

Revue : 3 Approches pour la Représentation d'Objets déformables en 2D

Florian Joyeux^{1,2} Benoit Telle¹ Salah Bourennane²

¹ 3DFeel Pôle d'Activités d'Aix en Provence Tech'Indus
D 645 rue Mayor de Montricher
13854 Aix en Provence Cedex 3

² Institut Fresnel UMR 6133
Domaine Universitaire de Saint Jérôme
13397 Marseille Cedex 20

florian.joyeux@ec-marseille.fr benoit.telle@3DFeel.com salah.bourennane@fresnel.fr

Résumé

Cet article présente 3 approches pour la représentation d'objets déformables en 2D. Il s'agit d'une étude bibliographique en vue du développement d'un système de détection de la main et reconnaissance de sa pose à partir d'images 2D en niveaux de gris pour l'élaboration d'une interactivité Homme/Machine. Ce document référence et décrit les différentes approches apportées par la littérature afin de déterminer quelle technique sera la plus adaptée à notre système. Les approches référencées sont classées suivant 3 types de méthodes :

- les approches géométriques.
- les approches variationnelles.
- les approches statistiques.

Mots clefs

Modèle de forme, objets déformables, représentation, images 2D, contour.

1 Introduction

Cet article est une étude bibliographique présentant différentes approches de représentation des objets déformables en 2D, basées sur leurs contours. Cette étude (évidemment non exhaustive) s'inscrit dans le cadre du développement d'un système de détection de la main et de reconnaissance de sa pose à partir d'images 2D en niveaux de gris pour l'élaboration d'un bureau virtuel (interface Homme/Machine), ce qui induit l'impératif du temps réel. La difficulté de ce problème tient surtout de la complexité et de la déformabilité de l'objet d'étude. En effet, rien que sous l'aspect statique, le nombre de cas est important : il faut compter sur la diversité des paramètres de caractérisation d'une main (main potelée, main fine, couleur de la peau...), le nombre de postures que nous effectuons naturellement (main au repos, pointage, ombres chinoises...) ainsi que les variations de forme pour chaque posture (translation, rotation, changement d'échelle, doigts plus ou moins écartés...), sans oublier les variations de luminosité, le bruit ou tout autre événement extérieur

pouvant perturber le traitement. Le choix du modèle est donc des plus importants quelque soit le niveau du traitement, que ce soit en général pour la détection (Cootes et al. [1]), le suivi d'objet (Heap [2]) ou la reconnaissance (la majorité des travaux présentés dont Heap [2], Grenander et al. [3], Freeman et al. [4,5]...). Le but de cet article a donc été de référencer les différentes techniques de représentation des objets déformables en 2D afin d'avoir les éléments pour en choisir une pour l'élaboration de notre système.

Compte tenu de la difficulté des problèmes liés au traitement des objets déformables, mais aussi de la quantité d'applications potentielles (reconnaissance des visages [6], suivi d'objets [2], interface Homme/Machine [5]...), de nombreuses approches ont été développées associées à de nombreux modèles de représentation de l'objet considéré. Ces modèles sont souvent spécifiques à l'application élaborée ainsi qu'à l'objet d'étude¹. Dès lors, nous avons donc tenté de classer ces modèles, nous avons donc dégagé 3 types d'approches pour la représentation d'objets déformables en 2D : les approches géométriques, variationnelles et statistiques.

2 Modèles géométriques

Les modèles géométriques consistent à décrire les contours de l'objet par une série d'éléments simples qui peuvent se déformer et/ou s'articuler les uns par rapport aux autres.

2.1 Primitives 2D

Les modèles 2D sont donc l'assemblage de primitives 2D simples telles des droites, des polygones, des cercles. Le

¹Et donc aussi au degré de déformabilité de l'objet considéré. Pour reprendre l'exemple de la main, il s'agit d'un objet déformable car l'une et l'autre ne sont pas semblables, mais aussi parce que nos doigts bougent. Un objet déformable est donc tout autant un objet statique dont la forme varie d'un individu à l'autre qu'un objet bien défini et qui se meut.

