

# Evaluation de la confiance des minuties d'une empreinte digitale

Christophe CHARRIER

Normandie Univ., UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen

christophe.charrier@unicaen.fr

**Résumé** – Les systèmes de reconnaissance biométriques par empreinte digitale sont généralement basés sur une étape d'appariement de deux modèles de minuties. Chaque minutie représente un point particulier de l'empreinte. Les applications embarquées ont en général une capacité mémoire restreinte, et ne peuvent contenir qu'un nombre limité de minuties. La vitesse de reconnaissance des empreintes digitales étant proportionnelle au nombre de minuties comparées, une nouvelle mesure de confiance est proposée pour chaque minutie du modèle. Cet indice prédit l'importance et l'utilité de chaque minutie par rapport aux autres minuties du modèle. Cette mesure a été appliquée dans le cas d'utilisation de la réduction de gabarit sur les bases de données Fingerprint Verification Competition (FVC) et SFINGE et a démontré sa capacité à atteindre des performances élevées.

**Abstract** – Fingerprint-based biometric recognition systems are usually based on a matching step of two minutiae patterns. Each minutiae represents a particular point of the fingerprint. Embedded applications usually have a limited memory capacity, and can only contain a limited number of minutiae. Since the speed of fingerprint recognition is proportional to the number of minutiae compared, a new confidence measure is proposed for each minutia of the model. This index predicts the importance and usefulness of each minutiae relative to the other minutiae in the model. This measure has been applied in the use case of template reduction on the Fingerprint Verification Competition (FVC) and SFINGE databases and has been shown to achieve high performance.

## 1 Introduction

De nos jours, l'omniprésence des appareils électroniques tels que les smartphones, les tablettes, etc. nécessite des moyens d'accéder facilement aux données ou aux applications quel que soit l'appareil utilisé. L'une des solutions déployées repose sur l'utilisation d'une modalité biométrique, comme le visage ou les empreintes digitales la plupart du temps. Même si les systèmes de reconnaissance biométrique basés sur les empreintes digitales atteignent un niveau de confiance élevé, certaines questions reposant sur l'évaluation de la qualité des empreintes digitales restent difficiles à résoudre. Pour que le système biométrique accorde ou non l'accès à la personne qui présente son empreinte digitale au capteur, cette dernière est tout d'abord analysée afin d'obtenir un nouveau template (ou modèle) de minuties (qui sont des points d'intérêt de l'empreinte). Une étape d'appariement est ensuite opérée. Cette étape consiste à trouver l'alignement [1] des minuties entre le modèle de référence de l'empreinte digitale et le modèle candidat.

Au cours de la dernière décennie, de nombreuses études ont été menées pour évaluer l'impact de la qualité de l'image sur les performances des systèmes de reconnaissance des empreintes digitales. Il est, aujourd'hui, communément admis par la communauté que la précision de tout processus de reconnaissance des empreintes digitales sans utilisation de la qualité peut être affectée. Ainsi, une fois que l'image acquise est considérée comme étant de qualité suffisante, un extracteur de minuties est appliqué et le modèle obtenu est également considéré comme étant de qualité équivalente afin de faciliter le processus de mise en correspondance.

Le modèle extrait est composé d'un ensemble de points spécifiques appelés minuties. Chaque minutie suit la norme ISO

Compact Card [2] et est codée selon le quadruplet  $(x_i, y_i, T_i, \theta_i)$ , où  $(x_i, y_i)$  représente les coordonnées spatiales de la minutie dans l'image,  $T_i$  son type (bifurcation, crête, terminaison, etc.) et  $\theta_i$  l'orientation (par rapport à la ligne de crête).

Le type des minuties se décline en trois catégories [2] : le type 0 correspondant à la fin de l'arête, le type 1 à une bifurcation de l'arête et le type 2 pour tous les autres types.

Néanmoins, une question demeure : lorsqu'on considère un modèle d'empreinte digitale (généralement obtenu à l'aide d'un extracteur de minuties), peut-on considérer que toutes les minuties extraites ont le même niveau de qualité ? En d'autres termes, les minuties extraites ont-elles le même niveau d'importance ? Nous pouvons reformuler cette question en abordant la question de la confiance que l'on peut avoir vis-à-vis d'une minutie, en intégrant un concept d'utilité.

dans la littérature, il existe très peu d'approches d'évaluation de la qualité qui prennent en compte simultanément les informations sur les minuties. On peut citer NFIQ [3] et MINDTCT [4]. MINDTCT fournit une mesure de fiabilité qui calcule un score de qualité en fonction de la carte de qualité [3] et des statistiques sur les pixels voisins. NFIQ, qui est un algorithme d'évaluation de la qualité des images d'empreintes digitales [5], calcule un ensemble de caractéristiques de qualité et les utilise pour prédire la qualité de l'image d'empreintes digitales. Il est à noter cependant que la mesure NFIQ fournit le nombre de minuties à différents niveaux de qualité et non un score de qualité pour chaque minutie.

