

# Reconnaissance biométrique sans contact de la main intégrant des informations de forme et de texture

Julien Doublet<sup>1</sup>, Marinette Revenu<sup>2</sup>, Olivier Lepetit<sup>1</sup>

1: France Telecom, 42 rue des coutures, B.P. 6243, 14066 CAEN Cedex 4, France

2: GREYC – ENSICAEN, 6 Bd Maréchal Juin, 14050 CAEN Cedex, France

E-mail: [julien.doublet@orange-ft.com](mailto:julien.doublet@orange-ft.com)

## Résumé

*La reconnaissance biométrique de la main a été développée avec succès pour l'authentification ou l'identification biométrique. Dans ce papier, nous proposons une méthode originale de reconnaissance biométrique combinant des informations de couleur, de texture et de forme. Tout d'abord, la segmentation intègre les composantes couleurs de la peau et un modèle de forme. Ensuite, le processus d'authentification fusionne par convolution les caractéristiques géométriques des doigts et la texture de la paume analysée. Les résultats expérimentaux montrent le bien fondé de cette approche avec un taux d'erreur d'authentification inférieur à 2% pour une population de 16 individus.*

## Mots clefs

Reconnaissance de la main, biométrie, processus de fusion, détection de la peau, modèle de forme actif.

## 1 Introduction

La biométrie joue un rôle de plus en plus important dans les systèmes d'authentification et d'identification. Les processus de reconnaissance biométrique permettent la reconnaissance d'individus en se basant sur les caractéristiques physiques ou comportementales. Différentes technologies ont été développées telles que: les empreintes digitales, l'iris, le visage, la voix, la signature et la main. Cette dernière méthode s'appuie sur une étude de la forme de la main et de la texture de la paume. Elle présente de nombreux avantages par rapport aux autres technologies. Premièrement, le système de capture est moins coûteux que celui pour la reconnaissance d'iris, les caractéristiques de la main sont plus nombreuses que celles des empreintes digitales et elles peuvent être déterminées avec des images à faible résolution. De plus, ce système est bien accepté par les utilisateurs et la main laisse peu de traces contrairement à un système basé sur l'empreinte digitale. Ce type de système est ainsi prôné par la Commission Nationale de l'Informatique et des Libertés.

L'architecture des systèmes biométriques est organisée en 4 étapes: l'acquisition des données, le traitement du signal

pour l'amélioration des données et la segmentation de la modalité analysée, l'extraction des caractéristiques et la comparaison avec une ou plusieurs références. Dans notre système, l'étape 2 correspond à la détection de la main dans une image complexe. Les méthodes de Skin Blob Tracking, de contours actifs, de Mean Shift ou de condensation [1] sont couramment utilisées dans les systèmes d'interactions homme-machine. Ces processus présentent cependant deux contraintes majeures incompatibles avec une reconnaissance biométrique. Ils fournissent une détection peu précise de chaque doigt de la main et ils nécessitent une séquence d'images. Pour augmenter la qualité de la segmentation nous proposons une méthode combinant les informations de couleur de peau et de forme de main.

La troisième étape consiste en une caractérisation de la forme et de la paume de la main. De nombreux travaux ont été réalisés ces dernières années. La forme de la main peut être décrite par la largeur de la paume et la forme des doigts. Chaque doigt est défini classiquement par sa longueur, un ensemble de largeurs et la courbure de son extrémité. Initialement, la paume de la main a été caractérisée par ses lignes [2] en utilisant une méthode proche de la reconnaissance des empreintes digitales. D'autres systèmes utilisent l'information globale de la paume de la main [3-5] et ils présentent de meilleures performances en reconnaissance. Pour aboutir à des systèmes plus fiables, il est possible de fusionner les décisions obtenues à partir de la paume et de la forme de la main [6].

Dans ce travail, nous proposons une méthode de reconnaissance de la main sans contact. La section 2 introduit le processus de détection de la main. La section 3 décrit l'extraction des caractéristiques de la forme et de la paume de la main ainsi que la méthode de fusion de ces caractéristiques. Les sections 4 et 5 présentent les résultats expérimentaux et donnent les conclusions, respectivement.

## 2 Segmentation de la main dans un fond complexe

Dans un système de reconnaissance biométrique, un processus de segmentation précis et rapide doit être élaboré. Dans cette section, nous décrivons rapidement notre méthode de détection de la main basée sur une

combinaison d'une modélisation de la couleur de la peau et d'un modèle de forme [7].

