

Reconnaissance de gestes par vision monoculaire temps réel : application à la formation des *chargés de manœuvres* pour la conduite des ponts polaires

Gesture recognition by monocular vision : application to the training course for the polar crane driving in nuclear plant

T. Chateau¹

F. Jurie¹

R. Marc²

¹ Lasmea,

² EDF

Lasmea, UMR6602 du CNRS/Université Blaise Pascal, 63172 Aubiere Cedex, France
thierry.chateau@lasmea.univ-bpclermont.fr

Résumé

Nous présentons une application à la formation des *chargés de manœuvres* pour la conduite des ponts polaires dans les réacteurs de centrales nucléaires. Le cœur de cette application repose sur l'étude d'un système de reconnaissance de gestes humains par vision couleur monoculaire. Pour cela, nous proposons de suivre la trajectoire des poignets et de la tête du stagiaire, puis de détecter, à partir de l'analyse de ces trajectoires, le type de geste réalisé, parmi dix huit gestes possibles. Nous proposons une approche originale pour la détection de gestes cyclique, basée sur un formalisme flou. La solution proposée est évaluée, à partir d'une base formée par sept personnes différentes.

Mots Clef

Reconnaissance de gestes, théorie des possibilités, logique floue, réalité virtuelle, suivi d'objets.

Abstract

The application described here is about gesture recognition. The system is used into an application of training course for the polar crane driving in nuclear plant. Real time gesture tracking provides information to an avatar which is seen by all the actors into the virtual nuclear plant reactor. We propose a color based approach to track both head and shoulders into the video. The trajectories are then analyzed in order to detect the current gesture from 18 possible gestures. An original solution for cyclic gesture detection is also presented.

Keywords

gesture recognition, virtual reality, theory of possibility, fuzzy logic, tracking.

Introduction

Le bâtiment du réacteur des centrales nucléaires est équipé d'un pont polaire (pont de levage) servant à transporter des charges lourdes. La conduite de ce pont est assurée par plusieurs personnes. D'une part, le *pontier*, assis en hauteur dans une cabine (à 20m de hauteur), est aux commandes du pont. D'autre part, un *chargé de manœuvres* guide ce dernier, à partir du sol, par un langage gestuel. Dans le cadre de la formation de son personnel, EDF souhaite mettre en place un système basé sur la réalité virtuelle. La figure 1 présente le scénario de cette formation. Le stagiaire, équipé

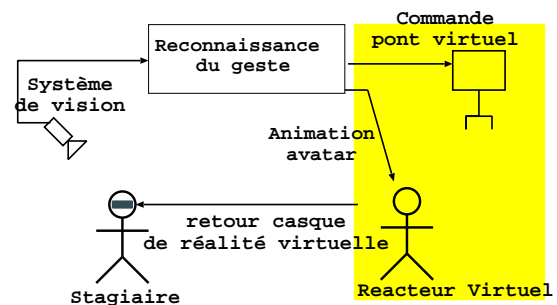


FIG. 1 – Système de reconnaissance par apprentissage

d'un casque de réalité virtuelle, sera immergé dans le réacteur virtuel d'une centrale nucléaire. Il effectue des gestes de commandement, observés par une caméra et analysés par un système de reconnaissance. Le geste reconnu est alors transmis à un moteur de rendu temps réel qui anime le pont roulant avec l'action correspondante.

Dans [6], une solution de calcul de la pose 3D du stagiaire, sous certaines conditions, est exposée ; de même que la reconnaissance du geste effectué à partir d'une comparaison des trajectoires des poignets dans le plan image par rapport à une base de trajectoires. L'étude présentée ici porte sur le module de reconnaissance et étudie une méthode basée sur une modélisation linguistique des gestes.

