

# **Représentations hiérarchiques et discriminantes pour la reconnaissance des formes, l'identification des personnes et l'analyse des mouvements dans les séquences d'images**

**par Nicolas THOME**  
Thèse de doctorat en informatique  
sous la direction de Serge MIGUET  
Par décision du jury, cette thèse n'est pas diffusable  
Soutenue le 11 juillet 2007

Devant un jury composé de : Serge MIGUET, professeur à l'université Lyon 2 Jenny BENOIS-PINEAU, professeur à l'université Bordeaux 1 Monique THONNAT, directeur de recherches à l'INRIA de Sophia-Antipolis Jean-Michel JOLION, professeur à l'INSA de Lyon Patrick PÉREZ, directeur de recherche à l'INRIA de Rennes Jean-Baptiste DUCATEZ, industriel



# Table des matières

<b>Résumé .</b>	1
<b>Summary . .</b>	3
<b>1. Introduction générale . .</b>	5
1.1. Contexte et historique des travaux de thèse . .	5
1.2. Organisation et contributions principales . .	5
1.2.1. Analyse du mouvement humain dans les séquences d'images . .	5
1.2.2. Lecture de plaques minéralogiques . .	5
<b>2. Lecture de plaques minéralogiques . .</b>	7
2.1. Problématique de la lecture de plaques minéralogiques par analyse de la vidéo . .	7
2.1.1 Approches récentes . .	7
2.2 Méthode proposée . .	7
2.2.1 Contribution . .	7
2.3 Détection de la plaque . .	7
2.3.1 Etat de l'art . .	8
2.3.2. Comparaison et approche retenue . .	8
2.3.3 Pré traitements . .	8
2.3.4 Algorithme de détection de la plaque . .	8
2.3.5 Extraction d'un ensemble de candidats pour la plaque . .	8
2.3.6 Suivi de la plaque . .	8
2.3.7 Post traitements . .	8
2.4 Binarisation de l'image de plaque extraite . .	8
2.4.1 Critère de détection de régions homogènes . .	8
2.4.2 Résultats . .	8
2.5 Extraction des caractères . .	8
2.5.1 Stratégie proposée . .	8
2.5.2 Etiquetage en composantes connexes . .	8

2.5.3 Connaissance <i>a priori</i> sur les caractères individuels . . . . .	9
2.5.4 Contraintes globales sur les caractères . . . . .	9
2.5.5 Analyse du rapport d'aspect . . . . .	9
2.5.6 Utilisation des projections d'histogrammes . . . . .	9
2.5.7 Détection de caractères fragmentés . . . . .	9
2.5.8 Discussion . . . . .	9
2.6 Reconnaissance des caractères . . . . .	9
2.6.1 Approche proposée . . . . .	9
2.6.2 Classification statistique . . . . .	9
2.6.3 Analyse structurelle . . . . .	9
2.7 Moyennage Temporel . . . . .	9
2.7.1 Détection et suivi d'un véhicule unique . . . . .	9
2.7.2 Mise à jour dynamique des reconnaissances au cours du temps . . . . .	9
2.8 Feedback entre les niveaux . . . . .	10
2.9 Résultats . . . . .	10
2.9.1 Performances de la détection de plaques . . . . .	10
2.9.2 Performances de la reconnaissance de plaques . . . . .	10
2.10 Conclusions et perspectives . . . . .	10
<b>3 Analyse du mouvement humain . . . . .</b>	<b>11</b>
3.1. Introduction . . . . .	11
3.2 La détection d'humains dans la vidéo . . . . .	11
3.2.1 Méthodes Bottom/Up . . . . .	11
3.2.2 Approches Top/Down . . . . .	11
3.3 Suivi d'humains dans la vidéo . . . . .	11
3.3.1 Méthodes 2D vs 3D . . . . .	12
3.3.2 Suivi avec/sans détection . . . . .	12
3.3.3. Stratégies de recherche . . . . .	12
3.3.4 Représentation adoptée et caractéristiques mises en correspondance . . . . .	12
3.4 Détection, étiquetage des membres et suivi d'humains . . . . .	12