Dans cet article, nous définissons une mesure de confiance d'une minutie uniquement basée sur le template de l'empreinte digitale, sans aucune référence à l'image associée.

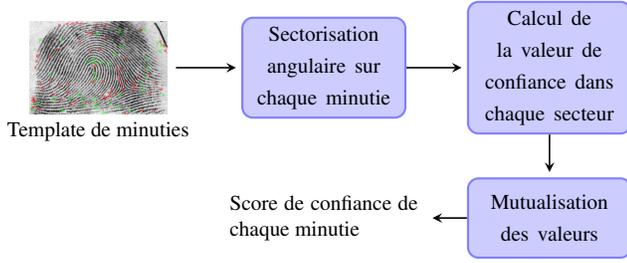


FIG. 1: *Synopsis de la méthode de calcul de l'indice de confiance.*

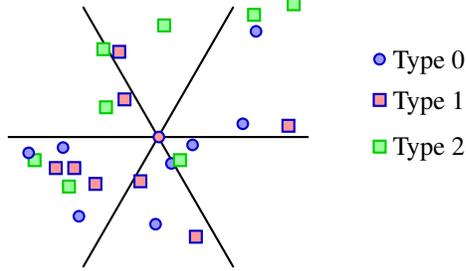


FIG. 2: *Illustration de la répartition des types de minuties autour d'une minutie donnée (disque rouge) par secteur*

## 2 Indice de confiance des minuties

Tel qu'illustré dans la figure 1, l'indice de confiance MiCI proposé est calculé comme suit : Tout d'abord, l'une des minuties est sélectionnée comme point de référence. La région de l'empreinte digitale, en considérant le point de référence comme origine du repère, est ensuite décomposée en  $s$  secteurs de même angle. Un indice de confiance est ensuite calculé pour chaque secteur. L'indice de confiance globale de la minutie de référence est finalement calculé à partir de valeurs calculées pour chacun des secteurs angulaires  $s$ .

### 2.1 Décomposition sectorielle

Soit un modèle de minuties  $T_i$  d'une empreinte digitale qui contient l'ensemble des minuties  $V_{up}$  extraites de l'image de l'empreinte digitale

$$V_{up} = \{m_i\}_{i=1}^n \quad (1)$$

où  $n$  est le nombre total de minuties de l'ensemble  $V_{up}$ .

$$m_i = \{(x_i, y_i), \theta_i, T_i\} \quad (2)$$

La  $i^{me}$  minutie est désignée par  $m_i$  où  $(x_i, y_i)$ ,  $\theta_i$  et  $T_i$  représentent respectivement les coordonnées, l'orientation et le type de la minutie  $i$ .

Au cours de cette étape, l'une des minuties  $m_i$  de  $V_{up}$  est sélectionnée comme minutie de référence  $m_{ref}$ . Le template est ensuite divisé en  $s$  secteurs de largeur angulaire égale autour de la minutie de référence, dans le sens anti-horaire, et  $V_{s_j}$  représente l'ensemble des minuties dans le secteur  $s_j$ .

Pour chaque ensemble  $V_{s_j}$ , un sous-ensemble  $V_{T_i, s_j}$  est généré par type  $T_i$  de minuties. Ainsi, à la fin de cette étape,

trois sous-ensembles sont générés par secteur angulaire (un par type).

La figure 2 illustre la décomposition sectorielle appliquée autour d'une minutie en six secteurs. Pour chaque secteur, on observe les trois différents types de minuties rencontrés, ce qui nous permet de générer un sous-ensemble de minuties par type.

Cette décomposition sectorielle est appliquée pour chaque minutie de  $V_{up}$ .

