

Implémentations matérielles d'un système biométrique bimodal

FAN YANG¹, VINCENT BROST¹, AUDREY POINSOT¹, DOMINIQUE GINHAC¹, ZHICHUN MU²

LE2I-CNRS 5158 Laboratory, University of Burgundy, 21078 Dijon France¹

Automation department, University of Science and Technology of Beijing, 100083 Beijing China²

fanyang@u-bourgogne.fr, vincent.brost@u-bourgogne.fr, audrey.poinsot@u-bourgogne.fr, mu@ies.ustb.edu.cn

Résumé - La biométrie est une application majeure de la reconnaissance de formes. Dans cet article, nous présentons les expériences d'implémentation matérielle d'une chaîne de reconnaissance biométrique basée sur la paume et le visage. Les traitements utilisés sont adaptatifs, peu coûteux et ils respectent des contraintes de faible connaissance a priori. D'ailleurs, le système proposé est totalement sans contact, apportant confort et hygiène à l'utilisateur. L'extraction de paramètres par filtrage de Gabor est étudiée dans le cas de la paume et du visage. La fusion de deux modalités testée et validée sur les deux bases respectivement composées de 130 et de 119 personnes nous confirme de très bonnes performances de robustesse. Les implémentations matérielles ont été réalisées sur les deux types de composants électroniques DSP et FPGA. Nous pouvons reconnaître une personne parmi 100 en moins de 100 ms, ceci permet d'atteindre l'impression d'instantanéité que nous cherchons à obtenir pour des applications grand public.

Abstract - Biometric science is an important application of pattern recognition. This paper proposes an adaptive, contactless, biometric system which combines two modalities: palmprint and face. The processing chain has been designed to overcome embedded system constraints and small sample set problem. After a palmprint is extracted from a hand image, Gabor filters are applied to both the palmprint and face in order to extract parameters, which are then used for classification. Fusion possibilities are also discussed and tested using two multimodal biometric databases of 130 and 119 persons. High recognition performance has been obtained. Hardware implementations have been realized using two electronic components: DSP and FPGA. We can perform the recognition processing (1 person among 100) in less of 100 ms and, this is suitable for real-time, general-public biometric applications.

1 Introduction

La biométrie est une branche de la reconnaissance de formes qui suscite toujours plus d'intérêt. Les systèmes biométriques permettent, en effet, d'identifier automatiquement les personnes à l'aide de caractéristiques physiques ou comportementales telles que le visage, l'empreinte, la signature ou encore la démarche. Au vu des besoins grandissants en matière de sécurité, leur fiabilité est un enjeu essentiel.

Chaque caractéristique biométrique possède ses avantages et ses inconvénients. L'utilisation d'une seule modalité implique parfois certains désagréments auxquels il faut faire face comme la mauvaise qualité ponctuelle des données, la non universalité de la caractéristique choisie, la faiblesse des résultats ou encore la possibilité de l'usurpation d'identité. C'est pourquoi de plus en plus de recherches se focalisent sur la fusion de plusieurs sources d'information biométrique : la biométrie multimodale permet de pallier certains inconvénients de la mono modalité, mais surtout d'augmenter les performances [1].

Le but de notre travail est de proposer un système biométrique multimodal destiné à des applications grand public, qui respecte des contraintes de confort et d'embarquabilité (calculs peu coûteux, robustesse, temps réel, utilisation réduite de mémoire, faible encombrement, basse consommation, ...). Dans ce contexte, le problème de la faible connaissance a priori

devient incontournable (« small sample set problem ») : afin de réduire l'utilisation de mémoire et la durée d'acquisition, seuls quelques exemples peuvent être capturés lors de la phase d'enregistrement de l'utilisateur.

