

Who_Is: système d'identification des visages humains

Mohamed Tayeb Laskri, Djallel Chefrour

► **To cite this version:**

Mohamed Tayeb Laskri, Djallel Chefrour. Who_Is: système d'identification des visages humains. Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées, INRIA, 2002, 1, pp.39-61. hal-01261701

HAL Id: hal-01261701

<https://hal.inria.fr/hal-01261701>

Submitted on 25 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1. Introduction

De nos jours, ils existent des ordinateurs capables d'effectuer des tâches, qui faites par l'homme, demandent de l'intelligence. Malgré tout ce progrès, la machine est toujours située à des années lumières du cerveau humain – si comparaison a lieu entre les deux –, ceci est dû principalement à la complexité des tâches que nous cherchons à automatiser et à notre méconnaissance de cet organe mystérieux qu'est le cerveau et des mécanismes qui régissent son activité (l'activité cérébrale).

La reconnaissance des visages humains [4] [16] [7] [32] [12] [23] [30] [33] est considérée actuellement comme l'une des tâches les plus importantes pour relever le défi en reconnaissance des formes [18] [19] [11]. La facilité et la précision avec lesquelles nous pouvons identifier nos amis et nos ennemis, même dans des conditions défavorables, sont des capacités des plus étonnantes du système visuel humain, le but de la reconnaissance des visages est de concevoir des systèmes informatiques capables d'égaliser les êtres humains dans ce domaine.

S'investir dans le domaine de la reconnaissance des visages est sans doute motivé par la multiplicité et la variété des champs d'applications de celui-ci (télé-surveillance et vérification d'identités, chirurgie plastique, langage des signes...). L'intérêt pour ce domaine s'explique aussi par le fait que par rapport aux autres filières de la reconnaissance des formes (reconnaissance de l'écriture, de la voix, etc.), la reconnaissance des visages n'a pas atteint un niveau aussi avancé que ces autres filières. Ceci s'explique avant tout par la complexité de la forme en question (le visage) comparée avec les autres formes.

Pour notre système *WHO_IS*, la description du visage est faite par un modèle géométrique basé sur un ensemble de points caractéristiques extraits à partir de l'image de l'individu, ce modèle permet de représenter le visage de chaque personne par un vecteur caractéristique, la méthode de reconnaissance utilisée est celle des K-PPV (K Plus Proches Voisins).

2. Reconnaissance des visages humains

La Reconnaissance des Visages Humains a pris de l'ampleur vers la fin des années soixante dix pour devenir un domaine de recherche très actif [7]. Le système WISCAD réalisé en 1980 fut l'un des premiers systèmes de reconnaissance de visages, ses inventeurs ont adopté l'approche connexionniste [10]. Depuis, plusieurs approches ont été proposées. *L'identification* qui revient à retrouver l'identité d'une personne à partir d'une vue du visage, est une tâche cognitive complexe, les systèmes qui cherchent à la

simuler sont appelés : les systèmes biométriques [15]. La reconnaissance d'un visage est une tâche de mémoire épisodique, elle revient à décider si ce visage a été vu précédemment .

2.1. Apport des sciences cognitives

Comme dans d'autres disciplines, beaucoup de chercheurs intéressés par la reconnaissance des visages se sont inspirés du système visuel humain pour réaliser des systèmes automatiques capables d'égaliser les capacités humaines dans ce domaine. Comprendre le fonctionnement du système visuel de l'homme, est une tâche qui incombe aux neuropsychologues et aux cognitivistes. Ils essaient de trouver des réponses à des questions du type :

- Comment le système visuel humain traite-t-il les visages ?
- Comment est-il stocké en mémoire et comment un visage devient familier ?
- Comment se généralise le processus de reconnaissance pour de nouvelles images ?
- Y a-t-il, dans le cerveau, un mécanisme spécifique pour la reconnaissance des visages ou est-ce qu'il s'agit d'un seul et d'un même mécanisme pour tout type de forme ?
- ...

Bien que la plupart de ces questions attendent toujours une réponse, certains progrès ont été réalisés. En effet, depuis l'arrivée des nouvelles technologies d'imagerie cérébrale, les recherches ont réellement pu progresser sur une meilleure connaissance du système visuel humain [26] :

- Certains chercheurs comme Shepherd [24] avancent que la reconnaissance des visages ne semble pas être une tâche spécialisée dans le cerveau, mais emploie plutôt des mécanismes utilisés pour des tâches générales de reconnaissance d'objets, alors que d'autres comme Moses [16] estiment qu'on ne peut pas appliquer au visage les techniques de R.F. employés pour d'autres objets, à cause des grandes différences dans les natures mêmes de ces formes.
- Les mécanismes dans le système visuel humain pour la reconnaissance de l'identité et pour la reconnaissance de l'expression sont complètement séparés. Ceci implique qu'il est possible de reconnaître un individu indépendamment de son humeur [21] [2].
- Les caractéristiques les plus importantes utilisées pour l'identification des visages humains semblent être [9] : la région des yeux et des sourcils, la bouche et la lèvre inférieure, le nez, le menton.

2.2. Processus de reconnaissance des visages

Tout processus automatique de reconnaissance des visages doit prendre en compte plusieurs facteurs qui contribuent à la complexité de sa tâche, car le visage est une entité dynamique qui change constamment sous l'influence de plusieurs facteurs. La Figure 1. illustre la démarche générale adoptée pour réaliser de tels systèmes.

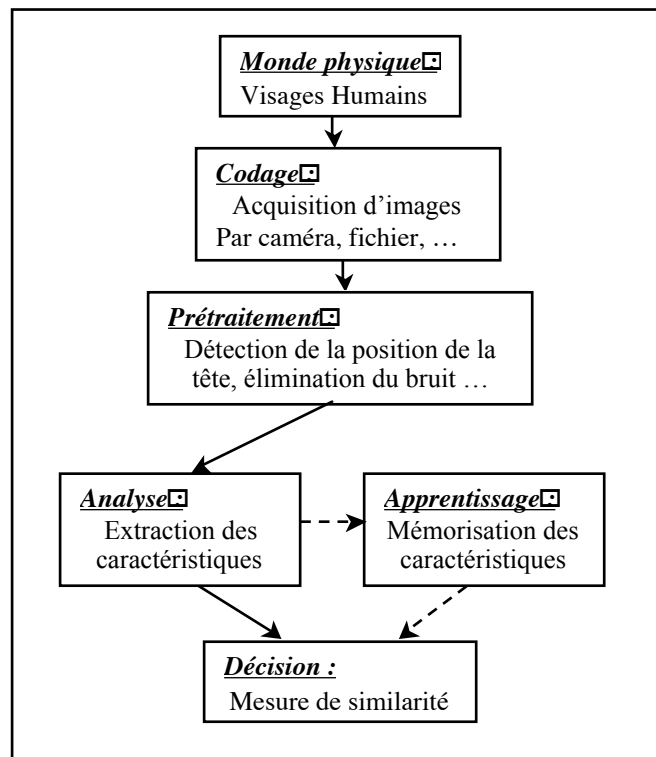


Figure 1. Schéma général d'un système de reconnaissance des visages

Dans le *monde physique*, il y a trois paramètres à considérer : *L'éclairage*, la *variation de posture* et *l'échelle*. La variation de l'un de ces trois paramètres peut conduire à une distance entre deux images du même individu, supérieure à celle séparant deux images de deux individus différents.

