C., Anzai, A., Ohzawa, I., & Freeman, R.D. (1995). Receptive field structure in the visual cortex : does stimulation induce plasticity. *Proceedings of the National Academy of Science USA*, 92, 9682-9686.
- DeBonet, J.S. & Zaidi, Q. (1997). Comparison between spatial interactions in perceived contrast and perceived brightness. *Vision Research*, 37, 1141-1155.
- DeCharms, R.C., & Zador, A. (2000). Neural representation and the cortical code. *Annual Review of Neuroscience*, 23, 613-647.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193-222.
- DeValois, R.L., Albrecht, D.G., & Thorell, L.G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22, 545-549.
- DeValois, R.L., & DeValois, K.K. (1988). *Spatial vision*. Oxford : Oxford University Press.
- DeWeerd, P., Desimone, R., & Ungerleider, L.G. (1998). Perceptual filling-in : a parametric study. *Vision Research*, 38, 2721-2734.
- DeWeerd, P., Gattass, R., Desimone, R., & Ungerleider, L.G. (1995). Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature*, 377, 731-734.
- Dresp, B. (1997). On « illusory » contours and their functional significance. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 16, 489-518.
- Dresp, B., & Bonnet, C. (1991). Psychophysical evidence for low-level processing of illusory contours and surfaces in the Kanizsa square. *Vision Research*, 31, 1813-1817.
- Dresp, B., & Bonnet, C. (1993). Psychophysical measures of illusory form perception : further evidence for local mechanisms. *Vision Research*, 33, 759-766.
- Dresp, B., & Grossberg, S. (1997). Contour integration across polarities and spatial gaps : From local contrast filtering to local grouping. *Vision research*, 37, 913-924.
- Dresp, B., & Grossberg, S. (1999). Spatial facilitation by color and luminance edges : boundary, surface, and attentional factors. *Vision Research*, 39, 3431-3443.
- Durgin, F.H., Tripathy, S.P., & Levi, D.M. (1995). On the filling-in of the visual blind spot : some rule of thumb. *Perception*, 24, 827-840.
- Eckstein, M.P., Shimozacki, S.S., & Abbey, C.K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision*, 2, 25-45.
- Enns, J.T., Di Lollo, V. (2000). What new in visual masking ? *Trends in Cognitive Sciences*, 4, 345-352.
- Enroth-Cugell, C., & Robson, J.G. (1966). The contrast sensitivity of retinal ganglion

- cells of the cat. *Journal of Physiology*, 187, 517-522.
- Eysel, U.T., & Schweigart, G. (1999). Increased receptive field size in the surround of chronic lesions in the adult cat visual cortex. *Cerebral Cortex*, 9, 101-109.
- Farid, H. (2002). Temporal synchrony in perceptual grouping : a critique. *Trends in Cognitive Sciences*, 7, 284-288.
- Felleman, D.J., & Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1-47.
- Field, D.J., Hayes, A., & Hess, R.E. (1993). Contour integration by the human visual system : evidence for a local « association field ». *Vision Research*, 33, 173-193.
- Fitzpatrick, D.J. (2000). Seeing beyond the receptive field in primary visual cortex. *Current Opinion in Neurobiology*, 10, 438-443.
- Fitzpatrick, D., Lund, J.S., & Blasdel, G.G. (1985). Intrinsic connections of the macaque striate cortex : afferent and efferent connections of lamina 4c. *Journal of neuroscience*, 5, 3329-3349.
- Fitzpatrick, D., Usrey, W.M., Schofield, B.R., & Einstein, G. (1994). The sublimar organization of corticogeniculate neurons in layer 6 of macaque striate cortex. *Visual Neuroscience*, 11, 307-315.
- Friedman, H.S., Zhou, H., & von der Heydt, R. (1999). Color filling-in under steady fixation : behavioral demonstration in monkeys and humans. *Perception*, 28, 1383-1395.
- Geisler, W.S. (1989). Sequential ideal-observer analysis of visual discrimination. *Psychological Review*, 96, 267-314.
- Geisler, W.S., Perry, J.S., Super, B.J., & Gallogly, D.P. (2001). Edges co-occurrence in natural images predicts contour grouping performance. *Vision Research*, 41, 711-724.
- Gerrits, H.J.M., & Timmerman, G.J.M.E.N. (1968). The filling-in process in patients with retinal scotoma. *Vision Research*, 9, 439-442.
- Gerrits, H.J.M., & Vendrik, A.M. (1970). Simultaneous contrast, filling-in process and information processing in man's visual system. *Experimental Brain Research*, 11, 411-430.
- Ghose, G.M., & Maunsell, J.R. (2002). Attentional modulation in visual cortex depends on task timing. *Nature*, 419, 616-620.
- Giersch, A., Humphreys, G.W., Boucart, M., & Kovacs, I. (2000). The computation of occluded contours in visual agnosia: Evidence for early computation prior to shape binding and figure-ground coding. *Cognitive Neuropsychology*, 17, 731-759.
- Gilbert, C.D. (1995). Dynamic properties of adult visual cortex. In Gazzaniga M. (Ed.), *The cognitive neuroscience*. Cambridge, MA : MIT Press.
- Gilbert, C.D. (1996). Plasticity in visual perception and physiology. *Current Opinion in Neurobiology*, 6, 269-274.
- Gilbert, C.D., & Wiesel, T.N. (1979). Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. *Nature*, 280, 120-125.
- Gilbert, C.D., & Wiesel, T.N. (1989). Columnar specificity of intrinsic horizontal and

- corticocortical connections in cat visual cortex. *Journal of Neuroscience*, 9, 2432-2442.
- Gilbert, C.D., & Wiesel, T.N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Research*, 30, 1689-1701.
- Gilbert, C.D., & Wiesel, T.N. (1992). Receptive field dynamics in adult primary visual cortex. *Nature*, 356, 150-152.
- Gilbert, C.D., Das, A., Ito, M., Kapadia, M., & Westheimer, G. (1996). Spatial integration and cortical dynamics. *Proceedings of the National Academy of Science of the United States of America*, 93, 615-622.
- Ginsburg, A.P. (1986). Spatial filtering and visual form perception. In K.R. Boff, L. Kaufman, & J.P. Thomas (eds.), *Handbook of human perception and performance*, vol. 2. New-York : Wiley.
- Gold, J., Murray, R.F., Bennett, P.J., & Sekuler, A.B. (2000). Deriving behavioural receptive fields for visually completed ? *Current Biology*, 10, 663-666.
- Gouras, P. (1968). Identification of cone mechanisms in monkey ganglion cells. *Journal of Physiology*, 199, 533-547.
- Gray, C.M., Konig, P., Engel, A.K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338, 334-337.
- Green, D.M., & Swets, J. (1966). *Signal detection theory and Psychophysics* New-York : Robert E. Krieger Publishing Company.
- Grosf, D.H., Shapley, R.M., & Hawken, M.J. (1993). Macaque V1 neurones can signal « illusory » contours. *Nature*, 365, 550-552.
- Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception and Psychophysics*, 55, 48-120.
- Grossberg, S. (2000). The complementary brain : a unifying view of brain specialization and modularity. *Trends in Cognitive Science*, 4, 233-246.
- Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception : Boundary Completion, illusory figures, and neon color spreading. *Psychological Review*, 92, 173-211.
- Grossberg, S., & Todorovic', E. (1988). Neural dynamics of 1-D and 2-D brightness perception : a unified model of classical model and recent phenomena. *Perception and Psychophysics*, 43, 241-277.
- Grossberg, S., Mingolla, E., & Ross, W.D. (1997). Visual brain and visual perception : how does the cortex do perceptual grouping ? *Trends in Neuroscience*, 20, 106-117.
- Gyoba, J. (1997). Loss of a forest : perceptual fading and filling-in of static texture patterns. *Perception*, 26, 1317-1320.
- Haith, M.M., Bergman, T., & Moore, M.J. (1977). Eye contact and face scanning in early infancy. *Science*, 198, 853-855.
- Hendry, S.H.C., & Reid, R.C. (2000). The koniocellular pathway in primate vision. *Annual Review of Neuroscience*, 23, 127-153.

-
- Hess, R.F., Beaudot, W.H.A. , & Mullen, K.T. (2001). Dynamics of contour integration. *Vision Research*, 41, 1023-1037.
- Hess, R., & Dakin, S. (1997). Absence of contour linking in peripheral vision. *Nature*, 390, 602-604.
- Hess, R., & Dakin, S. (1999). Contour integration in the peripheral field. *Vision Research*, 39, 947-959.
- Hess, R., Dakin, S., & Field, D.J. (1998). The role of « contrast enhancement » in the detection and appearance of visual contours. *Vision Research*, 38, 783-787.
- Hess, R., & Field, D. (1999). Integration of contours : new insights. *Trends in Cognitive Science*, 3, 480-486.
- Hirsch J.A., & Gilbert, C.D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *Journal of Neuroscience*, 11, 1800-1809.
- Hirsch, J., DeLaPaz, R., Relkin, N.R., Victor, J., Kim, K., Li, T., Borden, P., Rubin, N., & Shapley, R. (1995). Illusory contours activate specific regions in human visual cortex : evidence from functional magnetic resonance imaging. *Proceedings of the National Academy of Science of the United States of America*, 92, 6469-6473.
- Huang, X., MacEvoy, S.P., & Paradiso, M.A. (2002). Perception of brightness illusions in the macaque monkey. *Journal of Neuroscience*, 22, 9618-9625.
- Hubel, D. (1994). *L'œil, le cerveau et la vision*. Paris : Belin.
- Hubel, D., & Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in two non-striate visual areas of the cat. *Journal of Physiology*, 160, 106-154.
- Hubel, D., & Wiesel, T. (1968). Receptive fields, and functional architecture of monkey striate cortex. *Journal of Physiology*, 195, 215-243.
- Hung, C.P., Ramsden, B.M., Chen, L.M., & Roe, A.W. (2001). Building surfaces from borders in area 17 and 18 of the cat. *Vision Research*, 41, 1389-1407.
- Kapadia, M.K., Ito, M., Gilbert, C.D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context : parallel studies in human observers and in V1 of alert monkeys. *Neuron*, 15, 843-856.
- Kapadia, M.K., Westheimer, G., & Gilbert, C.D. (1999). Dynamics of spatial summation in primary visual cortex. *Proceedings of the National Academy of Science of the United States of America*, 96, 12073-12078.
- Kapadia, M.K., Westheimer, G., & Gilbert, C.D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Physiology*, 87, 2048-2062.
- Kaplan, E ; (2003). The M, P and K pathways of the primate visual system. In L.M. Chalupa & J.S. Werner (Eds), *The visual neurosciences*. Cambridge, MA : MIT Press.
- Kaplan, & Shapley, R. (1986). The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Science of the United States of America*, 83, 2755-2757.
- Kasamatsu, T., Polat, U., Pettet, M.W., Norcia, A.M. (2001). Colinear facilitation promotes reliability of single cell responses in cat striate cortex. *Experimental Brain*

Research, 138, 163-172.

- Kastner, S., Nothdurft, H.C., & Pigarev, I.N. (1997). Neuronal correlates of pop-out in cat striate cortex. *Vision Research*, 37, 371-376.
- Kessen, W., Salapatek, P., & Haith, M. (1972). The visual response of the human newborn to linear contour. *Journal of Experimental Child Psychology*, 13, 9-20.
- Kinoshita, M., & Komatsu, H. (2001). Neural representation of the luminance and brightness of a uniform surface in the macaque primary visual cortex. *Journal of Physiology*, 86, 2559-2570.
- Knierim, J.J., & Van Essen, D.C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, 67, 961-980.
- Knoblauch, K. (1999). Vision des couleurs chez les primates. *Primatologie*, 2, 29-59.
- Knoblauch, K., & Shevell, S.K. (2003). Color appearance. In L.M. Chalupa & J.S. Werner (Eds.), *The visual neurosciences*. Cambridge (MA) : MIT Press.
- Knoblauch, K. & Yssaad-Fesselier, R. (2000). Stimulus Features used during learning of an unknown visual signal (s). *Investigative Ophthalmology & Visual Science*, 41, s46.
- Komatsu, H., & Murakami, I. (1994). Behavioral evidence of filling-in at the blind spot of the monkey. *Visual Neuroscience*, 11, 1103-1113.
- Komatsu, H., Kinoshita, M., & Murakami, I. (2000). Neural responses in the retinotopic representation of the blind spot in the macaque V1 to stimuli for perceptual filling-in. *Journal of Neuroscience*, 20, 9310-9319.
- Komatsu, H., Kinoshita, M., & Murakami, I. (2002). Neural Responses in the primary visual cortex of the monkey during perceptual filling-in at the blind spot. *Neuroscience Research*, 44, 231-236.
- Komatsu, H., Murakami, I., & Kinoshita, M. (1996). Surface representation in the visual system. *Cognitive Brain Research*, 5, 97-104.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical region involved in perceiving object shape. *Journal of Neuroscience*, 20, 3310-3318.
- Kovacs, I. (1996). Gestalten of today : early processing of visual contours and surface. *Behavioural Brain Research*, 82, 1-11.
- Kovacs, I., & Julesz, B. (1993). A closed curve is much more than incomplete one : effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Science of the United States of America*, 90, 7495-7497.
- Kovacs, I., & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature*, 370, 644-646.
- Krauskopf, J. (1963). Effect of retinal image stabilization on the appearance of heterochromatic targets. *Journal of the Optical Society of America*, 53, 764-778.
- Lamme, V.A. (1995). The neurophysiology of figure-ground segmentation in primary visual cortex. *Journal of Neuroscience*, 15, 1605-1615.
- Lamme, V.A.F., & Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience*, 23, 571-579.
- Lamme, V.A.F., & Spekreijse, H. (1998). Neuronal synchrony does not represent texture segregation. *Nature*, 396, 362-366.