À noter que la majorité des modèles sont développés pour des applications médicales, nous n'avons pu tous les rassembler dans cette revue. Nous avons donc focalisé notre recherche autour de la main et présenté des modèles susceptibles de décrire notre objet d'étude.

contour peut aussi être modélisé par des courbes ou encore des polynômes implicites. Par exemple, Hill et Taylor [7] traitent des échogrammes du coeur et localisent le ventricule gauche (souvent difficile à capturer, même pour un expert) à partir d'un modèle décrit en Figure 1 variant selon 6 paramètres de forme contrôlant les ratios de taille et les mouvements du contour. À ces paramètres s'ajoutent 4 paramètres de position (translation, rotation, changement d'échelle). Ces paramètres définissent alors des points de contrôle, la forme est ensuite générée par une courbe passant en ces points.

Le système développé par Hill et Taylor a des performances très honorables compte tenu de la difficulté du problème. Mais une telle méthode n'est pas adaptée pour des objets telle une main où certaines parties s'occulent selon son orientation.

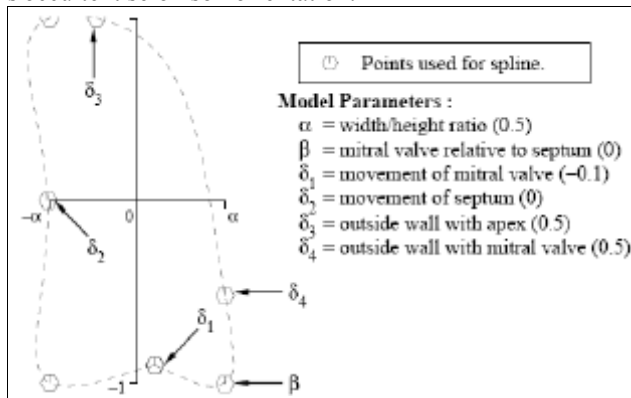


Figure 1 - Modèle 2D de forme du ventricule gauche.

2.2 Modèles par projection

Pour palier au problème des auto-occlusions, les modèles développés deviennent alors des projections des contours occultant de modèles 3D utilisant aussi des primitives géométriques simples (cylindres, sphères, parallélépipèdes...).

En ce qui concerne la main (par exemple), il existe de nombreux modèles 3D à partir desquels on peut représenter la main en 2D. Ces modèles peuvent être géométriques, surfaciques, squelettiques... et plus ou moins complexes selon le degré de vraisemblance que l'on veut atteindre. L'une des références en la matière est le modèle à 27 degrés de liberté de Regh [8] qui a été la source de nombreux autres modèles simplifiés (Delamarre [9], Ouhaddi [10], La Gorce [24]) ou sophistiqués (modèle "metaball" de Pavone [11] prenant en compte les couches musculaires et la peau). On peut aussi souligner le modèle de Stenger et al. construit à l'aide de quadriques ([12]).

Ces modèles sont en général développés à partir d'études biomécaniques et anthropométriques de la main (Wagner [13], Ouhaddi [10]) afin de déterminer les contraintes de taille des éléments ainsi que leurs articulations les uns par rapport aux autres. Dans un premier temps, il faut donc initialiser les paramètres de forme (taille, épaisseur, dimensions des doigts...) afin d'adapter au mieux le

modèle à la réalité. Puis, en ce qui concerne les variations dynamiques, la main est un objet à 28 degrés de liberté, mais certaines articulations des modèles sont négligées ou contraintes² abaissant ainsi la complexité jusqu'à 21 degrés de liberté pour des résultats tout à fait probant.

Le principal défaut de ces modèles est qu'ils sont lourds à manipuler et demandent de grandes ressources en termes de temps de calculs afin d'avoir des résultats satisfaisant. Un autre défaut repose aussi dans le manque de généralité d'une telle approche. En effet, un modèle construit ainsi ne rend compte que des connaissances listées que l'on a de la forme considérée., ce modèle peut donc omettre certains paramètres aboutissant ainsi à des résultats inattendus. Il manque de généralité au sens aussi où il n'est pas facilement adaptable et qu'un nouveau modèle doit être développé pour chaque nouvelle application.

3 Modèles variationnelles

Les modèles variationnels décrivent le contour de la forme à partir d'une forme simple à laquelle ils appliquent des déformations de plus en plus locales pour générer la forme considérée.