Le *codage* consiste en l'acquisition d'image et sa digitalisation, il comporte un risque de bruit et donne lieu à une représentation 2-D (la matrice des niveaux de gris) pour un objet 3-D (le visage). Peu de méthodes 3-D ont été proposées [17] [5], mais cela devrait

changer dans le futur proche avec l'invention des Range-Cameras qui peuvent mesurer la topologie 3-D du visage.

Dans le *prétraitement* il faut éliminer le bruit par des techniques de traitement et de restauration d'images et procéder à une localisation de la tête dans l'image, cette opération est très complexe, beaucoup supposent que l'image contient un visage et que l'arrière plan est neutre. La restauration d'images ou l'élimination du bruit consiste à compenser les dégradations connues ou estimées et rétablir la qualité initiale de l'image [1] [6].

Dans l'étape *Analyse* (appelée aussi indexation, représentation, modélisation ou extraction de caractéristiques), il faut extraire de l'image les informations qui seront sauvegardées en mémoire pour être utilisées plus tard dans la phase Décision. Le choix de ces informations utiles revient à établir un modèle pour le visage, elles doivent être discriminantes et non redondantes.

L'apprentissage consiste à mémoriser les représentations calculées dans la phase Analyse pour les individus connus.

La décision : Pour estimer la différence entre deux images, il faut introduire une mesure de similarité.

2.3. Les Méthodes Existantes

De nombreuses approches ont été proposées, elles se répartissent en deux catégories □ Les méthodes globales et les méthodes géométriques.

2.3.1. Les Méthodes Géométriques

On les appelle aussi les méthodes à traits, à caractéristiques locales, ou analytiques. L'analyse du visage humain est donnée par la description individuelle de ses parties et de leurs relations. Ce modèle correspond à la manière avec laquelle l'être humain perçoit le visage, c'est à dire, à nos notions de traits de visage et de parties comme les yeux, le nez, la bouche, etc. La plus part des travaux réalisés se sont concentrés sur l'extraction des traits à partir d'une image du visage et sur la définition d'un modèle adéquat pour représenter ce visage. Un certain nombre de stratégies *automatiques* et *semi-automatiques* ont modélisé et classé les visages sur la base de distances normalisées et angles entres points caractéristiques. La stratégie est qualifiée d'automatique ou non selon que la phase d'extraction des points est faite par l'ordinateur ou qu'elle est assistée par un opérateur. Cette phase constitue l'étape clé du processus, car la performance du système entier dépend de la précision avec laquelle les informations utiles sont extraites [29] [13] [30] [8].

L'avantage de ces méthodes est qu'elles prennent en compte la particularité du visage en temps que forme naturelle à reconnaître, en exploitant les résultats de la recherche en neuropsychologie et psychologie cognitive sur le système visuel humain. La difficulté éprouvée quand il s'agit de prendre en considération plusieurs vues du visages ainsi que le manque de précision dans la phase "extraction" des points constituent leur inconvénient majeur.

2.3.2. Les Méthodes Globales

Cette classe regroupe les méthodes qui mettent en valeur les *propriétés globales* de la forme. Le visage est traité *comme un tout*. Parmi les approches les plus importantes réunies au sein de cette classe, on trouve :

L'Approche connexionniste [28] [33] [2] [23]. L'inconvénient de cette approche est que l'apprentissage est long. A ce jour, il n'est pas claire comment de tels systèmes vont s'étendre à des problèmes plus larges, il faut prendre en compte un grand nombre d'individus.

L'Approche ACP (ou Les Visages Propres), son but est de capturer la variation dans une collection d'images de visages et d'utiliser cette information pour coder et comparer les visages (en termes mathématiques : trouver les vecteurs propres de la matrice de covariance de l'ensemble des images de visages). Le nombre possible de visages propres peut être approximé en utilisant seulement les meilleurs visages propres qui correspondent aux plus grandes valeurs propres [27]. Cette approche rencontre le problème du coût des calculs élevé et celui de la détermination du nombre de visages propres utiles [9].

Les *Cartes d'Iso densité (Isodensity Maps)* inspirées de la géologie. Une carte d'isodensité est un ensemble de courbes fermées obtenues par le traçage des frontières des régions de même densité après quantification de l'image en niveaux de gris [MAT 95].

L'Approche Stochastique, Samaria [20] avance que, quand les images frontales sont balayées de haut en bas il y a un ordre naturel dans lequel les caractéristiques apparaissent, et ce fait peut être modélisé d'une manière pratique en utilisant un modèle caché de Markov (HMM : Hidden Markov Model).

L'approche statistique et l'approche probabiliste : Yang et Ahuja [31] présente une méthode pour détecter des visages humains à partir d'images en couleur. Un modèle de la couleur de la peau humaine basé sur une analyse statistique multivariante est construit pour capturer les propriétés chromatiques, et dans [32] Yang, Ahuja et Kriegman présentent une méthode de probabilité pour détecter les visages humains qui utilisent un mélange d'analyseurs de facteur.

3. Le Système WHO_IS

Les enquêteurs sont les utilisateurs typiques de WHO_IS, il peut les assister dans l'identification de suspects en disposant d'une base d'images des visages des personnes recherchées. WHO_IS reçoit en entrée l'image du visage de 3 façons possibles :

- Par appareil photo numérique (présence de l'individu à reconnaître).
- Par scanner (photographie de l'individu).
- Ou à partir d'un fichier graphique.

L'image acquise est ensuite affichée sur l'écran pour que l'opérateur puisse à l'aide de la souris localiser les points caractéristiques choisis dans le modèle. Ensuite, WHO_IS effectuera la reconnaissance et affichera le résultat obtenu.

Le système WHO_IS possède les caractéristiques suivantes :

- Ne traite que la vue de face du visage. (Ne gère pas le problème de variation de posture).
- Invariant à la présence de moustaches et de barbes, et indifférent aux expressions faciales (tant que la bouche n'est pas totalement couverte ni trop ouverte).
- Invariant au port de chapeau, casquette, foulard, de lunettes (tant que sa ne cache pas les yeux). Le maquillage n'a pas d'influence non plus.
- Pas de contraintes imposées à l'environnement (Changement d'éclairage est sans influence, et l'arrière plan peut être quelconque).
- Invariant au changement d'échelle (taille de l'image et distance entre le visage et l'objectif de l'appareil photo).
- L'image peut être en niveaux de gris ou contenir un nombre quelconque de couleurs.
- Le problème posé par la vieillesse peut être contourné en procédant à une mise à jour du contenu de la base périodiquement ou à chaque identification correcte.

