Reconnaissance automatique de visages

Charles Beumier, Marc Acheroy

Ecole royale militaire
Chaire d'électricité
Avenue de la Renaissance, 30 B1000 Bruxelles
http://www.elec.rma.ac.be/index.elec.html

7 mars 1996

1 Introduction

La reconnaissance automatique des visages connaît un intérêt croissant à en voir le nombre d'articles parus ces dernières années [1]. Plusieurs raisons peuvent expliquer ce phénomène, bien que le sujet ait toujours suscité l'intérêt des services de sécurité ainsi que des centres de recherche analysant le fonctionnement du cerveau.

Les demandeurs de systèmes de sécurité abandonneraient volontiers les cartes magnétiques ou codes d'identification (échangeables) au profit d'identification à base biométrique, intimement et indéfiniment liée à l'individu. C'est aussi la solution naturelle par excellence, utilisée par les êtres humains eux-mêmes. Mais les problèmes ne sont pas résolus pour autant, car nombre de caractéristiques biométriques sont difficilement exploitables ou occasionnent de l'inconfort lors de leur utilisation (les empreintes digitales nécessitent un contact physique, l'analyse du fond de l'oeil dérange beaucoup de sujets). L'analyse des visages est la base la mieux acceptée, probablement parce qu'elle nous est la plus familière, mais les traitements s'y rapportant sont aussi très délicats si l'on veut conserver un minimum de convivialité pour les individus. Ainsi la luminosité, le positionnement et l'attitude du sujet sont autant de paramètres qui devront rentrer en ligne de compte dans l'établissement des méthodes à mettre en oeuvre. Dans ce contexte, il semble évident que l'on attende encore aujourd'hui un minimum de coopération de la part des individus.

Une grande source d'information pouvant aider la conception d'un système automatique de reconnaissance de visages concerne l'étude de la capacité d'un être humain pour cette tâche. Il est clair que le besoin de reconnaissance des humains entre eux a créé, si pas l'organe (Y a-t-il une zone du cerveau dédicacée aux visages [3]?), leur surprenante capacité, principalement eu égard aux variations pouvant survenir dans le "signal" (lumière, maquillage, chevelure, expression, ...). Bien que les mécanismes biologiques ne soient pas encore complètement expliqués, on présente le processus de traitement des visages par le cerveau humain comme une double activité [2]. D'une part, le visage est localisé et globalement évalué en terme de forme, quantité et couleur de cheveux, ... C'est un processus rapide, s'intéressant aux caractéristiques extérieures du visage et peu discriminant, séparant les individus en grandes classes. D'autre part, le milieu du visage (yeux, bouche et nez, par ordre d'importance) est analysé afin d'en extraire des caractéristiques fines et dynamiques (expression). L'utilisation de cette deuxième activité est prédominante en cours de communication.

2 Méthodologies

Plusieurs approches ont été considérées pour la reconnaissance de visages [1]. On peut classer les méthodes selon l'information qu'elles exploitent (face, profil) ou l'approche poursuivie (comparaison globale, avec ou sans transformation; analyse de caractéristiques; modèles déformables).

La vue de face concerne la majorité des systèmes. C'est aussi le mode le plus naturel pour les êtres humains. On y retrouve le principal canal de communication constitué des yeux et de la bouche, acteurs d'expression. L'analyse débute habituellement par une localisation des yeux ou par une recherche de la tête (ovoïde). Les deux yeux définissent un repère permettant de réduire dans certaines limites les problèmes de translation, d'échelle et de rotation.

La vue de profil, bien que moins utilisée, permet une localisation rapide et efficace de points de référence nécessaires à la comparaison. L'exploitation, se basant principalement sur l'information géométrique (de volume) et non radiométrique (de couleur) est moins sensible à la luminosité et les profils présentent peu de variation aux rotations limitées de la tête.