3.1 Descripteurs spectraux

Scott [14] propose une méthode de modélisation de forme par une suite de fonctions trigonométriques :

$$\begin{aligned} x &= x_0 + \sum_n a_n \sin(n\theta + \phi_n) \\ y &= y_0 + \sum_n b_n \sin(n\theta + \psi_n) \end{aligned} \quad (1)$$

On ne considère alors qu'un nombre n de fréquences, ne considérant que les principales variations. La forme ainsi générée par (1) est fonction des paramètres a_n , b_n , ϕ_n et ψ_n à partir desquels on peut définir les descripteurs de Fourier. Ces descripteurs peuvent être aussi définis à partir d'autres signatures de forme, Zhang et al. [15] ont mené une étude comparant pour la reconnaissance les différents descripteurs de Fourier construits à partir des représentation complexe, distance au centre, courbure et fonction angulaire cumulative.

L'avantage de ce modèle est que la forme générée est infiniment déformable³. Ces descripteurs de Fourier ont donc déjà été largement utilisés pour de nombreuses applications allant du codage de forme (Chellappa et Bagdazian [16]) à la reconnaissance (Raubert [17]).

Toutefois, ce modèle n'est pas approprié pour la description de formes générales. En effet, par l'utilisation des fonctions de base et d'un nombre fini de termes, on ne peut qu'approximer un coin carré. De plus, les variations

²Chacun connaît dans son entourage une personne capable de réaliser des poses "impossibles" avec ses mains. Ce sont les déplacements des articulations liées à ce genre de geste qui sont généralement omis.

³En faisant varier les paramètres et selon le nombre de termes utilisés, on génère différentes formes. N'introduisant pas d'information sur la forme générée, on peut donc modéliser l'objet que l'on veut.

entre la forme générée et les paramètres de la suite trigonométrique ne sont pas linéaires.

Sur le même principe, des descripteurs basés sur les ondelettes ont été proposés (Tieng et Boles [18]), avec toutes les améliorations connues relatives à cette nouvelle forme de décomposition.

3.2 FEM

Les FEM⁴ (Pentland et Sclaroff [6]) ne considèrent qu'un seul exemple de forme et l'objet comme étant une entité physique avec ses propriétés de rigidité et d'élasticité. Ce corps élastique peut être représenté par un ensemble de n noeuds, une matrice de masse $\underline{M}$ et une matrice de rigidité $\underline{K}$. En 2D, ces matrices sont de dimensions $2n \times 2n$. L'analyse modale permet ensuite le calcul d'un ensemble de modes de vibration par résolution du problème aux valeurs propres suivant :

$$\underline{K} \underline{\Phi} = \underline{M} \underline{\Phi} \underline{\Omega}^2 \quad (2)$$

où $\underline{\Phi} = (\underline{\Phi}_1 | \underline{\Phi}_2 | \dots | \underline{\Phi}_{2n})$ est une matrice $2n \times 2n$ de vecteurs propres représentant les différents modes et $\underline{\Omega}^2 = \text{diag}(\omega_1, \omega_2, \dots, \omega_{2n})$ est une matrice diagonale de valeurs propres associées à chaque vecteur propre (ω_i est la fréquence de vibration associée au mode représenté par $\underline{\Phi}_i$). Si $\hat{x} = (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n)^T$ est le vecteur représentant les points de l'objet d'origine de la famille de formes, une nouvelle forme est alors générée par :

$$\underline{x} = \hat{x} + \underline{\Phi} \underline{u} \quad (3)$$

où $\underline{u}$ est un ensemble de paramètres. L'énergie de déformation du i ème mode est proportionnelle à ω_i^2 , les premiers modes correspondent à des vibrations de grande échelle, les derniers modes induisent des variations plus localisées. On peut voir l'effet des premiers modes sur la Figure 2.

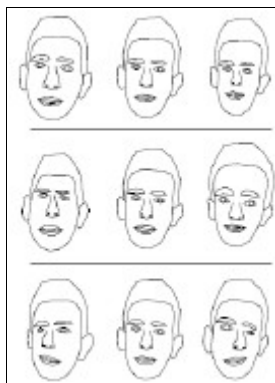


Figure 2 - Effet des trois premiers modes FEM sur le visage.

Une forme est ainsi considérée comme une combinaison linéaire de vibrations appliquées à un objet origine de la famille de formes. La qualité du modèle dépend ensuite du nombre de points utilisés pour la représentation (le

nombre de modes étant directement lié à ce nombre de points) ainsi que de leur répartition, mais aussi des matrices de masse et de rigidité représentant le matériau considéré. Ce sont en effet ces dernières qui définissent les modes de vibrations et doivent donc rendre compte localement des bonnes propriétés de déformabilité du matériau, auquel cas ces modes en deviennent arbitraires et non-représentatifs des variations de la famille de formes. Cette méthode possède aussi l'avantage que le modèle est relativement simple à construire et permet une représentation compacte d'une famille de formes. L'extension à des dimensions supérieures est de plus immédiate.

