

Détection et localisation de visage pour un système grand public d'authentification de visage implantable sur FPGA

L. PIERREFEU, J. JAY.

Laboratoire Hubert Curien (UMR CNRS 5516), Département Informatique et Image.
18 Rue du Professeur Benoît Lauras 42000 Saint-Étienne

Tél : 04 77 91 57 80 - Fax : 04 77 91 57 81 – lionel.pierrefeu@univ-st-etienne.fr

Résumé – Le projet dans lequel s'inscrit ce travail consiste en la réalisation d'un système mono puce grand public d'authentification de visage dont le but est le remplacement de la saisie du mot de passe sur des périphériques mobiles. La localisation du visage est une étape nécessaire dans cette chaîne de traitements afin de réduire de manière significative le temps d'exécution du système. La solution proposée utilise un réseau neuronal (RBF) pour une localisation basée sur la détection des contours du visage. Elle répond à un cahier des charges précis offrant une bonne adéquation entre complexité et performances.

Abstract – In this paper we present a part of a low cost mono chip system used as a classifier for a human face authentication. The aim of this application is to replace password identification on mobile devices. Face detection is an important step for time consuming minimization. We present a neural network based solution for face edge detection, adapted with an architecture implantation on hardware platform (methodology AAA).

1. Introduction :

Le travail présenté dans cet article est un sous-ensemble d'un système embarqué de traitement d'image dont le but est d'authentifier un visage. Il est destiné à des applications grand public afin de replacer la saisie du mot de passe sur des périphériques mobiles (téléphones portables, organiseurs,...). Le travail consiste donc à la sélection, au développement et à l'optimisation de la chaîne de traitement en vue d'une implantation matérielle (méthodologie Adéquation Algorithme Architecture).

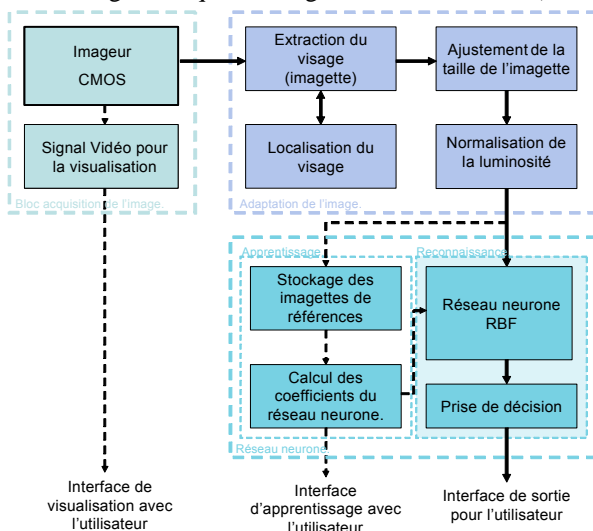


Fig. 1 : Diagramme complet du système.

Dans le cadre de ce système d'authentification de visage, une étape importante consiste à localiser dans l'image d'entrée le visage de la personne à identifier (figure 1). Deux alternatives s'offre à nous. La première, repose sur le parcours de l'intégralité de l'image avec l'ensemble de la chaîne de traitement destiné à authentifier le visage, soit l'ensemble des traitements prévus pour l'adaptation de l'image (correction des défauts et mise en forme de l'image) [1] suivi de l'algorithme d'authentification de visage (réseau neuronal) [2][3]. L'inconvénient de cette méthode est qu'elle nécessite un temps de calcul important. La seconde solution consiste à localiser au préalable le visage de la personne à l'aide d'un algorithme rapide, puis de tester uniquement la zone détectée avec la chaîne d'authentification.

2. Contexte d'application :

Afin de d'obtenir un algorithme adapté à la chaîne de traitement, il est nécessaire au préalable de définir les contraintes du système. Le traitement doit présenter une faible complexité afin de rester implantable sur une structure matérielle pour un faible coût (temps de calcul et nombre de cellules logiques) tout en présentant de bonnes performances afin de limiter le nombre de non détections ou de fausses détections lors de l'utilisation du système. Le traitement doit aussi être suffisamment robuste afin de fonctionner malgré de fortes variations de l'image d'entrée (éclairage et arrière-plans non contrôlés) dues à une acquisition des images non contrôlée. De plus l'imageur

utilisé pour cette acquisition est un capteur CMOS bas coût (images en niveau de gris, faible résolution, bruit, ...). En contre partie, il n'est nécessaire de détecter qu'un unique visage dans l'image, présenté de face avec une faible variation sur l'orientation ($\pm 15^\circ$) et un facteur d'échelle limité ($\pm 10\text{cm}$).