La première partie de ce document dresse un état de l'art sur la reconnaissance de gestes par vision artificielle, en traitant, d'une part, l'aspect modélisation du geste, et d'autre part, l'aspect reconnaissance. Dans une deuxième partie, nous exposons la solution retenue, utilisant les positions des deux poignets et de la tête, repérées dans l'image par une méthode de suivi d'objets colorés par filtrage particulaire, couramment utilisée en traitement d'images. Une approche originale concernant la détection de gestes cycliques est proposée, à partir d'une formalisation du problème par des ensembles flous. La troisième partie présente les performances de la méthode, évaluée sur un corpus de test de sept personnes différentes. Dans une dernière partie, nous concluons et proposons des perspectives possibles pour améliorer le système de reconnaissance.

1 La reconnaissance de gestes par vision artificielle

La reconnaissance de gestes est un domaine de recherche largement abordé en vision artificielle. Classiquement, la réalisation de cette tâche est découpée en deux étapes, illustrées sur la figure 2 :

- **La représentation ou modélisation du geste** qui permet de passer des données brutes à une représentation particulière du geste (ou de la forme), généralement sous la forme d'un vecteur de paramètres.
- **La décision ou reconnaissance du geste** qui décide, si possible, de la classification du geste.

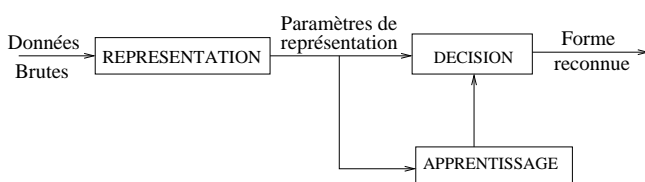


FIG. 2 – Synoptique classique d'un système de reconnaissance.

Il existe plusieurs difficultés inhérentes au problème de reconnaissance de gestes :

- **La variabilité du signal dans le temps** : la durée d'un geste et sa dynamique varient à chaque exécution, même si le geste est toujours effectué par la même personne.
- **La variabilité du signal dans l'espace** : de la même façon, un geste peut varier, à chaque exécution, dans chacune de ses dimensions de liberté.
- **La segmentation** : pour reconnaître une séquence de gestes, il est souvent indispensable de pouvoir segmenter automatiquement cette séquence pour étudier chaque

geste l'un après l'autre.

Il existe de nombreuses techniques permettant de reconnaître des gestes en résolvant, plus ou moins bien, une ou plusieurs difficultés ci-dessus.

1.1 La modélisation d'un geste

Les gestes peuvent être modélisés de différentes façons. Une représentation au moyen de prototypes est la plus simple. Il s'agit de stocker les données captées en entrée sans les modifier et de les utiliser dans la phase de décision. La sensibilité au bruit de cette méthode peut être atténuée par une approche statistique lors de l'apprentissage des gestes. Rubine [23] propose d'utiliser des caractéristiques géométriques et cinématiques calculées à partir de la trajectoire d'un geste de référence pour modéliser ce dernier (distances, vitesse, accélérations,...). Dans [5], Campbell et al. rappellent les avantages et inconvénients de chacun de ces paramètres, dont le choix est souvent effectué de manière empirique. Notons également les travaux de Woo et al. [26] qui proposent une méthode rapide pour calculer les principaux paramètres géométriques et cinématiques d'un geste. Ils modélisent également le mouvement par des caractéristiques «qualitatives» en se basant sur le modèle de Laban [18]. Hienz et al. [15] utilisent aussi ce type de paramétrage pour décrire la langue des signes.

Une autre méthode de modélisation des gestes, inspirée des langages formels consiste à décrire, de manière linguistique chaque geste en fonction d'informations issues d'une analyse bas niveau des données capteurs

1.2 La reconnaissance d'un geste

Une fois le geste modélisé, il s'agit de l'analyser pour prendre une décision concernant sa nature. L'approche la plus simple consiste à comparer le prototype du geste réalisé avec ceux d'une base de gestes. Les fonctions de mesure les plus couramment employées sont alors la distance de Hausdorff, la somme des valeurs absolues des différences ou la somme des carrés des différences. Davis et al. [8] proposent un système de reconnaissance basé sur une comparaison de prototypes du type moments calculés sur des pseudo-images décrivant le mouvement.