3.1 Architecture de WHO_IS

La Figure 2 montre le schéma bloc du système WHO_IS avec ses deux modes de fonctionnement : Apprentissage et Identification. Il est constitué de trois parties, les sous-systèmes de : traitement d'images, d'identification et de base de données. Le dernier sous-système est intégré dans le deuxième.

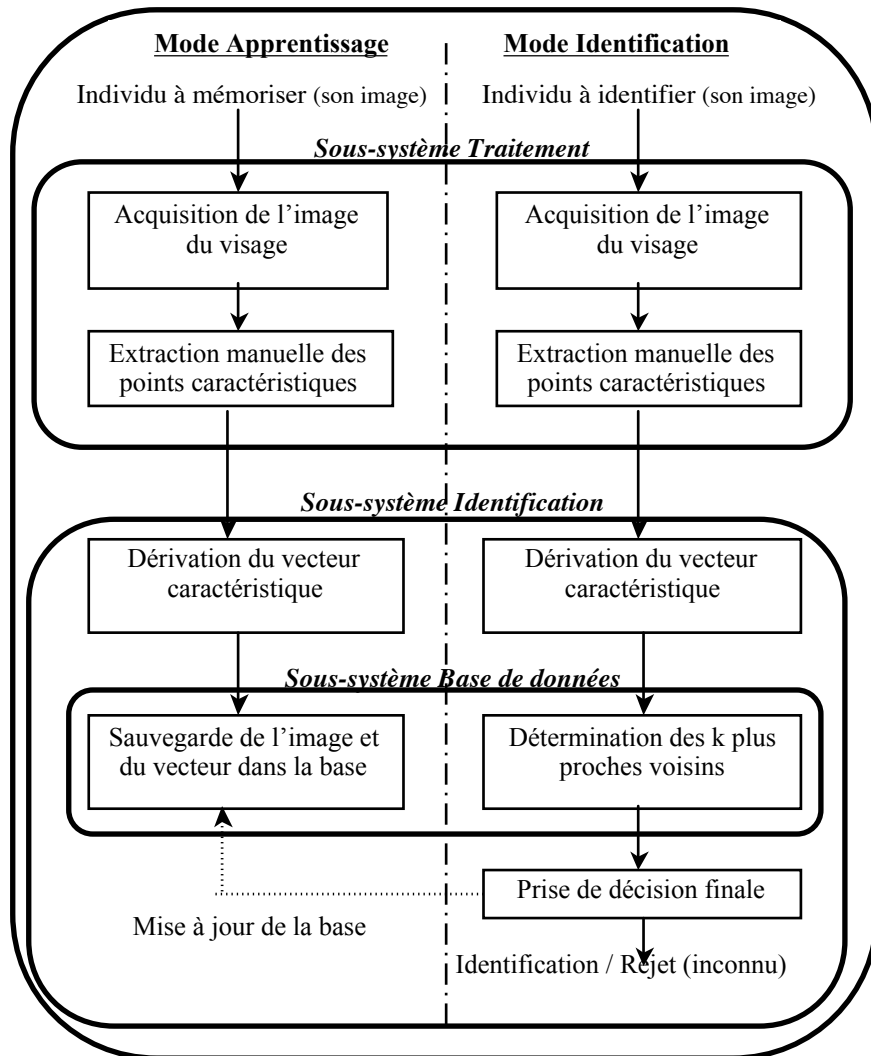


Figure 2. Schéma bloc du système WHO_IS

3.1.1. Sous système de traitement d'images

L'image contient en chaque point l'intensité lumineuse perçue par un appareil photo (ou tout autre système d'acquisition) dont le rôle est de procéder à une conversion analogique-numérique pour rendre cette image exploitable par la machine. Cette

numérisation se fait en deux étapes. D'abord un *échantillonnage* du support de l'image la remplace par une grille de points dans un système d'axes, ensuite une *quantification* de l'intensité lumineuse est réalisée. Ceci constitue la phase "Codage" du processus de reconnaissance. Toutes les images qu'on a utilisées ont été obtenues avec un appareil photo numérique, elles sont de très bonne qualité et nous avons jugé qu'il est inutile de développer des procédures d'élimination de bruit puisqu'il n'y en a pas, les performances de notre méthode sont pratiquement les mêmes dans un contexte d'utilisation plus variable comme avec des images provenant de photographies analogiques ou bien de fichiers graphiques car les points caractéristiques (Cf. Fig. 3) sont extraits par l'opérateur de façon manuelle. Le problème de localisation de la tête dans l'image n'est pas adressé étant donné que c'est l'utilisateur qui effectue l'extraction des points caractéristiques de façon manuelle. Ainsi la phase "Pré traitement" a été jugée inutile pour notre méthode.

Modèle géométrique de WHO_IS

Les points caractéristiques (à extraire) sont ceux définis par le modèle géométrique développé pour le système *WHO_IS*, et qui est basé sur celui proposé par Kamel et al. (1993) où ils avancent que 9 points bien choisis sont suffisants pour identifier un visage parmi un groupe de plus de 80 personnes mais que pour des bases de données plus larges le nombre de points caractéristiques doit être plus grand [13]. Cependant, le taux de reconnaissance correcte de leur système est 66%, ce qui est inférieur aux taux d'autres systèmes existants. Après avoir analysé le système de Kamel et al., on a conclu que l'erreur introduite au moment de l'estimation de l'orientation de la tête est la cause principale de la baisse du taux de reconnaissance. En effet, si par exemple la tête de l'individu est tournée vers la droite ce système estime que la distance AC reste approximativement invariante et que la correction concerne seulement les deux points B et D (voir Figure 3), ce qui n'est pas totalement vrai. En plus, le taux de reconnaissance peut être amélioré en augmentant légèrement le nombre de points caractéristiques (et donc la dimension du vecteur représentant l'individu) tout en veillant à ce que le système reste efficient.

Revenons maintenant à *WHO_IS*, dans un premier temps nous nous sommes intéressés à l'obtention d'un meilleur taux de reconnaissance pour l'identification d'un individu parmi un groupe plus important et avec un coût de calculs minimal, ceci dépend du choix des points caractéristiques nécessaires et par conséquent dépend des informations utiles dérivées de ces points et qui vont représenter l'individu en question. Bien que la résolution du problème de la variation de posture constitue une extension utile, *WHO_IS* ne traite que les vues de face et ce problème n'est pas adressé à ce stade du travail, mais fait plutôt partie des perspectives d'avenir en rajoutant la largeur du nez ainsi que l'angle de son inclinaison (vue de profil).