Pentland et Sclaroff ont ainsi représenté en 3D des visages humains en utilisant 30 modes. Une application a été réalisée par Karaolani et al. [19] pour modéliser des objets en 2D.

Ces modèles variationnels sont donc très appropriés pour modéliser des objets déformables, mais pas des objets articulés, contrairement aux modèles géométriques. Une utilisation combinée est donc envisageable pour modéliser un objet du type de la main.

4 Modèles statistiques

Les modèles statistiques reposent en général sur l'étude de la distribution de points caractéristiques⁵ de l'objet.

4.1 Approche statistique générale

De nombreux travaux (Grenander et al. [3], Mardia et al. [20], Cootes et al. [1]) ont étudié la distribution d'un ensemble de points caractéristiques d'un objet. En règle générale, l'objet est représenté par une série de points labellisés du contour⁶ selon un modèle préalablement construit. Ces points caractéristiques peuvent être de 3 types :

- Points marquant des parties de l'objet avec signification dépendante (centre d'un oeil, ...).
- Points marquant des parties de l'objet avec signification indépendante (plus haut point d'une main dans une orientation particulière, ...).
- Points d'interpolation.

Par exemple, on peut remarquer sur le modèle de la main proposé par Heap [2] (Figure 3), chaque bout de doigt est caractérisé par 3 points relatant sa courbure, chaque vallée est caractérisée par 3 points relatant sa courbure et les vallées et bouts de doigt sont reliés par 2 points d'interpolation (sauf entre le pouce et l'index).

Goodall [21] discute alors de l'utilisation des méthodes de Procrustes pour l'estimation d'une forme moyenne ainsi que d'une matrice de covariance entre les points caractéristiques, ainsi que de l'évaluation des différences entre ensembles de formes.

⁵Où tout autre éléments caractéristiques de l'objet, comme nous le verrons par la suite...

⁶Appelés "landmark points".

⁴Méthodes aux éléments finis.

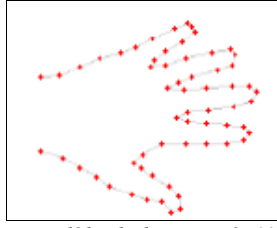


Figure 3 - Modèle de la main à 61 points.

La forme est ainsi représentée à partir de la forme moyenne, les points peuvent varier généralement selon une loi normale avec une matrice de covariance prédéfinie.

4.2 PDM et HPDM

Le PDM⁷ (Cootes et al. [1]) repose sur une représentation analogue de la forme et utilise une PCA⁸ afin de simplifier la structure de la matrice de covariance. Ce modèle de forme a été introduit à la base de l'ASM⁹, technique de segmentation utilisant le principe des contours actifs. Cette méthode était utilisée pour des objets peu déformables (non articulés) telle une main dans une pose spécifique (main au repos), puis elle a été adaptée par Heap [2] à des objets articulés, dont la main (HPDM¹⁰).

La première étape de construction du PDM consiste à labelliser un ensemble d'apprentissage¹¹ selon un modèle préalablement déterminé (cf. 4.1). À la différence donc du FEM qui génère des variations d'une seule forme, le PDM étudie les variations d'un ensemble de formes. Cette labellisation est manuelle et doit être précise autant que possible. Une forme est donc représentée par un vecteur de dimension $2n$: $\underline{x}_i = (x_{i1}, x_{i2}, \dots, x_{in}, y_{i1}, y_{i2}, \dots, y_{in})^T$. Après alignement de la base d'apprentissage (par des méthodes de Procrustes), le développeur dispose d'un ensemble de N vecteurs correspondant à N formes alignées : $\{\underline{x}_i\}_{i=1..N}$. Ces N vecteurs appartiennent tous à un espace de dimension $2n$, chaque forme y étant représenté par un point. Cet espace est appelé espace des formes possibles. En considérant un point de cet espace correspondant à une forme spécifique, tout point suffisamment proche de ce point génère une forme semblable à cette dernière, à savoir un objet de la même famille. Les points du précédent ensemble de N formes alignées sont donc tous quasi-similaires (pour des variations de formes pas trop importantes) et forment donc un ellipsoïde. Il est ensuite possible par PCA de déterminer le centre de gravité de cet ellipsoïde ainsi que ces axes principaux.