3. Techniques de localisation :

La littérature scientifique propose de nombreux algorithmes pour la détection de visage. Nous pouvons distinguer différentes approches basées :

- sur les caractéristiques invariantes du visage (couleur de la peau, forme elliptique du visage...) [4][5]
- sur la connaissance du visage et le positionnement relatif des différentes composantes du visage (les yeux, la bouche...) [6][7]
- sur l'analyse des composantes principales du visage par des méthodes de classification [8][9]
- sur de corrélations de modèles de visage [10][11][12].

Après avoir effectué une étude rapide des différents algorithmes et méthodes associées, nous avons conclu qu'ils n'étaient pas parfaitement adaptés à notre cas (complexité ou contraintes trop importantes). Nous proposons donc une technique basée sur une comparaison de la scène avec des modèles de visage. Le principe de la méthode proposée (figure 3) repose sur la recherche d'objets ayant les caractéristiques d'un visage, basée sur la détection des contours. Pour cette recherche, nous utilisons le même algorithme que pour l'authentification du visage : un réseau neurone RBF (Radial Basis Function). Avant de détailler un peu plus la méthode proposée, nous allons rapidement présenter le fonctionnement du réseau neurone.

4. Réseau neurone RBF :

Le réseau de neurone RBF (figure 2) s'apparente aux réseaux de type perceptron multicouches [2]. C'est un interpolateur universel basé sur des combinaisons de fonctions radiales. Il se compose de 3 couches, dont chacune d'elle est entièrement connectée à la suivante :

- La couche d'entrée : transfère directement les pixels de l'image à traiter (composantes du vecteur d'entrée).
- La couche RBF : chacun des neurones sont paramétrés à l'aide de deux variables, le vecteur de référence et le rayon (influence de la fonction radiale). La sortie du neurone est fonction de la similarité entre les pixels présents à l'entrée et du vecteur de référence, le tout filtré par la fonction radiale.
- La couche de sortie : la sortie du neurone est une combinaison linéaire des sorties de la couche RBF multipliés par le poids de leur connexion respective. En fonction du résultat fourni, une décision concernant la reconnaissance est prise.

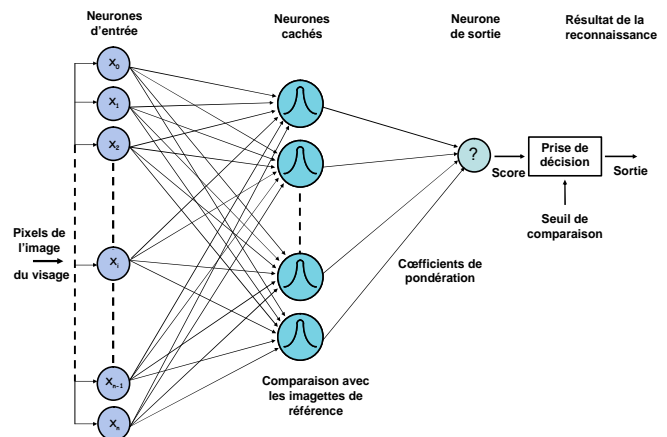


Fig. 2 : Réseau neurone RBF

5. Méthode proposée :

5.1 Principe :

Le système repose sur l'utilisation de réseaux de neurones. Les réseaux parcourent l'image d'entrée préalablement filtrée. Pour chaque position, la similarité est calculée avec les images fournies en référence (masques). L'analyse des résultats permet ensuite de localiser la zone de l'image la plus similaire à un visage.

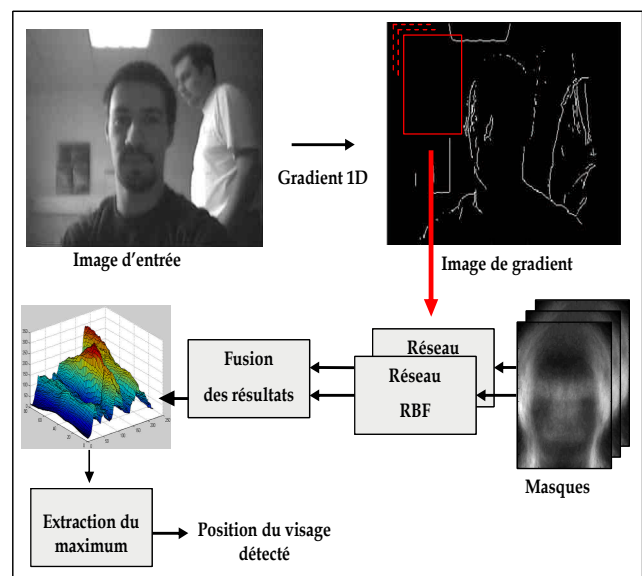


Fig. 3 : Diagramme de l'algorithme de localisation du visage.