Le choix des points caractéristiques est basé sur les résultats des recherches de la psychologie cognitive concernant le système visuel humain. Notamment le fait que la région des yeux, celle du nez et celle de la bouche soient importantes pour l'identification, les points qu'on utilise couvrent l'ensemble des trois régions, ils sont au nombre de 12 : 10 points (A à J) extraits manuellement et 2 autres (S et M) déduits à partir des 10 premiers. Nous avons réutilisé les 9 points du modèle de Kamel et al. plus 3 points que nous avons rajoutés. Il s'agit, tel qu'illustré par la Figure 3, de :

- Les quatre coins des yeux : A, B, C, et D.
- Le point S centre de gravité de A, B, C et D (milieu de la tête).
- Le point G à la base du nez.
- Les points extrême droit et extrême gauche du nez : E et F.
- Le point H milieu du contour supérieur de la lèvre supérieure.
- Les deux coins de la bouche I et J.
- Le M à l'intersection des droites (SG) et (IJ) représentant le milieu de la bouche.

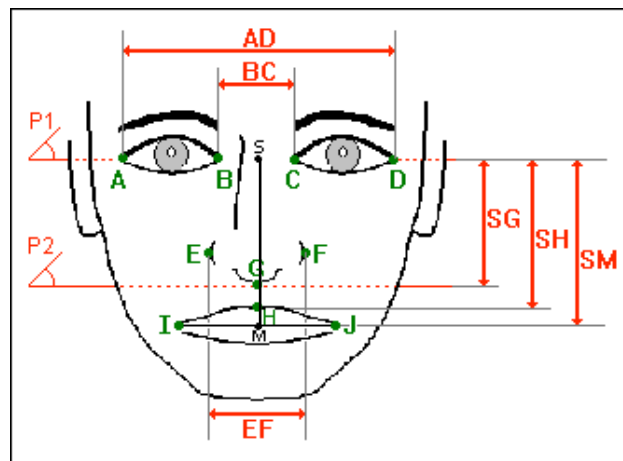


Figure 3. *Modèle géométrique du visage dans le système WHO_IS.*

A part le fait que la bouche de l'individu ne doive pas être trop ouverte, ce qui correspond aux conditions normales, les points choisis permettent au modèle d'être invariant aux expressions du visage (sourire, colère, tristesse...), ceci s'explique par le fait que la plus part de ces points sont liés à la structure osseuse sous-jacente (le crâne) et qu'on a ignoré les caractéristiques qui varient selon l'expression du visage comme la forme des sourcils et la largeur de la bouche.

Après acquisition, l'image est affichée à côté d'une "liste des caractéristiques" contenant les 10 points ainsi que 2 autres variables C1 et C2, l'utilisateur muni d'un dispositif de pointage (e.g. la souris) localise avec précision le point de l'image du visage dont le nom est sélectionné dans la liste, à chaque fois que les coordonnées d'un point sont récupérés le système sélectionne automatiquement le point suivant et ainsi de suite. Quand les deux points A et D sont situés, le système dessine tout au long de l'image le plan horizontal P1 (voir Figure 3) passant par ces deux points, le plan P2 parallèle à P1 apparaît suite au click localisant le point H. Ces 2 plans permettent la localisation de la position des oreilles (voir Figure 4), ils sont associés aux variables C1 et C2 dont les valeurs sont choisies par l'utilisateur parmi celles de la liste suivante :

- Le plan coupe (de manière franche) les oreilles.
- Le plan est tangent aux oreilles.
- Le plan ne coupe pas les oreilles.
- Information non disponible (oreilles cachées...)

La Figure 5 montre des exemples des plans horizontaux P1 et P2 tracés sur quelques images ainsi que les valeurs correspondantes de C1 et C2. Pour enlever l'ambiguïté dans le cas où la position d'une oreille par rapport à un plan serait différente de celle de l'autre oreille par rapport à ce même plan, on a défini la règle de priorité suivante : L'oreille coupée par un plan a la plus grande priorité, ensuite vient la tangente et puis la non coupée, l'utilisateur tient compte de l'oreille prioritaire et ignore l'autre.

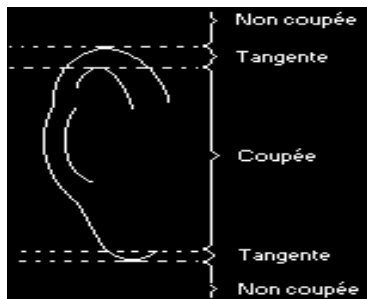


Figure 4. Position des oreilles par rapport aux plans P1 et P2.

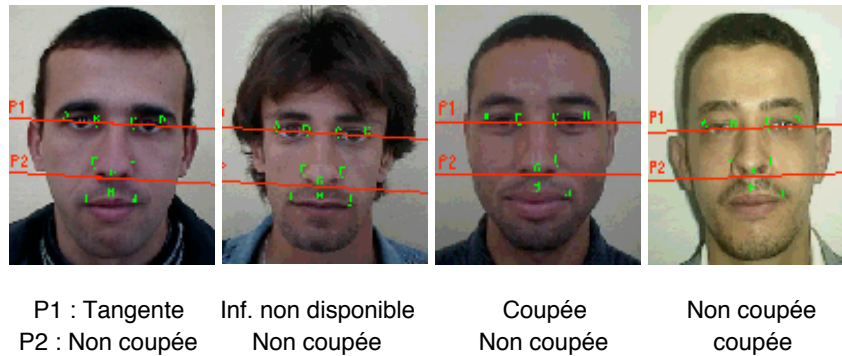


Figure 5. Exemples de positions d'oreilles par rapport aux plans P1 et P2

A n'importe quel moment l'utilisateur peut revenir en arrière à un élément de liste des caractéristiques déjà localisé pour rectifier la localisation précédente si nécessaire.

3.1.2. Sous-système d'identification

Les coordonnées des 12 points et les deux variables C1 et C2 seront utilisées dans le calcul du vecteur caractéristique qui représente l'individu, et qui sera sauvegardé dans la base du système pour être utilisé plus tard dans l'identification. Le Tableau 1 montre la codification adoptée pour ces variables C1 et C2 :

Position des oreilles par rapport aux plans P1/P2	Valeur de C1/C2
Information non disponible	0
Non coupées par le plan	1
Coupées par le plan	2
Tangentes au plan	3

Tableau 1. Codification de la position des oreilles.

En plus de C1 et C2 les autres composantes du vecteur caractéristiques sont des *distances normalisées* entre les 12 points extraits. Comme les informations choisies pour modéliser le visage doivent être discriminantes, compactes et non redondantes, nous avons sélectionné 6 parmi les 66 ($= C_{12}^2$) distances possibles entre les 12 points, à raison de 2 distances pour chaque région du visage :

- *Région des yeux* : Généralement l'œil gauche a la même taille que l'œil droit et donc les points A et B sont symétriques aux points C et D par rapport à S, pour éviter la redondance d'information ($AB = CD$ et $AC = BD$) nous n'avons utilisé que les 2 distances AD et BC pour coder la région des yeux.
- *Région du nez* : La hauteur du nez est représentée par la distance SG, et la largeur par la distance EF.
- *Région de la bouche* : Etant donné que la lèvre supérieure est liée à la mâchoire supérieure qui est immobile (contrairement à la mâchoire inférieure), on peut dire que la distance SH (voir Figure 3) est invariante, et de même pour la distance verticale SM puisque dans la plus part des cas la bouche n'est pas trop ouverte.