Une première approche possible est la comparaison globale des données après normalisation (en couleur ou géométrique). Les résultats sont modestes vu le nombre de sources de variations externes (dues à la prise de vues) et internes (dues à la personne). Ces méthodes mettent principalement en évidence les caractéristiques globales du visage (cheveux, forme, taille). Plusieurs variations de la comparaison globale ont été entreprises selon le mode de normalisation réalisée au préalable. On peut ainsi faire ressortir les hautes fréquences ou même décomposer l'image en parties (yeux, bouche, ...) qui seront chacunes comparées séparément. Le système qui semble le plus prometteur analyse l'image dans un espace d'images de références afin de lui attribuer une classe.

A l'opposé des techniques globales se trouvent les méthodes qui extraient des caractéristiques précises du visage. Toute information qu'elle soit de couleur, de forme ou de distance est intéressante à condition qu'elle soit reproductible et discriminante. Ces techniques nécessitent des programmes plus élaborés, dépendants de l'application, mais plus faciles à contrôler. La combinaison des valeurs caractéristiques (éventuellement assorties d'une probabilité) requiert à elle seule une étude statistique et des algorithmes adaptés.

Pour être complet, signalons l'existence de méthodes partageant les propriétés globales et locales. C'est ainsi le cas des modèles déformables qui cherchent à minimiser une fonction de coût (globale) par l'adaptation de paramètres (caractéristiques). Dans la même optique, les architectures dynamiques (Dynamic Link Architecture) visent à déformer une image avec un minimum de distorsion afin de faire coïncider des informations locales de l'image avec une image de référence.

3 La recherche à l'ERM

Les objectifs de la recherche au "Signal & Image Centre" de l'Ecole royale militaire - Chaire d'électricité (Prof. M. Acheroy) concernent le contrôle d'accès par traitement numérique (automatique) des images de visage. En 1990, Christiaan Perneel [4] appliqua un réseau de neurones ("Kohonen Map") à la reconnaissance de visages. Il conclut par la nécessité d'un prétraitement des images en termes d'échelle, de rotation et de luminosité, sans quoi l'apprentissage serait prohibitif en mémoire et en temps requis.

En 1991, une étude de faisabilité basée sur la vue de face a permis de dégager les problèmes inhérants à ce type d'approche: plusieurs degrés de rotation difficilement contrôlables, grande influence d'un changement d'illumination, difficulté d'analyser certaines caractéristiques, due à la résolution limitée (yeux) ou par manque de stabilité (bouche, cheveux). La comparaison des caractéristiques devrait alors tenir compte de leur probabilité de détection, valeur difficilement estimable.

Il est apparu plus opportun de prendre les images de profil [5], ce qui simplifia la détection du visage, offrit de bons points de référence (nez, menton, sourcils, bouche - voir Fig. 2) tout en conservant une attitude naturelle de la personne, les profils étant peu sensibles à la rotation. C'est le contour avant de la tête, du front à la gorge, qui a fait l'objet d'une analyse en termes de courbure et d'angle. Ainsi le contour est quantifié selon des caractéristiques locales de courbure et globales d'angle. La solution présente les avantages de simplicité et de rapidité tout en nécessitant un encombrement mémoire faible (300 bytes par image de référence) et permettant des comparaisons rapides.

Toutefois, le profil n'apporte qu'une information partielle du visage complet. On a donc pensé à exploiter la surface du visage aux fins d'analyse géométrique [6] (comparables à celles menées pour le profil). Il est clair que l'on conservera les avantages de stabilité d'un grand nombre de caractéristiques (basées sur le nez, le menton, le front, les joues) et que l'on dépendra peu (en théorie pas) de l'illumination. Mais les capteurs de surface sont chers, souvent inadaptés aux visages et plutôt lents. C'est pourquoi, on a évalué l'adéquation de la lumière structurée [7] qui, projetée sur le visage (voir Fig. 3), permet par triangularisation de recouvrer les coordonnées tridimensionnelles de points sur la surface. Elle est actuellement en cours de mise au point (conception de la diapositive, fixation des éléments - projecteur et caméra, étalonnage en vraie distance).