### 2.2 Indice de confiance sectoriel

L'objectif est de prendre en compte l'influence de chaque sous-ensemble sur la minutie  $m_{ref}$  pour lequel un indice confiance est calculé. Il est, en effet, communément admis que l'influence d'un sous-ensemble diminue si la distance à la minutie  $m_{ref}$  augmente. En outre, la taille du sous-ensemble a une influence sur l'utilité, *i.e.*, plus la cardinalité d'un sous-ensemble est faible, moins son influence est grande. À partir de ces deux hypothèses, un indice de confiance sectoriel (ICS) est élaboré et défini comme suit :

$$ICS_{m_{ref}, s} = \sum_{k=1}^3 p(T_k, s) * 1/\sqrt{d(m_{ref}, c(T_k, s))} \quad (3)$$

où

- $p(T_k, s)$  représente la probabilité d'apparition des minuties du type  $T_k$  dans le secteur  $s$ ,
- $d(m_{ref}, c(T_k, s))$  est la distance euclidienne entre le centroïde  $c(T_k, s)$  des minuties de type  $T_k$  dans le secteur  $s$  et la minutie de référence  $m_{ref}$ .

ICS détermine ainsi une information utile sur la distribution des minuties dans le secteur, ainsi que l'importance du secteur.

### 2.3 Indice de confiance global

Enfin, une stratégie de mise en commun des valeur ICS est appliquée pour calculer l'indice de confiance final des minuties (MiCI) de chaque minutie de référence :

$$MiCI_{m_{ref}} = \sum_{i=1}^s ICS_{m_{ref}, i} \quad (4)$$

où  $s$  représente le nombre de secteurs angulaires obtenus à partir du processus de décomposition des secteurs.

## 3 Résultats Expérimentaux

Afin d'évaluer les performances de la méthode proposée, nous considérons le problème de la réduction du gabarit. L'objectif de ce problème est de réduire le nombre de minuties du modèle de référence tout en préservant les performances des systèmes biométriques. Les cas d'usage sont nombreux, notamment lorsque les empreintes sont stockées dans les éléments sécurisés des cartes à puce, comme cela est notamment le cas pour les passeports biométriques.

### 3.1 Configuration expérimentale

#### 3.1.1 Bases de données d'empreintes digitales

Dans cette étude, les expériences sont menées avec les quatre bases de données FVC de différentes résolutions [6] : FVC 2002 DB1 et DB2, ainsi que FVC 2004 DB1 et DB2

Toutes les images d’empreintes digitales FVC sont capturées à l’aide d’un capteur optique. Chaque base de données contient 100 empreintes digitales et 8 échantillons pour chaque empreinte. Les scores intra-classe contiennent  $7 \times 100 = 700$  scores authentiques, et les scores inter-classes consistent en  $7 \times 99 \times 100 = 69300$  scores d’imposteurs pour l’ensemble de la base de données.

En outre, cinq bases de données d’images d’empreintes digitales synthétiques, générées avec SFINGE [7], seront utilisées pour évaluer les performances de l’indice MiCI, à savoir : SFINGE0 avec différentes images de qualité allant aléatoirement d’une qualité faible à une qualité très élevée, SFINGEA où les images sont de très haute qualité, SFINGEB avec des images de haute qualité, SFINGEC avec des images de qualité moyenne et SFINGED contenant des images de faible qualité.

### 3.1.2 Extracteur de minutiae et appariement

Les templates de minuties ont été extraits avec MINDTCT [4]. Cet extracteur fournit une métrique de fiabilité attribuée à chaque minutie détectée dans l’image d’entrée. Pour chaque minutie de l’image un score de qualité est calculé entre 0 et 1, où 0 (resp. 1) correspond à une minutie de faible (resp. haute) qualité.

Deux algorithmes d’appariement des empreintes digitales ont également été utilisés : Bozorth3 [8] et MCC [9].

### 3.1.3 Méthodes de réduction de gabarit

Les performances de la méthode proposée sont comparées à celles issues de deux stratégies de réduction de template communément utilisées, à savoir, la méthode de troncature, utilisée par nombre de systèmes biométriques commerciaux et la méthode du barycentre proposée par le NIST [10].

### 3.1.4 Indice de performance

Il existe de nombreuses mesures pour évaluer les performances d’un système biométrique. Parmi toutes ces mesures, l’AUC (*Area-under-Curve*) de la courbe ROC (*Receiver Operating Characteristics*) est utilisée. L’AUC est un indice quantitatif dérivé de la courbe ROC [11]. La valeur AUC peut être considérée comme une mesure de classement, basée sur des comparaisons par paire entre les classifications de deux classes. En d’autres termes, la valeur AUC est égale à la probabilité qu’un classificateur classe une instance positive choisie au hasard avec une probabilité plus élevée qu’une instance négative choisie au hasard. Ainsi, elle peut être considérée comme un critère global de performance. Dès lors, si l’on considère deux algorithmes et leurs valeurs AUC associées, le meilleur algorithme est celui dont la valeur AUC est inférieure à celle des autres algorithmes sur une plage suffisamment large de valeurs de seuil.