-
- Lamme, V.A.F., & Spekreijse, H. (2000). Contextual modulation in primary visual cortex and scene perception. In Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge, MA: MIT Press.
- Lamme, V.A.F., Rodriguez-rodriguez, V., & Spekreijse, H. (1999). Separate processing dynamic for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex*, 9, 406-413.
- Lamme, V.A.F., Supér H., & Spkreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology*, 8, 529-535.
- Lamme, V.A.F., Zipser, K., & Spekreijse, H. (2002). Masking interrupts figure-ground signals in V1. *Journal of Cognitive Neuroscience*, 14, 1044-1053.
- Lee, B.L. (1996). Receptive Field structure in the primate retina. *Vision Research*, 36, 631-644.
- Lee, B.B., Kremers, J., & Yeh, T. (1998). Receptive fields of primate retinal ganglion cells studied with a novel technique. *Visual Neuroscience*, 15, 161-175.
- Lee, B.B., Pokorny, J., Smith, V.C., Martin, P.R., & Valberg, A. (1990). Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observer. *Journal of Optical Society of America*, 7, 2223-2236.
- Lee, T.S., & Nguyen, M. (2001). Dynamics of subjective contour formation in the early visual cortex. *Proceedings of the National Academy of Science of the United States of America*, 98, 1907-1911.
- Legge, (1978). Sustained and transient mechanisms in human vision. *Vision Research*, 18, 69-81.
- Leshner, G.W. (1995). Illusory contours : toward a neurally based perceptual theory. *Psychonomic Bulletin and Review*, 2, 279-321.
- Levi, D. M. & Klein, S. A. (1985). Vernier acuity, crowding and amblyopia. *Vision Research*, 25, 979-991.
- Levi, D.M. & Klein, S.A. (2002). Classification images for detection and position discrimination in the fovea and parafovea. *Journal of Vision*, 2, 46-65.
- Levitt, J.B., & Lund, J.S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature*, 387, 73-76.
- Li, Z. (1999). Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proceedings of the National Academy of Science of the United States of America*, 96, 10530-10535.
- Li, W., & Gilbert, C.D. (2002). Global contour saliency and local colinear interactions. *Journal of Neurophysiology*, 88, 2846-2856.
- Livingstone, M., & Hubel, D. (1988). Segregation of form, movement, and depth : anatomy, physiology, and perception. *Science*, 240, 740-749.
- Lou, I. (1999). Selective peripheral fading : evidence for inhibitory sensory effect of attention. *Perception*, 28, 519-526.
- Lu, Z.L., & Doshier, B.A. (1998). External noise distinguishes attention mechanisms. *Vision Research*, 38, 1183-1198.

- Lund, J.S., Yoshioka, T., & Levitt, J.B. (1993). Comparison of intrinsic connectivity in different areas of macaque monkey cerebral cortex. *Cerebral Cortex*, 3, 148-162.
- MacEvoy, S.P., & Paradiso, M. (2001). Lightness constancy in primary visual cortex. *Proceedings of the National Academy of Science of the United States of America*, 98, 8827-8831.
- MacEvoy, S.P., Kim, W., & Paradiso, M. (1998). Integration of surface information in primary visual cortex. *Nature Neuroscience*, 1, 616-620.
- Macknik, S.L., & Livingstone, M.S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience*, 1, 144-149.
- Macknick, S.P., Martinez-Conde, S., & Haglund, M. (2000). The role of spatiotemporal edges in visibility and visual masking. *Proceedings of the National Academy of Science of the United States of America*, 97, 7556-7560.
- MacMillan, N.A., & Creelman, C.D. (1991). *Detection theory : a user's guide*. New-York : Cambridge University Press.
- Magnussen, S., Spillman, L., Sturzel, F., & Werner, J. (2001). Filling-in of the foveal blue scotoma. *Vision Research*, 41, 2961-2967.
- Malach, R., Amir, Y., Harel, M., & Grinvald, A. (1993). Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. *Proceedings of the National Academy of Science of the United States of America*, 90, 10469-10473.
- Marr, D. (1982). *Vision : a computational investigation into the human representation and processing of visual information*. San Francisco : W.H. Freeman.
- Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London*, 207, 187-217.
- Marr, D., & Nishihara, H.K. (1978). Visual information processing : artificial intelligence and the sensorium of sight. *Technology Review*, 81, 2-23.
- Martin, K.A.C. (2002). Microcircuits in visual cortex. *Current Opinion in Neurobiology*, 12, 418-425.
- Maslan, R.H. (2001). The fundamental plan of the retina. *Nature Neuroscience*, 4, 877-886.
- Magnussen, S., Spillman, L., Sturzel, F., & Werner, J. (2001). Filling-in of the foveal blue scotoma. *Vision Research*, 41, 2961-2967.
- Mendola, J.D., Dale, A.M., Fischl, B., Liu, A.K., & Tootell, R.B. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, 19, 8560-8572.
- Merigan, W.H., & Maunsell, J.H. (1990). Macaque vision after magnocellular lateral geniculate lesion. *Visual Neuroscience*, 11, 155-163.
- Merigan, W.H., Katz, L.M., & Maunsell, J.H. (1990). The effects of parvocellular lateral geniculate lesions on the acuity and the contrast sensitivity of macaque monkeys. *Journal of Neuroscience*, 11, 994-1001.
- Miller, K.D., Pinto, D.J., & Simons, D.J. (2001). Processing in layer 4 the neocortical circuit : new insights from visual and somatosensory cortex. *Current Opinion in*

- Neurobiology*, 11, 488-497.
- Murakami, I., Komatsu, H., & Kinoshita, M. (1997). Perceptual filling-in at the scotoma following a monocular retinal lesion in the monkey. *Visual Neuroscience*, 14, 89-101.
- Motoyoshi, I. (1999). Texture filling-in and texture segregation revealed by transient masking. *Vision Research*, 39, 1285-1291.
- Nakayama, K., He, Z.J., & Shimojo, S. (1995). Visual surface representation : A critical link between lower-level and higher-level vision. In Kosslyn, S.M., & Osherson (Eds.) *An invitation to cognitive science : visual cognition*. Cambridge, MA : M.I.T. Press.
- Navon, D. (1977). Forest before trees : the precedence of global features in visual perception. *Cognitive Psychology*, 9, 353-383.
- Neri, P., & Heeger, D.J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience*, 5, 812-816.
- Neri, P., Parker, A.J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, 401, 695-698.
- Nothdurft, H.C. (2002). Attention shifts to salient target. *Vision Research*, 42, 1287-1306.
- Nowak, L., & Bullier, J. (1997). The timing of information transfer in the visual system. In Kaas, J.H., Rockland, K.L., & Peters, A.L. (Eds.), *Extrastriate visual cortex*. New-York : Plenum Pub. Corp.
- Ogmen, H. (1993). A neural theory of retino-cortical dynamics. *Neural Networks*, 6, 245-273.
- Ohtani, Y., Okamura, S., Yoshida, Y., Toyama, K., & Ejima, Y. (2002). Surround suppression in the human visual cortex : an analysis using magnetoencephalography. *Vision Research*, 42, 1825-1835.
- Ostergaard, A.L., & Davidoff, J.B. (1985). Some effects of color on naming and recognition of objects. *Journal of Experimental Psychology : Learning, Memory and Cognition*, 3, 519-526.
- Palmer, S.E. (1999). *Vision science : from photon to phenomenology*. Cambridge, MA : MIT Press.
- Palmer, S.E. (2002). Perceptual organization in vision. In H. Pashler (Ed.), *Stevens' handbook of experimental psychology (3rd edition) Volume 1 : sensation and perception*. New-York : Wiley and Sons.
- Paradiso, M.A. (2000). Illuminating the dark corner. *Current Biology*, 10, r15-r18.
- Paradiso, M.A. (2002). Perceptual and neuronal correspondence in primary visual cortex. *Current Opinion in Neurobiology*, 12, 155-161.
- Paradiso, M.A., & Hahn, S. (1996). Filling-in percepts produced by luminance modulation. *Vision Research*, 36, 2657-2663.
- Paradiso, M. & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, 31, 1221-1236.
- Pasupathy, A., & Connor, C.E. (1999). Response to contour features in macaque area V4. *Journal of Neurophysiology*, 82, 2490-2502.
- Perry, V.H., Oehler, R., & Cowey, A. (1984). Functional lamination in the ganglion cell

- layer of the macaque's retina. *Neuroscience*, 25, 217-223.
- Pessoa, L. & Neumann ; H. (1998). Why does the brain fill-in ? *Trends in Cognitive Sciences*, 2, 422-424.
- Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex : contours bridging gaps. *Journal of Neuroscience*, 9, 1749-1763.
- Peterhans, E., & von der Heydt, R. (1991). Elements of form perception in monkey prestriate cortex. In Gorea, A., Frégnac, Y., Kapoula, Z., Findlay, J. (Eds.), *Representations of vision : trends and tacit assumption in vision research*. Cambridge : Cambridge University press.
- Pettet, M. W. (1999). Shape and contour detection. *Vision Research*, 39, 551-557.
- Pettet, M. W., & Gilbert, C.D. (1992). Dynamic changes in receptive-field size in cat primary visual cortex. *Proceedings of the National Academy of Science of the United States of America*, 89, 8366-8370.
- Pettet, M.W., McKee, , & Grywacz, N.M. (1998). Constraints on long-range interactions mediating contour-detection. *Vision Research*, 38, 865-879.
- Pinna, B. , Brelstaff, G., & Spillmann, L. (2001). Surface color from boundaries: a new 'watercolor' illusion. *Vision Research*, 20, 2669-2676.
- Pinker, S. (1984). Visual cognition : an introduction. In S. Pinker (Ed.), *Visual cognition*. Cambridge, MA : MIT Press.
- Pokorny, J., & Smith, V.C. (1997). Psychophysical signatures associated with magnocellular and parvocellular pathway contrast gain. *Journal of Optical Society of America*, 14, 2477-2486.
- Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels suppression and facilitation revealed by lateral masking experiments. *Vision Research*, 33, 993-999.
- Polat, U., & Sagi, D. (1994a). Spatial interactions in human vision : from near to far via experience-dependent cascades of connections. *Proceedings of the National Academy of Science USA*, 91, 1206-1209.
- Polat, U., & Sagi, D. (1994b). The architecture of perceptual spatial interactions. *Vision research*, 34, 73-78.
- Polat, U., & Tyler, C.W. (1999). What pattern the eye sees the best. *Vision Research*, 39, 887-895.
- Polat, U., Mizobe, K., Pettet, M.W., Kasamatsu, T., & Norcia, A.M. (1998). Colinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature*, 391, 580-584.
- Purves, D., Shimpi, A., & Lotto, R.B. (1999). An experimental explanation of the Corsweet effect. *Journal of Neuroscience*, 19, 8542-8551.
- Ramachandran, V.S. (1992a). Filling in the blind spot. *Nature*, 356, 115.
- Ramachandran, V.S. (1992b). Blind spots. *Scientific American*, 266, 86-91.
- Ramachandran, V.S. (1993). Filling in gaps in perception : scotomas and phantom limbs. *Current Directions in Psychological Science*, 2, 56-65.

-
- Ramachandran, V.S., & Gregory, R. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature*, *350*, 699-702.
- Ramachandran, V.S., Gregory, R.L., & Aiken, W. (1993). Perceptual fading of visual texture borders. *Vision Research*, *33*, 717-721.
- Ratliff, E.F. (1965). *Mach bands : quantitative stuides on neural networks in the retina*. San Franscisco : Holden Day.
- Ringach, D.L. (2002). Spatial structure and symetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of Neurophysiology*, *88*, 455-463.
- Rodieck, R.W. (1965). Quantitative analysis of cat ganglion cell response to visual stimuli. *Vision Research*, *5*, 583-601.
- Rodieck, R.W. (1998). *First step in seeing*. Sunderland : Sinauer.
- Roelfsema, P.R., & Singer, W. (1998). Feature article : detecting connectedness. *Cerebral Cortex*, *8*, 358-396.
- Rogers-Ramachandran, D.C., & Ramachandran, V.S. (1998). Psychophysical evidence for boundary and surface systems in human vision. *Vision Research*, *38*, 71-77.
- Reich, L.N., Levi, D.M., & Frishman, L.J. (2000). Dynamic random noise shrinks the twinkling aftereffect induced by artificial scotomas. *Vision Research*, *40*, 805 - 816.
- Rensink, R.A., & Enns, J.T. (1998). Early completion of occluded objects. *Vision Research*, *38*, 2489-2505.
- Rossi, A.F., & Paradiso, M.A. (1999). Neural correlates of brightness in the responses of neurons in the retina, LGN, and primary visual cortex. *Journal of Neuroscience*, *19*, 6145-6156.
- Rossi, A.F., Desimone, R., & Ungerleider, L.G. (2001). Contextual modulation in primary visual cortex of macaques. *Journal of Neuroscience*, *21*, 1698-1709.
- Rossi, A.F., Rittenhouse, C.D., & Paradiso, M.A. (1996). The representation of brighthness in primary visual cortex. *Science*, *273*, 1104-1107.
- Rovamo, J. & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, *37*, 495-510.
- Saarinen, J., & Levi, D.M. (2001). Integration of local features into global shape. *Vision Research*, *41*, 1785-1790.
- Sakaguchi, Y. (2001). Target/surround asymmetry in perceptual filling-in. *Vision Research*, *41*, 2065-2077.
- Salapatek, P., & Kessen, W. (1966). Visual scanning og triangles by the human newborn. *Journal of Experimental Child Psychology*, *3*, 155-167.
- Sandusky, A. & Ahumada, A.J. (1970). Contrast in detection with gated noise. *Journal of the Acoustical Society of America*, *49*, 1790-1794.
- Sceniak, M.P., Ringach, D.L., Hawken, M.J., & Shapley, R. (1999). Contrast's effect on spatial summation by macaque V1 neurons. *Nature Neuroscience*, *2*, 733-739.
- Schwartz, S.H. (1999). *Visual perception : a clinical orientation*. Stamford : Appleton & Lange.
- Schwartz, E.L., Desimone, R., Albright, T.D., & Gross, C.G. (1983). Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Science USA*,

80, 5776-5778.