Les méthodes de reconnaissance basées sur les réseaux connexionnistes sont également largement utilisées pour la reconnaissance de postures ou de gestes [14, 3]. Gravila [13] propose une méthode de comparaison dynamique entre des signaux temporels afin de reconnaître un geste parmi des modèles enregistrés dans une base par des personnes différentes.

Les modèles de Markov cachés (HMM), sont très utilisés dans la reconnaissance de la parole. Il est donc naturel de retrouver cette technique dans un grand nombre de travaux traitant de la reconnaissance des gestes [24, 2, 25, 4].

2 Solution proposée

Le but de cette étude est de reconnaître, à partir d'une séquence vidéo couleur monoculaire, les gestes réalisés par

un stagiaire, parmi un corpus de dix huit gestes. Seize de ces gestes sont représentés sur la figure 3. Tous ces gestes, à l'exception du geste d'arrêt et d'arrêt express peuvent être décrits par un état de chaque main, soit statique, soit cyclique. Une précision sur la localisation de chaque main durant l'exécution du geste renseigne alors sur sa nature. Pour localiser chaque main, l'image est découpée en zones d'intérêt, à partir d'une origine centrée sur la position de la tête, approche similaire à celle proposée par Marcel et al.[20]. Le découpage, en sept zones d'intérêt, est illustré sur la figure 4.

A l'exception des gestes d'arrêt et d'arrêt express, traités par des heuristiques non abordés dans cette étude, tous les gestes sont modélisés par une expression linguistique. Pour le geste de montée rapide, le modèle s'écrit : « *la main droite est en cycle dans la zone haute droite et la main gauche est statique sur la hanche gauche* ».

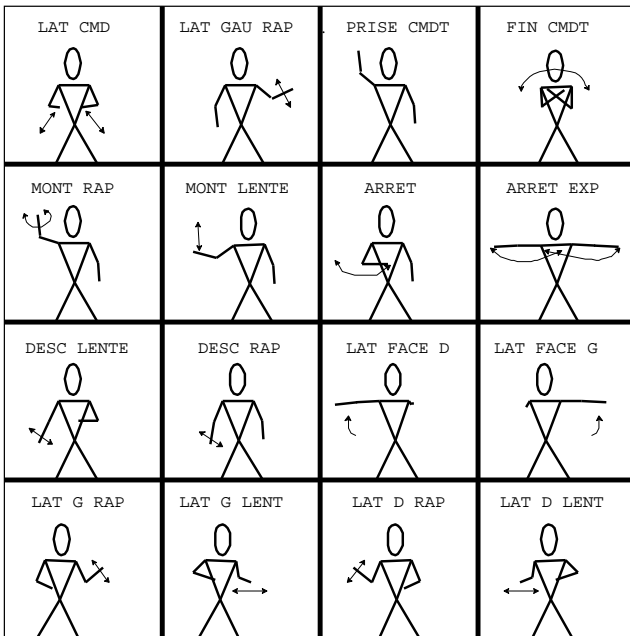


FIG. 3 – Corpus de gestes

La solution proposée se décompose en 3 parties, illustrées sur la figure 5. Dans la partie bas niveau, une méthode de suivi d'objets colorés, basée sur l'algorithme CONDENSATION est utilisée pour suivre les deux mains et la tête du stagiaire dans un flux vidéo monoculaire. Pour faciliter cette tâche, un équipement spécifique est utilisé (deux gants et un casque de couleur). Les positions des trois zones suivies sont alors transmises à un module d'extraction de primitives gestuelles liées à la trajectoire des poignets. A partir de ces informations moyen niveau, un module de reconnaissance permet d'une part, de segmenter une séquence de gestes en gestes élémentaires, et d'autre part, de reconnaître le geste effectué parmi une base de gestes modélisés de manière linguistique.