Dans cet article, nous proposons un système adaptatif, utilisant la paume et le visage. Les biométries basées sur la main demandent traditionnellement un contact avec le capteur, mais de plus en plus d'études s'attachent à la rendre plus confortable et plus hygiénique en supprimant cette obligation de contact. Dans les sections suivantes, nous décrirons d'abord la chaîne complète de traitement, puis nous présenterons les résultats expérimentaux d'implémentation matérielle de celle-ci. Nous allons finalement conclure et discuter de futur travail dans la section 4.

2 Système biométrique complet

La Figure 1 montre la structure de notre système bimodal [2] [3]. Nous travaillons sur des images de mains gauches acquises sur fond vert, sans aucun contact, et dans des conditions d'éclairage naturelles. Les seules contraintes imposées aux sujets ont été de placer leur main horizontalement, et de façon à ce que leurs doigts ne se touchent pas entre eux. Ce libre placement de la main induit des translations, des rotations et des changements d'échelle entre les images.

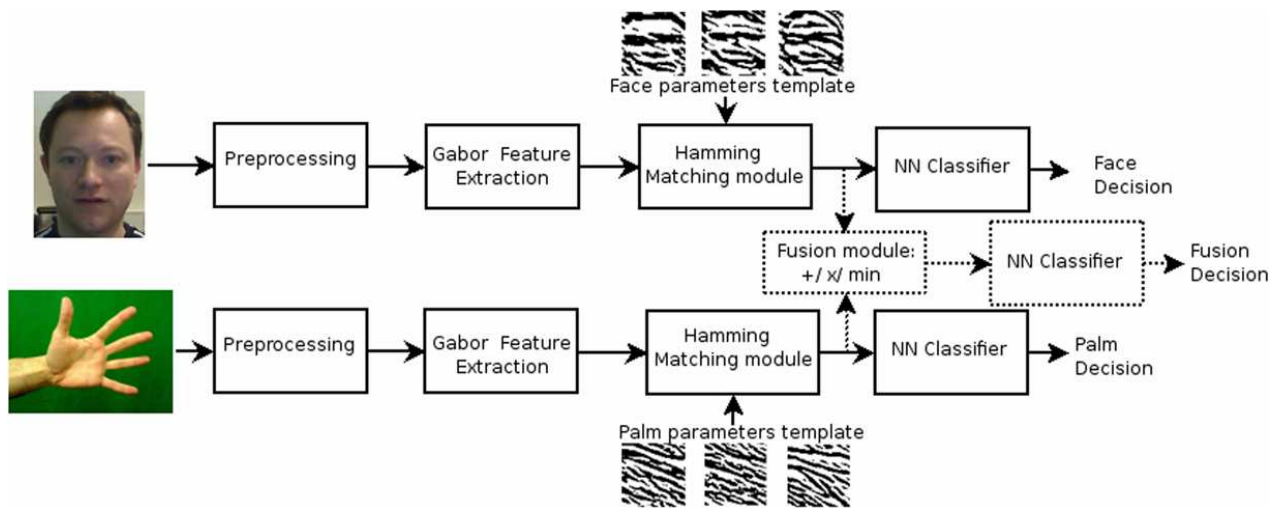


Figure 1 : Chaîne de traitement complète avec possible fusion des scores.

L'extraction de la paume est donc effectuée en plusieurs étapes. La main est d'abord localisée par une segmentation basée sur la couleur. La position des bouts des doigts est ensuite déterminée grâce au contour de la main, lui-même calculé par un algorithme de suivi de contour. Un carré de taille proportionnelle à la largeur de la main est enfin défini à partir de ces points de référence. L'image extraite est alors normalisée : elle subit une rotation et un sous échantillonnage pour atteindre la résolution fixe de 64×64 pixels.

La paume présente une texture riche en lignes et en stries qui permet de différencier les personnes. Au vu de ses performances et dans l'optique de réduire le temps de calcul, nous avons choisi un filtre de Gabor ellipsoïdal réel pour extraire les caractéristiques pertinentes de paume. Ce filtrage a été réalisé à l'aide d'une convolution de l'image de paume avec un masque de 9×9 dont les coefficients ont été choisis empiriquement (voir [3]). Ce dernier est suivi d'un seuillage du résultat visant à limiter la taille de la matrice de caractéristiques ainsi que le temps de calcul lors de la phase de comparaison.