La forme moyenne est déterminée par : $\bar{\underline{x}} = \frac{1}{N} \sum_{i=1}^N \underline{x}_i$

La matrice de covariance est : $\underline{\underline{S}} = \frac{1}{N} \sum_{i=1}^N \underline{dx}_i^T \underline{dx}_i$ où $\underline{dx}_i = \bar{\underline{x}} - \underline{x}_i$. Les axes de variations sont donnés par les vecteurs propres de cette matrice définis par :

$$\underline{\underline{S}} \underline{p}_k = \lambda_k \underline{p}_k \quad (4)$$

avec $\lambda_{k+1} \geq \lambda_k$ et $\underline{p}_k^T \underline{p}_k = 1$. L'importance de ces variations est donnée par la valeur propre associée. Afin de simplifier le modèle, les variations les moins significatives sont négligées, on ne choisit alors que t modes pour représenter l'objet. Une forme est donc ainsi définie :

$$\underline{x} = \bar{\underline{x}} + \underline{P} \underline{b} \quad (5)$$

où $\underline{P} = (\underline{p}_1 | \underline{p}_2 | \dots | \underline{p}_t)$ et $\underline{b} = (b_1, b_2, \dots, b_t)^T$ est un vecteur de poids. Comme le montre la Figure 4, on peut atteindre une forme en jouant sur les coefficients du vecteur de poids $\underline{b}$. Ces variations doivent bien entendu être normées. Cootes et al. [1] utilisent une borne égale à 3 fois l'écart type ($|b_k| \leq 3\sqrt{\lambda_k}$) mais suggèrent aussi l'utilisation de la distance de Mahalanobis ($D_m^2 = \sum_{k=1}^t (\frac{b_k^2}{\lambda_k}) \leq D_{max}^2$).



Figure 4 - Effet de la variation du vecteur de poids sur une main dans une pose spécifique.

Heap [2] a repris par la suite le PDM pour l'appliquer à plusieurs poses de la main et a observé qu'alors, le modèle pouvait générer des formes complètement absurdes. Cela est dû au fait que l'ensemble de formes ne définit plus un ellipsoïde dans l'espace des formes possibles, mais qu'il parcourt un chemin. Ce phénomène montre la non linéarité de la variation de la forme par articulation. Heap a donc développé le HPDM consistant à caractériser ce chemin par des PCA locales.

La principale difficulté de la construction du PDM tient donc dans l'acquisition et le traitement de la base d'apprentissage. De plus, si cette base n'est pas suffisamment représentative, le PDM ne permettra pas de variations suffisantes de la forme. Ayant donc des propriétés complémentaires, Martin et al. [22] ainsi que Cootes et Taylor [23] ont combiné FEM et PDM.

Comme dit précédemment, les modèles statistiques peuvent aussi s'intéresser à la distribution d'autres éléments que les points caractéristiques du contour. Par exemple, Freeman et al. [4] passent en revue de

⁷Modèle de distribution de points.

⁸Analyse en composantes principales.

⁹Modèle actif de forme ou "Smart Snakes".

¹⁰Modèle de distribution de points hiérarchique.

¹¹Constitué d'éléments de la même famille de l'objet.

nombreuses méthodes pour contrôler une télévision par les gestes de la main et proposent une approche originale par l'étude de la distribution en orientations (Figure 5).

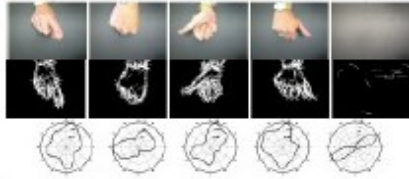


Figure 5 - Histogrammes d'orientations.

5 Conclusion

Nous avons présenté 3 types d'approches de représentation d'objets déformables en 2D : géométrique, variationnelle et statistique. Les modèles géométriques par projections sont simples à développer mais lourds à manipuler et peuvent manquer de généralité. Les modèles variationnels sont adaptés aux objets déformables, mais pas aux objets articulés. Le PDM/HPDM est plus adapté pour la modélisation de la main, mais sa construction est difficile et doit être représentative.