## 2.1 Modélisation de la couleur de la peau

Contrairement aux modèles basés sur les classifieurs de Bayes ou les modèles de mélanges de Gaussiennes [8], la teinte de peau est modélisée par apprentissage supervisé. Pour obtenir un bon compromis entre la vitesse d'exécution et la précision de la détection, nous utilisons un réseau de neurones (RN). Les entrées du RN sont composées par trois neurones, une pour chaque composante couleur des pixels dans le domaine RGB. La sortie du RN est la probabilité qu'un pixel soit un pixel de peau. L'entraînement du réseau à partir d'une base de données de pixels de peau et de fond dans le domaine RGB permet la modélisation de la couleur de peau. Parallèlement, une analyse en composantes principales [9] sur les pixels de peau de la base établit un domaine couleur spécifique à la peau.

## 2.2 Segmentation de la peau

Dans la phase de détection, le RN calcule la probabilité que chaque pixel soit un pixel de peau. Ce processus construit la carte des probabilités (Fig. 1). Pour obtenir une segmentation proche du temps réel, un processus multi-résolution construit la carte de probabilités. Un seuillage et un renforcement du contraste sont appliqués sur cette carte pour obtenir une image binarisée.

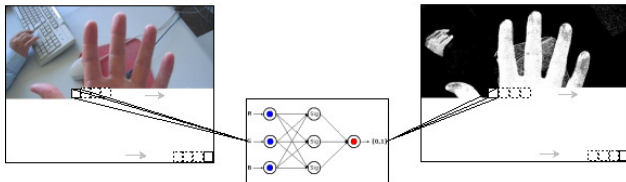


Figure 1 - Calcul de la carte des probabilités

## 2.3 Segmentation de la main par modèle de forme

La segmentation par la couleur de la peau ne peut pas effectuer de façon robuste la tâche de détection de la main. Un modèle spécifique de forme active [10] est défini pour résoudre ce problème. Les deux difficultés majeures dans ces modèles actifs de forme sont l'initialisation du contour qui doit être proche de la forme à rechercher et la convergence du modèle dans la phase de détection.

Classiquement, la forme à détecter est définie par une série de points: les landmarks. Dans la phase d'apprentissage de la forme, la forme moyenne et les variations du contour sont calculées par analyse en composantes principales sur une base de mains annotées par ces landmarks. Dans la phase de segmentation, le contour est initialisé par les points caractéristiques de la

main: les cinq points au bout des doigts et les quatre points situés dans la vallée entre deux doigts adjacents. Ces points sont calculés à partir de la carte de l'image binarisée par analyse du contour. Deux autres points sont automatiquement ajoutés près du poignet à partir de ces points. Les autres landmarks définissant plus précisément la forme de la main sont disposés entre ceux-ci. Ainsi le modèle  $X$  est définie par les 11 points initiaux et les  $N$  points intermédiaires disposés entre ceux-ci. On obtient  $X$  par  $X = (X[0], \dots, X[11 + N \times 5 \times 2 - 1])$  où  $X[i]$  est le  $i$ ème landmark. Après la phase d'initialisation, le modèle est déformé. Pour maîtriser le problème de la divergence du modèle qui ne suit pas les contours réels de la main, un poids est appliqué aux déformations pour limiter les contraintes de formes [7]. Pour que le terme gradient ne fasse intervenir que les contours de la main et limite ainsi les possibilités de divergence de la forme, il est déterminé dans l'espace de couleur de peau par l'algorithme de Di Zenzo [11]. Ce gradient est ensuite pondéré par le coefficient des pixels de la carte de probabilités. Les expérimentations montrent qu'un bon compromis entre le temps d'exécution et la précision de la détection est obtenu en fixant  $N$  à 12. Le processus de détection complet est illustré à la Figure 2.

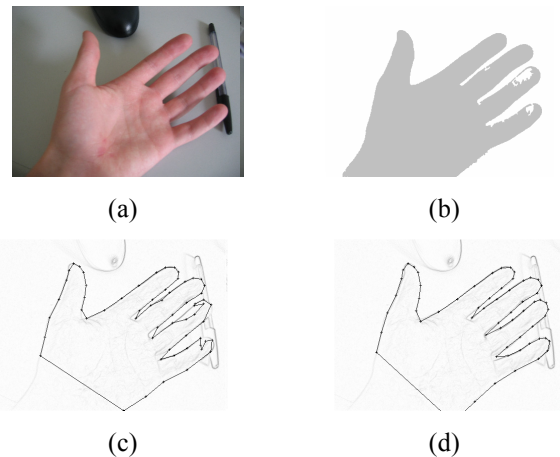


Figure 2 - Processus complet: (a) image originale (b) segmentation par la peau (c) contour initial (d) contour final

## 3 Extraction des caractéristiques

Cette section explicite les caractéristiques de la paume et de la forme de la main ainsi que la méthode de fusion de ces informations pour une reconnaissance biométrique performante. Dans la section 3.1, l'extraction de la paume et la spécification de la texture par un filtre de Gabor sont décrites. Les caractéristiques de la forme de la main sont définies à partir de ses contours. Finalement, la fusion de toutes les caractéristiques par un processus de convolution est détaillée dans la section 3.2.

