

# Capture infra-rouge et fusion multimodale de scores dans le projet VINSI : Vérification d'Identité Numérique Sécurisée Itinérante

Bernadette Dorizzi\*, Sonia Garcia-Salicetti\*, Walid Hizem\*, Lorene Allano\*, Yang Ni\*\*

\*GET/INT, département EPH, 9 rue Charles Fourier, 91011 Evry  
{Prenom.Nom@int-edu.eu}

\*\*Société NIT, {yang.ni@new-imaging-technologies.com}

## I. Introduction

### I.1 Objectifs du projet

Des besoins nouveaux en appareils de contrôle d'identité portables, communicants et sécurisés voient le jour aujourd'hui. En effet la suppression des contrôles aux frontières intérieures conduit à multiplier les vérifications itinérantes ; les contrôles aériens doivent même parfois se faire en passerelle, nécessitant là aussi des dispositifs nomades. Dans le contexte de la mise en place des visas et passeports biométriques, le document d'identité contient des éléments biométriques de la personne: ses empreintes digitales et son visage. C'est donc à partir de ces éléments que l'identité pourra être vérifiée, en plus des informations textuelles classiques (nom, prénom etc.).

Un autre contexte d'utilisation d'un tel terminal nomade est celui du paiement sécurisé en ligne ou par carte à puce où des outils d'identification forte allant au delà du mot de passe ou du code pin seraient appréciés. La aussi des éléments biométriques pour caractériser la personne sont envisagés dans un futur proche.

Le projet VINSI consiste à mettre au point un système de vérification d'identité biométrique multimodale en référence à un document électronique stocké dans une base de données distante, garantissant des conditions de sécurité adéquates. Le système a pour vocation d'être porté sur un terminal nomade pour des vérifications en condition de mobilité.

Ce projet de trois ans a regroupé des partenaires industriels (Thalès, Trusted logic pour les études de sécurité) et académiques (GET/INT) et a conduit à la réalisation d'un démonstrateur porté sur le terminal nomade ARTEMA de Thalès (voir Figure 1).

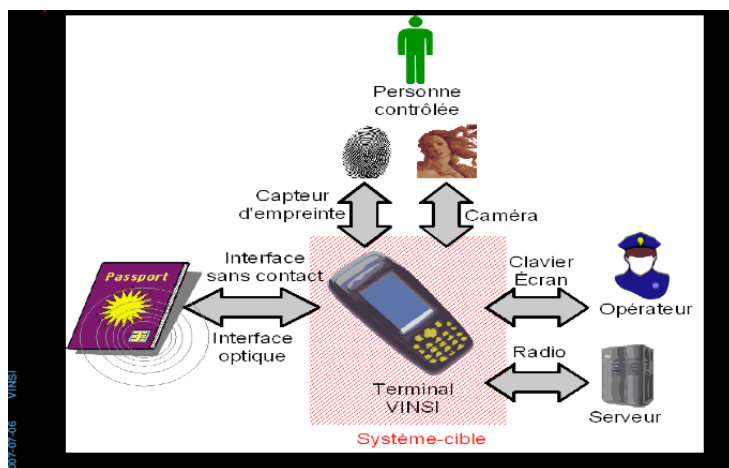


Fig.1: Schéma de fonctionnement du système VINSI

Dans cet article, nous nous contenterons de détailler deux innovations résultant de ce projet, réalisées

par le GET/INT, à savoir la proposition d'un nouveau capteur d'images de type proche infra-rouge qui permet d'acquérir des images de visage de qualité stable, indépendamment des variations d'illumination qui peuvent par exemple se produire lorsqu'on fait des acquisitions en extérieur. Dans un deuxième temps, nous détaillerons notre stratégie de fusion multimodale originale qui permet de combiner les scores d'empreintes digitales et de visage en assurant un excellent compromis entre performance et coût.

## **I. 2 Motivations**

La Biométrie apparaît comme un moyen de pallier les déficiences des systèmes de sécurité traditionnels, en particulier en permettant d'insérer dans un document d'identité une ou plusieurs caractéristiques personnelles d'un individu (images d'empreintes digitales, du visage etc.), faisant ainsi le lien entre ce que l'individu prétend être et ce qu'il possède. De ce fait, l'usurpation d'identité est rendue plus difficile.