Références

- [1] T. Cootes, C. Taylor, D. Cooper et J. Graham, Active shape models – their training and application. *Computer Vision and Image Understanding*, 61(1) : 38-59, 1995.
- [2] A.J. Heap, Learning deformable shape models for object tracking. University of Leeds, thèse soutenue en septembre 1997.
- [3] U. Grenander, Y. Chow et D.M. Keenan, Hands. A pattern theoretic study of biological shapes. Springer-Verlag New-York Inc., 1991.
- [4] W.T. Freeman, D.B. Anderson, P.A. Beardsley, C.N. Dodge, M. Roth, C.D. Weissman, W.S. Yezauris, H. Kage, K. Kyuma, Y. Miyake et K. Tanaka, Computer vision for interactive computer graphics. *IEEE Computer Graphics and Applications*, 18(3) : 42-53, Mai-Juin 1998.
- [5] W.T. Freeman, C.D. Weissman, Television control by hand gestures. *IEEE Int. Workshop on Automatic Face and Gesture Recognition* : 179-183, 1995.
- [6] A. Pentland et S. Sclaroff, Closed-form solutions for physically based modeling and recognition. *IEEE Trans. PAMI*, 13(1) : 715-729, 1991.
- [7] A. Hill et C.J. Taylor, Model based image interpretation using genetic algorithms. *Image Vision Comput.*, 10(1) : 295-300, 1992.
- [8] J.M. Rehg et T. Kanade, Visual Analysis of High DOF Articulated Objects with Application to Hand Tracking. School of Computer Science, Carnegie Mellon University, thèse soutenue en avril 1995.
- [9] Q. Delamarre, Modélisation de la Main pour sa Localisation dans une Séquence d'Images. *Rapport Technique INRIA* n° 0198, décembre 1996.
- [10] H. Ouahdi et P. Horain, Conception et ajustement d'un modèle 3D articulé de la main. *Actes des 6èmes Journées de Travail du GT Réalité Virtuelle*, pages 83-90, Issy-les-Moulineaux, mars 1998.
- [11] J.P. Pavone, Modélisation et suivi de la main. Stage de DEA SIC Image-Vision, septembre 2001.
- [12] R. Cipolla, B. Stenger, A. Thayananthan et P.H.S. Torr, Hand tracking using a quadric surface model and bayesian filtering. Dans *Lecture Notes in Computer Science, Mathematics of Surfaces*, pages 129-141, Springer Berlin / Heidelberg, 2003.
- [13] C. Wagner, The pianist's hand : anthropometry and biomechanics. *Ergonomics*, 31(1) : 97-131, 1988.
- [14] G.L. Scott, The alternative Snake – And other animals. Dans *Proceedings, 3rd Alvey Vision Conference*, pages 341-347, Cambridge, 1987.
- [15] D. Zhang and G. Lu, A comparative study on shape retrieval using Fourier descriptors with different shape signatures. *Proc. of International Conference on Intelligent Multimedia and Distance Education* : 1-9, Juin 2001.
- [16] R. Chellappa et R. Bagdazian, Fourier coding of image boundaries. *IEEE Trans. PAMI*, 6(1) : 102-105, 1984.
- [17] T.W. Rauber, Two-dimensional shape description. *Technical Report: GR UNINOVA-RT-10-94*, University Nova de Lisboa, Portugal, 1994.
- [18] Q.M. Tieng et W.W. Boles, Recognition of 2D object contours using the wavelet transform zero-crossing representation. *IEEE Trans. on PAMI*, 19(8), 1997.
- [19] P. Karaolani, G.D. Sullivan, K.D. Baker et M.J. Baines, A finite element method for deformable models. *Proceedings of the Fifth Alvey Vision Conference*, pages 73-78, Reading, 1989.
- [20] K.V. Mardia, J.T. Kent et A.N. Walder, Statistical shape models in image analysis. *Proceedings of the 23rd Symposium on the Interface*, pages 550-557, Seattle, 1991.
- [21] C. Goodall, Procrustes methods in the statistical analysis of shape (with discussions). *J.R. Stat. Soc. B.*, 53 : 285-339, 1991.
- [22] J. Martin, A. Pentland et R. Kikinis, Shape analysis of brain structures using physical and experimental modes. *MIT Media Lab. Perceptual Computing Technical Report n° 276*, Janvier 1994.
- [23] T.F. Cootes et C.J. Taylor, Combining point distribution models with shape models based on finite element analysis. *Image and Vision Computing*, 13(5) : 403-410, juin 1995.
- [24] M. de la Gorce et N. Paragios, Fast Dichotomic Multiple Search Algorithm for Shortest Circular Path. *18th IARP International Conference on Pattern Recognition*, 2006.