### 3.1 Calcul des caractéristiques de la forme et de la paume

Après avoir détecté la main, il est nécessaire de localiser la paume indépendamment de la distance entre la main et le capteur. La détermination de la paume s'appuie sur la dimension de la main et la méthode de localisation de paumes décrite dans [3]. Dans cet article, deux valeurs sont fixées: la distance entre les points  $O1$  et  $O2$  et la taille de la paume  $||A1A2||$  (Fig. 3). Ces valeurs, constantes dans les systèmes de reconnaissance classiques, sont ici définies suivant la taille de la main. Elles sont déterminées à partir de la largeur de la main calculée par la distance Euclidienne entre les points  $X[L1]$  et  $X[L2]$  où  $L1$  et  $L2$  sont des indices fixés après expérimentations à 30 et 125. Ainsi,  $||O1O2||$  et  $||A1A2||$  sont définis par:

$$||O1O2|| = \alpha ||X[L1]X[L2]|| \quad [1]$$

$$||A1A2|| = \beta ||X[L1]X[L2]|| \quad [2]$$

Où  $\alpha$  et  $\beta$  sont les coefficients de dimensions choisis à 1/10 et 2/3, respectivement. Ensuite la paume est redimensionnée à une taille fixe  $M \times M$  où  $M$  est fixé à 100.

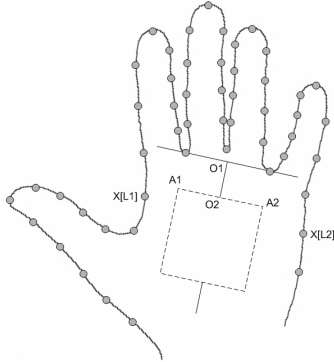


Figure 3 - Extraction de la paume avec  $N=4$

La paume extraite contient des lignes principales qui peuvent être déterminées par une méthode spécifique d'extraction [2]. Ces lignes ne sont pas propres à chaque individu, il est donc nécessaire d'utiliser les lignes secondaires de la main. Ces lignes plus fines ne peuvent pas être extraites de la paume avec des images à faible résolution, ainsi une caractérisation globale de la paume est préférable.

Différentes méthodes permettent d'obtenir les caractéristiques de la paume: l'analyse en ondelettes, la transformée de Fourier, l'Analyse en Composantes Principales, le filtre de Gabor... Grace à ses bonnes performances en reconnaissance de l'iris et de la paume et à ses qualités propres: localisation précise en temps/fréquence et robustesse aux variations de contraste et de luminosité, nous avons utilisé un filtre de Gabor. Différentes implémentations de ce filtre existent. Dans [3], un filtre de Gabor 2D dans le domaine complexe est

utilisé. Pour limiter le temps de calcul et la taille des caractéristiques, le filtre dans le domaine réel décrit dans [12] est employé:

$$G(x, y) = \exp\left[-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right] \times \cos\left(2\pi \frac{x'}{\lambda} + \varphi\right) \quad [3]$$

$$x' = (x - \xi) \cos \theta - (y - \eta) \sin \theta$$

$$y' = (x - \xi) \sin \theta - (y - \eta) \cos \theta$$

Où le couple  $(\xi, \eta)$  définit le centre de la fonction,  $\theta$  contrôle l'orientation de la fonction,  $\sigma$  est la déviation standard de l'enveloppe gaussienne,  $\gamma$  est le ratio de l'aspect visuel fixé à 0.5,  $\lambda$  est la période de l'onde et  $\varphi$  est la phase.  $\sigma$  est défini par le rapport constant  $\sigma/\lambda=0.56$  [12]. Pour plus de robustesse à la luminosité, le filtre est centré au point  $(0,0)$  en utilisant pour un filtre de taille  $(2k+1)^2$  la formule:

$$\Omega(x, y) = G(x, y) - \frac{\sum_{i=-k}^k \sum_{j=-k}^k G(i, j)}{(2k+1)^2} \quad [4]$$