- Shepherd, J.W., Davies, G.M., & Ellis, H.D. (1981). Studies of cue saliency. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces*. London : Academic Press.
- Sekuler, A.B., & Palmer, S.E. (1992). Perception of partly occluded objects : a microgenetic analysis. *Journal of Experimental Psychology : General*, 121, 95-111.
- Sekuler, A., Gold, J.M., Gaspar, C.M., & Bennett, P. (2001). The efficiency of upright and upside-down face for identification. *Investigative Ophthalmology and Visual Sciences (Suppl.)*, 42.
- Sekuler, A.B., Gold, J.M., Murray, R.F., & Bennett, P.J. (2000). Visual completion of partly occluded objects : insights from behavioral studies. *Journal of Neuro-Ophthalmology*, 23, 165-168.
- Sheth, B.R., Sharma, J., Rao, S.C., & Sur, M. (1996). Orientation maps of subjective contours in visual cortex. *Science*, 274, 2110-2115.
- Shimojo, S., Kamitani, Y., & Nishida, S. (2001). Afterimage of perceptually filled-in surface. *Science*, 293, 1677-1680.
- Shore, D.I., & Enns, J.T. (1997). Shape completion times depends on the size of the occluded region. *Journal of Experimental Psychology : Human, Perception and Performance*, 23, 980-998.
- Sillito, A.M., & Jones, H.E. (1996). Context dependent interactions and visual processing in V1. *Journal of Physiology (Paris)*, 90, 205-209.
- Sillito, A.M., Jones, H.E., Gerstein, G.L., & West, D.C. (1994). Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature*, 369, 479-482.
- Silveira, L.C.L., & Perry, V.H. (1991). The topography of magnocellular projecting ganglion cells in the primate retina. *Neuroscience*, 40, 217-237.
- Singer, W. (1999). Time as coding space ? *Current Opinion in Neurobiology*, 9, 189-194.
- Singer, W., & Gray, C.M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, 18, 555-586.
- Solomon, J.A. (2002). Noise reveals visual mechanisms of detection and discrimination. *Journal of Vision*, 2, 105-120.
- Solomon, J.A., & Pelli, D.G. (1994). The visual filter mediating letter identification. *Nature*, 369, 395-397.
- Somers, D.C., Nelson, S.B., & Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience*, 15, 5448-5465.
- Spillman, L. (1997). Guest editorial : colour in a larger perspective the rebirth of gestalt psychology. *Perception*, 26, 1341-1352.
- Spillman, L., & De Weerd, P. (2003). Surface completion : psychophysics and neurophysiology of texture fading and filling in. In L. Pessoa & P. De Weerd (Eds.), *Filling in : from perceptual completion to skill learning*. Oxford : Oxford University Press.
- Spillman, L., & Kurtenbach, A. (1992). Dynamic noise backgrounds facilitate target

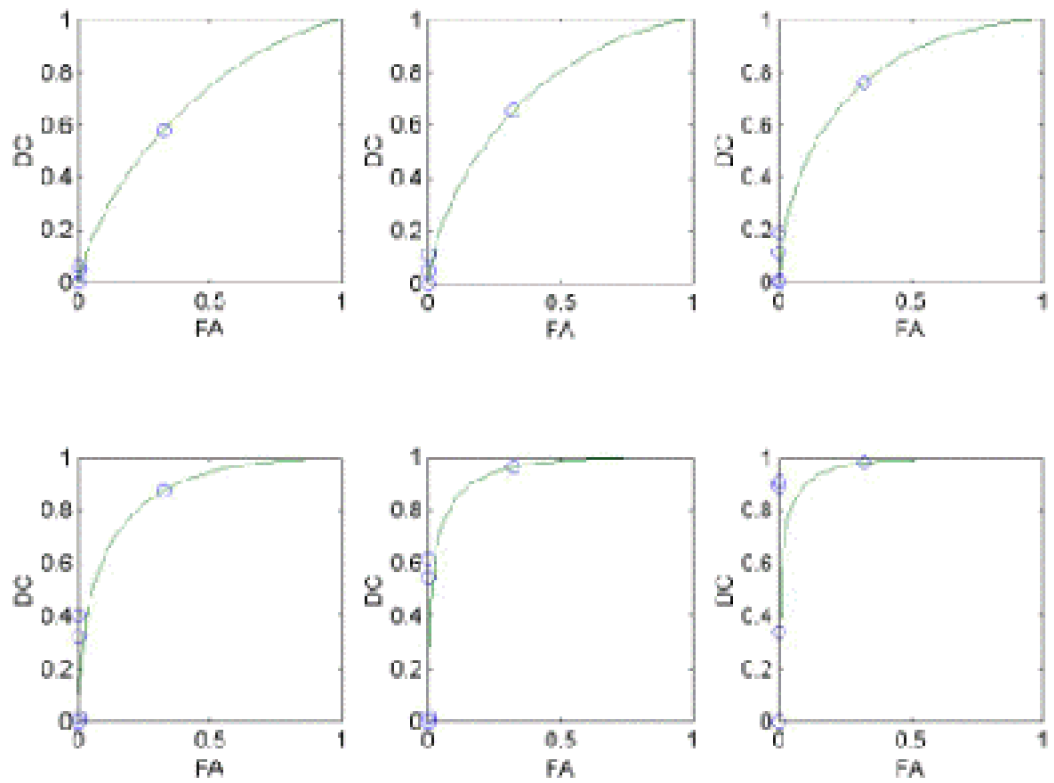
- fading. *Vision Research*, 32, 1941-1946.
- Spillman, L., & Werner, J.S. (1996). Long-range interactions in visual perception. *Trends Neurosciences*, 19, 428-434.
- Stoper, A.E., & Mansfield, J.G. (1978). Metacontrast and paracontrast suppression of a contourless area. *Vision Research*, 18, 1669-1674.
- Stuart, G., McAnally, K.I., & Castles, A. (2001). Can contrast sensitivity functions in dyslexia be explained by inattention rather than a magnocellular deficit ? *Vision Research*, 41, 3205-3211.
- Sturzel, F., & Spillman, L. (2001). Texture fading correlates with stimulus salience. *Vision Research*, 41, 2969-2977.
- Sugita, Y. (1999). Grouping of image fragments in primary visual cortex. *Nature*, 401, 269-272.
- Tjan, B.S., Braje, W.L., Legge, G.E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, 35, 3053-3069.
- Thomas, J. P. & Knoblauch, K. (1998) What do viewers look for when detecting a luminance pulse. *Investigative Ophthalmology and Visual Sciences (Suppl.)*, 39.
- Thompson, E., Noë, A., & Pessoa, L. (1999). Perceptual completion : a case of study in phenomenology and cognitive science. In Petitot, J. Roy, J-M., Pachoud, B., & Varela, F. (Eds.) *Naturalizing phenomenology : issues in contemporary phenomenology and cognitive science*. Stanford : Stanford University Press.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520-522.
- Tieger, T., & Ganz, L. (1979). Recognition of faces in the presence of two-dimensionnal sinusoidal masks. *Perception & Psychophysics*, 26, 163-167.
- Tootell, R.B.H., Hadjikhani, N.K., Vanduffel, W., Liu, A.K., Mendola, J.D., Sereno, M.I., & Dale, A.M. (1998). Functionnal analysis of primary visual cortex (V1) in humans. *Proceedings of the National Academy of Science of the UnitedStates of America*, 95, 811-817.
- Turvey, M.T. (1973). On peripheral and central processes in vision : inferences from an information processing analysis of masking with patterned stimuli. *Psychological Review*, 81, 1-52.
- Ullman, S. (1984). Visual routines. *Cognition*, 18, 97-159.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549-586). Cambridge, MA: MIT Press.
- Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature*, 394, 179-182.
- Vidnyanszky, Z., Papathomas, T.V., & Julesz, B. (2001). Contextual modulation of orientation discrimination is independent of stimulus processing time. *Vision Research*, 41, 2813-2817.
- Von der Heydt, R. (1995). Form analysis in visual cortex. In Gazzaniga M. (Ed.), *The cognitive neuroscience*. Cambridge, MA : MIT Press.

- Von der Heydt, R., & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex: lines of pattern discontinuities. *Journal of Neuroscience*, *9*, 1731-1748.
- Zucker, S.W., & Davis, S. (1988). Points and endpoints : a size /spacing constraint for dot grouping. *Perception*, *17*, 229-247.
- Wade, N.J. (1998). *A natural history of vision*. Cambridge (MA) : MIT Press.
- Walls, G. (1954). The filling-in process. *American Journal of Optometry*, *31*, 329-340.
- Wandell, B. (1995). *Foundations of vision*. Sunderland, MA : Sinauer.
- Watanabe & Rodieck (1989). Parasol and midget ganglion cells of the primate retina. *Journal of Comparative Neurology*, *289*, 434-454.
- Watson, A.B. (1979). Probability summation over time. *Vision Research*, *19*, 515-522.
- Watson, A.B., & Pelli, D. (1983). QUEST : a bayesian adaptative psychometric method. *Perception & Psychophysics*, *33*, 113-120.
- Watson, A.B. & Rosenholtz, R. (1997). A Rorschach test for visual classification strategies. *Investigative Ophthalmology and Visual Sciences (Suppl.)*, *38*.
- Welchman, A.E., & Harris, J.M. (2003). Is neural filling in necessary to explain the perceptual completion of motion and depth information ? *Proceedings of the Royal Society of London*, *270*, 83-90.
- Wilson, H.R., Levi, D., Maffei, L., Rovamo, J., & DeValois, R. (1990). In L. Spillman & J.S. Werner (Eds.), *The neurophysiological foundations*. San Diego : Academic Press.
- Wörgötter, F., & Eysel, U.T. (2000). Context, state and the receptive fields of striatal cortex cells. *Trends in Neuroscience*, *23*, 497-503.
- Yen, S. C., & Finkel, L.H. (1998). Extraction of perceptually salient contours by striate cortical networks. *Vision Research*, *38*, 719-741.
- Ziebell, O., & Nothdurft, H.C. (1999). Cueing and pop-out. *Vision Research*, *39*, 2113-2125.
- Zipser, K., Lamme, V.A.F., & Schiller, P.H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, *16*, 7376-7389.
- Zucker, S.W., & Davies, S. (1988). Points and endpoints : a size spacing constraint for dot grouping. *Perception*, *17*, 229-247.

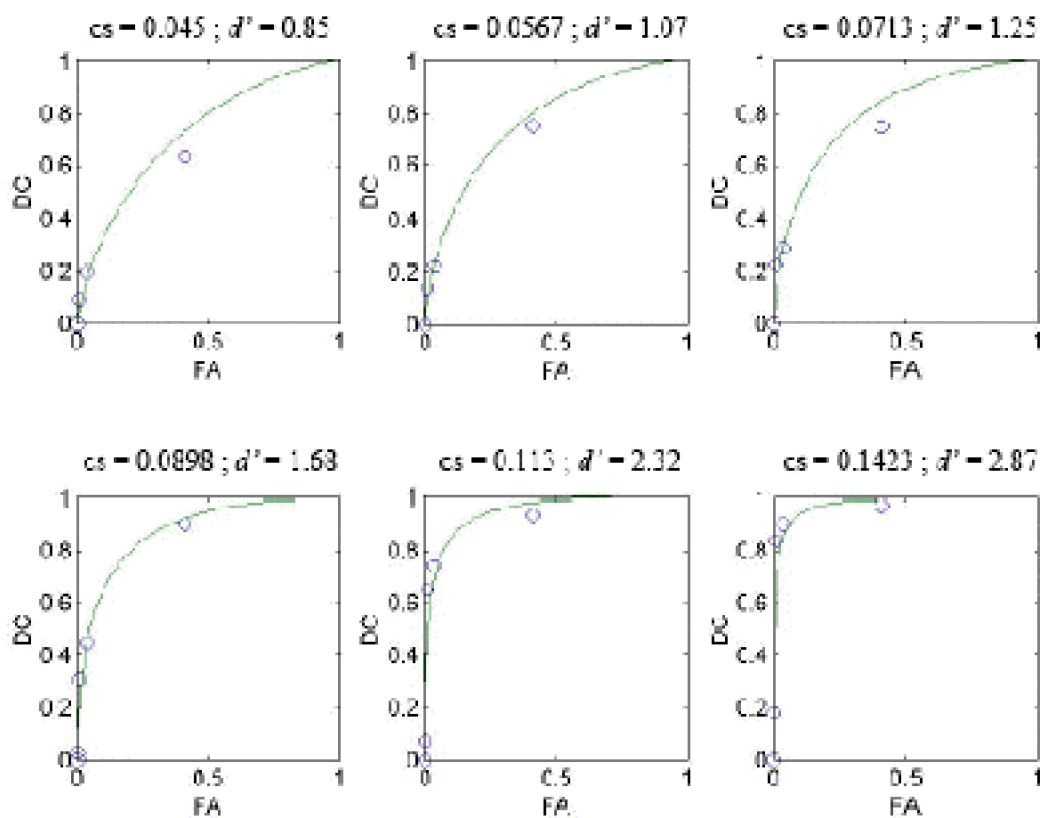
Annexes

Annexe 1 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour un signal de 3 deg et un temps de présentation de 400 ms.