3.4.1 Approches Top/Down . . . . .	12
3.4.2 Approches Bottom/Up . . . . .	12
3.5 Interprétation du mouvement humain dans la vidéo . . . . .	12
3.5.1 Méthodes discriminantes et génératives . . . . .	12
3.5.2 Identification d'humains par reconnaissance du mouvement . . . . .	12
3.6 Approche Retenue . . . . .	12
<b>4 Le suivi d'humains par modèle d'apparence articulé . . . . .</b>	13
4.1 Extraction d'objets . . . . .	13
4.1.1 Etiquetage des pixels mobiles . . . . .	13
4.1.2 Suppression des ombres . . . . .	13
4.1.3 Post traitements . . . . .	13
4.2 Mise en correspondance dynamique de régions . . . . .	13
4.2.1 Tirer profit de la cohérence temporelle . . . . .	14
4.3. Formation d'un modèle d'apparence articulé pour le suivi . . . . .	14
4.4 Détection et étiquetage des membres . . . . .	14
4.4.1 Généralités sur les graphes . . . . .	14
4.4.2 Approches existantes . . . . .	14
4.4.3 Génération d'un graphe modélisant le squelette humain . . . . .	14
4.4.4 Génération d'un graphe image . . . . .	14
4.4.5 Mise en correspondance de graphes . . . . .	14
4.4.6 Résultats de l'étiquetage . . . . .	14
4.5 Mise à jour du modèle d'apparence de chaque membre . . . . .	14
4.6 Utilisation du modèle d'apparence pour la reconnaissance . . . . .	14
4.7 Suivi lors des occlusions . . . . .	15
4.8 Résultats . . . . .	15
4.9 Conclusions et perspectives . . . . .	15
<b>5 Interprétation du mouvement humain : application à la détection de chutes . . . . .</b>	17
5.1 Contexte applicatif . . . . .	17
5.2. Etat de l'art . . . . .	17

5.3. Approche proposée . . . . .	17
5.4. Discriminer les poses debout/couché dans chaque image . . . . .	18
5.4.1 Extraction de caractéristiques . . . . .	18
5.4.2 Rectification métrique de l'image . . . . .	18
5.4.3 Relier les caractéristiques aux poses . . . . .	18
5.4.4 Simulation numérique . . . . .	18
5.5 Analyse de la séquence des poses avec un HHMM . . . . .	18
5.5.1 Architecture du modèle . . . . .	18
5.5.2 Paramètres du HHMM . . . . .	18
5.6 Résultats . . . . .	18
5.7 Conclusions et discussion . . . . .	18
<b>6 Conclusions et perspectives . . . . .</b>	<b>19</b>
6.1 Synthèse des travaux entrepris . . . . .	19
6.1.1 L'importance des approches hiérarchiques . . . . .	19
6.1.2 L'invariance et la spécificité des caractéristiques extraites . . . . .	19
6.1.3 La combinaison horizontale des stratégies . . . . .	19
6.1.4 Les approches génératives et discriminantes, et la coopération verticale . . . . .	19
6.2 Limites et perspectives de recherche . . . . .	20
<b>A Les descripteurs de Fourier . . . . .</b>	<b>21</b>
A.1 Description par tangente . . . . .	21
A.2 Représentation complexe . . . . .	21
A.3 Extension pour gérer des courbes non connexes . . . . .	21
<b>B Méthodes de classification statistique pour la reconnaissance des formes . . . . .</b>	<b>23</b>
B.1 Les Réseaux de Neurones artificiels . . . . .	23
B.1.1 Des réseaux de neurones biologiques aux Perceptrons multicouches . . . . .	23
B.1.2 Les Perceptrons multicouches . . . . .	23
B.2 Combinaison de classificateurs élémentaires : la méthode ADABOOST . . . . .	23
B.2.1 Les arbres de décision et le boosting . . . . .	24
B.2.2 Introduction au boosting . . . . .	24