Ainsi les caractéristiques de la paume de la main sont obtenues par le résultat de la convolution de l'image de la paume avec ce filtre de Gabor robuste par:

$$C = I * \Omega \quad [5]$$

Où  $*$  est l'opérateur de convolution. Pour compléter la représentation, les caractéristiques de la forme de la main sont extraites. Il s'agit des largeurs et des longueurs des doigts. Les longueurs et les largeurs sont approximées par la distance entre les points du modèle de main et sont définies par:

$$L[i] = d(m(X[H_{2i}], X[H_{2i+2}]), X[H_{2i+1}])$$

$$l[i][j] = d(X[H_{2i+1} - j], X[H_{2i+1} + j]) \quad [6]$$

Où  $d(a, b)$  est la distance Euclidienne entre les points  $a$  et  $b$ ,  $m(a, b)$  retourne le milieu du segment  $||ab||$ ,  $L[i]$  est la longueur du  $i$ ème doigt,  $l[i][j]$  est la  $j$ ème largeur du  $i$ ème doigt et  $H_i=i(N+1)$  est l'indice du  $i$ ème point caractéristique de la main (Fig. 4).

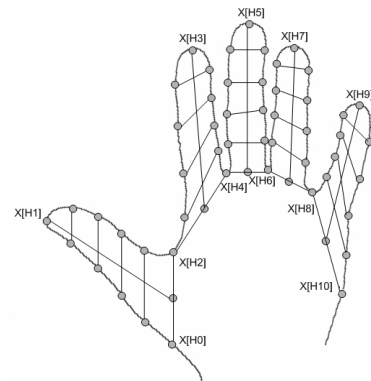


Figure 4 - Extraction des données de formes avec  $N=4$

### 3.2 Fusion de la texture et de la forme

Dans les systèmes biométriques, trois méthodes de combinaison de données sont utilisées. La combinaison peut être effectuée à la représentation en regroupant les caractéristiques extraites, au niveau de la comparaison ou au niveau des décisions. Notre fusion est basée sur la convolution de la texture de la paume avec les caractéristiques de la forme de la main afin d'ajouter un facteur géométrique à la texture. Le résultat de cette convolution est binarisé pour limiter la taille des caractéristiques et les temps de calcul dans la phase de comparaison. Le processus complet est défini par:

$$S(x, y) = b(C(x, y) * H) \quad [7]$$

Où  $b(x) = 0$  si  $x < 0$  et  $b(x) = 1$  sinon,  $*$  est l'opérateur de convolution et  $H$  est un filtre de taille  $5 \times 5$  correspondant aux caractéristiques de formes. Il peut être explicité par:

$$H(x, y) = H'(x, y) \times \frac{1}{\sum_{i=0}^4 \sum_{j=0}^4 H'[i][j]} \quad [8]$$

Où  $H' =$

L[0]	I[0][1]	I[0][2]	I[0][3]	I[0][4]
L[1]	I[1][1]	I[1][2]	I[1][3]	I[1][4]
L[2]	I[2][1]	I[2][2]	I[2][3]	I[2][4]
L[3]	I[3][1]	I[3][2]	I[3][3]	I[3][4]
L[4]	I[4][1]	I[4][2]	I[4][3]	I[4][4]

Le coefficient de normalisation de  $H$  permet d'être robuste à la distance entre la main et le système d'acquisition. La fusion par convolution des données caractéristiques augmente l'unicité de la main. Elle permet en effet de distinguer des personnes ayant des formes de main ou des paumes très similaires. L'extraction de la texture seule et la combinaison de la texture et de la géométrie de la main sur une paume sont illustrées à la Figure 5.

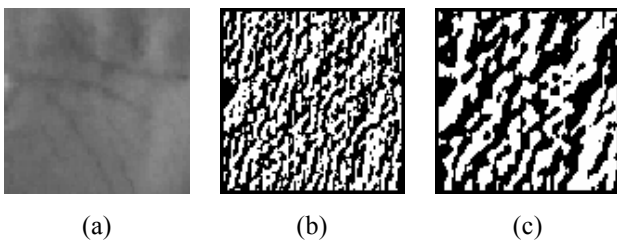


Figure 5 – Extraction des caractéristiques (a) paume extraite, (b) texture de la paume, (c) convolution de la texture et de la forme

## 4 Expérimentations

Dans cette section, nos résultats expérimentaux sont indiqués pour valider cette approche convolutionnelle. Premièrement, la méthode de comparaison des mains est

indiquée. Ensuite, la base de données de mains et l'évaluation des performances du système sont décrites.