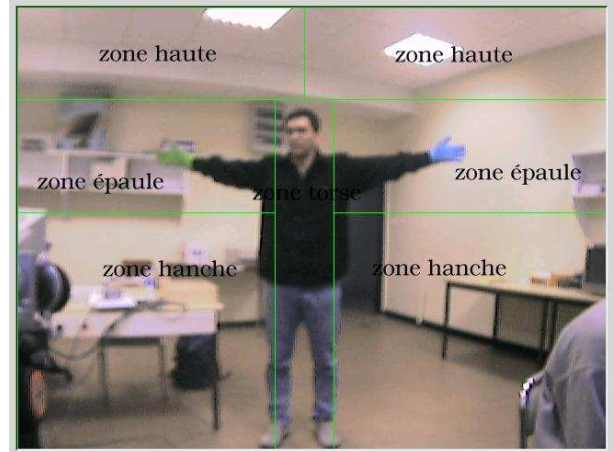


FIG. 4 – Découpage de l'image en sept zones d'intérêts

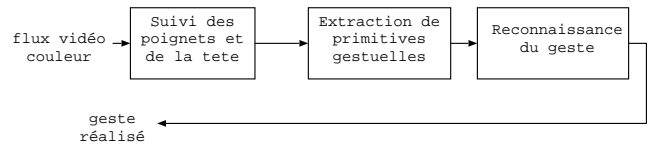


FIG. 5 – Synoptique du système

2.1 Le suivi des poignets et de la tête par traitement d'images

L'approche retenue pour suivre les deux poignets et la tête est dérivée des travaux de Perez et al. [21] sur le *tracking* d'objets couleur non rigides. C'est une approche utilisant l'algorithme CONDENSATION proposé par Isard et Black [19], et repris dans de nombreux travaux [16, 12, 17].

Le problème du suivi d'un objet dans une image peut être décrit par la recherche de la densité de probabilité a *posteriori* $p(\mathbf{x}_t | \mathbf{z}_{1:t})$, à partir de la densité de probabilité $p(\mathbf{z}_{1:t} | \mathbf{x}_t)$; où \mathbf{x}_t et le vecteur d'état composé des paramètres du modèle à suivre et $\mathbf{z}_{1:t} = \{\mathbf{z}_1, \dots, \mathbf{z}_t\}$ est le vecteur d'observation qui regroupe l'historique des mesures. Le suivi s'effectue alors en deux étapes. La première phase consiste en une prédiction de $p(\mathbf{x}_t | \mathbf{z}_{1:t-1})$ à partir de $p(\mathbf{x}_{t-1} | \mathbf{z}_{1:t-1})$ et d'une densité de transition $p(\mathbf{x}_k | \mathbf{x}_{k-1})$. La deuxième phase utilise la règle de Bayes afin de mettre à jour les $p(\mathbf{x}_t | \mathbf{z}_t)$ en fonction des nouvelles mesures.

Dans un cas général, les densités ne sont pas gaussiennes, l'estimation de la solution peut passer par l'utilisation de méthodes de Monte-Carlo.

La simulation de Monte-Carlo effectuée à l'aide d'un filtre à particules se base sur une densité de probabilité a *posteriori* estimée par un ensemble de particules pondérées $\{(\mathbf{s}_k^{(i)}, \pi_k^{(i)}) \mid i = 1, \dots, N_F\}$. Une présentation détaillée du filtrage particulaire est donnée dans [1]. L'application la plus connue de cette méthode en traitement d'images est l'algorithme CONDENSATION. Il se décompose en cinq étapes :

1. Durant l'initialisation, un jeu de particules est généré,

en fonction de la valeur initiale du vecteur d'état \mathbf{x}_0 . La valeur du poids associé à chaque particule est fixé à l'inverse du nombre de particules.

2. Une prédiction de l'état de chaque particule est alors calculée, en fonction d'un modèle d'évolution. Dans le cas du suivi d'objets dans une séquence d'images, soit on considère un modèle simple (position constante, vitesse constante, accélération constante, etc.), soit le modèle d'évolution est construit à partir d'un apprentissage.
3. Le poids associé à chaque particule est alors mis à jour en fonction des mesures.
4. Le calcul de l'espérance du jeu de particules fournit une estimation de l'état du système.
5. Un nouveau jeu de particules est choisi. Il s'agit d'un tirage au sort avec remise, de N_F particules parmi le jeu courant. Plus une particule du jeu courant a un poids élevé, plus elle est susceptible d'être choisie plusieurs fois ; de même, une particule de poids faible a de fortes chances d'être éliminée. Une fois le nouveau jeu de particules construit, le système boucle à l'étape 2.