Ainsi, dans sa totalité, le vecteur caractéristique est :

$$V [C1 C2 AD BC EF SG SH SM]$$

3.1.3. Sous-système de base de données

Pour que le système puisse identifier une personne, il doit avoir sauvegardé son vecteur caractéristique ainsi que son image, cette sauvegarde peut être faite dans un fichier. Chaque enregistrement du fichier contient alors : l'identité de la personne, le vecteur caractéristique calculé et une référence au fichier graphique contenant l'image de cette personne. Cependant la structure de fichier a plusieurs inconvénients, particulièrement le fait qu'elle est non évolutive dans le sens où l'application qui l'utilise demeure fermée sur elle-même et difficile à maintenir ou à modifier pour prendre en compte de nouveaux besoins comme par exemple de nouvelles informations concernant la personne enregistrée (*erg.* âge, adresse, fonction). Pour cette raison nous avons choisi le concept des bases de données pour permettre à *WHO_IS* de rester ouvert (i.e. modification facile de la structure de la base et possibilité de partage de celle-ci avec d'autres applications).

Un système base de données est composé d'un SGBD et d'une base de données. La base de *WHO_IS* est une base relationnelle exploitée via le SGBD fourni par le système d'exploitation sous-jacent. Elle est constituée d'une seule relation appelée *Features* qui a 10 attributs dont le premier est le code de la personne qui joue le rôle de clé de relation, le second est le nom de la personne, et les 8 restants correspondent aux 8 éléments du vecteur caractéristique. Le Tableau 2 illustre quelque tuples de la relation avec les noms des 10 attributs dans la première ligne du tableau.

L'image est stockée dans un fichier à part en même temps que le tuple correspondant à la personne enregistrée, le nom du fichier image est construit à partir de la valeur de l'attribut *PersonID* (clé de la relation) ce qui permet de le retrouver une fois qu'on a la clé. Ce fichier a un format compressé qui permet d'économiser l'espace de stockage sur la mémoire de masse.

On ne peut pas parler de bases de données sans parler de l'administrateur de la base. Dans le cas de *WHO_IS* le rôle de l'administrateur est de gérer l'ensemble d'apprentissage, c'est à dire de décider quelles sont les personnes que le système est sensé connaître et qu'il faut enregistrer dans la base. Les autres utilisateurs de *WHO_IS* peuvent procéder à des identifications, mais ils ne peuvent pas changer le contenu de la base en ajoutant un individu ou en supprimant un autre.

3.2. Méthode de reconnaissance

Avec le système *WHO_IS* on attribue une identité à un individu ou on le déclare inconnu. La méthode de reconnaissance adoptée est celle des *k plus proches voisins (K-PPV)*. La distance utilisée est celle de *la somme des valeurs absolues* appelée aussi *distance de Minkowski de degré 1* ou *distance City-Block*, elle est équivalente, dans \mathbb{R}^6 , aux autres distances (e.g. la distance euclidienne,...) mais elle nécessite moins de calculs (à ne pas confondre avec les distances AD, BC,...,SM).

Avec la méthode des K-PPV il faut calculer la distance entre le vecteur caractéristique de la personne à identifier et chaque vecteur de la base, si le nombre N de personnes dans la base est très grand, le coût des calculs sera élevé. Pour remédier à ce problème le système sélectionne un sous-ensemble E de M individus parmi ceux enregistrés dans la base, E est appelé *ensemble des candidats à l'identification*, Les éléments de E sont les vecteurs W_i de la base dont les composantes C1 et C2 sont nulles (non déterminées) ou égales à celles du vecteur caractéristique V de l'individu test (à identifier).

La sélection est effectuée en envoyant une requête SQL du type SELECT au SGBDE

```
SELECT * FROM Features WHERE
  (EarsVsP1 =V.C1) OR (EarsVsP1=0) AND ((EarsVsP2=V.C2) OR
  (EarsVsP2=0))
```

Si C1 et C2 sont nulles alors M est égal à N (E = toute la base) ; Sinon, si on considère que les 3 valeurs possibles pour chacune des variables C1 et C2 sont équiprobables, et sachant qu'il y a $3^2 = 9$ combinaisons de valeurs, alors on peut dire qu'en moyenne M est égal à $N / 9$. Dans le cas où E serait vide (M=0) la procédure d'identification prend fin : l'individu test est inconnu, dans le cas contraire on passe à l'étape suivante comme le montre l'organigramme de la Figure 6.

L'étape suivante est le calcul des M distances D_i (distances City-Block) entre le vecteur caractéristique $V(c1, c2, v_1, v_2, \dots, v_6)$ de l'individu test et chaque élément $W_i(c1^i, c2^i, w^i_1, w^i_2, \dots, w^i_6)$ de l'ensemble E des candidats à l'identification : $D_i = \sum_j |v_j - w^i_j|$ où w^i_j et $v_j, j=1..6$, sont les distances normalisées du modèle.

Quand une distance D_i est calculée, on la compare avec un seuil Th , si elle dépasse ce seuil alors le candidat "i" de E ne ressemble pas à l'individu à identifier et il est donc ignoré. Le seuil Th constitue au fait la distance maximale qui peut exister entre deux images de visage du même individu, théoriquement cet écart doit être nul mais en pratique ceci n'est pas vrai à cause des variations des paramètres suivants :

- L'éclairage.
- Le vieillissement.
- Le port de lunettes, de barbes, de moustaches...
- La perte d'information due à la représentation 2D (qui est l'image) du visage qui est un objet 3D [5], cette perte est celle des 3^{èmes} coordonnées des points du modèle correspondant à la 3^{ème} dimension (épaisseur de la tête).
- La variation de posture : Même si on ne considère que les vues de faces il est pratiquement impossible, dans des conditions normales, de prendre deux photos d'une même personne avec la même position de la tête ou le même angle de prise de vue, même si on peut fixer l'appareil photo ce n'est pas toujours évident pour la tête, ceci donne lieu à une petite déformation des distances (composantes de V).

PersonId	PersonName	EarsVs		distAD	distBC	distEF	distSG	DistSH	DistSM
		P1	P2						
.....
31	MOUNIR H.	2	1	0,57010	0,19452	0,26937	0,34999	0,40566	0,43905
32	KHALED F.	2	1	0,58550	0,22271	0,24543	0,33397	0,42294	0,43067
33	HASSEN T.	2	1	0,60649	0,22625	0,25031	0,29966	0,41515	0,41791
34	DJAMEL C.	2	1	0,54071	0,20462	0,28266	0,34750	0,43958	0,43942
35	SABRI S.	1	2	0,55215	0,20042	0,23599	0,29182	0,42103	0,46322
.....

Tableau 2. Extrait de la table de la relation Features.