5.2 Description de l'algorithme :

L'image d'entrée est sous échantillonnée afin de réduire le nombre de pixels à traiter. Les contours présents dans l'image sont mis en évidence par un calcul de gradient. Pour cela un filtre de type Laplacien à une dimension est utilisé. Il permet une intégration rapide sur une structure matérielle et le calcul s'effectue en temps réel

(cadence du capteur). L'image de contours obtenue est ballayée afin d'extraire une zone de l'image de la taille du visage qui doit être authentifié. Les imagerie extraites en chacune des positions sont ensuite comparées à des images de « contours types » de visages (masques) par deux réseaux RBF câblés en parallèles. L'utilisation de cette paire de réseau permet d'augmenter la fiabilité du système. Chacun des réseaux possède une fonction particulière :

- Le premier réseau de neurones utilise la totalité des pixels des images des masques, ce qui permet une détection de visages sur fond complexe.
- Le second réseau de neurones utilise uniquement les pixels clairs des masques, ce qui offre une détection des visages sur des fonds unis de manière plus efficace.

Les résultats des deux réseaux neurones sont ensuite fusionnés. Pour cela, les sorties des deux réseaux sont multipliées, puis normalisées. La position testée fournissant le résultat le plus important correspond à la position du visage la plus probable dans l'image.

5.3 Méthode d'obtention des masques :

Le fonctionnement des réseaux RBF étant basé sur un calcul de distance entre les vecteurs de référence (les masques dans notre cas) et les images d'entrée, un bon fonctionnement de cet algorithme passe donc par une sélection minutieuse de ces images de masques.

Cette sélection peut s'avérer complexe : prise en compte des variations de la masse corporelle des différentes personnes et des variations d'échelle du visage. Afin d'obtenir cet ensemble de masques, nous avons calculé les contours moyens normalisés en amplitude et en nombre, obtenus par accumulation de nombreuses images de contours de visages (figure 4). Ce travail a été réalisé à partir de la base de données ORL (série de 10 images pour 40 sujets) [13].

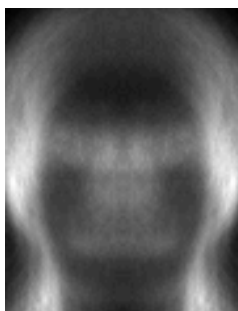


Figure 4 : Exemple de masque calculé.

5.4 Résultats :

Afin d'évaluer les performances du système de localisation, une étude statistique a été menée sur la base de données VALID [13] composée de plus de 500 images. De par les nombreuses variations des conditions d'acquisition (exemple figure 5), cette base d'images

présente des images très proche de celles de l'application ciblée et constitue donc un excellent échantillon de test pour notre algorithme.



Figure 5 : Exemple d'images de la base VALID.

Pour déterminer la qualité des résultats obtenus avec notre algorithme, une comparaison est effectuée entre les positions estimées des visages et les positions fournies par l'algorithme (figure 6). Sur cette figure, sont montrés les écarts de positionnement pour chaque image. En abscisse est indiqué le numéro de l'image testée et en ordonnée, l'écart calculé selon une distance Euclidienne sur la composante xy (en pixels). Hormis une vingtaine d'images (5% de l'ensemble), l'écart obtenu reste inférieur à 20 pixels.

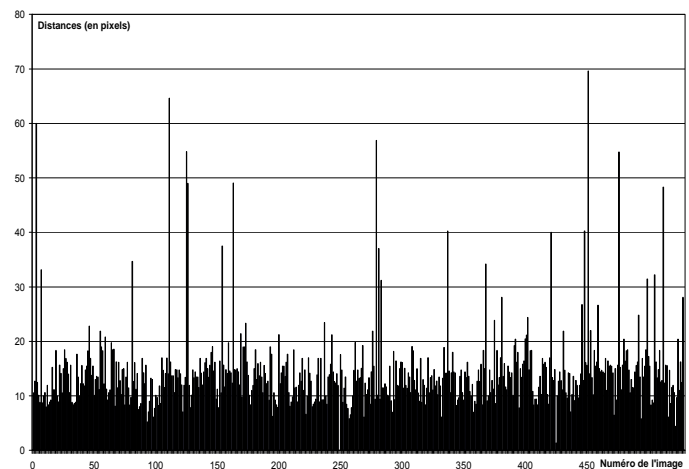


Figure 6 : résultats des localisations de visage pour la base VALID.