B.2.3 Le premier algorithme de boosting . . . . .	24
B.2.4 Le boosting probabiliste et l'algorithme ADABOOST . . . . .	24
B.2.5 L'utilisation du boosting . . . . .	24
B.2.6 En résumé . . . . .	24
<b>C Filtrage de Kalman et filtrage particulaire . . . . .</b>	<b>25</b>
C.1 Le filtre de Kalman . . . . .	25
C.1.1 Optimalité du filtre de Kalman . . . . .	25
C.1.2 Mise en œuvre de l'algorithme . . . . .	25
C.1.3 Extensions du filtre . . . . .	25
C.2 Pourquoi le filtrage de Kalman n'est pas toujours adapté au suivi visuel . . . . .	25
C.3 Le filtrage particulaire . . . . .	26
C.3.1 Résultats et performances . . . . .	26
<b>D Modèles deMarkov . . . . .</b>	<b>27</b>
D.1 Exemple introductif . . . . .	27
D.2 Modèles de Markov observables . . . . .	27
D.3 Modèles de Markov cachés . . . . .	27



## Résumé

L'analyse automatique de vidéos est un domaine qui a connu un essor fulgurant au cours des dix dernières années. L'augmentation exponentielle de la puissance des machines ayant ouvert la porte à des possibilités de traitements en temps réel des séquences d'images, les demandes pour des systèmes d'interprétation automatiques des événements d'une scène se sont multipliées. Pourtant, le niveau de confiance atteint par les systèmes limite encore la diffusion des solutions à grande échelle. En effet, bien que l'analyse de séquences d'images ou la vision artificielle correspondent à des thématiques de recherches dont l'initiation remonte au début des années 1980, les performances des méthodes sont encore loin de concurrencer les capacités d'interprétation du système visuel humain. Mettre au point des algorithmes robustes pour analyser des flux vidéos et détecter des événements dans l'objectif d'apporter une description sémantique d'une scène donnée correspond donc aujourd'hui encore à un problème largement ouvert. D'un côté, les images contiennent une sur-abondance d'informations, et la première difficulté à surmonter pour analyser les données consiste à définir un ensemble de caractéristiques pour en extraire le contenu pertinent. D'autre part, on se trouve confronté au problème du fossé sémantique qui apparaît entre les descripteurs bas niveau calculés et l'ensemble des représentations du monde et des concepts humains très abstraits. Nous proposons d'aborder le problème de l'analyse des séquences d'images autour de deux thèmes principaux : la lecture de plaques minéralogiques et l'analyse du mouvement humain, en nous focalisant sur l'interprétation de vidéos monoculaires et dans un contexte algorithmique proche du temps réel. La lecture de plaques minéralogiques est une problématique qui a suscité l'intérêt de la communauté scientifique et du monde industriel depuis longtemps, à la fois pour les thématiques de recherches soulevées par les questions de reconnaissance de formes que pour les applications pratiques qui en découlent. Une quantité importante de solutions logicielles est actuellement commercialisée sur le marché, à des prix abordables et présentant des taux de reconnaissance impressionnantes. Pourtant, les performances observées dans les situations réelles sont nettement dégradées en pratique par un ensemble de conditions d'observation difficiles : éclairage variable, plaques sales, distance et orientations relatives de la caméra par rapport au véhicule, etc. La méthode que nous proposons se décompose autour des étapes suivantes : localisation de la plaque, binarisation, extraction des caractères, reconnaissance et moyennage temporel. Le cœur de la technologie, constitué par la reconnaissance optique de caractères, est pris en charge par un réseau de neurones hiérarchique qui analyse la forme des descripteurs de Fourier des contours extraits sur chaque image binaire. Le découplage des deux niveaux logiques du classifieur permet un apprentissage fiable qui limite au maximum les phénomènes de surapprentissage, et rend possible la reconnaissance de caractères altérés : contours internes bouchés, topologie inexacte etc. D'autre part, nous mettons en place à plusieurs points critiques de l'algorithme une coopération verticale entre niveaux logiques. Celle-ci permet à une couche donnée d'analyser la pertinence de l'interprétation effectuée aux niveaux inférieurs, offrant la possibilité de suivre plusieurs hypothèses en cas d'ambiguités. Malgré la simplicité apparente de la problématique et la nature relativement standard des caractères à identifier, l'identification des véhicules par l'interprétation du contenu de leur plaque minéralogique nécessite la mise en place de stratégies évoluées. A l'opposé, l'analyse du mouvement humain apparaît d'emblée comme un problème difficile. La nature complexe et articulée du corps humain, les auto-occlusions entre membres, la forte variabilité de l'apparence ou la faible observabilité des degrés de liberté en profondeur sont autant d'éléments qui rendent la détection, le suivi et l'interprétation du comportement des personnes dans les séquences d'images difficiles. La question a pourtant été intensivement