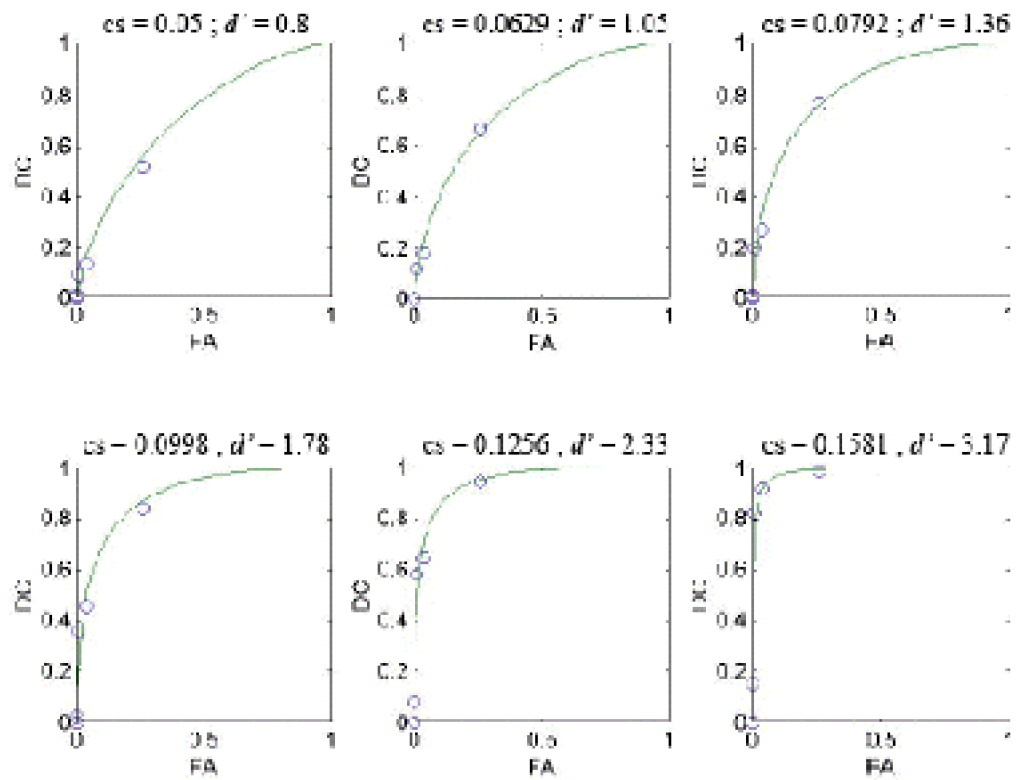
Les Figures présentées dans cette annexe représentent les courbes ROC déterminées pour chaque observateur et pour chaque condition expérimentale. Chaque annexe comporte 6 figures (une figure pour chaque contraste du signal ou cs), la première courbe ROC (en haut à gauche) concerne le contraste du signal le plus faible, et la dernière courbe (en bas à droite) concerne la valeur la plus élevée de contraste du signal. La courbe est déterminée par les réponses données par l'observateur (les cercles qui sont positionnées sur la courbe). Le taux de détection correcte (DC) est placé en ordonné et le taux de fausse alarme (FA) est placé en abscisse.



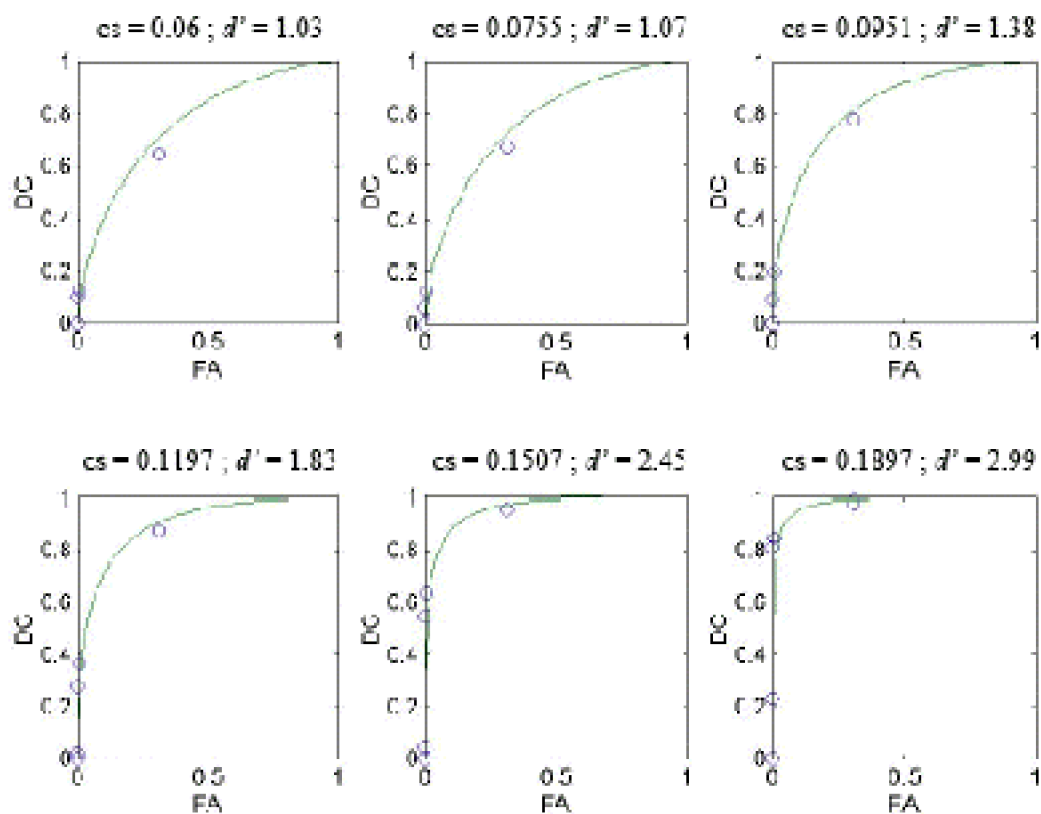
Annexe 1.1 : Courbes ROC de l'observateur FD pour un contraste du bruit de 0.5



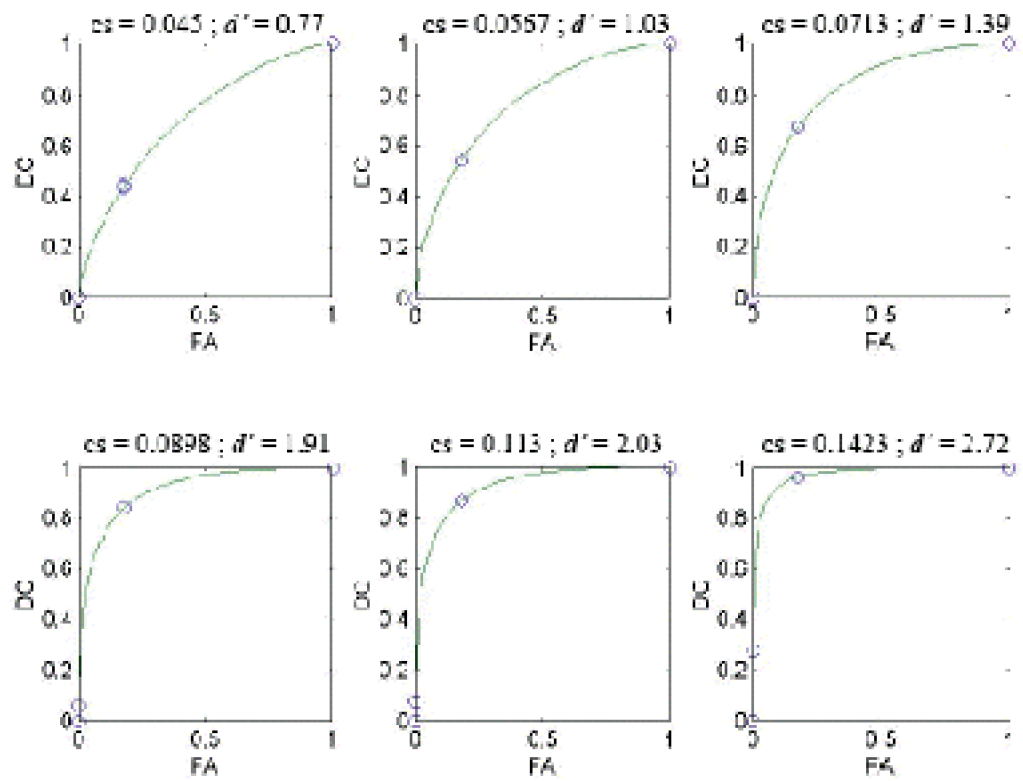
Annexe 1.2 : Courbes ROC de l'observateur FD pour un contraste du bruit de 0.6.



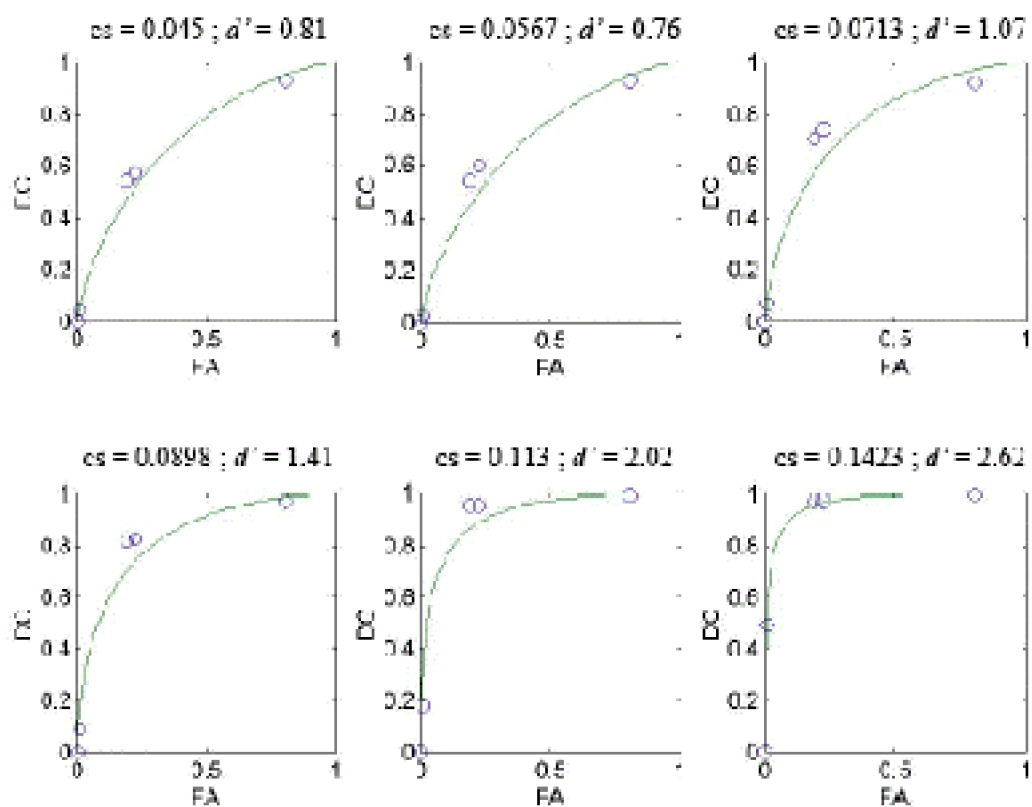
Annexe 1.3 : Courbes ROC de l'observateur FD pour un contraste du bruit de 0.7.



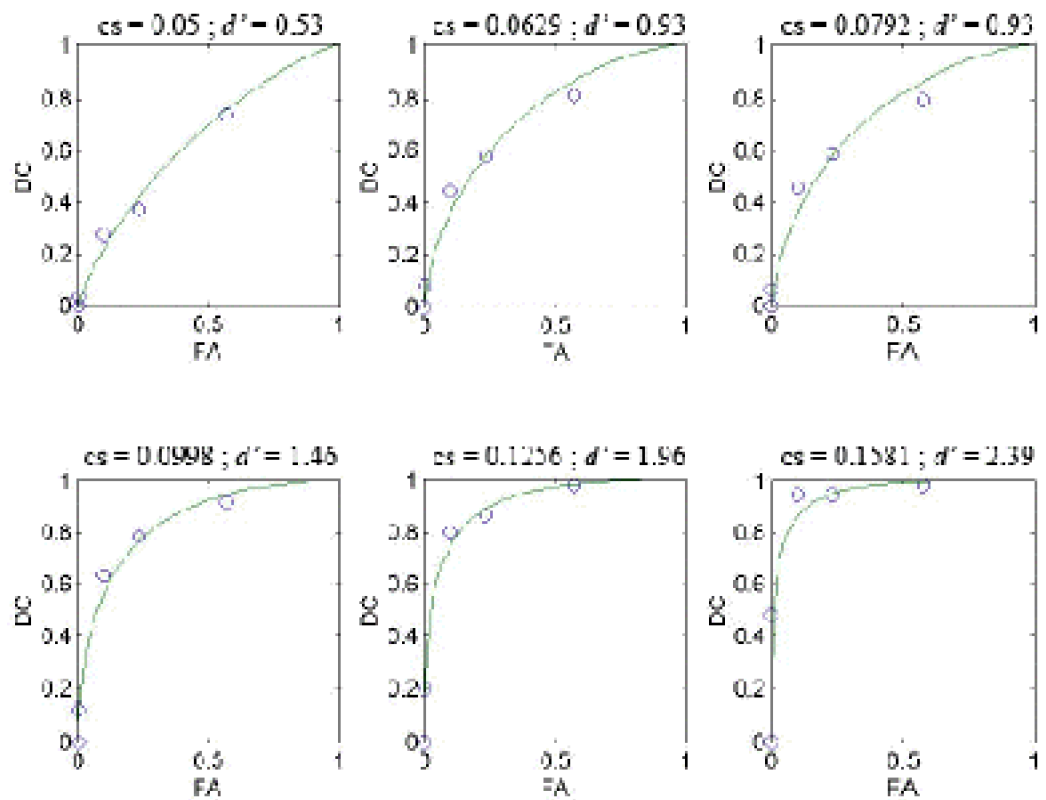
Annexe 1.4 : Courbes ROC de l'observateur FD pour un contraste du bruit de 0.8.