## 3.2 Résultats

### 3.2.1 Nombre optimal de secteurs

Afin de déterminer le nombre optimal de secteurs nécessaires pour obtenir les meilleures performances, l’indice MiCI est appliqué en utilisant 4, 6, 8 et 10 secteurs. Nous réduisons le modèle d’empreinte digitale en éliminant les 5, 10 et 15 minuties ayant les scores MiCI les plus faibles. Nous limitons le nombre

		# Sectors	#Minuties retirées		
			5	10	15
Bases de données réelles	FVC2002DB1	Sans réduction	0.015	0.015	0.015
		Troncature	0.024	0.038	0.066
		Centroid	0.021	0.032	0.050
		MINDTCT	0.022	0.029	0.050
		MiCI	<b>0.020</b>	<b>0.026</b>	<b>0.047</b>
	FVC2002DB2	Sans réduction	0.018	0.018	0.018
		Troncature	0.024	0.033	0.055
		Centroid	0.022	0.029	0.053
		MINDTCT	0.022	0.027	0.047
		MiCI	<b>0.010</b>	<b>0.021</b>	<b>0.025</b>
	FVC2004DB1	Sans réduction	0.034	0.034	0.034
		Troncature	0.041	0.055	0.096
		Centroid	0.045	0.061	0.089
		MINDTCT	0.040	0.056	0.080
		MiCI	<b>0.039</b>	<b>0.048</b>	<b>0.078</b>
	FVC2004DB2	Sans réduction	0.036	0.036	0.036
		Troncature	<b>0.036</b>	0.044	0.058
		Centroid	0.042	0.057	0.082
		MINDTCT	0.039	0.048	0.061
		MiCI	0.037	<b>0.042</b>	<b>0.051</b>
Bases de données synthétiques	SFINGE0	Sans réduction	0.027	0.027	0.027
		Troncature	0.048	0.093	0.161
		Centroid	0.041	0.074	0.130
		MINDTCT	0.044	0.078	0.141
		MiCI	<b>0.040</b>	<b>0.065</b>	<b>0.113</b>
	SFINGEA	Sans réduction	0.017	0.017	0.017
		Troncature	0.032	0.069	0.118
		Centroid	0.030	0.064	0.097
		MINDTCT	0.031	0.067	0.101
		MiCI	<b>0.029</b>	<b>0.048</b>	<b>0.093</b>
	SFINGEB	Sans réduction	0.020	0.020	0.020
		Troncature	0.036	0.072	0.121
		Centroid	0.028	0.060	0.101
		MINDTCT	0.032	0.065	0.109
		MiCI	<b>0.024</b>	<b>0.033</b>	<b>0.066</b>
	SFINGEC	Sans réduction	0.023	0.023	0.023
		Troncature	0.044	0.076	0.121
		Centroid	0.042	0.070	0.117
		MINDTCT	0.041	0.069	0.117
		MiCI	<b>0.041</b>	<b>0.067</b>	<b>0.116</b>
SFINGED	Sans réduction	0.038	0.038	0.038	
	Troncature	0.057	0.089	0.153	
	Centroid	0.056	0.077	0.147	
	MINDTCT	0.055	0.080	0.151	
	MiCI	<b>0.051</b>	<b>0.075</b>	<b>0.140</b>	

TAB. 1: Valeurs AUC des méthodes de réduction de template calculées sur les bases de données, en utilisant Bozorth3.

maximum de minuties retirées à 15 étant donné que dans les applications opérationnelles, le nombre de minuties stockées est habituellement fixé à 48 et que les extracteurs de minuties utilisés fournissent jusqu'à 61 minuties.

Nous effectuons ensuite l'appariement à l'aide des modèles d'empreintes réduits et nous calculons la valeur AUC associée.

Les résultats obtenus montrent que le nombre de secteurs minimisant les valeurs AUC est atteint lorsque 8 secteurs sont utilisés pour calculer l'indice de confiance des minuties MiC, et ce quelle que soit la base de données utilisée. On observe que la base de données utilisée n'a pas d'influence sur le nombre de secteurs pour effectuer la décomposition, puisque la décomposition en 8 secteurs donne toujours les meilleurs résultats.