étudiée et pour schématiser, on peut dire qu'il existe deux écoles différentes. Les premières approches, dites discriminantes, reposent sur une stratégie Bottom Up et sont la plupart du temps applicables en temps réel. Elles consistent proposer des méthodes où le suivi est initié par une étape de détection. Les secondes approches, dites génératives, consistent à modéliser de manière explicite la structure articulée humaine pour chercher à identifier la position des différents membres au cours du temps. Ces stratégies Top Down reposent la plupart du temps sur une initialisation manuelle, et la poursuite simultanée de plusieurs cibles est rarement garantie. L'approche que nous proposons se situe en quelque sorte à un niveau intermédiaire. Dans un premier temps, nous proposons une approche de détection et de suivi assez basique, reposant sur une extraction des régions en mouvement qui sont ensuite dynamiquement mises en correspondance au cours du temps. Cette partie de l'algorithme, qui ne tient pas compte de l'aspect déformable de la projection du corps humain dans l'image, permet cependant un traitement rapide et rend possible la détection de situations de fusion ou de séparation de régions. Ce premier niveau d'analyse est générique et est applicable à un type d'objet quelconque. Dans les cas où le suivi peut être assuré sans ambiguïté, nous passons à une échelle plus fine, en proposant une méthode pour détecter et étiqueter les différentes parties visibles du corps de chaque personne à partir de la silhouette extraite en mouvement. Cette étape est menée à bien par une technique de mise en correspondance entre un graphe calculé à partir du squelette 2D image et un modèle 3D du squelette humain. L'ensemble des candidats pour les membres est identifié en introduisant une information a priori qui fixe des contraintes d'assemblage, utilise uniquement l'information morphologique et topologique de la silhouette, et est donc applicable dans des conditions générales, indépendamment du point de vue, de la pose de la personne, de la géométrie ou de l'apparence des membres. Elle permet à chaque instant de capturer un ensemble de caractéristiques de forme, de couleur et de texture sur chaque membre, conduisant à la mise à jour dynamique d'un modèle d'apparence articulé pour chaque personne suivie. Ce dernier sera ensuite utilisé dans des situations difficiles pour effectuer une identification des personnes et assurer une poursuite robuste du suivi. L'approche propose donc de mettre à jour une caractéristique dont le contenu informatif sera ensuite rétro propagé au niveau de la détection région pour contraindre le suivi. L'approche dans son ensemble permet un fonctionnement temps réel, et utilise un niveau de détail adaptatif en fonction des difficultés rencontrées lors des différentes étapes. En ce qui concerne l'analyse du mouvement humain, nous proposons une approche dédiée à la détection de chutes dont la contribution se décompose en deux points. D'une part, nous proposons une étude de la pertinence d'un détecteur de verticalité calculé à partir d'une caractéristique image pour discriminer les postures couché et debout. Les propriétés établies permettent de montrer qu'on peut toujours garantir une identification correcte pour les poses debout, mais qu'un certain nombre de postures couchées ne seront pas détectées, et que ceci dépend directement de l'orientation relative de l'axe tête pied par rapport au plan image de la caméra. On montre alors qu'il est possible de lever l'ensemble des ambiguïtés pour les postures couchées en ajoutant une caméra supplémentaire dont le traitement est indépendant de la première. Le second point central de la détection de chutes repose sur l'utilisation d'un Modèle de Markov Caché Hiérarchique (HHMM) pour interpréter la séquence angulaire observée. L'analyse temporelle qui est effectuée à deux échelles temporelles différentes permet une sensibilité suffisante pour détecter les mouvements brusques comme les chutes, tout en assurant une robustesse au bruit par une vérification haut niveau. Les différentes approches mises au point, aussi bien au niveau de la lecture de plaque que de l'analyse du mouvement humain ont montré la qualité et l'importance des approches discriminantes en terme de rapidité de calcul, et souligné l'importance des rétroactions génératives pour contraindre l'interprétation dans les situations équivoques.