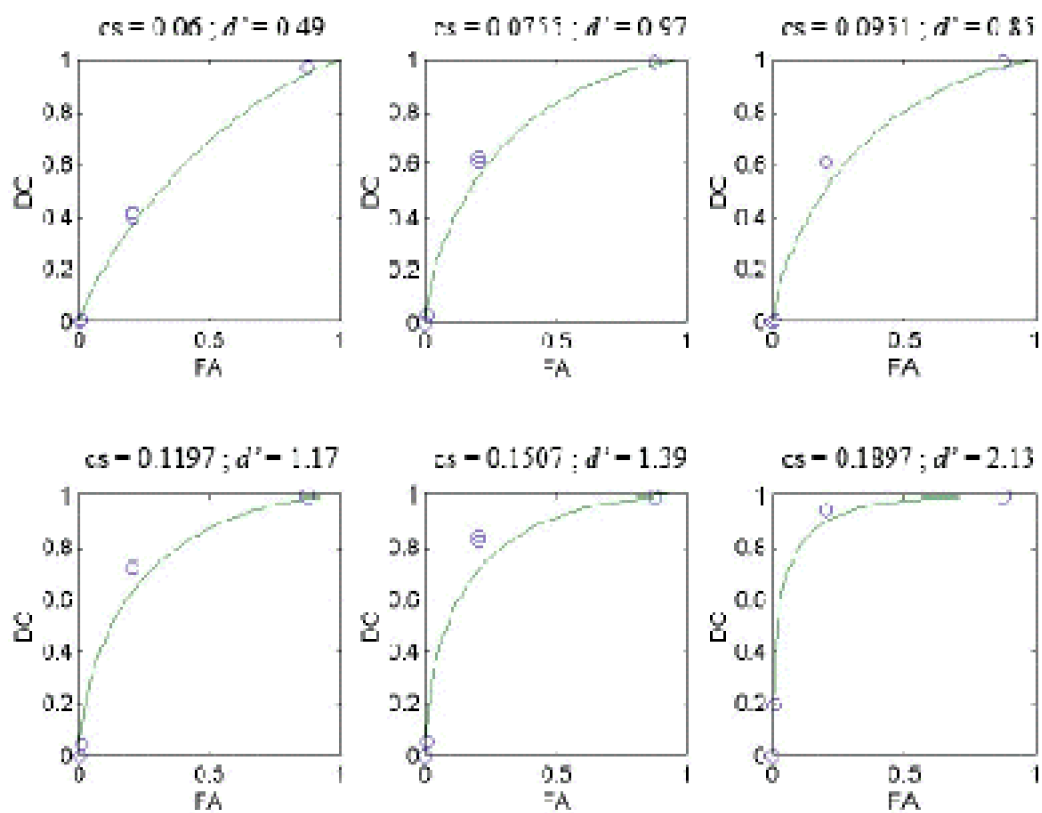
Annexe 1.5 : Courbes ROC de l'observateur SD pour un contraste du bruit de 0.5.



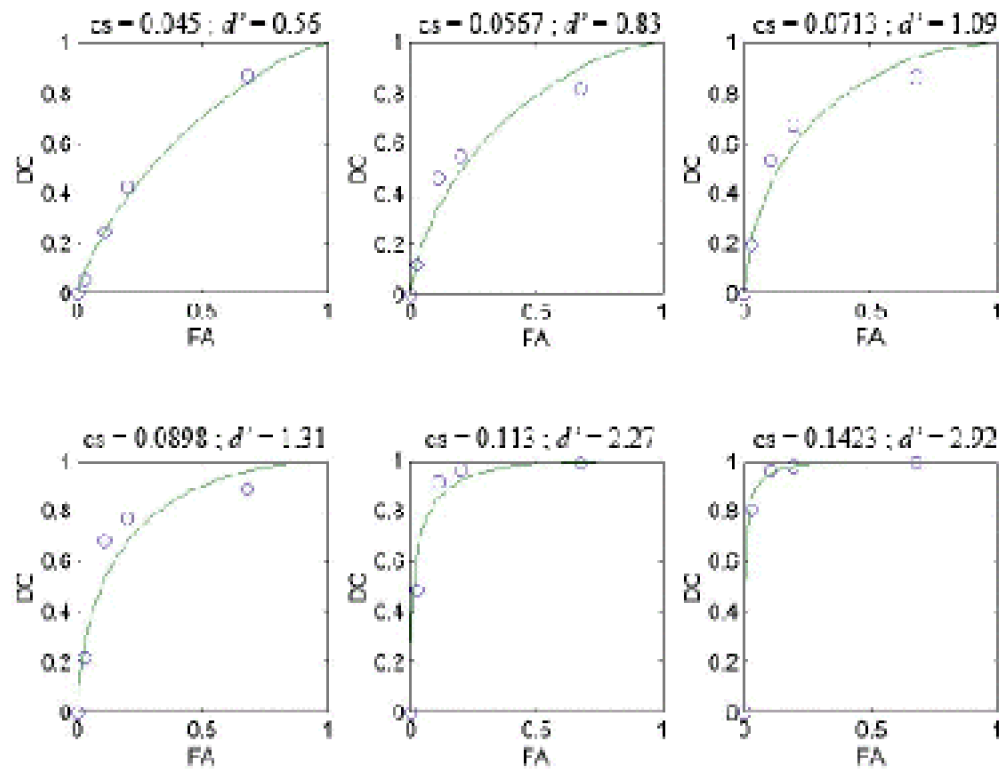
Annexe 1.6: Courbes ROC de l'observateur SD pour un contraste du bruit de 0.6.



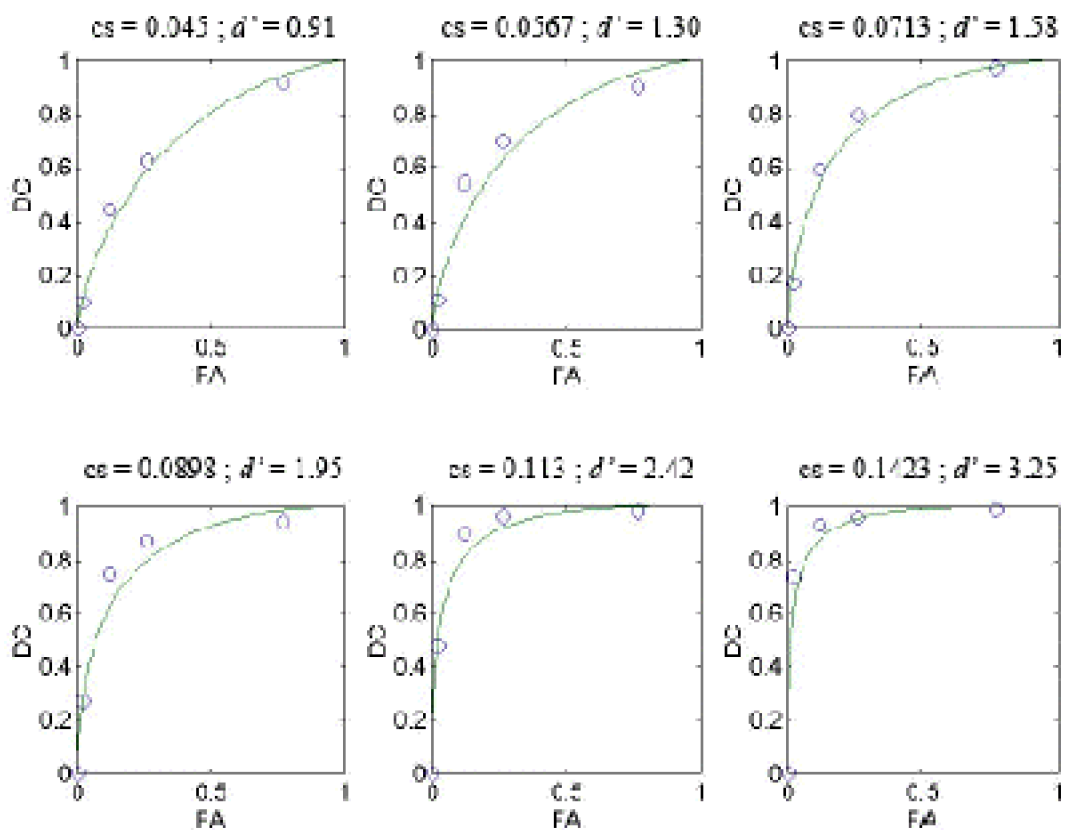
Annexe 1.7: Courbes ROC de l'observateur SD pour un contraste du bruit de 0.7.



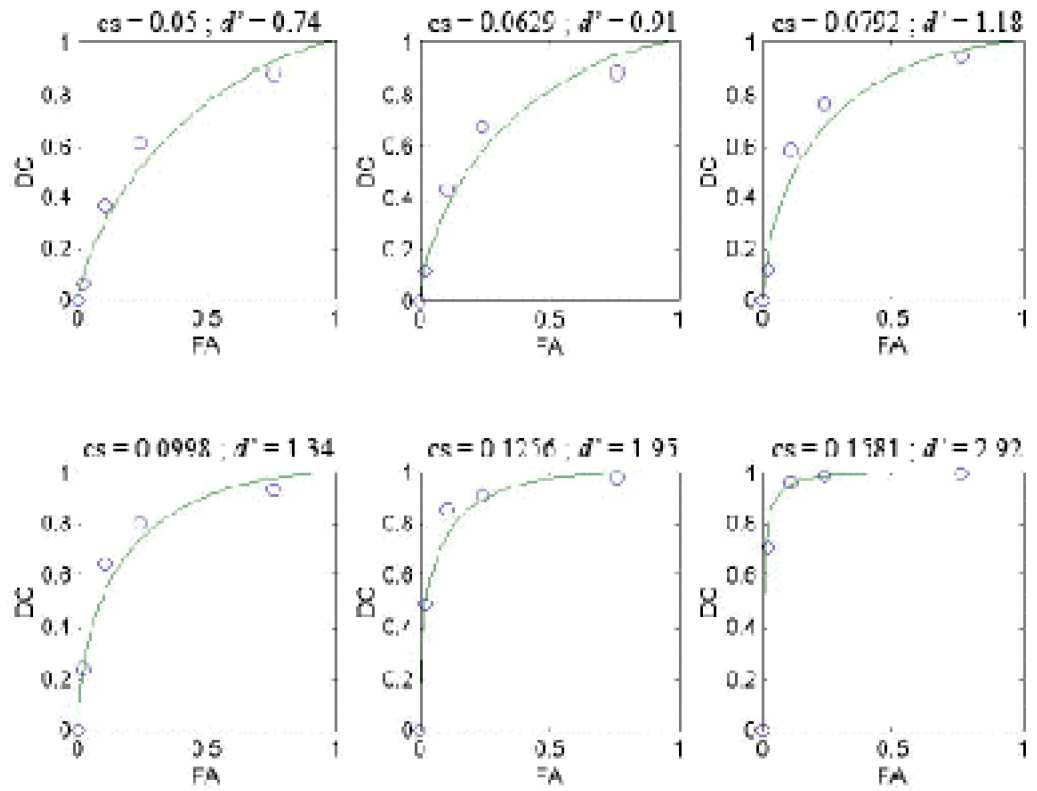
Annexe 1.8: Courbes ROC de l'observateur SD pour un contraste du bruit de 0.8.



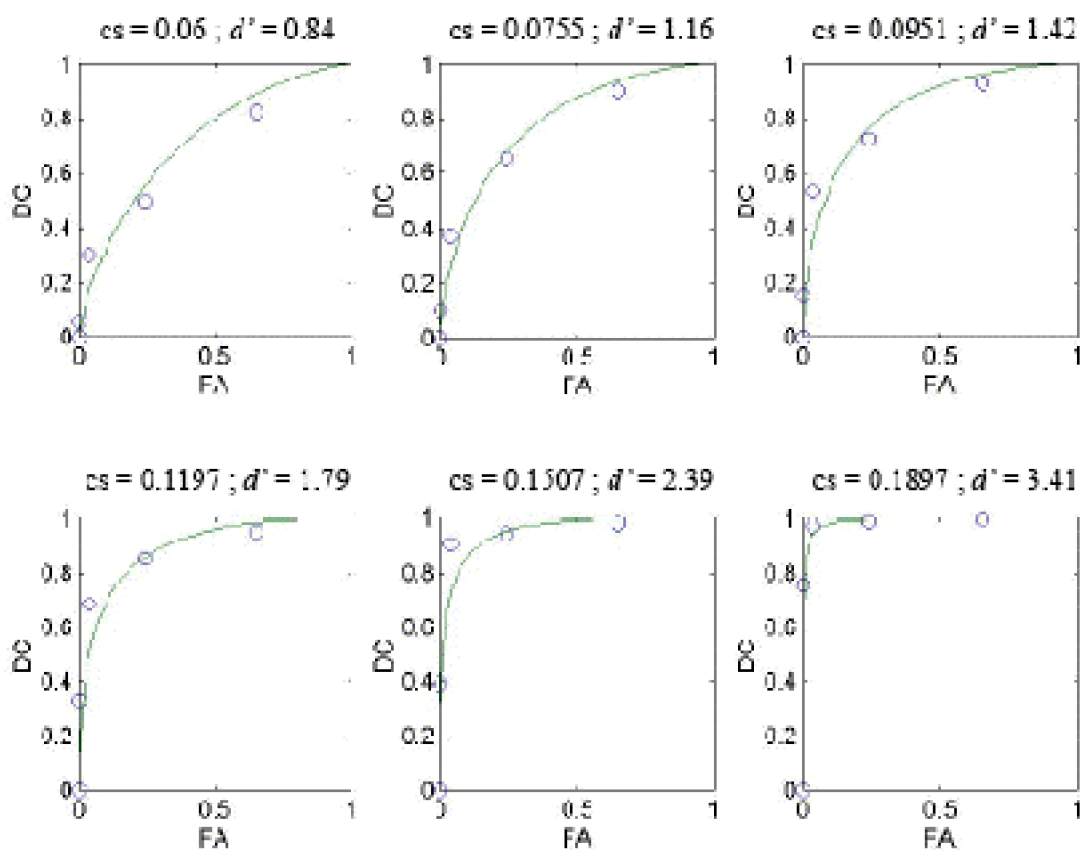
Annexe 1.9: Courbes ROC de l'observateur SH pour un contraste du bruit de 0.5.