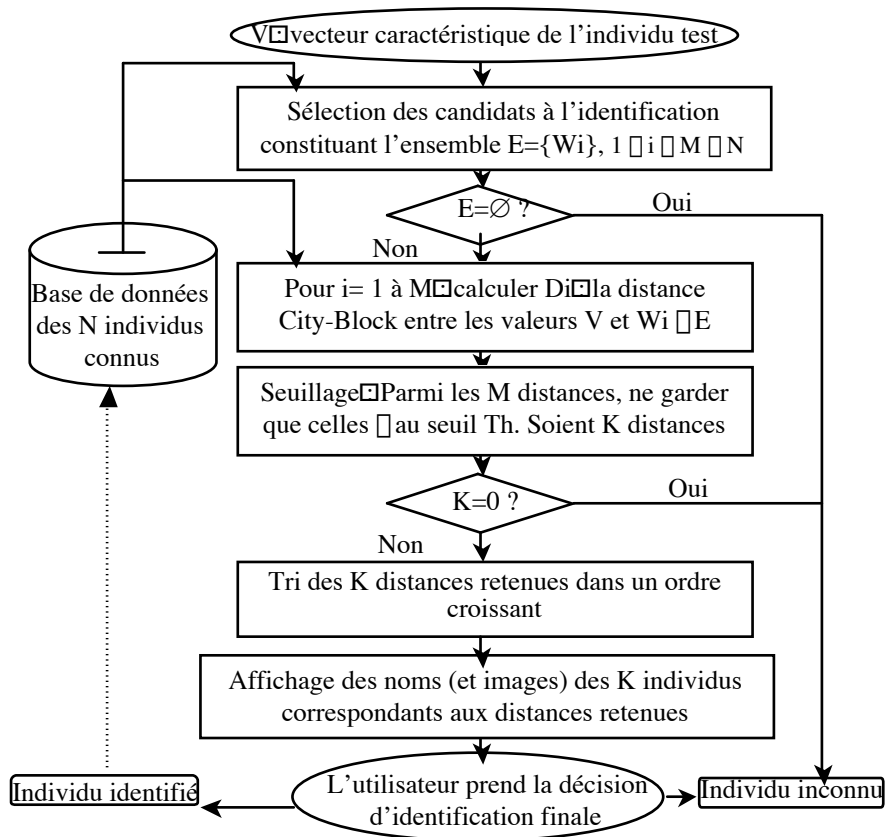


Figure 6. Organigramme de la méthode de reconnaissance de WHO_IS

La valeur du seuil Th est déterminée empiriquement en analysant les distances City-Block calculées (chacune) entre deux images du même individu sur un échantillon de 100 personnes. Les variations des paramètres n'ont pas une grande influence sur les résultats de l'identification puisque entre deux images du même individu (majorée par Th) est généralement inférieure à celle entre deux images de deux individus différents. Parmi tous les individus enregistrés dans la base, les K individus retenus après le seuillage sont ceux qui ressemblent à l'individu test, le nombre K est inférieur ou égal à la taille M de l'ensemble E des candidats à l'identification, il dépend de la valeur du seuil Th . Si K est nul alors l'individu à identifier est un inconnu, sinon on procède au tri des K distances retenues dans un ordre croissant.

3.3. Décision d'identification finale

Les identités des K individus qui ressemblent le plus à l'individu test sont affichées sur écran dans une liste ordonnée en commençant par le plus ressemblant. L'utilisateur peut alors, en sélectionnant un élément de liste, visionner l'image de cette personne et décider s'il s'agit de l'individu à identifier ou non. Si ce dernier ne figure pas dans la liste alors il s'agit d'un inconnu (non enregistrée dans la base) qui ressemble au K candidats de la liste. Dans le cas contraire (identification correcte) l'utilisateur peut effectuer une mise à jour de la base en remplaçant l'ancienne image de l'individu identifié par la nouvelle image (i.e. remplacer l'image qui a servi pour l'apprentissage par celle qui a servi pour l'identification). Le fait d'effectuer cette mise à jour périodiquement (e.g. chaque 6 mois) permet au système d'être invariant aux changements du visage dus à la vieillesse (effet du temps).

4. Résultats et Discussions

L'évaluation d'un système de reconnaissance des visages humains revient à déterminer, par un test, le *taux de reconnaissance correcte*, qui est la probabilité avec laquelle une personne enregistrée dans la base du système est correctement reconnue en donnant son nom. Dans le cas de *WHO_IS* on a utilisé pour le test 200 images de visage pour 100 personnes (des étudiants universitaires), c'est un échantillon plus large que ceux des systèmes étudiés [10] [27] [29] [13]. Pour chaque individu on a pris deux photos (vue de face) : Une pour l'apprentissage et l'autre pour l'identification. Dans l'ensemble la base était constituée de personnes moustachues et sans moustaches, imberbes et avec barbe, chauves et avec longue chevelure. Les photos ont été prises de jour et de nuit (utilisation de flash), en plein air et dans des environnements clos (salles, bureaux ...), avec variation des expressions faciales et changement de l'éclairage pour la même personne, pour celles qui portaient des lunettes transparentes on a pris une photo avec et une autre sans lunette.

Dans un premier temps ce test a servi pour la détermination de la valeur du seuil Th qui représente la distance maximale pouvant exister entre deux images du même individu. Et puis, on a déterminé le taux de reconnaissance correcte de *WHO_IS*. Les résultats obtenus sont résumés par le Tableau 3 et le graphe de la Figure 7.

Les valeurs de la moyenne et de l'écart-type de la distance City-Block entre chaque individu et lui-même sont 0,03928619 et 0,013926 respectivement. La faible valeur de l'écart-type montre que les distances (pour les 100 individus, chacun avec lui-même) sont regroupées au tour de leur moyenne et que généralement il n'y a pas une grande dispersion, la Figure 7 apporte une confirmation visuelle à cette constatation. On peut conclure que la distance séparant deux images du même individu ne dépend pas de

l'échantillon considéré mais elle est fonction des conditions de fonctionnement du système et des paramètres comme l'éclairage, le port de lunette, la variation de posture, la présence de barbes, de moustaches, la perte d'informations relative à la 3^{ème} dimension...

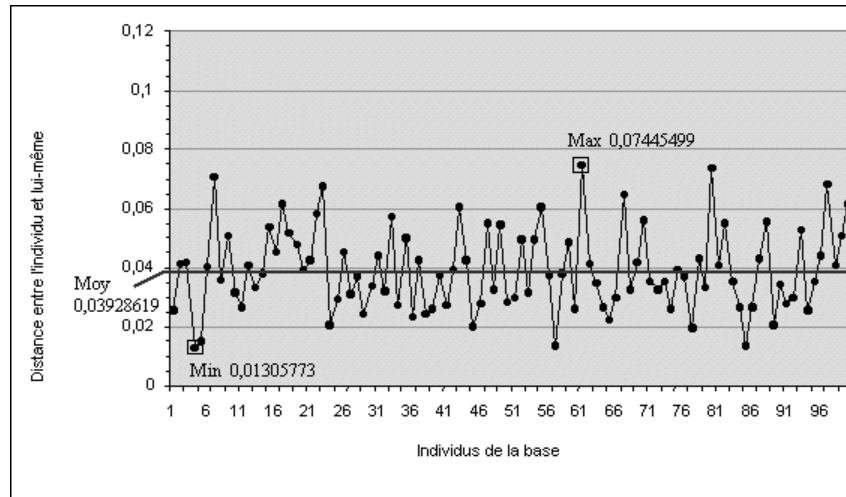


Figure 7. Détermination de la valeur du seuil Th .