### 4.1 Comparaison de mains

Pour la comparaison, les caractéristiques de la main  $S$  sont assimilées à une matrice. Une méthode classique de comparaison de matrices binaires est ainsi appliquée: la distance de Hamming normalisée. Cette distance est une comparaison pixel par pixel et elle donne une réponse normalisée entre 0 et 1, 0 étant la correspondance parfaite. Elle est définie pour deux caractéristiques de mains  $X$  et  $Y$  par:

$$D_0(X, Y) = \frac{\sum_{i=1}^M \sum_{j=1}^M X(i, j) \oplus Y(i, j)}{M^2} \quad [9]$$

Où  $\oplus$  est l'opérateur ou-exclusif. La segmentation de la main ainsi que la détection de la paume n'étant pas parfaites, une tolérance en translation et en rotation est appliquée au processus de comparaison. Cette comparaison souple s'exprime pour deux caractéristiques de mains  $X$  et  $Y$  par:

$$D_r(X, Y) = \min_{|s| < S, |t| < T, |a| < A} D_0(R(T(X, s, t), a), Y) \quad [10]$$

Où  $T(X, s, t)$  est la translation de l'image  $X$  horizontalement par  $s$  et verticalement par  $t$  et  $R(X, a)$  est la rotation de l'image  $X$  par un angle  $a$ . Les constantes  $S$  et  $T$  sont fixés à 2 pixels et  $A$  est limité à  $2^\circ$  afin de limiter le coût de calcul.

### 4.2 Évaluation des performances

Une base de données spécifique est élaborée pour valider notre approche. Toutes les images de la base ont été acquises avec une webcam Philips ToUcam Pro 740K à la résolution  $640 \times 480$  pixels. La base contient 160 images, certaines avec des mains possédant des bagues, provenant de 16 personnes. 10 images sont acquises pour chaque individu de la base.

Pour obtenir les coefficients optimaux et le meilleur filtre représentant la forme, un ensemble de paramètres est utilisé pour valider notre approche. Premièrement, un banc de filtre est créé pour déterminer le meilleur score de reconnaissance pour la paume seule. Chaque paume de la base de mains est comparée avec toutes les autres afin de déterminer ce score. Le taux de reconnaissance est le rapport entre le nombre de comparaisons incorrectes et le nombre total de comparaisons. Une comparaison est incorrecte si un utilisateur est accepté à tort ou si un utilisateur est rejeté à tort. Deux taux définissent ces erreurs: le taux de faux acceptés (FAR) défini par le rapport entre le nombre de personnes authentifiées à tort et le nombre de comparaisons et le taux de faux rejetés (FRR) défini par le rapport entre le nombre de personnes rejetées à tort et le nombre de comparaisons.

Les paramètres de tests pour le filtre de Gabor sont inspirés de [12], l'orientation du filtre est testée suivant huit valeurs  $\theta=22.5^\circ, \theta=45^\circ, \dots, \theta=180^\circ$ , trois valeurs sont utilisées pour la fréquence spatiale  $\lambda=5.47, \lambda=8.20$  et  $\lambda=10.93$ , le couple  $(\xi, \eta)$  et la phase  $\phi$  sont fixés à  $(0,0)$  et  $0$  respectivement, tandis que la déviation standard  $\sigma$  est déterminée suivant le rapport  $\sigma/\lambda=0.56$ . Le meilleur taux de comparaison présenté par la courbe de la Figure 7 est obtenu avec les paramètres  $\theta=157.5^\circ, \lambda=10.93$  et  $\sigma=6.12$ . Avec ces coefficients, l'EER (Equal Error Rate) correspondant à FAR=FRR indique une erreur de 2.25%. Le filtre de Gabor complexe présenté en [3] avec les paramètres optimaux montrent des résultats similaires sur notre base de données. Pour diminuer les erreurs de reconnaissance, nous avons présenté une nouvelle méthode de fusion des caractéristiques de forme et de texture de la paume. Après expérimentations, le filtre optimal définissant la forme est seulement composé des largeurs des doigts. Ainsi, le filtre optimal H est défini par:

$$H(x, y) = l(x, y + 1) \times \frac{1}{\sum_{i=0}^4 \sum_{j=0}^4 l[i][j + 1]} \quad [11]$$

Les performances du système de reconnaissance globale sont bien augmentées entre la reconnaissance par la paume seule et la reconnaissance globale de la main (Fig. 6). En effet, l'EER est égal à 1.85% et un système plus sécurisé peut être défini avec FAR=10<sup>-3</sup>% et FRR=2.2%.