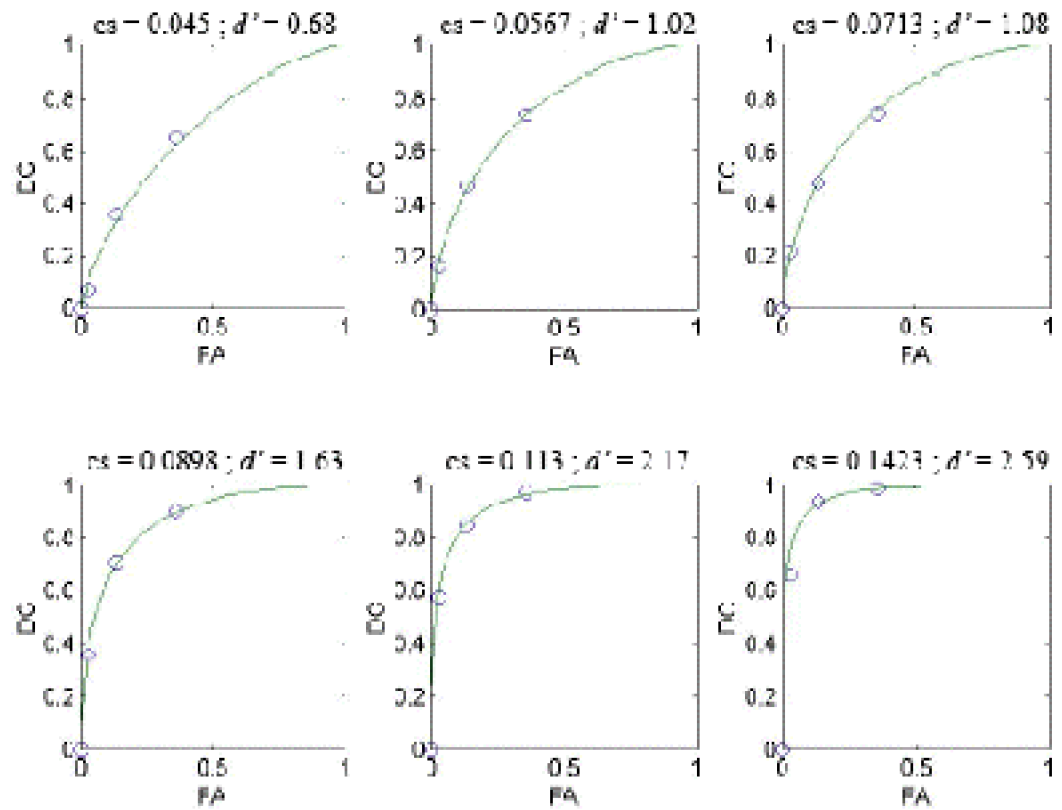
Annexe 1.10 : Courbes ROC de l'observateur SH pour un contraste du bruit de 0.6.



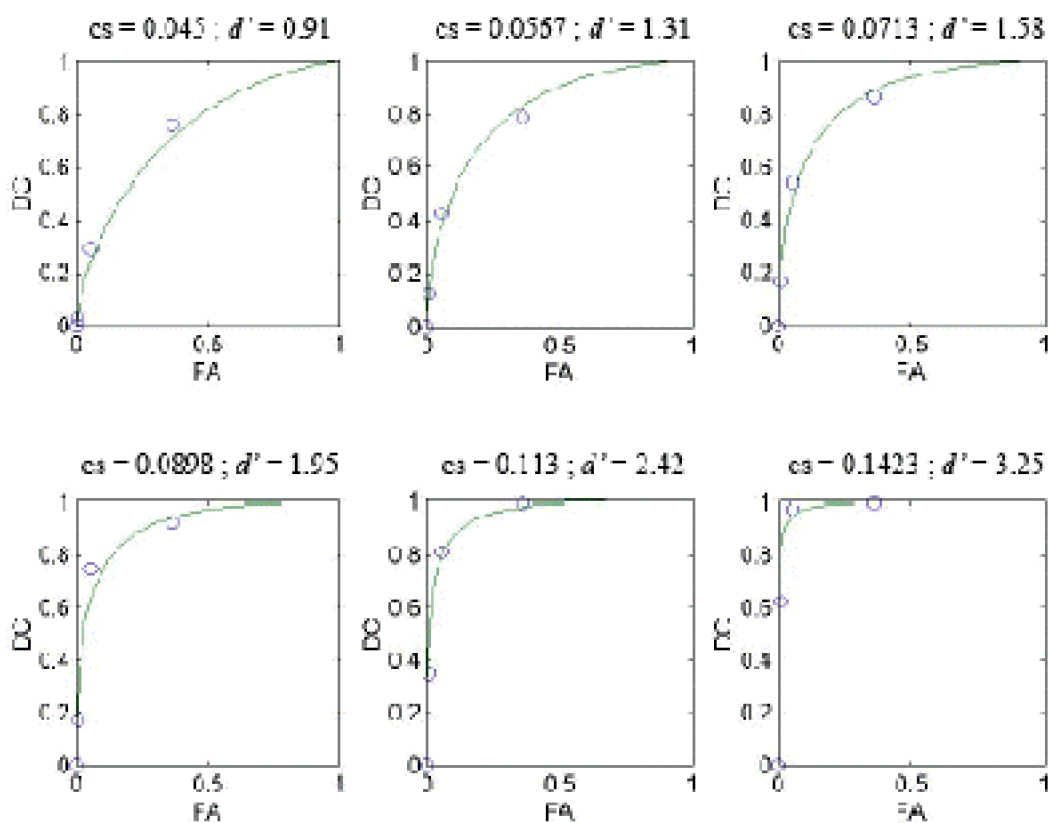
Annexe 1.11 : Courbes ROC de l'observateur SH pour un contraste du bruit de 0.7.



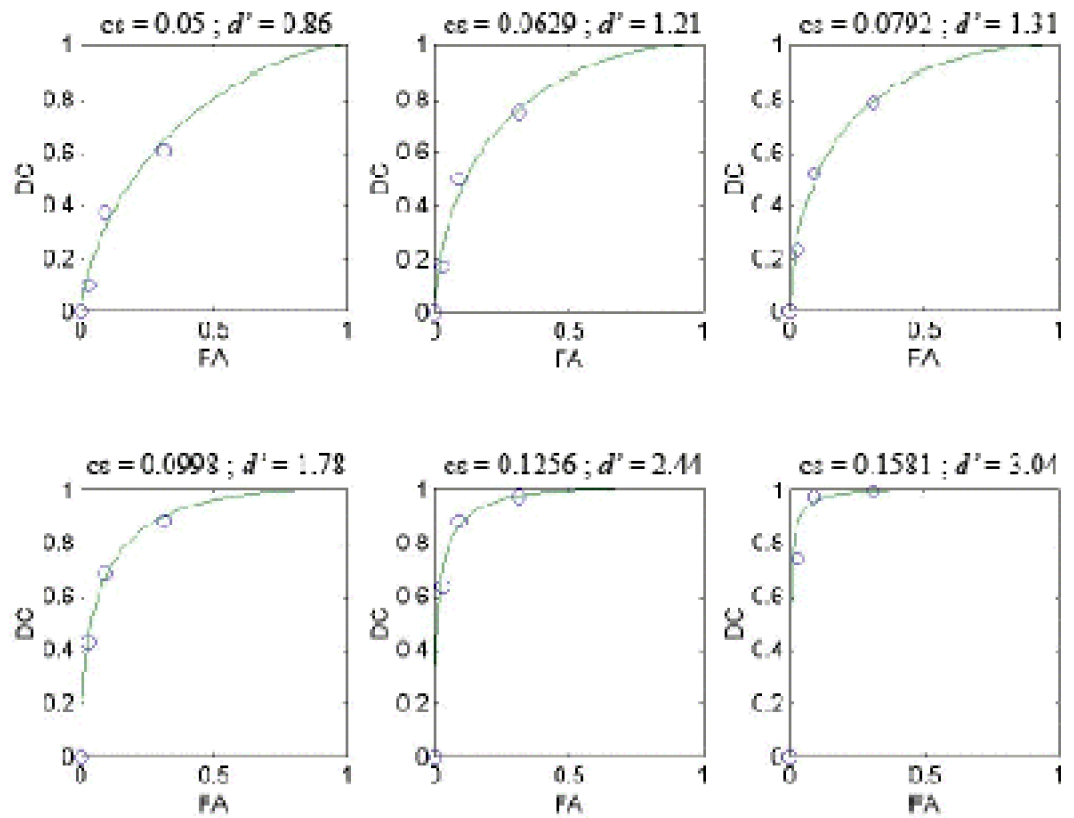
Annexe 1.12 : Courbes ROC de l'observateur SH pour un contraste du bruit de 0.8.



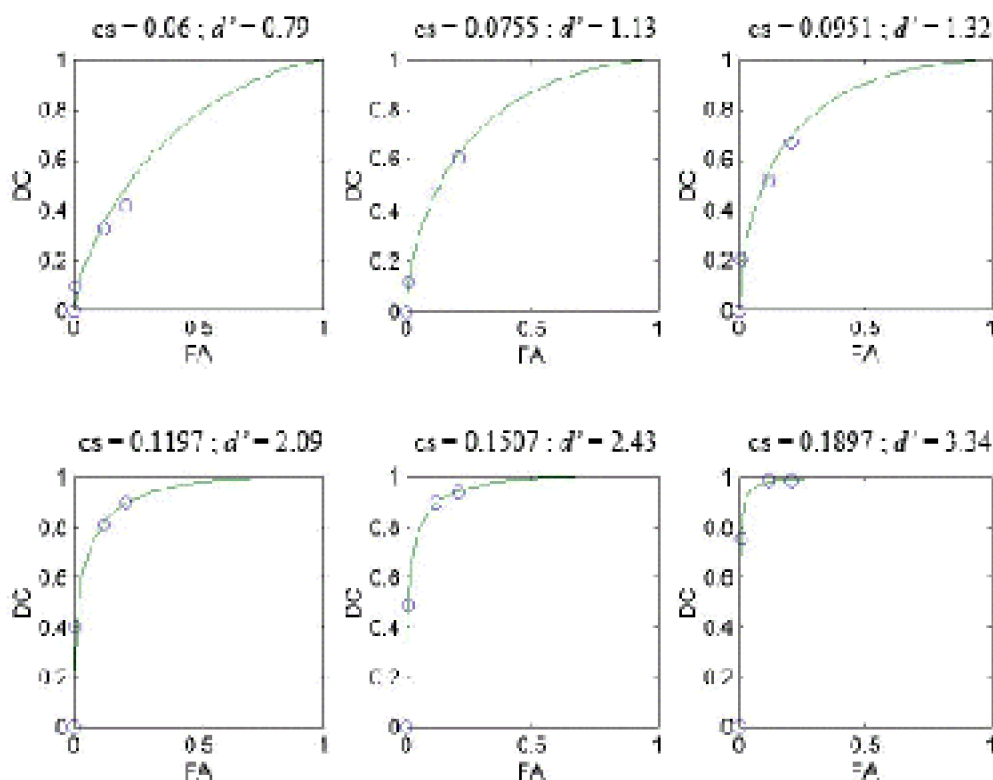
Annexe 1.13 : Courbes ROC de l'observateur VB pour un contraste du bruit de 0.5.



Annexe 1.14 : Courbes ROC de l'observateur VB pour un contraste du bruit de 0.6.

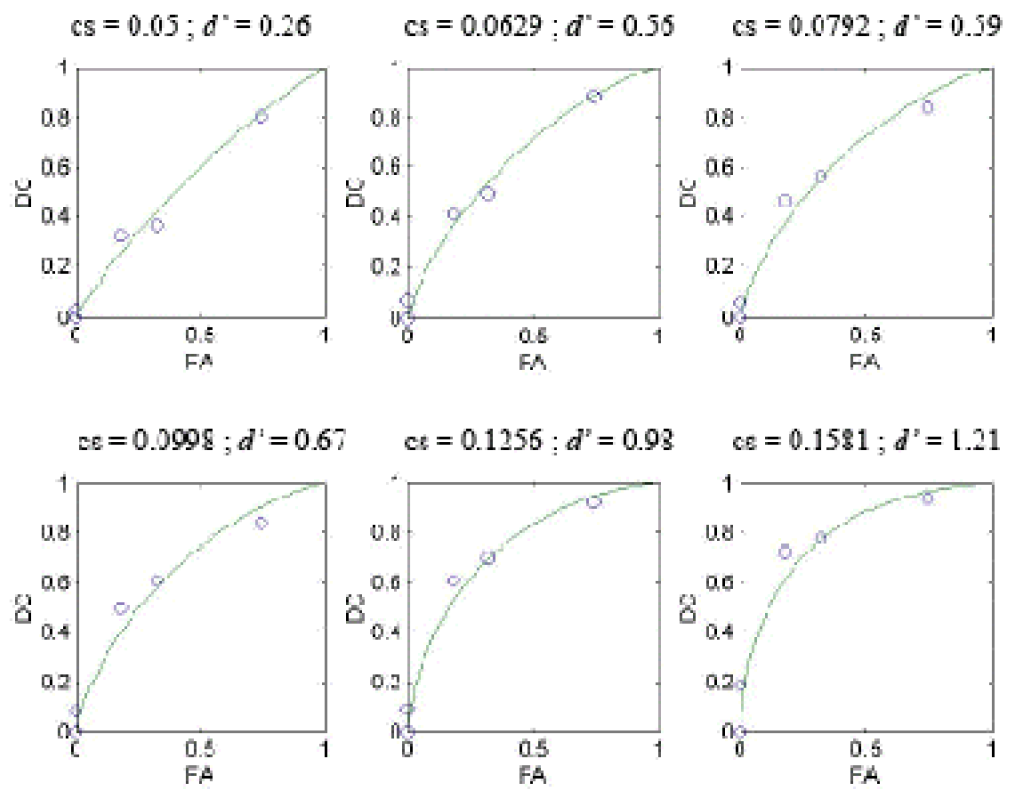


Annexe 1.15 : Courbes ROC de l'observateur VB pour un contraste du bruit de 0.7.