La mesure de similarités utilisée consiste en une distance de Hamming normalisée, qui est une comparaison pixel à pixel utilisant l'opérateur booléen *ou exclusif*. Comme l'extraction de l'image de paume n'est pas forcément idéale, nous avons introduit une tolérance en translation : nous calculons la distance pour plusieurs décalages (5 pixels de décalage maximal horizontalement et verticalement), puis nous choisissons la distance minimale.

Nous avons appliqué la même méthode de traitement sur la modalité de visage, sauf que celle-ci a été capturée avec un cadre fixe pour extraire directement la zone d'intérêt. En ce qui concerne la fusion de deux modalités, nous avons choisi la fusion des scores avec plusieurs types d'opération : l'addition, la multiplication des deux scores et la technique ABC (*Averaged Bayes Classifier*). Dans tous les cas, la décision finale est prise grâce à l'algorithme du plus proche voisin (NN). Notre système a été testé et validé sur deux bases de données différentes : la base uB composée de 130 sujets

et le couple de bases publiques AR (visage)/PolyU (paume avec contact) contenant 119 sujets [3].

Tab 1 : Performances de reconnaissance : ARR (Average Recognition Rate %)

Méthode	Base uB	Base AR/PolyU
Reconnaissance de visages	90.93	81.85
Reconnaissance de paumes	96.82	94.06
Fusion par la somme de scores	98.96	97.68

Le Tableau 1 montre un extrait des résultats obtenus avec ces deux bases de données. Il est à noter que seulement 2 exemples par modalité et par personne ont été utilisés pour ces tests. Ce tableau confirme que la paume est une modalité très performante, même utilisée seule, puisque avec seulement 2 exemples par personne on obtient une moyenne de 95.44% de bonne reconnaissance. Le visage n'atteint pas d'aussi bons résultats, mais ces derniers sont déjà élevés au vu du très faible coût calculatoire des algorithmes de traitement. Quand à la fusion, elle permet de toujours surpasser les performances de la paume.

3 Description des implémentations matérielles

Basé sur nos résultats d'évaluation algorithmique (voir [3] et le Tableau 1), nous avons choisi la configuration suivante pour les implémentations matérielles : le système bimodal est composé de 25 sujets avec 2 exemples d'apprentissage par individu. Au total, 100 images de base de comparaisons ont été sélectionnées et stockées.

3.1 Implémentation de la chaîne complète sur le DSP (Digital Signal Processor)

La chaîne complète de reconnaissance biométrique a été réalisée sur le DSP TMS320C6415 de Texas Instrument en utilisant l'environnement Code Composer Studio [4]. Ce processeur destiné au traitement du signal comporte deux chemins de données, une mémoire interne à 2 niveaux et divers périphériques d'interface.

Il possède plusieurs niveaux de pipelines (instruction fetch, dispatch and decode) pour traiter les instructions de type VLIW (Very Long Instruction Word) permettant d'effectuer des calculs sur 8 unités fonctionnelles simultanément. De plus, certaines unités autorisent l'exécution de multiples opérations de même type par cycle pour former une machine SIMD (Single Instruction Multiple Data).

Nous avons d'abord implémenté l'étape d'extraction de la paume avec différentes techniques d'optimisation. Celle-ci se décompose en plusieurs sous étapes : segmentation de la main, suppression de débris, détection de contours et normalisation de la paume (rotation, remise à l'échelle et conversion en niveau de gris). Ensuite, la paume extraite comme le visage a été soumise à l'étape de reconnaissance qui compare les traits biométriques à tester aux exemples d'apprentissage. Ceux-ci se composent de 100 images de la taille 56×56 pixels (voir Figure 2). Il est à noter que pour chaque paume/visage de test, $5 \times 5 = 25$ comparaisons ont été effectuées afin d'introduire certaines tolérances en translations verticales et horizontales.