Dans notre cas, un filtre à particules est utilisé par zone à suivre, avec un vecteur d'état \mathbf{x}_k^i défini à l'instant k pour la zone i par :

$$\mathbf{x}_k^i = (\mathbf{u}_k^i, s_k^i)^T \quad (1)$$

où \mathbf{u}^i représente les coordonnées du centre de la zone (carrée) à suivre et s_k^i la largeur de la zone. Le critère de mesure est basé sur la comparaison entre des histogrammes couleur [21].

2.2 Extraction des primitives gestuelles

Le but de ce module est de détecter l'état de chaque poignet, statique, en cycle ou autre. L'estimation de l'état statique d'un poignet est calculée à partir de la surface du rectangle englobant de la trajectoire du poignet dans l'image, sur une fenêtre temporelle.

Les mouvements cycliques sont très intéressants à étudier car il sont présents dans la plupart des mouvements humains de déplacement (marche, course). Certains travaux abordent le problème de l'estimation des mouvements périodiques, soit à partir d'une analyse de Fourier [22], soit à partir de méthodes basées sur la corrélation [7]. La méthode proposée ici utilise une formalisation totalement différente.

L'estimation de l'état cyclique d'un poignet est calculée à partir de quatre attributs issus de la trajectoire des mains dans le plan image :

- **Le nombre de passages au même endroit et dans le même sens, noté X_1** - pour chaque nouveau point de la trajectoire, une analyse effectuée sur une fenêtre temporelle estime si la main considérée est déjà passée par un point proche, et dans le même sens.

- **La longueur de la trajectoire entre chaque passage, notée X_2** - il s'agit du calcul de la distance parcourue sur une fenêtre temporelle.
- **La corrélation temporelle pour chaque cycle notée X_3** - c'est un critère de corrélation entre la durée de deux cycles consécutifs.
- **La différence spatiale entre les trajectoires de chaque cycle, notée X_4** - cette différence est estimée de manière très sommaire, par une distance Euclidienne, entre les points se situant à la mi-période de chaque cycle.

La solution proposée consiste à «diagnostiquer» si un poignet est dans un état cyclique à partir des valeurs des 4 attributs définis ci-dessus. Le terme diagnostic apparaît car la méthode utilisée ici reprend les travaux de Dubois et al . [10] [9] qui proposent une approche possibiliste générale du diagnostic causal. Le fondement de cette approche est basée sur la logique floue et la théorie des possibilités [27, 11]. Le raisonnement présenté ci-dessous est effectué pour chaque poignet. Soit \mathcal{D} l'ensemble des actions possibles : $\mathcal{D} = \{d_1, d_2\}$ où d_1 est un geste cyclique et d_2 est un geste non cyclique. Soit \mathcal{A} l'ensemble des 4 attributs observables : $\mathcal{A} = \{X_1, X_2, X_3, X_4\}$. Soit $d \in \mathcal{D}$ et $i \in \{1, \dots, 4\}$, π_d^i est la distribution de possibilités indiquant les valeurs plus ou moins possibles de l'attribut X_i en présence de l'action d . Soit U_i le domaine des valeurs de X_i , $\pi_d^i : U_i \rightarrow [0, 1]$. Les observations peuvent aussi être imprécises : $\mu_{\mathcal{O}_i} : U_i \rightarrow [0, 1]$, distribution de possibilité, indique les valeurs plus ou moins plausibles pour l'attribut X_i . Nous considérons que pour les quatre attributs constituant \mathcal{A} la mesure est précise : $\mu_{\mathcal{O}_i}(u_m) = 1$ pour u_m , singleton de U_i , valeur mesurée et $\forall u \in U_i, u \neq u_m, \mu_{\mathcal{O}_i}(u) = 0$. Nous noterons \mathcal{C}_i^d l'ensemble flou associé à la distribution de possibilité π_d^i . Elle représente le degré de présence d'une action en fonction de la valeur de l'attribut.