## Summary

The field of human motion analysis by means of computer vision technologies has known a huge increase over the last decade. Boosted by the new ability of the computers to process image sequences in real time, the requirement for systems being able to automatically monitor events in the world has been enlarged. Although these new solutions open promising applications, the level of confidence that is reached by the systems still prohibits their general broadcast. Image or video processing, artificial vision have known their infancy in the early eighties. However, after 25 years of research in these fields, the proposed technologies output performances that are still far from the capacity of the human brain visual interpretation. Developing robust algorithms to analyse image sequences is thus still an open problem. On the one hand, images contain an over-abundance of information's, and the first problem to face consists in identifying relevant features to extract a meaningful content. On the other hand, filling the semantic gap between low level images features and high level human concepts remains the most challenging task. We propose here to tackle the problem in the following aspects : Licence Plate Recognition and Human Motion Analysis, in a monocular and real time context. Licence Plate Recognition (LPR) has interested both the scientific community and the industrial world, because of the questions raised by the shape recognition as well as the number of interesting applications that can come from these solutions. A large amount of softwares is today proposed at very affordable costs and exhibiting impressive recognition rates. However, the recognition observed in real situations is sensibly decreased by various difficult acquisition conditions : variable lighting conditions, dirty plates, and variable distances and orientations between camera and vehicles. The proposed approach decomposes into the following steps : plate detection, binarization, character segmentation, shape classification and temporal averaging. The heart of the technology, corresponding to the Optical Character Recognition (OCR), is handled by a Hierarchical Neural Network which analyses the shape of the Fourier descriptors extracted from the contours of each binary image. Logically decoupling the two level of the Hierarchy makes it possible to reliably learn shape pattern during the training phase, limiting the risk of overfitting, and managing to recognize degraded characters (no inner contours, differences of topology). Moreover, we investigate the idea of back propagating the results of a given level to lower ones in order to disambiguate competitive hypotheses. Despite its apparent simplicity, LPR requires in fact to carry out sophisticated strategies. At the opposite, human motion analysis immediately appears to be a complicated task. The articulated composition of the human body, auto occlusions of the limbs, strong variability of their appearance or the weak observability in depth make the detection, tracking or motion analysis of humans difficult. The problem has however intensively been addressed, and we can classify the approaches in two main groups. Bottom Up approaches first rely on extracting low level image features without explicit use of a human model, assembly constraints arising in a second step. These methods most of the time manage real time processing. Top Down approaches explicitly model the human body, and search for the best parameters settings with respect to a given image to model measurement. These strategies essentially require manual initialization, and tracking several humans is not always satisfied. Our approach can be considered intermediate with respect to the former archetypes. Firstly, we extract moving regions from each image and track them over time using a simple dynamical matching approach. This step is applicable to any type of object and is very fast. Additionally, it makes it possible to detect split and merge situations. When the tracking can be successfully achieved by this strategy, we analyse the shape silhouette at a finer scale, aiming at detecting and labelling body parts. We propose to perform this task by matching an image graph generated