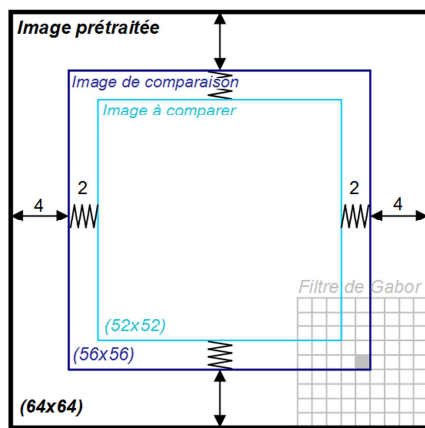


Figure 2 : Taille d'images durant le traitement : image prétraitée d'origine = 64×64 pixels, image d'exemple = 56×56 pixels à cause de l'effet de bord avec la convolution 9×9 , et image de test = 52×52 pixels avec 5 translations horizontales et verticales.

D'après nos résultats expérimentaux, la vitesse de ce prétraitement de paume varie entre 68 M et 75 M cycles avec une fréquence de 1 GHz pour le DSP. Ceci correspond à une durée de traitement comprise entre 68 ms et 75 ms. Cette étape peut très bien être réalisée pendant l'acquisition de visage. Le prétraitement de visage (remise à l'échelle et conversion en niveau de gris), tout comme la reconnaissance (comprenant la fusion par addition) ne demande que très peu de cycles. Au total, en incluant les prétraitements, l'identification d'un sujet sur un unique DSP a été réalisée en moins de 80 ms dans le cas où la base contient les exemples de 25 personnes. Les résultats de simulation nous montrent aussi une vitesse inférieure à 100 ms avec les exemples de 100 personnes, ceci permet d'atteindre l'impression d'instantanéité, que nous cherchons à obtenir pour des applications grand public.

3.2 Implémentation de la reconnaissance sur le FPGA (Field Programmable Gate Array)

Nous avons choisi le FPGA Virtex-5-XC5VFX70T de la société Xilinx [5] pour réaliser la seconde implémentation matérielle. A part des éléments logiques configurables de base (Slices), il contient, en particulier, 128 slices DSP48 permettant d'exécuter des algorithmes de traitement du signal (images) de manière massivement parallèle. Nous avons gardé la configuration précédente (25 personnes avec 2 exemples par individu) pour cette implémentation sur le FPGA. Pour l'instant, seul le traitement de reconnaissance (filtrages Gabor, comparaisons et fusion) a été testé car contrairement à l'extraction de la paume, la reconnaissance est un traitement générique qui pourrait être utilisé directement pour d'autres applications de reconnaissance de formes. D'ailleurs, le prétraitement de paumes contient moins de parallélisme intrinsèque (beaucoup de tests séquentiels, d'opérations à virgule fixe, etc), sa réalisation sur FPGA demande donc plus de temps de développement. Nous préférons le tester ultérieurement avec un outil de prototypage rapide mis en œuvre dans le cadre d'une autre thèse [6].