On peut aussi lire sur la courbe, la valeur empirique du seuil Th représentant la distance maximale qui peut exister entre deux images de visage du même individu., elle vaut approximativement 0,08 puisque la courbe apparaît comme majorée par la droite horizontale d'équation $y = 0,08$. Cette droite est située un peu au-dessus de la valeur maximale 0,07445499 ce qui procure une certaine marge de sécurité ($\approx 0,05$) qui évite au système le rejet d'une personne connue.

Le nombre K des individus de la base ressemblant le plus à l'individu (à identifier) test n'est pas déterminé à l'avance comme c'est le cas dans le système de Kamel *et al.*, mais il dépend de la valeur du seuil Th . Le Tableau 3 montre que dans 99% des cas l'individu test est dans les quatre premiers de la liste (des K -PPV) calculée par *WHO_IS*, dans 97% il est dans les trois premiers, dans 93% il est dans les deux premiers et enfin dans 86% des cas il est tête de liste (voir Tableau 3).

Ordre dans la liste des K-PPV	1 ^{er}	2 ^e	3 ^e	4 ^e	5 ^e
Nombre d'individus identifiés	86	7	4	2	1

Tableau 3. Résultats du test d'identification.

Si on omet le rôle de l'utilisateur dans la prise de décision finale, on peut dire que le taux de reconnaissance correcte de *WHO_IS* est 86%, les 14% qui séparent le système de la perfection peuvent être expliqués par les deux facteurs suivants :

L'erreur commise pendant l'extraction manuelle des points caractéristiques : Les coordonnées récupérées suite à un click de la souris sur un point donné de l'image ne sont pas toujours exactes car l'utilisateur peut ne pas cliquer sur le point en question mais juste à côté. Cette erreur inévitable va se propager jusqu'à l'étape de calcul de distances pour donner lieu à un infime écart.

Le fait que les caractéristiques choisies soient discriminantes n'exclue pas la possibilité qu'ils y aient quelques exceptions à la règle, la probabilité d'avoir des individus qui se ressemblent beaucoup est minime mais non nulle.

WHO_IS ne risque pas de faire une fausse reconnaissance en attribuant à une personne à identifier l'identité d'une autre personne car si celle-ci n'est pas parmi les K PPV l'utilisateur conclut qu'elle est inconnue.

5. Conclusion et Perspectives

La méthode utilisée dans ce travail a l'intérêt d'être simple et robuste, et les résultats obtenus fournissent globalement une solution de qualité.

Comparé à d'autres systèmes, *WHO_IS* est caractérisé par certains avantages :

1. Taille importante de l'échantillon par rapport à d'autres systèmes. Il est encore possible d'élargir la base du système en prenant en compte un grand nombre d'individus, pour cela il faut d'abord estimer le nombre de personnes au-delà duquel le taux de reconnaissances de *WHO_IS* va chuter considérablement.

2. Il n'y a pas de contraintes sur l'environnement comme c'est le cas pour d'autres systèmes (arrière plan neutre, distance fixe entre l'individu et la caméra, pas de variation de l'éclairage).

3. *WHO_IS* est invariant au type de coiffure, au port de chapeau, foulard, lunettes (tant que ça ne cache pas les yeux), à la présence ou absence de barbe, de moustaches et à la variation de l'expression du visage, à la couleur, ...

4. Le test d'identification est réalisé avec des images différentes de celles utilisées pour l'apprentissage.

Le développement d'une méthode mixte qui bénéficie des avantages de l'approche globale et de l'approche géométrique, tout en exploitant le domaine de la morphométrie en chirurgie maxillo-faciale pour retenir des invariants utilisables dans le contexte de la recherche de similarités entre faces, permettra sans doute d'augmenter le taux de reconnaissance tout en considérant les différentes situations d'identification.

Le travail présenté dans ce papier est considéré comme l'étape d'une étude approfondie consacrée à la reconnaissance des visages humains, qui nous permettra sans doute à mettre en valeur les idées élaborées suite à la réalisation du système WHO_IS. L'étude des différents travaux de ces dernières années notamment les travaux sur la détection des visages humains dans des conditions générales et la reconnaissance des différentes expressions du visage humain [2] [3] [8] [11] [30] [31] [32] [33] nous permettra d'augmenter les champs d'application de *WHO_IS* tout en l'orientant vers les dernières tendances du domaine que nous résumons en :

– L'exploitation des nouvelles technologies d'imagerie cérébrale et l'obtention d'une meilleure connaissance du système visuel humain pour tenir compte des lois essentielles de la mémoire humaine. En effet, la mémoire influence notre présent psychologique et ce qui sera perçu, représenté et stocké dans notre environnement. Un système cognitif ne peut émerger que du fonctionnement d'un système « à mémoire ».

– L'exploitation de la complémentarité des mécanismes cognitifs et des mécanismes cérébraux pour étudier plus particulièrement la perception catégorielle des identités faciales et des expressions faciales, la spécificité des processus engagés dans le traitement des visages, le processus d'apprentissage qui conduit à la familiarisation, la reconnaissance d'identités faciales et les mécanismes qui sous-tendent l'association entre un visage et un nom.

– La détection des visages humains à partir d'images en couleur : Un modèle de la couleur de la peau humaine permet de capturer les propriétés chromatiques basé sur une analyse statistique multivariante. Cette méthode segmente des images dans des régions homogènes, et elle extrait les régions de la peau. Ces régions sont cernées jusqu'à ce que la forme soit approximativement elliptique.

– La reconnaissance des expressions du visage humain (joie, peur, dégoût, tristesse, colère et surprise) à partir de séquences vidéo. Basé sur des considérations utilisant le 'Facial Action Coding System', des points sont choisis sur le visage des sujets. Dans un premier temps, ces points sont automatiquement suivis d'une image à l'autre, et sont retrouvés de manière automatique en cas de perte entre deux images consécutives. Par la suite, les données issues du suivi des points sont regroupées de manière locale (lèvres, sourcils). Ces groupes sont transmis à un ensemble de classificateurs par région du visage. Le résultat de chaque classificateur est enfin combiné pour donner le résultat

final de chacune des classes d'expressions. La méthode de classification utilisée est aussi connue sous le nom de 'data fusion'.

- ...