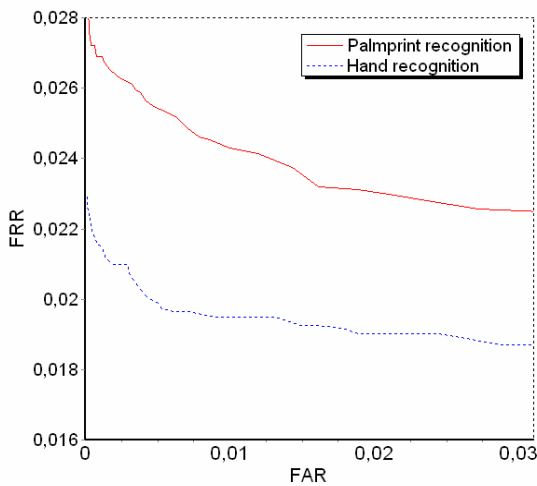


Figure 6 - Reconnaissance globale

Le processus de reconnaissance complet (segmentation, extraction des caractéristiques et comparaison) est effectué en moins de 1 seconde sur un Pentium M à 1.6GHz.

## 5 Conclusion

Dans ce papier, nous avons présenté une nouvelle méthode de reconnaissance biométrique de la main pour

un système sans contact. Tout d'abord, la segmentation de la main est expliquée. Elle est effectuée grâce à une intégration des composantes couleurs de la peau et un modèle de forme de main. Ensuite, le processus d'authentification par fusion est décrit. Il s'appuie sur une combinaison par convolution des données géométriques de la main et de la texture de la paume. Les caractéristiques de la paume sont déterminées par un filtre de Gabor 2D dans le domaine réel permettant une représentation compacte. Les données du contour de la main sont extraites du modèle de forme ayant permis la segmentation. Le processus complet est validé après expérimentations. Il présente un taux d'erreur de reconnaissance de 1.85% pour un temps d'exécution inférieur à 1 seconde.

Afin de gérer la rotation dans l'espace de la main, une méthode invariante aux perspectives et aux cisaillements devra être recherchée. De plus, la base de tests devra être diversifiée et complétée pour confirmer nos approches. Ces deux pistes font actuellement l'objet de notre travail.

## 6 Références

- [1] Y. Zhu. Hand Detection in an Active Vision System. Thèse, 2003.
- [2] N. Duta, A. K. Jain, K. V. Mardia. Matching of palmprints. Pattern Recognition Letters: 23(4), 477-485, Février 2002.
- [3] W. Kong, D. Zhang, W. Li. Palmprint feature extraction using 2-D Gabor filters. Pattern Recognition: 36(10), 2339-2347, Octobre 2003.
- [4] G. Lu, D. Zhang, K. Wang. Palmprint recognition using eigenpalms features. Pattern Recognition Letters: 29(9-10), 1463-1467, Juin 2003.
- [5] X. Q. Wu, K. Q. Wang, D. Zhang. Wavelet Based Palmprint Recognition. Conference on Machine Learning and Cybernetics, Beijing, Novembre 2002.
- [6] A. Kumar, C. M. Wong, C. Shen, A. K. Jain. Personal Verification using palmprint and hand geometry biometric. Proc. Intl. Conf. Audio Video based Biometric Personal Authentication, Washington D. C., Mars 1999.
- [7] J. Doublet, O. Lepetit, M. Revenu. Hand detection for contact less biometrics identification. Intl. Conf. Cognitive System with Interactive Sensors, Paris, Mars 2006.
- [8] L. Lucchese, S. Mitra. Color image segmentation: A state of the art survey. Proc. of the Indian National Science Academy (INSA-A), New Delhi, India: 67(2), 207-221, Mars 2001.
- [9] L. I. Smith. A tutorial on Principal Components Analysis. 2002.
- [10] T.F. Cootes, C.J. Taylor. Statistical models of appearance for computer vision. Technical report, University of Manchester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, Manchester M13 9PT, United Kingdom: 1999.
- [11] S. Di Zenzo. A note on the gradient of a multi-image. Computer Vision, Graphics and Image Processing: 33(1), 1986.
- [12] P. Kruizinga, N. Petkov. Nonlinear Operator for Oriented Texture. IEEE Trans. on Image Processing: 8(10), 1395-1407, 1999.