from the skeleton silhouette and a model graph derived from the articulated human body structure. Only morphological and topological information is encoded in the graph, making the approach independent of the pose of the person, the viewpoint, the geometry or the appearance of the limbs. The body parts labelling makes it possible to update an appearance model for each limb, capturing shape, color and texture properties. The appearance model is used in difficult situations to identify people. Thus, the feature extracted during each people tracking is back propagated to enforce the detection step. Concerning human motion analysis, we propose an approach dedicated to fall detection, whose contribution decomposes into two points. Firstly, we propose to analyse the relevance of a proposed verticality detector, computed from an image feature and that is devoted to discriminating standing from lying poses. We prove that we can ever correctly identify standing postures. This property does not remain for lying ones, but we point out the fact that the misclassification is directly related to the relative direction between the feet-head axis and the camera image plane. We then demonstrate that the ambiguities can be resolved by using an additional camera whose processing is independent of the first one. The second contribution corresponds to the application of a Hierarchical Hidden Markov Model (HHMM) to classify the pose sequences. The analysis is performed at two different temporal scales, making it possible to detect sudden changes as falls, as well as being robust to noise in fast motions by incorporating high level verification. The different proposed approaches for License Plate recognition or human motion analysis have proved the efficiency of Bottom Up approaches for a real time purpose, and pointed out the importance of generatives back propagations to revolve ambiguous situations.

# 1. Introduction générale

## 1.1. Contexte et historique des travaux de thèse

## 1.2. Organisation et contributions principales

### 1.2.1. Analyse du mouvement humain dans les séquences d'images

### 1.2.2. Lecture de plaques minéralogiques



## 2. Lecture de plaques minéralogiques

### 2.1. Problématique de la lecture de plaques minéralogiques par analyse de la vidéo

#### 2.1.1 Approches récentes

---

### 2.2 Méthode proposée

#### 2.2.1 Contribution

---

### 2.3 Détection de la plaque

en vertu de la loi du droit d'auteur.