Malgré une précision de l'algorithme moyenne la méthode présente des résultats suffisamment fiables (95% de bonne localisation). Cette erreur de positionnement (due à la forme des masques) impose à la chaîne d'authentification un parcours de quelques pixels autour de la zone localisée. Même si le taux de réussite approche les performances obtenues avec des algorithmes beaucoup plus complexes [10][11][12], les images de la base testée sont nettement moins complexes : un seul visage présent, une taille de visage relativement fixe et de rotation du visage faible. Néanmoins, la méthode visée ne devant pas nécessairement être robuste aux très fortes variations d'échelle et d'angle de rotation pour le visage, cet algorithme est bien adapté à notre application.

5.5 Résultats de l'implantation :

Cet algorithme a été implanté sur un StratixII (Altera). Au moyen de l'utilisation d'une version générique du réseau RBF, précédemment développée pour la chaîne d'authentification, ce traitement a été implanté en très peu de temps. Il est relativement simple, ne nécessitant qu'un faible nombre de cellules logiques et présentant une structure hautement parallélisable permettant une accélération importante du traitement lors de l'implantation matérielle [3]. Voici les résultats obtenus pour une image de dimension égale à 180 par 144 pixels :

- surface faible estimée à 700 CLB
- mémoire RAM 250 Ko
- fréquence max. 150 MHz

Le gain de temps avec l'utilisation de cet algorithme est important, nous obtenons un temps de traitement de 200ms (circuit cadencé à 50MHz).

6. Conclusion :

L'algorithme proposé est basé sur un réseau neurone RBF. Cette solution offre l'avantage d'offrir un très bon compromis entre les performances de la localisation, les ressources matérielles utilisées, un temps d'exécution compatible avec le cahier des charges et un temps de développement court. La partie localisation du système consomme des ressources matérielles et offre un temps d'exécution du même ordre de grandeur que la partie authentification.

Les résultats fournis par l'algorithme sont satisfaisant : le taux de réussite pour la localisation est proche des 95%. En contrepartie, du fait de la forme des masques, la localisation manque un peu de précision. Mais il peut tout de même être envisagé, pour un surcroît de coût matériel, de faire fonctionner en parallèle plusieurs intégrations du même traitement pour différentes configurations afin d'obtenir des performances plus importantes : une meilleure robustesse aux facteurs d'échelle par exemple. Afin d'augmenter la précision de la localisation, il pourrait être envisagé l'ajout d'une détection des yeux dans les zones présentant une forte similarité avec les contours d'un visage.

7. Remerciements :

Nous remercions tout spécialement la région Rhône-Alpes pour son aide et le financement de nos recherches.

8. Référence :

[1] L. Pierrefeu & Al., Normalisation de la luminosité d'image appliquée à un système d'authentification de visage, READ, 2005.

[2] L. Pierrefeu & Al., Auto-adjustable method for Gaussian width optimization on RBF neural network. Application to face authentication on a mono-chip system. IEEE IECON 2006.

[3] F. Yang and M. Paindavoine, Implementation of an RBF neural network on embedded systems: Real-time face tracking and identity verification. IEEE transaction on neural network, 2003.

[4] D. Saxe and R. Foulds, Toward robust skin identification in video images, Proceeding of international conference on automatic face and gesture recognition 1996.

[5] K. Sandeep et A.N. Rajagopalan, Human Face Detection in Cluttered Color Images Using Skin Color and Edge Information.

[6] F. Samaria & S. Young, HMM based architecture for face identification, Image and computer vision 1994.

[7] A. Nefian & M. Hayes, Hidden markov models for face recognition, International conference on acoustics, speech, and signal processing, ICASSPS. 1998.

[8] M. Turk & A. Pentland, Face recognition using eigenfaces, Proceedings of international conference on pattern recognition 1991.

[9] P. Belhumeur, & Al, Eigenfaces vs. Fisherfaces: Recognition using specific linear projection, Proceeding of fourth European conference on computer vision ECCV 1996

[10] H. Rowley, & T. Kanade, Neural Network-Based Face Detection, IEEE Transaction on pattern analysis and machine intelligence 1998.

[11] P. Viola & M. Jones, Robust Real-time Object Detection, International workshop on statistical and computational theories of vision 2001.

[12] C. Garcia & M. Delakis, Convolutional Face Finder: A Neural Architecture for Fast and Robust Face Detection. IEEE Transaction on pattern analysis and machine intelligence 2004

[13] Face Recognition Homepage: <http://www.face-rec.org/>