La figure 6 montre les ensembles flous retenus pour les actions «geste cyclique» et «geste non cyclique». Dans certains cas, une action est possible quelque soit la valeur d'un attribut. C'est le cas, par exemple, de l'attribut de corrélation temporelle qui ne renseigne pas sur l'action «geste non cyclique».

De manière générale, la compatibilité des actions avec les observations est donnée par :

$$\forall d \in \mathcal{D}, \mu_{\text{cons}}(d) = \min_{i=1}^n \sup_{u \in U_i} \min(\mu_{\mathcal{O}_i}(u), \pi_d^i(u)) \quad (2)$$

L'indice «cons» signifie *consistant* (cohérent). Dans cette approche, l'action en cours est alors estimée par :

$$\hat{d} = \arg(\sup_{d \in \mathcal{D}} \mu_{\text{cons}}(d)) \quad (3)$$

2.3 Reconnaissance du geste effectué

La traduction du modèle linguistique associé à chaque geste s'effectue de façon simple, à partir des distributions de possibilités calculées ci-dessus, pour l'action cyclique

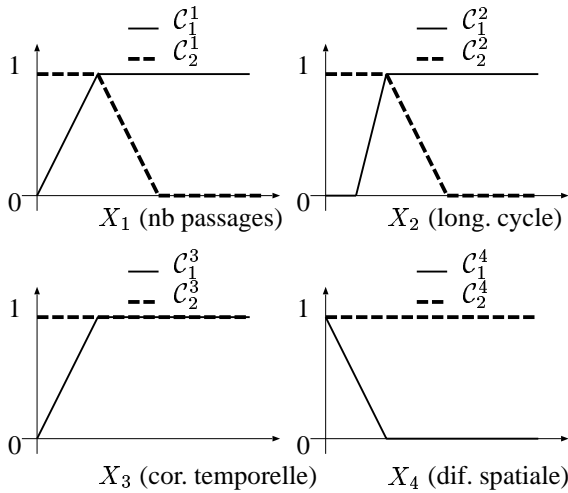


FIG. 6 – Ensembles flous associés aux actions «geste cyclique» et «geste non cyclique» pour chaque attribut

ou non cyclique, et de distributions binaires pour l'action statique et la zone d'évolution du cycle. Ainsi, une mesure de possibilité peut être associée au geste de montée rapide, et calculée par :

$$\Pi(\text{mr}) = \min[\mu_{\text{cons}}^{\text{dr}}(d_1), b_p^{\text{dr}}(z_1), b_s^{\text{ga}}, b_p^{\text{ga}}(z_4)] \quad (4)$$

avec : $b_p^{\text{dr}}(z_1)$ fonction booléenne vérifiant la position du poignet droit dans la zone z_1 (zone haute gauche), b_s^{ga} variable booléenne vraie si le poignet gauche est statique et $b_p^{\text{ga}}(z_4)$ fonction booléenne vérifiant la position du poignet droit dans la zone z_4 (zone hanche droite).

3 Résultats

La figure 7 montre un exemple de trajectoire réalisée par les poignets dans l'image, pour un geste de «montée rapide». On distingue, pour le poignet droit (gauche sur l'image), le geste cyclique réalisé. On distingue également la phase de préparation et de fin de geste.