### 3.2.2 Évaluation des performances

Avec l'algorithme de réduction de gabarit basé sur le MiCI, les minuties ayant les scores MiCI les plus faibles sont supprimés du gabarit. Nous procédons à l'élimination des minuties par pas de 5 en trois fois. Nous obtenons ainsi trois modèles réduits pour chaque modèle initial. Le premier modèle réduit englobe les minuties à l'exception des 5 minuties ayant les scores MiCI les plus faibles. Chaque modèle réduit obtenu suit le même processus deux fois. Au final, nous obtenons trois modèles d'empreintes digitales réduits de 5, 10 et 15 minuties respectivement.

La même stratégie est appliquée à la technique de réduction des modèles basée sur la qualité MINDTCT. Le processus de suppression des minuties est piloté par les scores de qualité MINDTCT les plus faibles.

Enfin, les valeurs AUC sont calculées pour tous les modèles générés.

Seuls, les résultats obtenus avec l'algorithme d'appariement Bozorth3 sont présentés dans le tableau 1, les résultats obtenus avec l'algorithme MCC ayant un comportement similaire.

Les valeurs AUC obtenues à partir de l'algorithme de réduction des gabarits basé sur les valeurs MiCI proposées sont les plus performantes par rapport aux trois schémas. Cela nous permet de conclure que la stratégie proposée pour évaluer la confiance d'une minutie est valable et prometteur. La réduction des modèles de minuties de 5 minuties n'a pas un impact aussi important sur les performances de l'appariement puisque les valeurs AUC des différentes méthodes sont approximativement égales à l'AUC obtenue lorsqu'aucune réduction n'est calculée sur le modèle.

Deux types de bases de données d'images d'empreintes digitales sont pris en compte pour effectuer la comparaison : 1) les empreintes digitales capturées à partir de doigts réels et 2) les empreintes digitales synthétiques.

Le tableau 1, qui présente les résultats obtenus à partir des empreintes digitales réelles et Bozorth3, montre que la différence des valeurs AUC entre l'approche basée sur MiCI et la deuxième meilleure approche est minime et varie en moyenne entre 0,001 (pour une étape de réduction de 5 minuties) et 0,004 (pour une étape de réduction de 15 minuties). Lorsque les empreintes synthétisées sont prises en compte, ces différences augmentent globalement et varient entre 0,0025 et 0,0158. Nous

observons une légère baisse de performance lorsque des bases de données synthétisées sont utilisées avec Bozorth3.

## 4 Conclusion

Dans ce papier, un nouvel indice de confiance des minuties (MiCI) qui est une valeur calculé sans référence à l'image et qui vise à attribuer un indice de confiance à chaque minutie a été proposé. Cet indice se base, entre autre, sur la distribution spatiale des minuties par type dans un voisinage proche. Les résultats obtenus permettent de valider l'approche dans le cas de la réduction de gabarit pour les applications embarquées. À l'avenir, l'étude la robustesse de l'indice MiCI pour différents types et différentes intensités de distorsion des empreintes digitales, telles que le flou, le recadrage, la rotation et les rides, sera réalisée.

## Références

- [1] H. Bischof H. Ramoser, B. Wachmann. Efficient alignment of fingerprint images. *International Conference on Pattern Recognition*, 3:748–751, 2002.
- [2] ISO/IEC 19795-2. information technology - biometric data interchange format - part 2: Finger minutiae data, 2004.
- [3] Watson C. Tabassi E., Wilson C. Fingerprint image quality. 7151, 8 2004.
- [4] K. Ko. User's guide to NIST biometric image software (NBIS). *NIST Interagency/Internal Report (NISTIR)-7392*, 2007.
- [5] Tabassi E. Schwaiger C. Nfiq 2.0 open source distribution. *NIST Gaithersburg*, 1 2014.
- [6] Maio Dario, Maltoni Davide, Cappelli Raffaele, Wayman Jim L, and Jain Anil K. FVC2004: Third fingerprint verification competition. In *Biometric Authentication*, pages 1–7. Springer, 2004.
- [7] D. Maio R. Cappelli and D. Maltoni. Sfinge: an approach to synthetic fingerprint generation. *International Workshop on Biometric Technologies (BT2004)*, pages 147–154, 2004.
- [8] P. A. Flanagan. NIST biometric image software (NBIS). 2010.
- [9] R. Cappelli, M. Ferrara, and D. Maltoni. Minutia cylinder-code: A new representation and matching technique for fingerprint recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32:2128–2141, 2012.
- [10] Grother N. Grother P., Salamon W. Interoperability of the iso/iec 19794-2 compact card and 10 iso/iec 7816-11 match-on-card specifications 11. *International Conference on Pattern Recognition*, 2007.
- [11] McNeil J. Hanley J. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology*, 1(143):29–36, 1982.