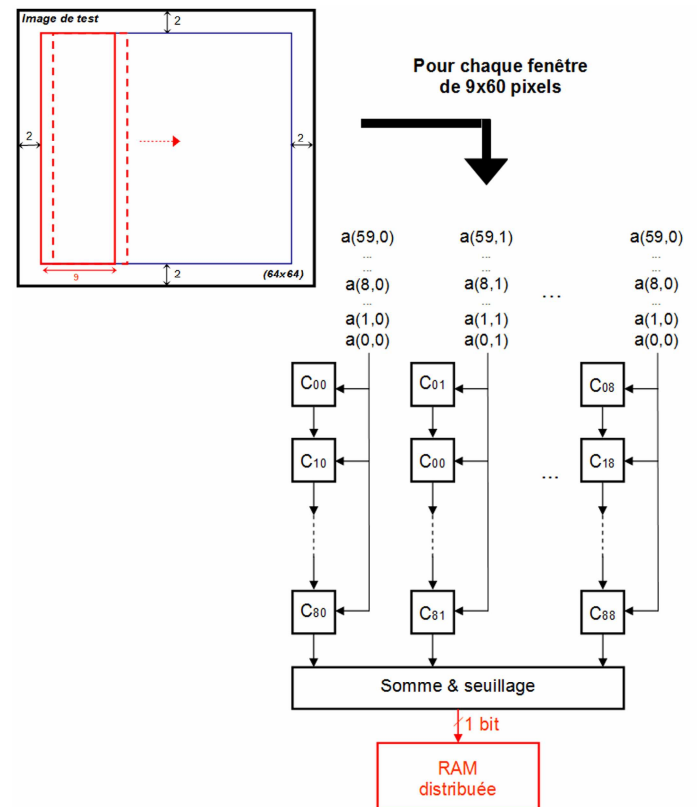


Figure 3 : Structure parallèle de la phase de filtrage de Gabor : l'image de paume (visage) est distribuée dans les fenêtres successives de 60×9 pixels, une architecture de 9 lignes x 9 colonnes de DSP48 Slices réalise 81 multiplications et accumulations simultanément.

Les Figures 3 et 4 illustrent la structure parallèle (décrite par le langage de bas niveau VHDL) pour les deux phases distinctes de la reconnaissance : réalisation de 81 multiplications et accumulations (convolution avec le masque 9×9) en un seul cycle et parallélisation