Différentes modalités biométriques peuvent être considérées. Ainsi des systèmes à base d'empreintes digitales ont été élaborés depuis déjà longtemps. Proposés tout d'abord dans le cadre policier pour la lutte contre la criminalité, ces systèmes qui utilisent principalement une image de l'empreinte obtenue à partir d'un encre, ont évolué vers une acquisition de l'image de l'empreinte réalisée à partir de capteurs, généralement optiques, permettant de récupérer une image numérisée de l'empreinte, qui peut alors être facilement manipulée de manière informatique. Différents types de capteurs et d'algorithmes ont été proposés et des compétitions internationales permettent de comparer les différents systèmes entre eux. Si les empreintes digitales sont considérées aujourd'hui comme une modalité suffisamment fiable pour envisager des déploiements sur de larges populations, il n'en reste pas moins des limites intrinsèques dont il est difficile de s'affranchir. Tout d'abord, une partie non négligeable de la population est résistante à la phase d'acquisition, ce qui signifie que l'image des empreintes de ces personnes n'est pas d'une qualité suffisante pour permettre une vérification automatique ultérieure. De plus, même si la phase d'acquisition a été réalisée de manière satisfaisante, des problèmes spécifiques peuvent rendre difficile l'identification par empreintes à un moment donné (présence de blessures, tâches, etc.). Enfin, bien que difficile, la falsification reste possible et des travaux récents ont montré que la plupart des capteurs se laissent piéger par l'usage de « faux doigts » (doigts en gélatine sur lesquels on a reproduit les empreintes récupérées à partir des traces laissées par une personne).

Ces réflexions conduisent à envisager l'utilisation conjointe d'une autre biométrie. Ainsi, les falsifications seront rendues plus difficiles (il faut copier deux caractéristiques propres de la personne plutôt qu'une) et la redondance d'information sur la personne permettra de pallier le déficit d'une des modalités.

Une idée naturelle est alors d'utiliser l'image numérisée du visage car nous sommes habitués depuis fort longtemps à utiliser des photographies de notre visage dans nos titres d'identité. C'est pourquoi la direction actuelle, au niveau titre d'identité, est d'enregistrer empreintes digitales et visage et de stocker ces informations dans le titre d'identité lui-même.

Malgré une recherche très active dans le domaine de la vérification par le visage, à ce jour, seuls les systèmes fonctionnant en conditions très restreintes (images de face, fonds uniformes, lumière contrôlée), identiques en phase d'acquisition et de test, conduisent à des performances acceptables. Les faux rejets restent cependant encore importants, à cause aussi des variations personnelles (lunettes, barbes, coupes de cheveux, couleur de cheveux, etc.). Il est difficile d'extraire de l'image du visage des invariants stables. C'est pourquoi l'identification et même la vérification d'identité par le visage, même si elle paraît « naturelle » et sera bien acceptée par la population, reste encore non appropriée de

manière automatique, pour de grandes populations. La situation s'empire encore si on envisage des variations d'illumination comme celle qui vont résulter d'une utilisation en extérieur d'un terminal nomade. L'image peut être tellement saturée qu'il ne sera pas possible de l'utiliser du tout. La direction de recherche que nous menons actuellement à l'INT consiste à s'affranchir des problèmes d'illumination à l'aide de capteurs spécifiques et en particulier de capteurs proche-infra rouge.

Outre les avantages déjà cités, les recherches actuelles sur la fusion de modalités par l'utilisation jointe des images de visage et d'empreintes, permettra de fiabiliser les performances des systèmes isolés. En effet, un certain nombre de résultats préliminaires permettent de montrer que la fusion multimodale permet d'améliorer les performances des systèmes individuels. Ceci dit, un certain nombre de questions restent encore en suspens. Quelle stratégie de fusion faut-il employer, séquentielle, alternative, conjointe? Et dans quel cas? Quels algorithmes de fusion sont susceptibles de fonctionner au mieux? Est-il nécessaire d'apprendre les différentes pondérations entre modalités et de les adapter aux différents environnements? En effet, si la vérification doit se faire en conditions de mobilité, les conditions de test peuvent varier considérablement par rapport aux conditions d'acquisition, et de ce fait, les stratégies de fusion devront être adaptées à ce cadre. Enfin, bien évidemment, la mise en oeuvre de biométrie multimodale pose des questions de coût, d'ergonomie, de miniaturisation, qu'il est nécessaire d'étudier dans le cadre de l'application spécifique envisagée.

Dans la suite de cet article, nous traitons d'abord (chapitre II) les aspects captures infra-rouge et traitement d'images associés et ensuite dans le chapitre III les problématiques de fusion de scores.