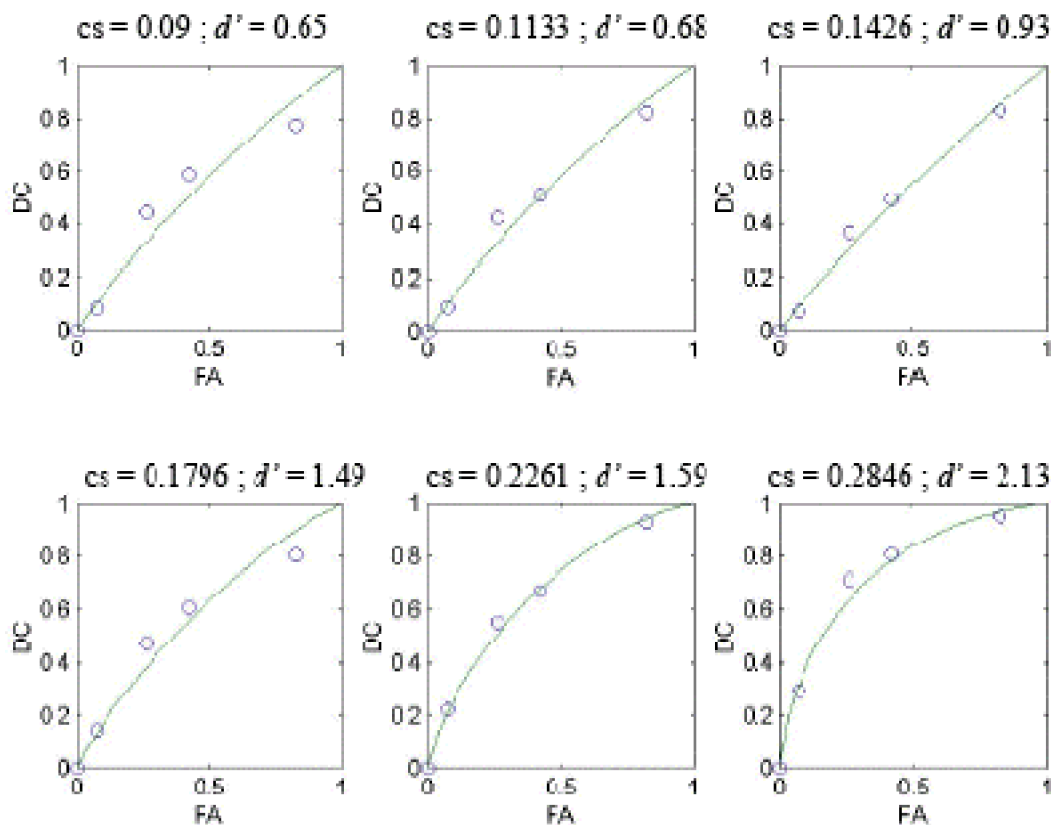


Annexe 1.16 : Courbes ROC de l'observateur VB pour un contraste du bruit de 0.8.

Annexe 2 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour un temps de présentation de 40 ms.

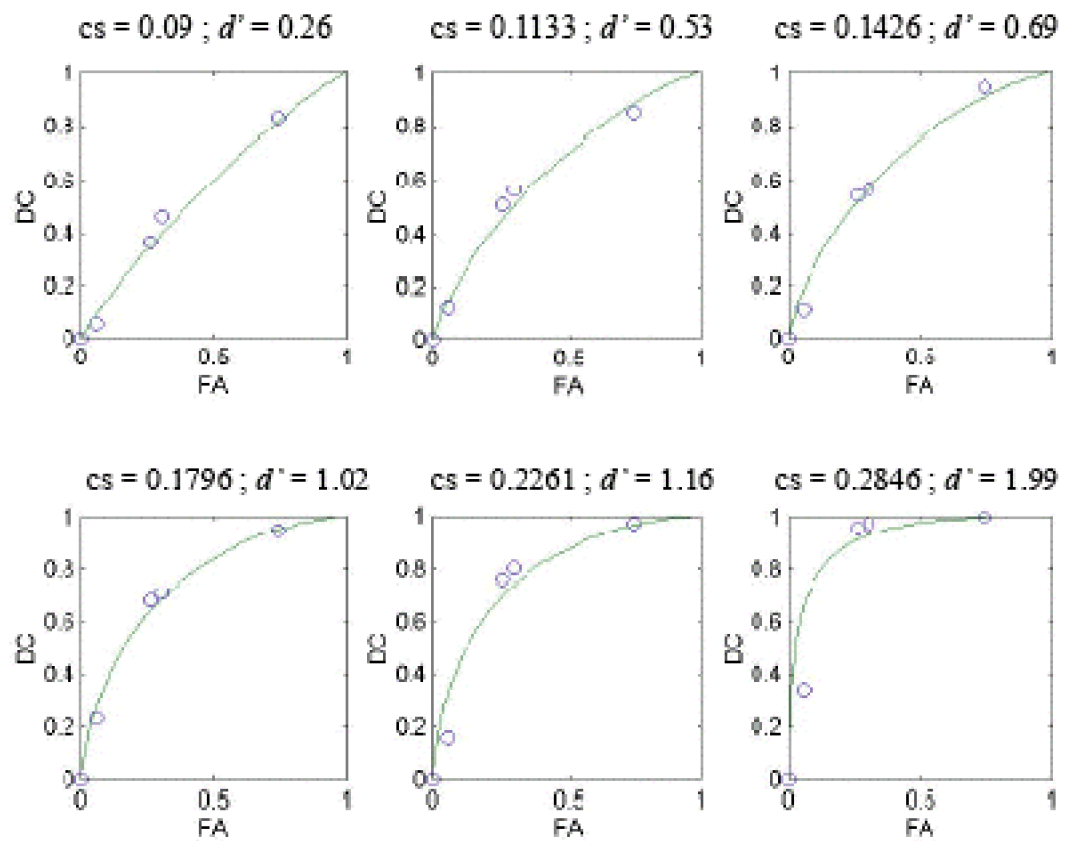


Annexe 2.1 : Courbes ROC de l'observateur FD.

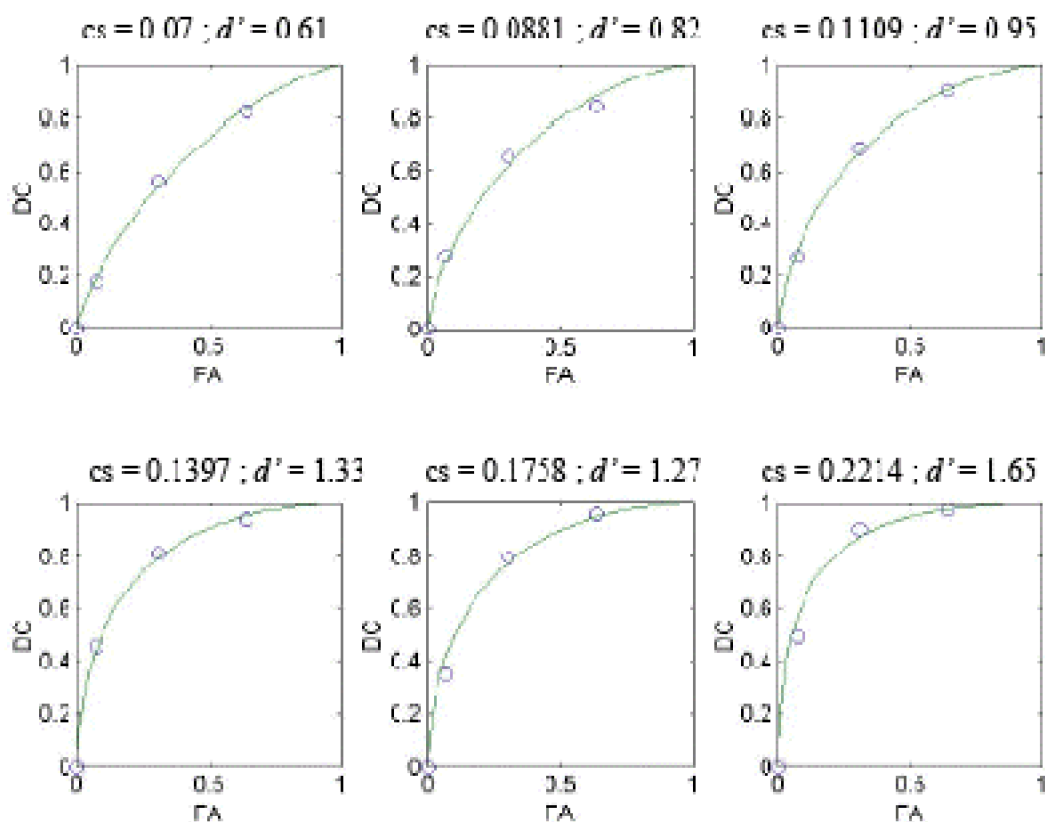


Annexe 2.2 : Courbes ROC de l'observateur SH.

Annexe 3 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour une taille du signal de 1 deg.

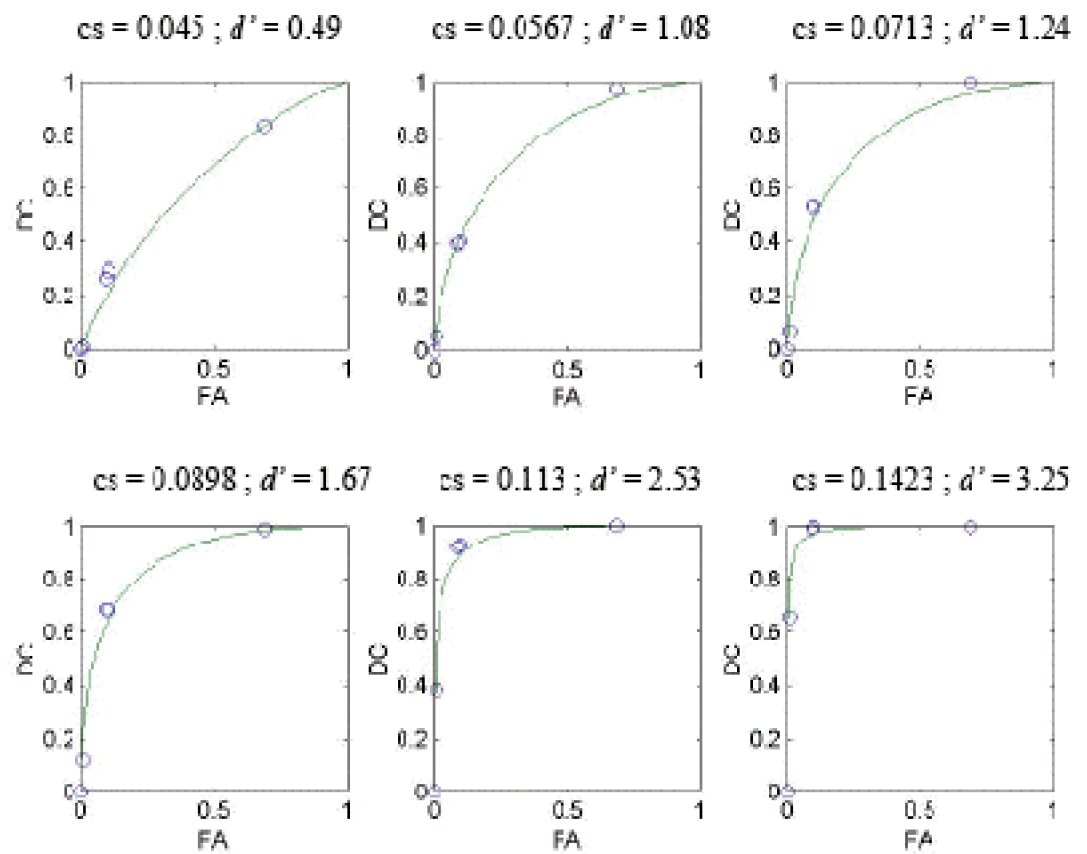


Annexe 3.1 : Courbes ROC de l'observateur FD.

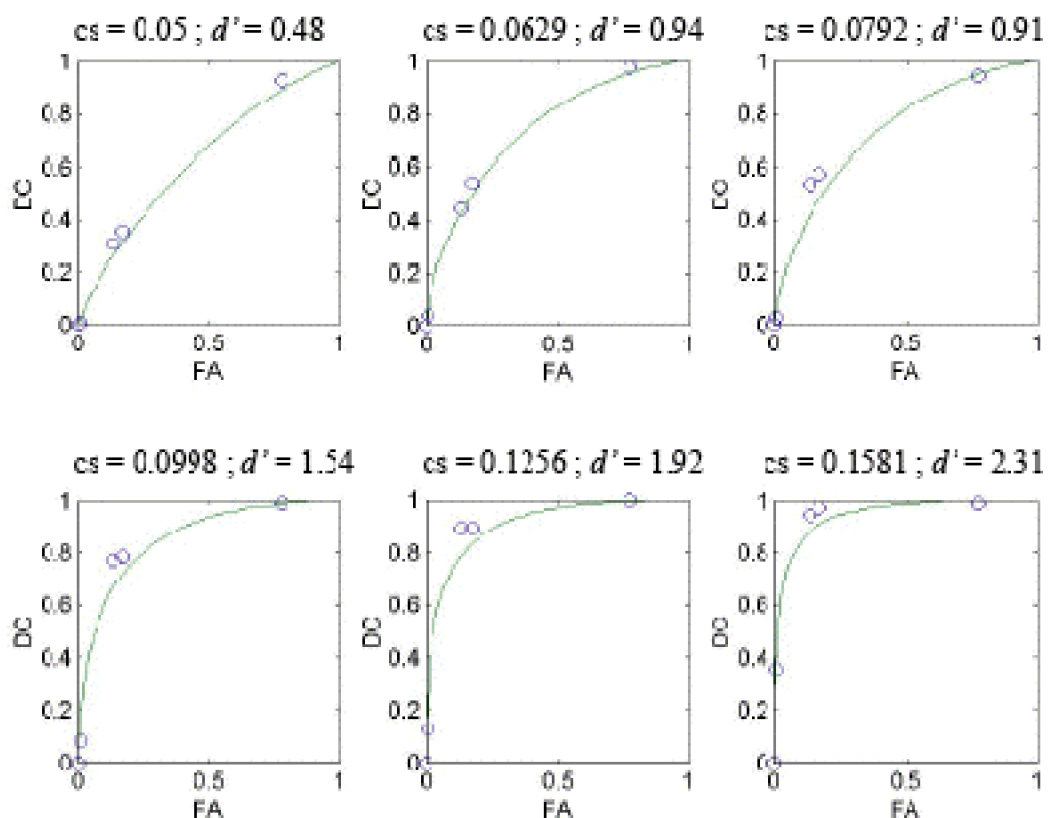


Annexe 3.2 : Courbes ROC de l'observateur VB.