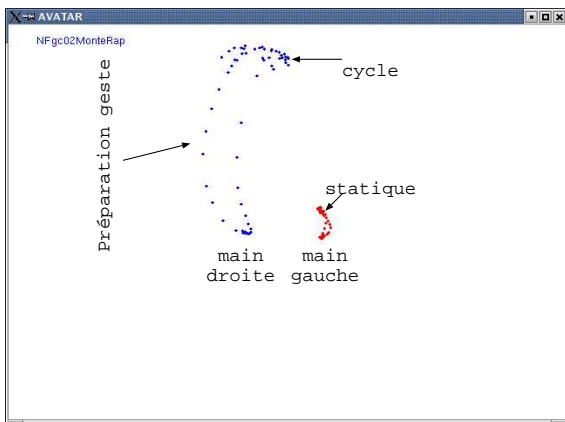


FIG. 7 – Exemple de trajectoire réalisée dans le plan image par les deux poignets pour un geste de «montée rapide»

Nombre de gestes	76
Gestes reconnus (%)	80%
Gestes non reconnus (%)	17.4%
Gestes mal reconnus (%)	2.6%

TAB. 1 – performance du système de reconnaissance

La méthode de reconnaissance de gestes a été testée sur corpus de gestes réalisé par sept personnes différentes. Chaque personne a réalisé une dizaine de gestes segmentés par un geste d'arrêt. Le tableau 1 présente les performances du processus de suivi et de reconnaissance.

80% des gestes réalisés sont correctement reconnus. Parmi les gestes mal reconnus, une distinction a été effectuée entre les gestes qui n'ont pas été reconnus comme un geste de la base et ceux qui ont été reconnus comme un autre geste de la base. Parmi les erreurs relevées, une part importante concerne la segmentation de la séquence en gestes de base. En effet, l'heuristique utilisée pour détecter le geste d'arrêt présente des défauts. La deuxième source d'erreur concerne l'affectation binaire d'une zone d'intérêt à la position de chaque poignet. Le passage d'une formalisation binaire à une formalisation basée sur la logique floue pourrait améliorer les résultats.

4 Conclusion

Nous avons présenté une solution de reconnaissance de gestes par vision monoculaire. Elle est basée sur le suivi de la tête et des poignets d'une personne, à partir de leurs caractéristiques couleur, par une technique de filtrage particulière. Chaque geste est modélisé par une expression linguistique dans laquelle la notion d'action statique ou cyclique est définie.

La détection de cycles dans la trajectoire des poignets s'effectue grâce à une modélisation du problème dérivée du diagnostic causal. Cette formulation offre de nombreux avantages tels que la prise en compte de l'imprécision et de l'incertitude dans les mesures images ou dans les relations causales entre ces mesures et la présence ou non d'une action cyclique. Une perspective possible concerne la prise en compte effective de l'imprécision des mesures images dans le système. Il suffirait alors d'agir sur l'ensemble flou décrivant les valeurs possibles de l'attribut considéré.

Pour l'instant seule la détection des mouvements cycliques est formalisée à partir de relations floues. Une généralisation de cette formalisation à la détection des mouvements statiques, et surtout à la localisation spatiale de chaque mouvement dans l'image pourrait être effectuée. Les apports d'une telle extension sur les performances du système devraient être importants.