## **II. Le capteur VINSI**

### **II.1 Description du capteur VINSI**

Le nouveau capteur VINSI nous permet d'acquérir à la fois une image en lumière visible et une image en infrarouge. En effet, nous avons exploité le principe d'acquisition d'un CCD interligne, pour capter en un premier temps une image visible et cela durant la trame paire et ensuite une image en infrarouge durant la trame impaire. Le temps entre les deux trames est de 20ms ; ce temps est assez faible et on peut considérer que l'acquisition se fait en simultané. La capture en infrarouge est réalisée grâce à la réduction du temps d'exposition et par l'activation d'un flash infrarouge<sup>1</sup>. Le flash est synchronisé avec le temps d'acquisition ; ainsi on économise la consommation et on peut pulser l'éclairage des LEDS. Le principe d'acquisition est présenté en Figure 2. La caméra développée au cours de ce projet est conçue à partir d'un module CCD interligne d'une résolution de 580x510. Après la capture de la vidéo, on sépare les deux trames Visible/Infrarouge.

Nous allons maintenant décrire la méthode originale de détection des points caractéristiques que nous avons mise en oeuvre pour les images proches infra-rouge issues de ce capteur. Nous avons évalué la précision de la détection sur une base de données de petite taille enregistrée avec le capteur VINSI.

---

<sup>1</sup> W. Hizem, E. Krichen, Y. Ni, B. Dorizzi, and S. Garcia-Salicetti, "Specific Sensors for Face Recognition", Int'l Conf. Audio- and Video-Based Biometric Person Authentication, 2006.

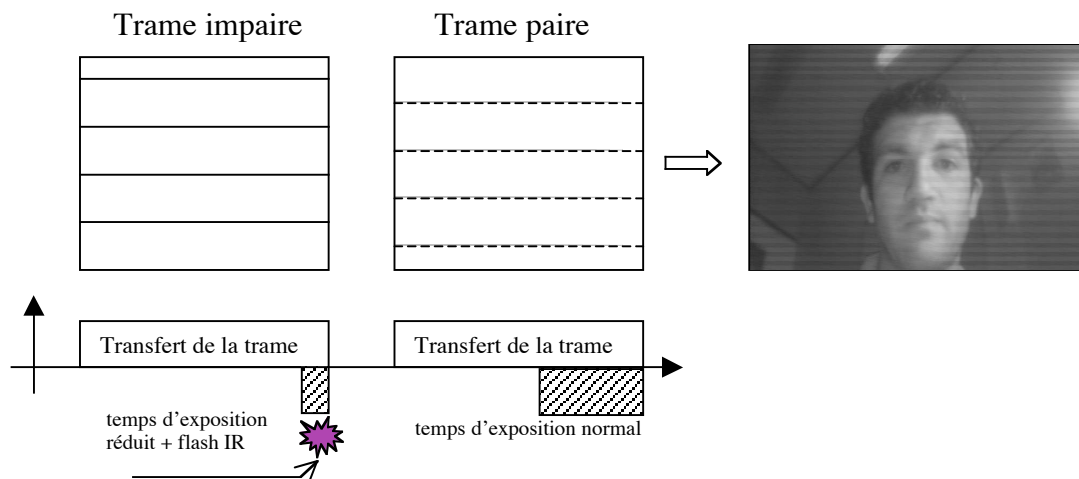


Fig. 2 : Principe d'acquisition de la caméra VINSI

## II.2 Détection des points caractéristiques :

La détection des points caractéristiques est très sensible à la variation de l'illumination, ce qui fait qu'elle n'est pas très fiable sur des images 2D en lumière visible<sup>2</sup>.

La stabilité des images en infrarouge vis à vis des variations de la lumière ambiante, nous permet de mettre au point des techniques robustes de détection de points caractéristiques qui serviront ensuite de point de départ pour les méthodes de reconnaissance associées. Nous nous sommes imposés une contrainte d'intégration sur chip qui ne nous permet pas d'utiliser des méthodes statistiques pour la détection bien que celles-ci correspondent à l'état de l'art<sup>3</sup>.

La stabilité de l'image nous permet d'avoir des contours stables. Nous allons exploiter cet avantage pour le développement de notre algorithme. L'application d'un filtre de détection de contour de 8 directions (voir Figure 3) permet d'accentuer la position des différents points caractéristiques du visage (les yeux, le nez, la bouche...). Une projection horizontale permet de détecter les ordonnées des différents points. Nous nous intéressons dans ce qui suit à la position des yeux, nez et bouche (les sourcils n'étant pas très stables d'une personne à une autre). La Figure 4 montre que les maxima des projections correspondent aux positions des points recherchés. Nous sélectionnons par la suite les zones de l'image qui correspondent à des maxima locaux dans la courbe de projection.