Annexe 4 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour une taille du signal de 6 deg.

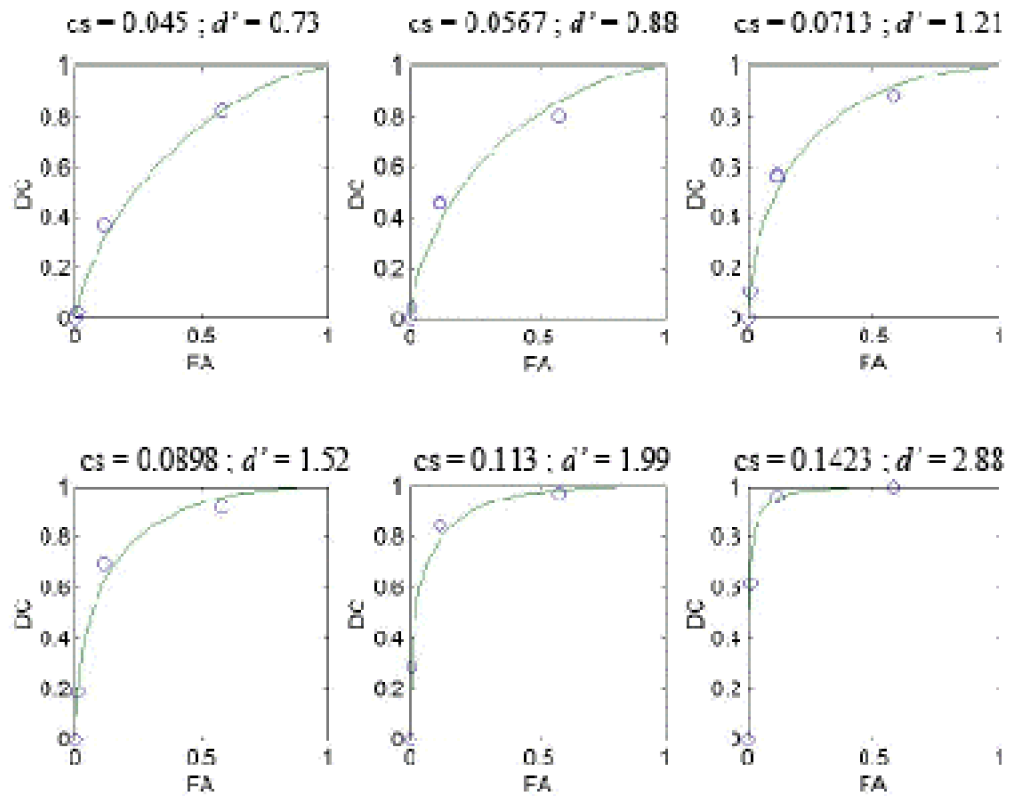


Annexe 4.1 : Courbes ROC de l'observateur FD.

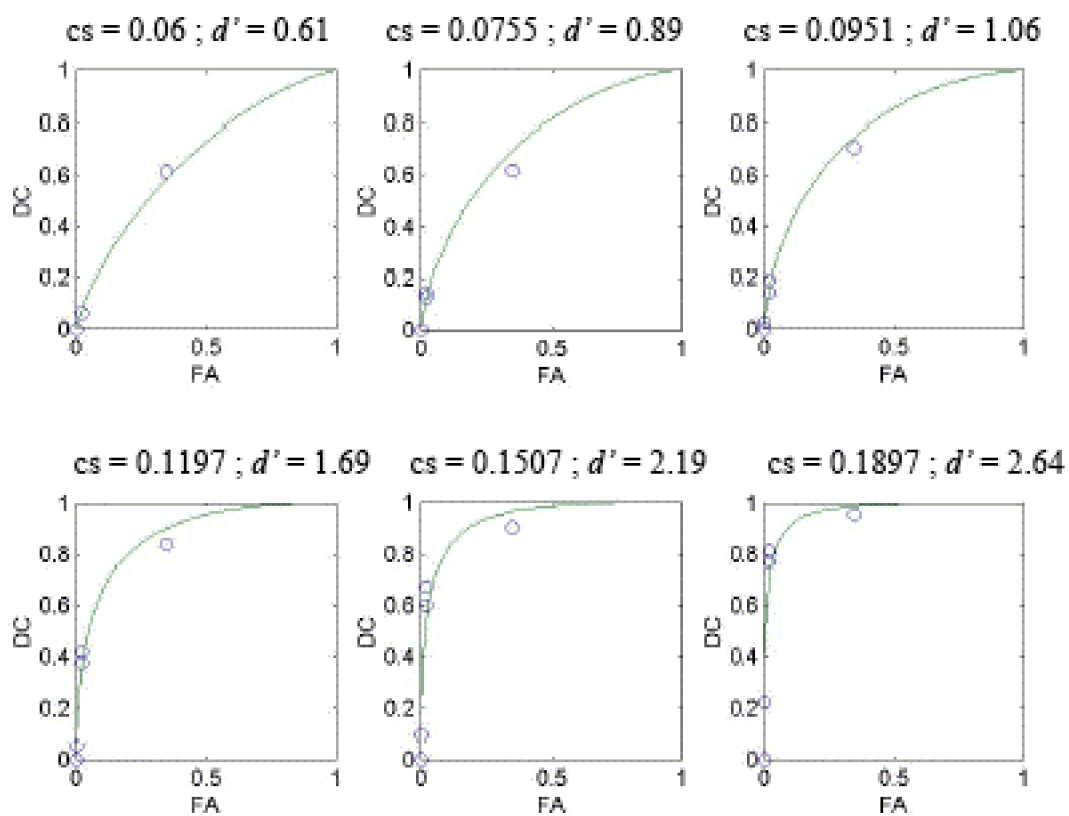


Annexe 4.2 : Courbes ROC de l'observateur SD.

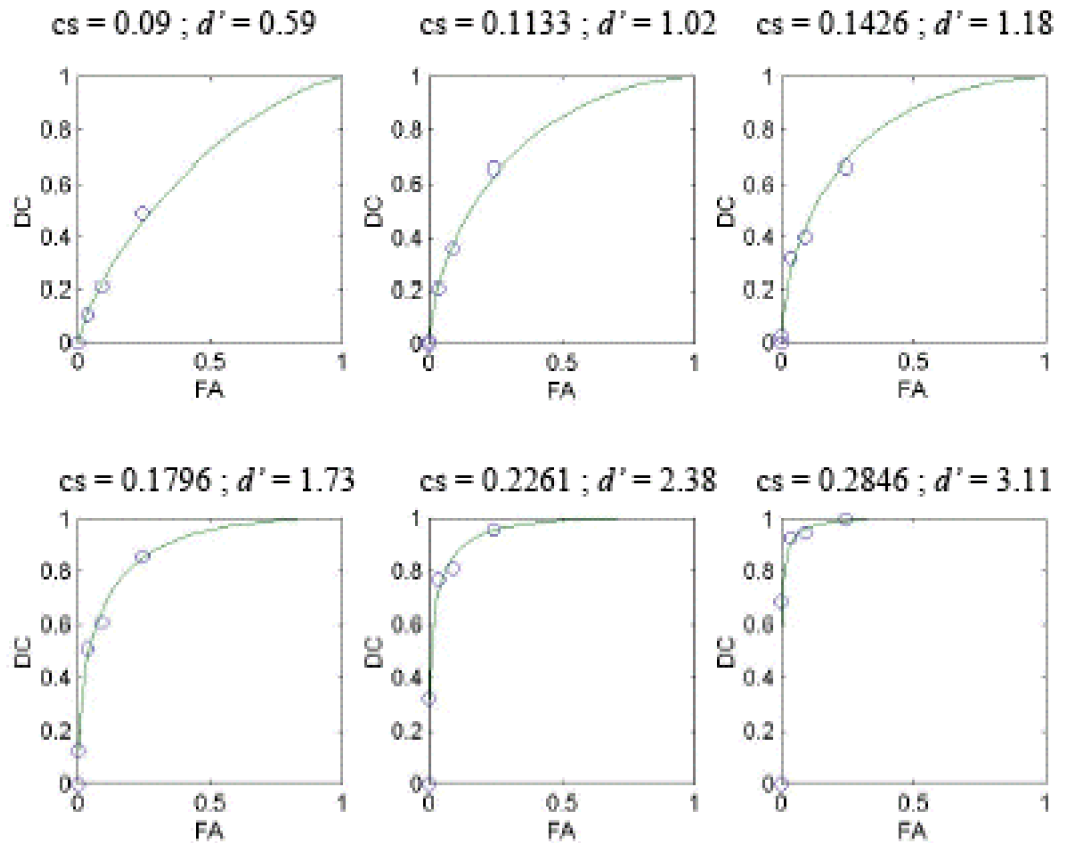
Annexe 5 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour une taille du bruit de 30 deg de l'observateur FD.



Annexe 6 : Représentation des courbes ROC réalisé dans l'expérience de détermination de seuil pour la tâche de détection où la consigne est imposée.



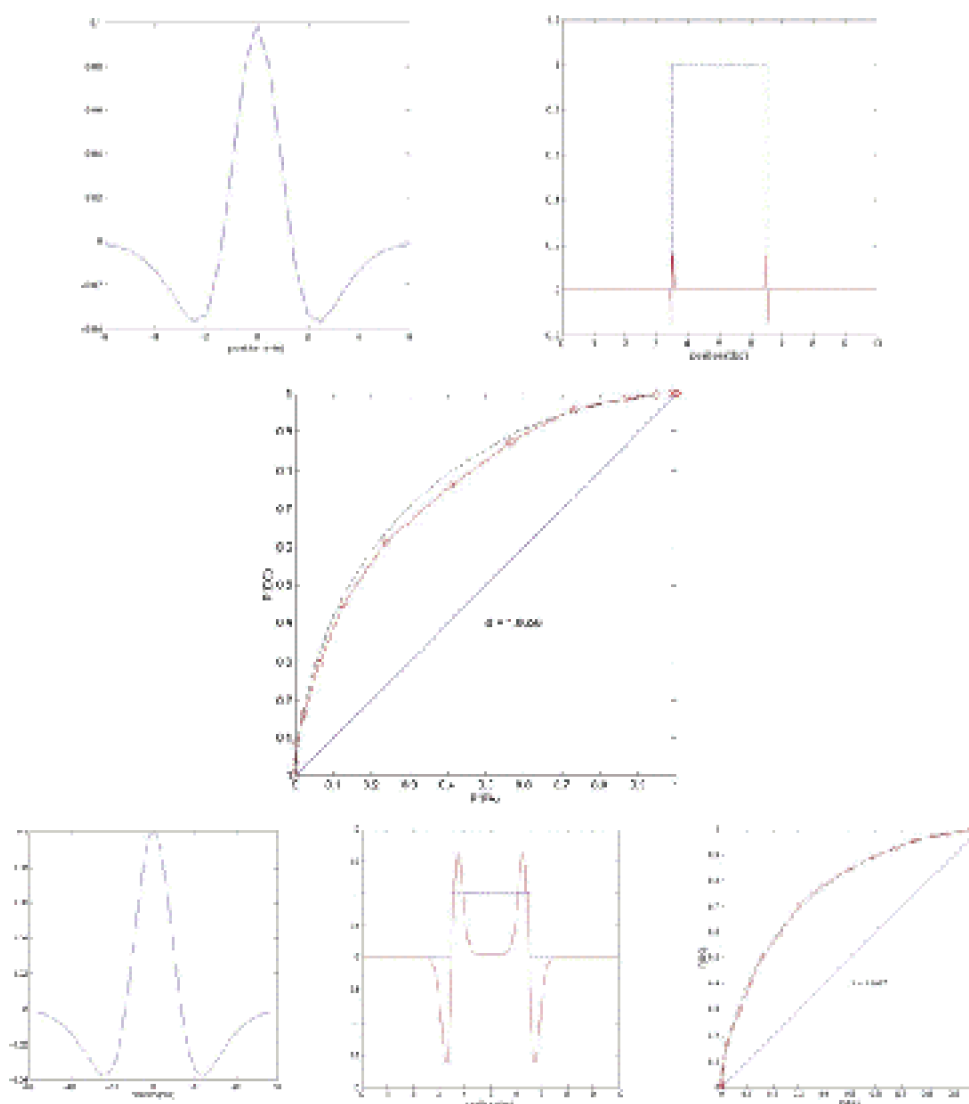
Annexe 6.1 : Courbes ROC de l'observateur FD.



Annexe 6.2 : Courbes ROC de l'observateur SG.

Annexe 7 : Résultats de l'expérience préliminaire pour l'observateur modèle.

Les trois figures représentent les résultats de l'expérience préliminaire de l'observateur modèle pour un centre/pourtour équilibré et de petite taille ($\square = 1$). Le graphe en haut à gauche présente la forme de la fonction DoG utilisée. La figure en haut à droite montre le prototype neuronal obtenu à la suite de la convolution entre le signal et la fonction DoG, le signal carré est représenté en pointillé. La figure du centre indique les résultats de l'indice d' , et présente l'ensemble des critères de réponse représenté dans un espace ROC. Les figures du bas représentent les résultats pour un centre/pourtour équilibré de grande taille ($\square = 10$).



Les résultats pour un centre/pourtour déséquilibré sont présentés en dessous, tout d'abord pour un pourtour inhibiteur représentant 80 % (les deux premières lignes) puis 50 % (les deux dernières lignes) de la zone excitatrice centrale. Les résultats pour une petite taille de champ récepteur ($\sigma = 1$) sont indiqués par les figures de gauche, et pour une grande taille de champ récepteur sont représentées par les figures de droite ($\sigma = 10$).