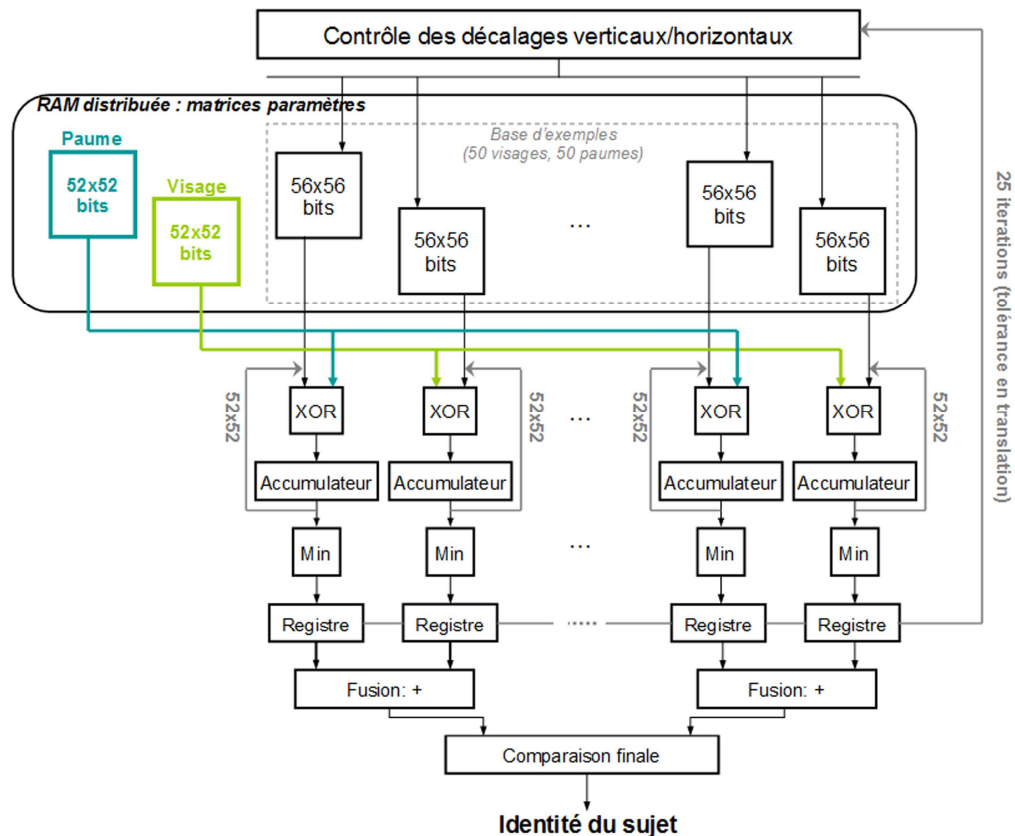


Figure 4 : Structure parallèle de la phase de comparaison/fusion : 100 distances de Hamming sont calculées en même temps (en 52 x 52 cycles) et ceci est répété 25 fois (correspondant aux 25 décalages).

massive de 100 comparaisons basées sur la distance de Hamming. Un cœur de processeur matériel programmé en Assembleur (PicoBlaze) contrôle et synchronise ces deux phases de traitement. Au niveau de ressources matérielles disponibles sur le FPGA, nous avons respectivement utilisé 89/128 (69.5%) des slices DSP48, 2/148 (1.4%) des Blocs RAM, et 8566/44800 (19.1%) des blocs logiques. Les 100 images d'exemple de 56x56 pixels ont été stockées dans les cellules de mémoire distribuée de FPGA. Nous avons obtenu la vitesse de 0.43 ms pour identifier une personne avec deux traits biométriques sur le FPGA (fréquence 250 MHz).

4 Conclusions et perspectives

Nous avons présenté un système biométrique sans contact qui combine deux modalités : paume et visage. Une chaîne complète de traitement depuis l'acquisition jusqu'à la décision a été développée et validée avec des contraintes d'applications grand public : petit coût, convivialité, robustesse, petit jeu d'exemples d'apprentissage disponible, hygiène. Nous avons aussi réalisé ce système bimodal sur les composants électroniques de type DSP et FPGA pour respecter les contraintes d'embarquabilité, de basse consommation, d'encombrement et de traitement en temps réel. Les résultats expérimentaux nous montrent que la chaîne complète peut être réalisée sur le DSP TMS320 C6415 en moins de 100 ms pour identifier une personne parmi 100 avec de très bonnes performances.

L'algorithme générique de reconnaissance n'utilise qu'un petit pourcentage de ressources matérielles

disponibles sur le FPGA. L'implémentation matérielle d'extraction de la paume sur ce composant est en cours de réalisation à l'aide d'un outil de prototypage rapide développé par notre équipe de recherche. L'objectif principal consiste à réduire le temps de développement des applications sur FPGA et à encourager les ingénieurs logiciels à accélérer les traitements grâce à des composants électroniques. En perspective, nous pensons aussi appliquer cette expérience conjointe logicielle/matérielle aux autres applications de reconnaissance de formes comme la catégorisation des expressions pour connaître la préférence olfactive des nouveau-nés à travers l'analyse des scènes de vidéo.

Références

- [1] A. K. Jain, A. Ross, and S. Pankanti, "Biometrics: a tool for information security", *IEEE Trans. Inf. Forensics Secur.* **1**(2), 125-143, 2006.
- [2] A. Poinot, F. Yang and M. Paindavoine, "Contactless Palmprint Recognition Using Gabor Filter", *Proceeding of the 24th International Conference on Computers and Their Applications (CATA)*, New Orleans, USA, pp. 25-30, April, 2009.
- [3] A. Poinot, F. Yang and M. Paindavoine, "Fusion de biométrie sans contact : paume et visage", *22^{ème} colloque Gretsi*, Dijon, France, Septembre, 2009.
- [4] Texas Instruments, <http://www.ti.com/>
- [5] ISE, 2009. ISE design, URL <http://www.xilinx.com/tools/designtools.htm>.
- [6] V. Brost, F. Yang, M. Pandaivone, and N. Farrugia, "Multiple modular VLIW processors based on FPGA," *Journal of Electronic Imaging*, *SPIE*, **16**(2), 023001:1-10, April-June 2007.