### **2.3.1 Etat de l'art**

---

### **2.3.2. Comparaison et approche retenue**

---

### **2.3.3 Pré traitements**

---

### **2.3.4 Algorithme de détection de la plaque**

---

### **2.3.5 Extraction d'un ensemble de candidats pour la plaque**

---

### **2.3.6 Suivi de la plaque**

---

### **2.3.7 Post traitements**

---

## **2.4 Binarisation de l'image de plaque extraite**

### **2.4.1 Critère de détection de régions homogènes**

---

### **2.4.2 Résultats**

---

## **2.5 Extraction des caractères**

### **2.5.1 Stratégie proposée**

---

### **2.5.2 Etiquetage en composantes connexes**

---

---

### **2.5.3 Connaissance *a priori* sur les caractères individuels**

---

### **2.5.4 Contraintes globales sur les caractères**

---

### **2.5.5 Analyse du rapport d'aspect**

---

### **2.5.6 Utilisation des projections d'histogrammes**

---

### **2.5.7 Détection de caractères fragmentés**

---

### **2.5.8 Discussion**

---

## **2.6 Reconnaissance des caractères**

### **2.6.1 Approche proposée**

---

### **2.6.2 Classification statistique**

---

### **2.6.3 Analyse structurelle**

---

## **2.7 Moyennage Temporel**

### **2.7.1 Détection et suivi d'un véhicule unique**

---

### **2.7.2 Mise à jour dynamique des reconnaissances au cours du temps**

---

## 2.8 Feedback entre les niveaux

## 2.9 Résultats

### **2.9.1 Performances de la détection de plaques**

### **2.9.2 Performances de la reconnaissance de plaques**

## 2.10 Conclusions et perspectives

## 3 Analyse du mouvement humain

### 3.1. Introduction

### 3.2 La détection d'humains dans la vidéo

#### 3.2.1 Méthodes Bottom/Up

---

#### 3.2.2 Approches Top/Down

---

### 3.3 Suivi d'humains dans la vidéo

### **3.3.1 Méthodes 2D vs 3D**

---

### **3.3.2 Suivi avec/sans détection**

---

### **3.3.3. Stratégies de recherche**

---

### **3.3.4 Représentation adoptée et caractéristiques mises en correspondance**

---

## **3.4 Détection, étiquetage des membres et suivi d'humains**

### **3.4.1 Approches Top/Down**

---

### **3.4.2 Approches Bottom/Up**

---

## **3.5 Interprétation du mouvement humain dans la vidéo**

### **3.5.1 Méthodes discriminantes et génératives**

---

### **3.5.2 Identification d'humains par reconnaissance du mouvement**

---

## **3.6 Approche Retenue**

## 4 Le suivi d'humains par modèle d'apparence articulé

### 4.1 Extraction d'objets

#### 4.1.1 Etiquetage des pixels mobiles

---

#### 4.1.2 Suppression des ombres

---

#### 4.1.3 Post traitements

---

### 4.2 Mise en correspondance dynamique de régions

#### **4.2.1 Tirer profit de la cohérence temporelle**

---

### **4.3. Formation d'un modèle d'apparence articulé pour le suivi**

## **4.4 Détection et étiquetage des membres**

#### **4.4.1 Généralités sur les graphes**

---

#### **4.4.2 Approches existantes**

---

#### **4.4.3 Génération d'un graphe modélisant le squelette humain**

---

#### **4.4.4 Génération d'un graphe image**

---

#### **4.4.5 Mise en correspondance de graphes**

---

#### **4.4.6 Résultats de l'étiquetage**

---

## **4.5 Mise à jour du modèle d'apparence de chaque membre**

## **4.6 Utilisation du modèle d'apparence pour la reconnaissance**

---

## **4.7 Suivi lors des occlusions**

## **4.8 Résultats**

## **4.9 Conclusions et perspectives**



## **5 Interprétation du mouvement humain : application à la détection de chutes**

### **5.1 Contexte applicatif**

### **5.2. Etat de l'art**

### **5.3. Approche proposée**

### **5.4. Discriminer les poses debout/couché dans**

## chaque image

### **5.4.1 Extraction de caractéristiques**

---

### **5.4.2 Rectification métrique de l'image**

---

### **5.4.3 Relier les caractéristiques aux poses**

---

### **5.4.4 Simulation numérique**

---

## **5.5 Analyse de la séquence des poses avec un HHMM**

### **5.5.1 Architecture du modèle**

---

### **5.5.2 Paramètres du HHMM**

---

## **5.6 Résultats**

## **5.7 Conclusions et discussion**

# 6 Conclusions et perspectives

## 6.1 Synthèse des travaux entrepris

### 6.1.1 L'importance des approches hiérarchiques

### 6.1.2 L'invariance et la spécificité des caractéristiques extraites

### 6.1.3 La combinaison horizontale des stratégies

### 6.1.4 Les approches génératives et discriminantes, et la coopération verticale

## **6.2 Limites et perspectives de recherche**

## A Les descripteurs de Fourier

### A.1 Description par tangente

### A.2 Représentation complexe

### A.3 Extension pour gérer des courbes non connexes



## **B Méthodes de classification statistique pour la reconnaissance des formes**

### **B.1 Les Réseaux de Neurones artificiels**

#### **B.1.1 Des réseaux de neurones biologiques aux Perceptrons multicouches**

---

#### **B.1.2 Les Perceptrons multicouches**

---

### **B.2 Combinaison de classificateurs élémentaires : la méthode ADABOOST**

## **B.2.1 Les arbres de décision et le boosting**

---

## **B.2.2 Introduction au boosting**

---

## **B.2.3 Le premier algorithme de boosting**

---

## **B.2.4 Le boosting probabiliste et l'algorithme ADABOOST**

---

## **B.2.5 L'utilisation du boosting**

---

## **B.2.6 En résumé**

---

# **C Filtrage de Kalman et filtrage particulaire**

## **C.1 Le filtre de Kalman**

### **C.1.1 Optimalité du filtre de Kalman**

---

### **C.1.2 Mise en œuvre de l'algorithme**

---

### **C.1.3 Extensions du filtre**

---

## **C.2 Pourquoi le filtrage de Kalman n'est pas toujours adapté au suivi visuel**

## C.3 Le filtrage particulaire

### C.3.1 Résultats et performances

---

## D Modèles deMarkov

### D.1 Exemple introductif

### D.2 Modèles de Markov observables

### D.3 Modèles de Markov cachés