Sensor Interoperability and Fusion in Fingerprint Verification: A Case Study using Minutiae- and Ridge-Based Matchers

F. Alonso-Fernandez [a], R. N. J. Veldhuis[b], A. M. Bazen[b], J. Fierrez-Aguilar[a] and J. Ortega-Garcia[a]

[a]Biometrics Research Lab.- ATVS, Escuela Politecnica Superior - Universidad Autonoma de Madrid
Avda. Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain
{fernando.alonso, julian.fierrez, javier.ortega}@uam.es

[b]University of Twente, 7500 AE Enschede, The Netherlands
{r.n.j.veldhuis, a.m.bazen}@utwente.nl

***Abstract*— Information fusion in fingerprint recognition has been studied in several papers. However, only a few papers have been focused on sensor interoperability and sensor fusion. In this paper, these two topics are studied using a multisensor database acquired with three different fingerprint sensors. Authentication experiments using minutiae and ridge-based matchers are reported. Results show that the performance drops dramatically when matching images from different sensors. We have also observed that fusing scores from different sensors results in better performance than fusing different instances from the same sensor[1].**

***Keywords*—Fingerprint, sensor interoperability, sensor fusion, minutiae, ridge.**

I. Introduction

Personal authentication in our networked society is becoming a crucial issue [1]. Due to its permanence and uniqueness, fingerprint recognition is widely used in many personal identification systems, not only in forensic environments, but also in a large number of civilian applications such as access control or on-line identification. Furthermore, due to the low cost and reduced size of new fingerprint sensors, several devices of daily use (i.e. mobile telephones, PC peripherals, etc.) already include fingerprint sensors embedded.

Several results related to information fusion for fingerprint verification have been presented [2-5]. However, only few papers have been focused on sensor fusion and interoperability [6-8]. In this paper, we study these two topics using minutiae and ridge-based matchers. The rest of the paper is organized as follows. Sensor interoperability and fusion topics are briefly addressed in Sects. II and III, respectively. Experiments and results are described in Sect. IV. Conclusions are finally drawn in Sect. V.

[1]This work has been carried out while F. A.-F. was guest scientist at University of Twente

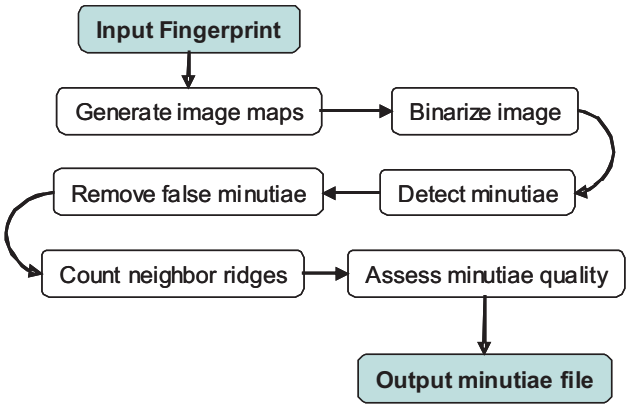

Fig. 1. Processing steps of the MINDTCT package of the NIST Fingerprint Image Software 2 (NFIS2).

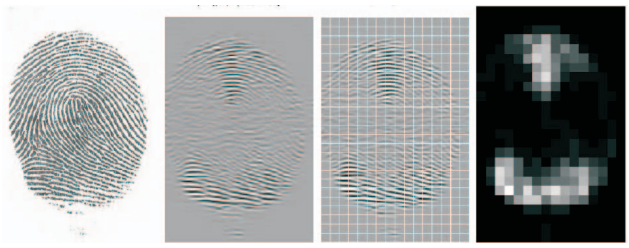

Fig. 2. Processing steps of the ridge based verification system. From left to right: original image, filtered image with filter orientation $\theta = 0$, tessellated image, and FingerCode.

II. Sensor interoperability

When a user interacts with a biometric system, a feature set is extracted from the raw data acquired by the sensor. This feature set is expected to be an invariant representation of the person. However, the feature set is sensitive to several factors [7]: i) changes in the sensor; ii) variations in the environment; iii) improper user interaction; or iv) temporary alterations of the biometric trait. Factors ii and iii can be eliminated with a

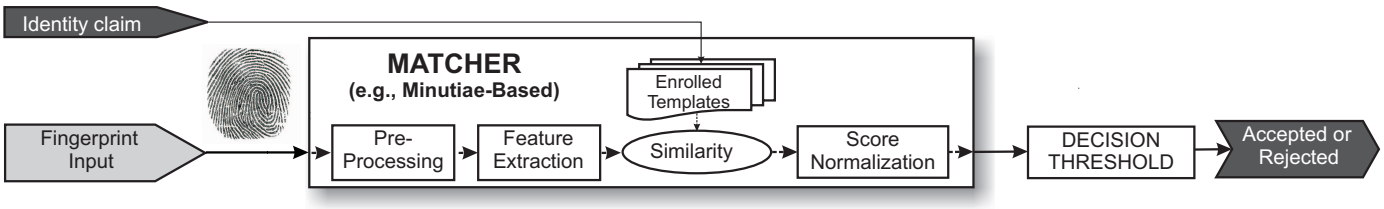

Fig. 3. Architecture of the proposed fingerprint verification system.

quality checking process while iv can be alleviated by using a periodic template update process, but the effect of changing the sensor has not been extensively studied.

Sensor interoperability in biometrics can be defined as the capability of a recognition system to operate with different sensors. Most biometric systems are designed under the assumption that the data to be compared are obtained from a unique sensor and are restricted in their ability to match or compare biometric data originating from different sensors. As a result, changing the sensor may affect the performance of the system, as demonstrated in several studies. Martin et al. [9] reported a significant difference in performance when different microphones are used during the training and testing phases of a speaker recognition system. Ross et al. [7] studied the effect of matching fingerprints acquired with two different fingerprint sensors, resulting in a significant drop of performance. Alonso et al. [10] studied the effect of matching two signatures acquired with two different Tablet PCs, resulting in a drop of performance when samples acquired with the sensor providing the worst signal quality are matched against samples acquired with the other sensor.