Références

- [1] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking. *IEEE Tran-*

- sactions on *Signal Processing*, 50(2) :174–188, February 2002.
- [2] M. J. Black and A. D. Jepson. A probabilistic framework for matching temporal trajectories : Condensation-based recognition of gestures and expressions. In *European Conf. on Computer Vision, ECCV-98*, volume 1406, pages 909–924, Freiburg, Germany, 1998. Springer-Verlag.
- [3] K. Boehm, W. Broll, and M. Sokolewicz. Dynamic gesture recognition using neural networks ; a fundament for advanced interaction construction. In *SPIES Conference on Electronic Imaging Science and Technology*, San Jose, California, USA, 1994.
- [4] M. Brand, N. Olivier, and A. Pentland. Coupled hidden markov models for complex action recognition. *IEEE CVPR97 Computer Vision and Pattern Recognition*, 1996.
- [5] L. Campbell, D. Becker, A. Azarbayejani, A. Bobick, and A. Pentland. Invariant features for 3d gesture recognition. Technical report, MIT MediaLab Perceptual Computing Section Technical Report, 1996.
- [6] T. Chateau, F. Jurie, R. Marc, and M. Dhome. Suivi et reconnaissance de gestes par vision monoculaire en temps réel : application à la formation des chargés de manoeuvres pour la conduite des ponts polaires. In *RFIA, Reconnaissance des formes et Intelligence Artificielle*, page (à paraître), Toulouse, France, January 2004.
- [7] R. Cutler and D. Larry. Robust periodic motion and motion symmetry detection. *Computer Vision and Pattern Recognition*, 2 :2615–2623, 13 June 2000.
- [8] J. Davis and A. Bobick. The representation and recognition of action using temporal templates. Technical Report 402, MIT MediaLab Perceptual Computing Section, 1997.
- [9] O. De Mouzon, D. Dubois, and H. Prade. Une approche possibiliste générale du diagnostic causal. In *Rencontres Francophones sur la logique floue et ses applications*, volume 1, pages 223–229, 1999.
- [10] D. Dubois, M. Grabish, and H. Prade. A general approach to diagnosis in a fuzzy setting. In *IFSA, International Fuzzy Systems Association Congress*, volume 2, pages 680–684, 1999.
- [11] D. Dubois and H. Prade. *Théorie des possibilités*. Masson, 1987.
- [12] E. Meier and F. Ade. Using the condensation algorithm to implement tracking for mobile robots. In *Third European Workshop on Advanced Mobile Robots, Eurobot99, IEEE*, pages 73–80, 1999.
- [13] D.M. Gavrila and L.S. Davis. 3-D model-based tracking of humans in action : a multi-view approach. In *CVPR, International Conference on Vision and Pattern Recognition*, pages 73–80, 1996.
- [14] P.A. Harling. Gesture input using neural networks. BSc degree in Computer Science, Dept of Computer Science, University of York, 1993.
- [15] H. Hienz and B. Bauer. HMM-based Continuous Sign Language Recognition using Stochastic Grammars. In *Human-Computer Interaction International Proceedings of Gesture Workshop, GW99, Lecture Notes in Artificial Intelligence*, volume 1739, pages 185–196, March 1999.
- [16] K. Nummiaro, E. Koller-Meier, and L. Van Gool. Object Tracking with an Adaptive Color-Based Particle Filter. In *symposium for Pattern Recognition of the DAGM*, September 2002.
- [17] E. B. Koller-Meier and L. Van Gool. Modeling and Recognition of Human Actions Using a Stochastic Approach. In *2nd European Workshop on Advanced Video-based Surveillance Systems AVBS'01*, London, 4 September 2001.
- [18] R. Laban. *Modern Educational Dance*. Trans-Atlantic Publications, 1988.
- [19] M. Isard and A. Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1) :5–28, 1998.
- [20] S. Marcel, O. Bernier, J.E. Viallet, and D. Collobert. Hand Gesture recognition using Input/Output Hidden Markov Models. In *FG'2000 Conference on Automatic Face and Gesture Recognition*, pages 456–461, March 2000.
- [21] P. Perez, C. Hue, J. Vermaak, and M. Gangnet. Color-Based Probabilistic Tracking. In *Computer Vision ECCV 2002*, volume 1, pages 661–675, May 2002.
- [22] R. Polana and R. Nelson. Low level recognition of human motion. In *Workshop on Non-rigid Motion*, pages 77–82, November 1994.
- [23] D. Rubine. *The automatic recognition of gestures*. PhD thesis, Carnegie Mellon University, 1991.
- [24] T. Starner and A. Pentland. Real-Time American Sign Language Recognition from Video Using Hidden Markov Models. Technical report, TR-375, MIT Media Lab, 1995.
- [25] A. D. Wilson and A. F. Bobick. Recognition and Interpretation of Parametric Gesture. In *ICCV International Conference on Computer Vision*, pages 329–336, 1998.
- [26] W. Woo, N. Kim, K. Wong, and M. Tadenuma. Sketch on dynamic gesture tracking and analysis exploiting vision-based 3d interface. In *SPIE PW-EI-VCIP'01*, volume 43, January 1.
- [27] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1 :3–28, 1978.