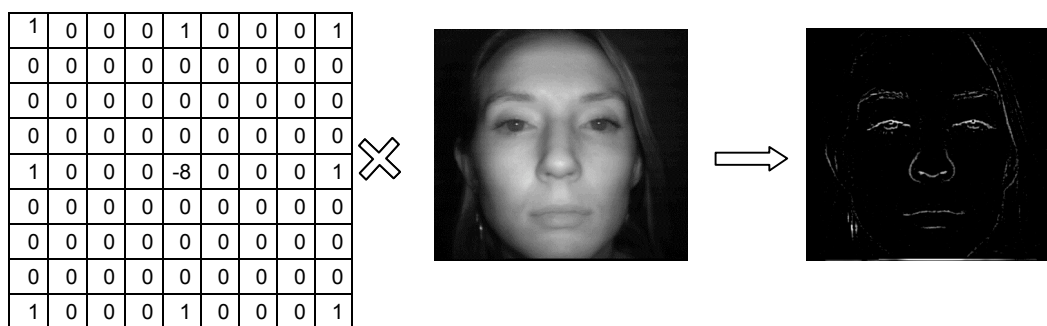


Fig. 3 : Détection des contours

<sup>2</sup> A.S. Georgiades and AL. "From Few to many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose". IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 23, No. 6, June 2001.

<sup>3</sup> P.Viola and M. Jones "Rapid Object Detection using a Boosted Cascade of Simple Features", ICVPR 2001,

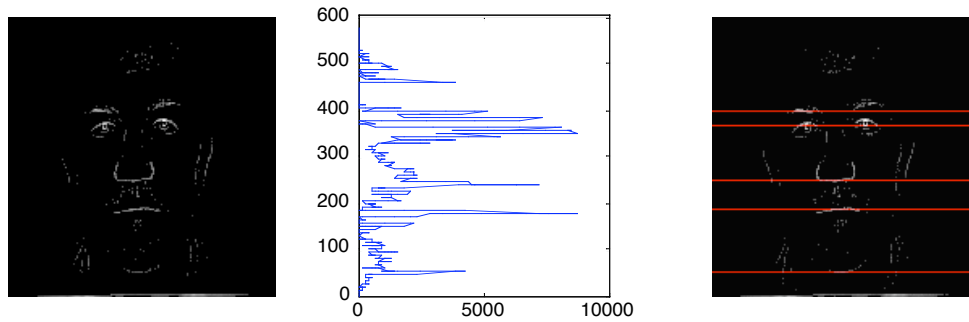


Fig. 4 : Image de contour et maxima locaux de la projection horizontale

Pour rechercher ces points nous avons développé une approche ascendante, c'est à dire que nous avons séquentiellement déterminés la bouche (élément très stable au niveau de la forme) puis le nez et enfin les yeux. Pour cela nous utilisons des techniques de morphologie mathématique associées à des projections horizontales et verticales. Pour plus de détails sur les algorithmes, le lecteur pourra se référer aux travaux récents de Hizem et al.<sup>4</sup>

## II.3 Evaluation de l'algorithme de détection.

### II.3.1 Description de la base INT-VINSI.

Cette base de donnée contient 25 personnes acquises en deux sessions sur plusieurs semaines. Dans chaque session, on dispose de trois vidéos correspondant à trois différents protocoles d'acquisition. Le premier protocole est acquis dans un environnement de bureau avec éclairage normal. Durant le second protocole, on ajoute un éclairage d'appoint latéral du côté gauche. Enfin, le troisième protocole est une acquisition qui a lieu près d'une fenêtre ; la vidéo acquise est sujette à des variations d'éclairage en fonction du temps. La Figure 5 présente un exemple pour une personne des images visibles et infra rouge obtenue lors d'une des sessions d'acquisition.

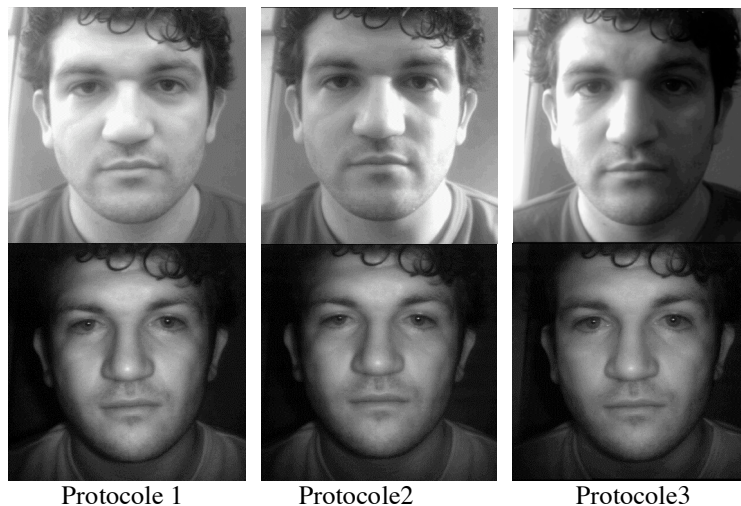


Fig. 5 : Exemple d'images de la base INT-VINSI en lumière visible (en haut) et en infra rouge (en bas) avec les trois protocoles d'éclairage précédemment décrits