Les activités de reconnaissance de visage suivent actuellement un projet européen ACTS ("m2vts"), dont l'objectif est d'apporter des solutions aux problèmes de contrôle d'accès aux bâtiments et aux sources d'information en fusionnant des systèmes se basant sur des modalités distinctes (principalement la parole et l'image). Dans un premier temps, le module d'analyse des profils servira à établir un prototype devant être opérationnel fin 1996. Ensuite, nous développerons la partie analyse des visages par acquisition de leur surface par lumière structurée. D'autres partenaires évaluent la vue de face et la synchronisation entre les mouvements de la bouche et la parole.

Figure 1: Image de profil

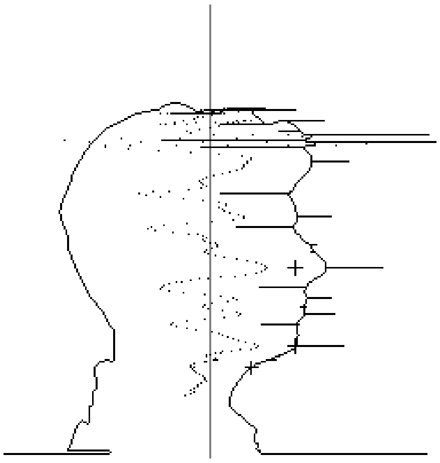

Figure 2: Courbure du contour et ses maxima

Références

[1] R. Chellappa, C.L. Wilson and S. Sirohey, "Human and Machine Recognition of Faces", In *Proceedings of the IEEE*, The institute of electrical and electronics engineers, inc., Vol. 83, No. 5, May 1995, pp. 705-740.

[2] J. Blanc-Garin, "Faces and non-faces in prosopagnosic patients", In *Aspects of Face Processing*, Doordrecht, Netherlands: Martinus Nijhoff Publishers, p. 274, 1986.

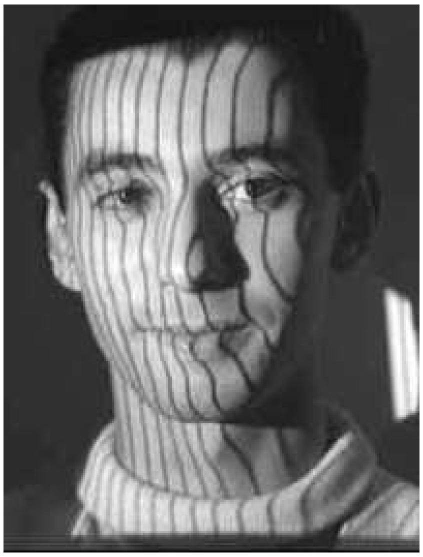

Figure 3: Image avec lumière structurée

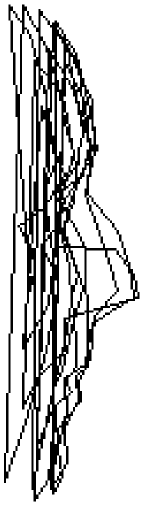

Figure 4: Reconstruction 3-D

[3] H.D. Ellis, M. Jeeves, F. Newcombe and A. Young, “Introduction to aspects of face processing: Ten questions in need of answers”, In *Aspects of Face Processing*, Doordrecht, Netherlands: Martinus Nijhoff Publishers, pp. 3-13, 1986.

[4] C. Perneel and M. Acheroy, “Face Recognition with Artificial Neural Networks”, *IMACS Annals on Computing and Applied Mathematics*, Proceedings MIMS2 ’90, Sept. 3-7, Brussels.

[5] C. Beumier, M.P. Acheroy, “Automatic Face Identification”, In *Applications of Digital Image Processing XVIII*, SPIE, vol. 2564, July 1995, pp. 311-323.

[6] G.G. Gordon, “Face recognition based on depth maps and surface curvature”, In *Geometric methods in Computer Vision*, SPIE, vol 1570, San Diego 1991.

[7] R.A. Jarvis, “A perspective on range finding techniques for computer vision”, In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-5, pp 122-139, March 1983.