Références Bibliographiques:

- [1] ACHOUR K., ZENATI N., LAGA H., : *Contribution of restoration of degraded images.* Actes du 4^e Colloque Africain sur la Recherche en Informatique CARI'98 (INRIA). Dakar, Sénégal, 12-15 octobre 1998.
- [2] BESSHO H., TWAI Y., YACHIDA M. : *Detecting Human Face and Recognizing Facial Expressions using potential Net.* Proceedings of the 15th International Conference on Pattern Recognition ICPR'2000, Vol. 1, p. 387-397, Barcelona, September 3-8, 2000.
- [3] BOUREL F. : *Reconnaissance robuste d'expressions sur visages humains.* Séminaire du L3I, Université la Rochelle , 2000.
<http://www-l3i.univ-lr.fr/~lmascari/seminaires.html>
- [4] BRUNNELLI R., POGGIO T. : *Faces Recognition : Features versus Templates.* IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 15, p. 1042-1052, October 1993.
- [5] CASS T.A. *Robust Affined Structure Matching for 3D Object recognition.* IEEE Transaction on Pattern Analysis and machine Intelligence, Vol. 20, Number 11, pp 1265-1270, November 1998
- [6] CHARBONNIER P., FERAU L.B., AUBERT G., BARLAU M. : *Deterministic Edge Preserving Regularisation in computer imaging.* IEEE Transaction on Image Processing, Vol. 6, Number 2, pp 298, February 1997.
- [7] CHELLAPPA R., WILSON C. L., and SIROHEY S.: *Human and machine recognition of faces: A survey.* Proceedings of the IEEE, p. 705-740, 1995.
- [8] CHIUNHSIUN L., KUO-CHIN F., : *Human Face Detection using Geometric triangle relationship.* 15th International Conference on Pattern Recognition ICPR'2000, Barcelona, September 3-8, 2000.
- [9] DAI D. Q., YUEN P. C. and FENG G. C.: *A multi-resolution decomposition method for human face recognition.* Proceedings/actes Vision Interface '98, p. 301-307, Vancouver, British columbia, June 1998.
- [10] H. D. ELLIS, M. JEEVES, F. NEWCOMBE and A. YOUNG: *Aspects of face.* Processing. Martinus Nijhoff Publishers, p. 426-441, Dordrecht, 1986.

- [11] GUO J. K., DOERMANN D., ROSENFELD A. : *Forgery Detection By local correspondence*. International Journal of Pattern Recognition and Artificial Intelligence, Vol. 15, No 4, p. 579, June 2001.
- [12] GROSS R., YANG J., WALBEL A. : *Growing Gaussian Mixture Models for Pose Invariant Face Recognition*. Proceedings of the 15th International Conference on Pattern Recognition ICPR'2000, Vol. 1, p. 1088-1091, Barcelona, September 3-8, 2000.
- [13] KAMEL M. S., SHEN H. C., WONG A. K. C. and CAMPEANU R. I.: *System for the recognition of human faces*. IBM Systems Journal, Vol. 32, No 2, p. 307-320, 1993.
- [14] MATSUHASHI S., NAKAMURA O. and MINAMI T.: *Human-face extraction using modified HSV color system and personal identification through facial image based on isodensity maps*. IEEE, vol. 2, p. 909-921, 1995.
- [15] MILLER B.: *Vital signs of identity (Biometrics)*. IEEE Spectrum, Vol. 31, Iss. 2, p. 22-30, February 1994.
- [16] MOSES Y.: *Face recognition: Generalization to novel images*. Thèse de PHD. The Weizmann Institute of Science, Israel, 1993. [Http://FTP.wisdom.weizmann.ac.il/pub/yael/th](http://FTP.wisdom.weizmann.ac.il/pub/yael/th)
- [17] NAGAMINE T., UEMURA Y. and MASUDA I.: *3D Facial image analysis for human identification*. IEEE Comput. Soc. Press, p. 324-331, 1992.
- [18] PAL S.K., MITRA S. : *Neuro-Fuzzy Pattern Recognition : Methods in Soft Computing*. John Wiley, N.Y., 1999.
- [19] PAL S.K., GHOSH A., KUNDU M.K. : *Soft computing for image processing*. Physica Verlag, Heidelberg, 2000.
- [20] SAMARIA F. S. and HARTER A. C.: *Parameterization of a stochastic model for human face identification*. IEEE Comput. Soc. Press, p. 138-180, 1994.
- [21] SANSONE S., TIBERGHIE G.: *Traitement de l'expression faciale et reconnaissance des visages : indépendance ou interaction ?* Revue Psychologie Française, N°39, p. 327-343, 1994.
- [22] SEITZ P. and BICHSEL M.: *The Digital Doorkeeper - Automatic face recognition with computer*. IEEE, p. 77-83, 1991.
- [23] SEHAD A., HOCINI H., HADID A., DJEDDI M., AMEUR S. : *Face recognition under varying views*. Actes du 5^e Colloque Africain sur la Recherche en Informatique CARI'2000 (INRIA), p. 375-382 Antananarivo, Madagascar, 14-19 octobre 2000.
- [24] SHEPHERD J., ELLIS H. D., DAVIES G.: *Perceiving and remembering face*. Academic Press, p. 171-197, 1987.
- [25] : SHIGEMURA T., MURAYAMA M., HONGO H., KATO K. et YAMAMOTO K. : *Estimating the face direction for the human interface*. Proceedings/acts Vision Interface' 98, p. 339-345, Vancouver, British Columbia, June 1998.

- [26] TIBERGIEN G., : *La mémoire oubliée*. Editions Pierre Mardaga, Liège, 1997.
- [27] TURK M. A. et PENTLAND A. P.: *Face recognition using eigenfaces*. IEEE Comput. Sco. Press, p. 586-591, June 1991.
- [28] VALENTIN D. et ABDI H. : *Modèles neuronaux, connexionnistes et numériques de la reconnaissance des visages*. It Psychologie française, 39(4), p. 357-392, 1994.
- [29] VEZJAK M., KOROSCEK J., GYERGYEK L., PAVESIC N., SAVSEK T., ERJAVC I. and GERE A.: *System for description and identification of individuals*. IEEE, Vol. 2, p. 1251-1254, 1991.
- [30] VOGEL S., SCHMID C., : *Face Detection on generic local descriptors and spatial constraints*. Proceedings of the 15th International Conference on Pattern Recognition ICPR'2000, Vol. 1, p. 1084-1087, Barcelona, September 3-8, 2000.
- [31] YANG M.H., AHUJA N., KRIEGMAN D., : *Detecting Human Faces in Color Images*. Proceedings of the 1998 IEEE International Conference on Image Processing (ICIP 98), pp. 127-139, Chicago, October, 1998.
- [32] YANG M.H., AHUJA N., KRIEGMAN D., : *Face Detection Using a Mixture of Factor Analyzers*. Proceedings of the 1999 IEEE International Conference on Image Processing (ICIP 99), Kobe, Japan, 1999.
- [33] YANG M.H., AHUJA N., KRIEGMAN D., : *Face Detection Using Mixtures of Linear Subspaces*. Proceedings of the fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG 2000), pp. 70-76, Grenoble, France, March, 2000.