<sup>4</sup> W. Hizem, Y. Ni and B. Dorizzi , « Near Infrared Sensing and Associated Landmark Detection for Face Recognition », SPIE-Journal of Electronic Imaging (vol. 17, No. 1) Jan-Mar 2008

### II.3.2 Performances

Pour mesurer la performance de notre approche, nous nous basons sur le critère donné par Jesorsky<sup>5</sup>. Celui-ci prend en considération la distance intra-oculaire :

$$d_{yeux} = \frac{\max(d_d, d_g)}{\|C_d - C_g\|},$$

Où  $C_d$  et  $C_g$  sont les centres des yeux (respectivement droit et gauche) marqués manuellement et  $d_d$ ,  $d_g$  sont les distances entre les centres des yeux (respectivement droit et gauche) marqués par l'algorithme et manuellement. Il faut noter que la position du centre de l'œil, marquée par l'opérateur peut être sujette à des erreurs.

Pour mesurer la performance de la détection du centre de la bouche et du nez, nous avons utilisé le même principe de mesure, comme suit :

$d_{nez} = \frac{d_n}{\|C_d - C_g\|}$ , où  $d_n$  est la distance entre le centre du nez marqué manuellement et par l'algorithme.

$d_{bouche} = \frac{d_b}{\|C_d - C_g\|}$ , où  $d_b$  est la distance entre le centre de la bouche marqué manuellement et par l'algorithme.

Jesorsky et al montrent que pour avoir une bonne détection des yeux il faut que  $d_{yeux}$  soit inférieure à 0.25 ; au delà le centre trouvé n'est plus dans l'oeil. Ce seuil peut varier selon l'algorithme de reconnaissance ou de vérification utilisé ultérieurement. Sadeghi et al.<sup>6</sup> montrent expérimentalement que  $d_{yeux}$  doit être inférieure à 0.05 pour une bonne qualité de vérification en utilisant l'algorithme SVM (Support Vector Machine).

Les performances sur la base INT-VINSI sont présentées sur la Figure 6:

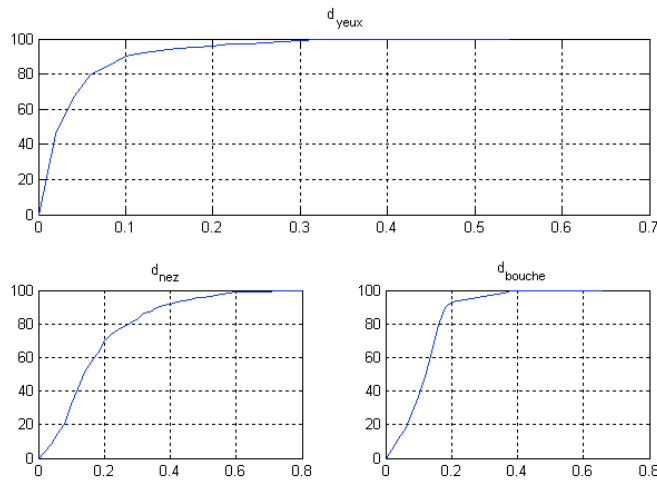


Fig. 6: Histogramme cumulatif de  $d_{yeux}$ ,  $d_{nez}$ ,  $d_{bouche}$  sur la base INT-VINSI

Sur cette base, nous obtenons des résultats de détection corrects car la qualité des images est bonne (le changement de la lentille et l'augmentation de la puissance du flash infrarouge améliorent sensiblement

<sup>5</sup> O. Jesorsky, K.J Kirchberg and R.W Frisholz, "Robust Face Detection Using the Hausdorff Distance" Proc Int'l Conf. Audio- and video-Based Biometric Person Authentication, pp.90-95,2001

<sup>6</sup> M. Sadeghi, J. Kittler, A. Kostin, and K. Messer, "A Comparative Study of Automatic Face Verification Algorithms on the BANCA Database," Proc. Fourth Int'l Conf. Audio- and Video-Based Biometric Person Authentication, pp. 35-43, 2003.

la qualité du capteur par rapport à des versions précédentes du prototype). Nous obtenons ainsi un taux de détection de 80% pour une erreur ( $d_{yeux}$ ) de 0.05.

## II.4 Conclusion :

Les résultats présentés ici sont encore préliminaires. La base de donnée est de très petite taille mais les résultats sont néanmoins très bons surtout si l'on considère que la détection des points caractéristiques est impossible sur un grand nombre d'images en lumière visible, celles qui sont saturées par une trop forte illumination.

Cet algorithme va nous servir pour mettre au point une méthode de vérification d'identité basée sur ces images proche infra-rouge que nous souhaitons, dans la mesure du possible intégrer dans le capteur.

## III. Multimodalité empreintes/visage dans le projet VINSI

### III.1 Introduction

Dans le cadre de l'application de vérification nomade développée dans le projet VINSI, la multimodalité peut jouer plusieurs rôles. En effet, lorsque l'on combine l'empreinte qui est très performante, avec le visage qui l'est moins, l'objectif principal n'est pas uniquement l'amélioration des performances. L'amélioration des performances sera en effet assez faible par rapport aux coûts engendrés par la multiplication des modalités. Outre le coût, le temps d'acquisition ainsi que le temps de traitement sur la plateforme mobile sont des enjeux très importants dans une telle application. L'apport essentiel de la multimodalité sera d'une part l'amélioration ponctuelle de la détection d'imposture et d'autre part la proposition d'une alternative aux cas dans lesquels l'empreinte ne peut être utilisée comme moyen de vérification. Dans cette étude, nous allons proposer une stratégie permettant de tirer profit des intérêts de la multimodalité tout en réduisant les inconvénients (coût, temps et complexité de traitement).

### III.2 Données utilisées pour l'étude des stratégies de fusion multimodales

La base BIOMET<sup>7</sup> a été choisie pour évaluer le système d'empreinte car elle possède des images acquises à la fois avec un capteur optique et un capteur capacitif. En effet, dans l'application VINSI, les images de référence stockées sur le document d'identité sont acquises avec un capteur optique, alors que le capteur intégré à la plateforme est un capteur capacitif. La base BIOMET permet donc d'évaluer le système dans des conditions proches de celles de l'application.

La base FRGCv2<sup>8</sup> a été choisie en visage. Cette base contient des variations d'illumination importantes et des images floues comme c'est le cas dans le cadre applicatif de VINSI.

Le système d'empreinte utilisé est le système NFIS2<sup>9</sup> (NIST Fingerprint Image Software) du NIST (National Institute of Standards and Technology). Le système de Visage a été développé par l'INT (Institut National des Télécommunications) et est basé sur une méthode de PCA (Principal Component Analysis)<sup>10</sup>.

---

<sup>7</sup> S. Garcia-Salicetti, C. Beumier, G. Chollet, B. Dorizzi, J. Leroux-Les Jardins, J. Lunter, Y. Ni, D. Petrovska-Delacretaz, "BIOMET: a Multimodal Person Authentication Database Including Face, Voice, Fingerprint, Hand and Signature Modalities", *Audio and Video-Based Biometric Person Authentication, 4<sup>th</sup> International Conference (AVBPA'03), Lecture Notes in Computer Science (LNCS)*, Vol. 2688, pp. 845-853, Guildford, Angleterre

<sup>8</sup> <http://www.frvt.org/FRGC/>

<sup>9</sup> <http://www.itl.nist.gov/iad/894.03/fing/fing.html>

<sup>10</sup> M. Anouar Mellakh, D. Petrovska-Delacretaz and B. Dorizzi, "Using Signal Residual Information of Eigenfaces for PCA Face Space Dimensionality Characteristics", Hong Kong ICPR 2006

L'analyse des distributions des scores fournis par les deux systèmes permet de constater que les performances des deux systèmes sont très différentes. Les distributions de scores client et imposteur sont bien séparées pour le système d'empreinte ; ceci va permettre dans la plupart des cas la prise d'une bonne décision par le système d'empreinte considéré seul. En revanche, les distributions de scores du système visage sont très mêlées, ce qui induira beaucoup d'erreurs de décision si le système de visage est considéré seul (voir Figure 7).

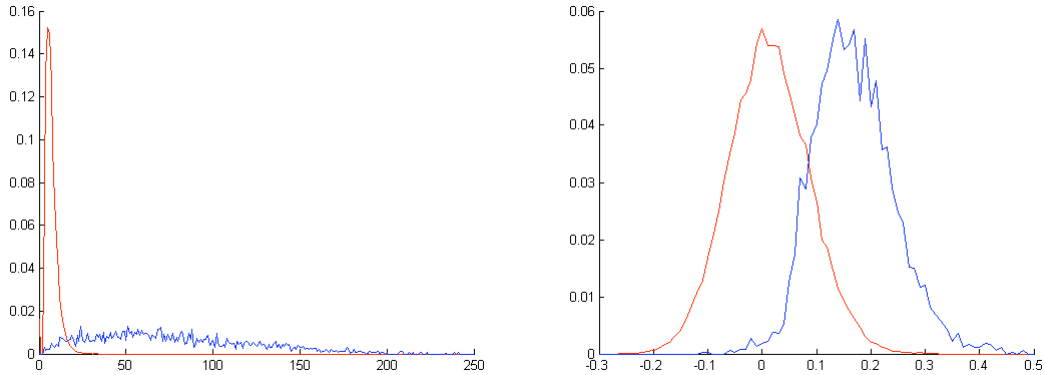


Fig. 7 : Distribution des scores des systèmes empreinte (gauche) et visage (droite)

### III.3 Protocole d'évaluation des stratégies de fusion multimodales

Pour réaliser les expériences de fusion nous avons considéré une base d'individus virtuels obtenus en combinant une empreinte de la base BIOMET avec un visage de la base FRGC. Nous avons choisi cette combinaison de bases car nous ne possédons pas de base biométrique contenant pour une même personne à la fois des empreintes et des visages dans des conditions proches de l'application VINSI. Les scores fournis par les systèmes d'empreinte et de visage ont permis de constituer 124 individus virtuels. Il est légitime dans le cadre de l'application VINSI, de construire une base virtuelle pour évaluer le système multimodal, étant donné que les modalités empreinte et visage peuvent être considérées comme étant des modalités décorrélées<sup>11</sup>.

### III.4 Méthodes globales de fusion

Il existe deux types de méthodes de fusion globales : les méthodes de combinaison des scores après une étape de normalisation et les méthodes de classification par apprentissage.

Nous avons testé deux méthodes de normalisation de scores avant de les combiner par une moyenne, la normalisation MinMax<sup>12</sup> et la normalisation via les probabilités a posteriori. Nous avons aussi testé deux méthodes de modélisation bidimensionnelle des scores, l'estimation des densités 2D par hypothèse gaussienne et par GMM<sup>13</sup>. Ces méthodes sont décrites en détail dans un travail précédent<sup>14</sup>. Nous avons également mis en oeuvre une méthode de combinaison des scores par une somme pondérée. Pour cette méthode, les scores sont d'abord normalisés par la normalisation MinMax avant d'être pondérés. Les poids de pondération sont obtenus en minimisant le taux d'erreur HTER sur la

<sup>11</sup> B. Dorizzi, S. Garcia-Salicetti, L. Allano. Multimodality in Biosecure : Evaluation on Real vs. Virtual Subjects. In Proc. of ICASSP. 14-19 May 2006. Toulouse, France.

<sup>12</sup> M. Indovina, U. Uludag, R. Snelick, A. Mink, A. Jain, "Multimodal Biometric Authentication Methods: A COTS Approach", Workshop on Multimodal User Authentication (MMUA), Santa Barbara, CA, December 2003.

<sup>13</sup> D.A. Reynolds and R.C. Rose, "Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models", IEEE Transactions on Speech and Audio Processing, Vol. 3, No. 1, Jan 1995.

<sup>14</sup> L. Allano, A. C. Morris, H. Sellahewa, S. Garcia-Salicetti, J. Koreman, S. Jassim, B. Ly-Van, D. Wu, B. Dorizzi, Non intrusive multi-biometrics on a mobile device: a comparison of fusion techniques. In Proc of SPIE 2006, conference on Biometric Techniques for Human Identification III, Orlando, Florida, USA, 17-21 April, 2006

base d'apprentissage de fusion par une recherche exhaustive des poids de façons à ce que leur somme soit égale à 1.

Nous avons choisi de présenter uniquement les résultats de la méthode de la somme pondérée qui est la méthode simple de fusion qui a donné les meilleurs résultats dans nos tests sur la base concernée (voir Table 1).

### **III.4 Stratégies de fusion multimodales locales**

Sur la figure 7 (gauche) qui représente la distribution des scores d'empreinte, on remarque que l'on peut distinguer 2 zones dans lesquelles le système d'empreinte a des comportements différents. Une zone « sûre » pour les scores d'empreinte supérieurs ou égaux à 35, et une zone « d'indécision » pour les scores d'empreinte inférieurs à 35.

La stratégie que nous proposons consiste à utiliser des méthodes de fusion et de décision différentes selon la zone dans laquelle se trouve le score du système d'empreinte.

La zone « sûre » correspond à 80% des scores clients et 0% des scores imposteurs. Dans cette zone le système d'empreinte ne fait pas d'erreurs. La zone « d'indécision » correspond à 20% des scores clients et 100% des scores imposteurs. Dans cette zone le système d'empreinte commet toutes ses erreurs à cause du fort recouvrement des distributions client et imposteur.

Dans la zone « sûre », le système d'empreinte seul suffit pour avoir de bonnes performances. L'utilisation du système d'empreinte seul à un avantage pour l'application: la réduction du temps de calcul, du temps d'acquisition et du coût. En effet, dans ce cas, on n'a pas besoin d'acquérir le visage et de calculer le résultat du système de visage.

En revanche, il a un inconvénient non négligeable dans le contexte applicatif VINSI car il ne permet pas de fiabiliser la détection d'imposteurs. En effet, la multimodalité permet d'améliorer la détection d'impostures en multipliant les contrôles. Le fait de n'utiliser que le système d'empreinte sans aucun contrôle du visage enlève toute possibilité de détection d'imposture par le système de fusion. Cependant, dans l'application VINSI, on pourrait trouver un bon compromis lorsqu'on utilise uniquement le système automatique d'empreinte en demandant un contrôle visuel, et non plus automatique (par le système), du visage par l'agent de contrôle ; ce qui ne demande ni acquisition ni traitement de données supplémentaire.

Dans la zone d'indécision du système d'empreinte, on va essayer de tirer parti de la fusion en appliquant une stratégie de fusion particulière à cette zone.

La mise en oeuvre de la fusion à cette zone d'indécision peut être faite de 2 façons :

- Utilisation d'une méthode de fusion globale, par exemple la somme pondérée après normalisation MinMax, mais en utilisant en aval un seuil spécifique pour prendre la décision lorsque le score issu du système d'empreinte se trouve dans cette zone.
- Utilisation d'une méthode de fusion spécifique lorsque le score issu du système d'empreinte se trouve dans cette zone : ainsi, les paramètres de cette méthode seront appris uniquement sur des données correspondant à cette zone (somme pondérée locale dans la Table 1). Dans ce cas particulier, la méthode de fusion peut être, si besoin, plus complexe.

Les résultats sur notre base de données sont présentés dans la Table 1.

<b>%</b> (en moyenne sur 10 tirages apprentissage/test)	<b>Somme pondérée globale (seuil de décision global)</b> Poids empreinte= 0.8 Poids visage = 0.2	<b>Somme pondérée globale et seuil de décision spécifique à la zone d'indécision</b>	<b>Somme pondérée locale et seuil de décision spécifique à la zone d'indécision</b> Poids empreinte dans la zone =0.5    Poids visage dans la zone = 0.5	<b>GMM pour la fusion dans la zone d'indécision (seuil local)</b>
<b>EER (%)</b>	<b>2.49</b>	<b>2.49</b>	<b>2.47</b>	<b>2.40</b>
<b>Amélioration relative de l'EER par rapport au système d'empreinte (EER = 4.59 %)</b>	-45.7%	-45.7%	-46.2%	-47.7%
<b>Amélioration relative de l'EER par rapport au système de visage (EER = 13.87 %)</b>	-82%	-92%	-82.2%	-82.7%

Table 1 : Performance des systèmes de fusion locaux et globaux

Cette stratégie locale donne des résultats équivalents à la meilleure méthode globale qui est la somme pondérée. Le GMM local améliore même un peu les résultats sans ce que cela puisse être considéré comme vraiment significatif.

Mais le réel intérêt de cette stratégie locale est de garder le même niveau de performance en diminuant le temps et le coût d'acquisition et de traitement pour une partie des tests. Dans notre cas, pour 80% des clients, l'acquisition et le traitement de l'empreinte suffisent à la décision. Si on souhaite réduire les impostures intentionnelles on peut alors associer à cette méthode un contrôle visuel du visage de la part de l'agent.

### III.5 Conclusion

Au cours de cette étude, nous avons observé un apport substantiel de la fusion : 40% à 45% d'amélioration relative par rapport au système d'empreinte pour la stratégie de fusion globale et jusqu'à 48% pour la stratégie de fusion locale.

Lorsque les systèmes monomodaux ont des performances très différentes comme c'est le cas entre le système d'empreinte et le système de visage, l'intérêt de la fusion ne réside pas seulement dans l'amélioration des performances mais aussi :

- D'une part, dans le fait de disposer d'une alternative à la modalité « principale » (la modalité donnant les meilleurs résultats) en cas d'impossibilité d'acquisition (inexistence ou qualité trop faible)
- D'autre part, dans la possibilité d'utiliser la fusion (la combinaison) de ces 2 modalités pour améliorer la détection d'imposture.

Cependant l'utilisation systématique de deux modalités alourdit les processus d'acquisitions et de traitement mais nous avons montré qu'une stratégie locale donne le même niveau de performance que les stratégies globales en diminuant le coût et le temps d'acquisition et de traitement : 80% des clients sont reconnus avec l'empreinte seulement.