Recent progress has been made in the development of common data-exchange formats to facilitate the exchange of feature sets between vendors. The sensor interoperability problem is being addressed by standardization bodies. In 2002, the INCITS M1 Biometrics committee [2] was formed by ANSI and also, the Sub-Committee 37 was formed by the Joint Technical Committee 1 [3] of ISO/IEC, including Working Groups related to biometric technical interfaces and data exchange formats. Regarding fingerprints, their standardization activities have resulted in the ANSI-INCITS 378 [11] and the ISO/IEC 19795-2 standards, both for minutiae-based templates. However, little effort has been invested in the development of algorithms to alleviate the problem of sensor interoperability. Some approaches to handle this problem are given in [7]. One example is the normalization of raw data and extracted features. Interoperability scenarios should also be included in vendor and algorithm competitions, such as in the Minutiae Interoperability Exchange Test - MINEX [8]. The MINEX evaluation is intended to assess the viability of the INCITS 378 templates as the interchange medium for fingerprint data. The MINEX evaluation reported different trials using two variants of the INCITS-378 format implemented by 14 vendors. Proprietary minutiae-based templates were also included in the evaluation. A number of interesting conclusions were extracted from this evaluation: i) proprietary templates always perform better than standard ones; ii) some template generators produce standard templates that are matched more accurately than others and some matchers compare more accurately than others, but the leading vendors in generation are not always the leaders in matching and vice-versa; and iii) performance is sensitive to the quality of the dataset, both in proprietary and standard templates.

[2]http://m1.incits.org/

[3]www.jtc1.org

III. Fusion of sensors

Multibiometric systems refer to biometric systems based on the combination of a number of instances, sensors, representations, units and/or traits [12]. Several approaches for combining the information provided by these sources have been proposed in the literature [13], [14]. However, fusion of data from different sensors has not been extensively analyzed. Chang et al. [15] studied the effect of combining 2D and 3D images acquired with two different cameras for face recognition. Marcialis et al. [6] reported experiments on fusing the information provided by two different fingerprint sensors. Alonso et al. [10] studied the effect of combining the signatures acquired with two different Tablet PCs. Fusion of sensors offers some important potentialities [6]: i) the overall performance can be improved substantially, ii) population coverage can be improved by reducing enrollment and verification failures and iii) it may discourage fraudulent attempts to spoof biometric systems, since deceiving a multisensor system by submitting fake fingers would require different kinds of fake fingers for each sensor. But there are some drawbacks as well: the cost of the system may be higher and more user cooperation is needed.

IV. Experiments

A. Fingerprint matchers

In the experiments reported in this paper, we use both the minutiae-based NIST Fingerprint Image Software 2 (NFIS2) [16] and the ridge-based fingerprint matcher [17] developed in the Biometrics Research Lab. at Universidad Autonoma de Madrid, Spain.

For minutiae extraction with NFIS2, we have used the MINDTCT package, sketched in Fig. 1. For fingerprint matching, we have used the BOZORTH3 package, which computes a similarity matching score s_m between the minutiae from a template and a test fingerprint. We normalize s_m into the

Fig. 4. Fingerprint samples of two different users of the database. Fingerprint images are plotted for the same finger for i) Atmel thermal (left), ii) Digital Persona optical (upper right) and iii) Polaroid optical (lower right).

$[0, 1]$ range by $\tanh(s_m/c_m)$, where c_m is a normalization parameter chosen heuristically to evenly distribute the impostor and score distributions into $[0, 1]$. For detailed information of MINDTCT and BOZORTH3, we refer the reader to [16]. We have also used the automatic quality assessment software included in the NIST Fingerprint Image Software 2 [16], [18]. This software computes the quality of a given fingerprint based on the minutiae extracted by the MINDTCT package. A fingerprint is assigned one of the following quality values: 5 (poor), 4 (fair), 3 (good), 2 (very good) and 1 (excellent).

The ridge-based matcher uses a set of Gabor filters to capture the ridge strength as described in [2]. The variance of the filter responses in square cells across filtered images is used as feature vector. This feature vector is called FingerCode because of the similarity to previous research works [2]. The automatic alignment is based on the system described in [19]. A dissimilarity matching score s_r is then computed as the Euclidean distance between the two aligned FingerCodes. No image enhancement is explicitly performed, but it is implicitly done during the Gabor filtering stage since Gabor filters are known to be appropriate to remove the noise and preserve true ridge/valley structures [20]. The output score s_r is normalized into a similarity score in the $[0, 1]$ range by $exp(-s_r/c_r)$, where c_r is a normalization parameter chosen heuristically to evenly distribute the impostor and score distributions into $[0, 1]$. The processing steps of the ridge based verification system are shown in Fig. 2.

In this paper we focus on fingerprint verification using these matchers. The system architecture of a fingerprint verification application is depicted in Fig. 3.

B. Database and protocol

A fingerprint database has been acquired at the University of Twente using three different sensors: i) Atmel (sweeping thermal), with an image size of 360×800 pixels; ii) Digital Persona UareU (optical), with an image size of 500×550 pixels; and iii) Polaroid (optical), with an image size of 300×302 pixels. From now on, they will be referred to as *sensor*1 (Atmel Sweep), *sensor*2 (Digital Persona) and *sensor*3 (Polaroid). For our experiments, we have used a subcorpus of 100 fingers. For each finger, 12 impressions with each sensor have been acquired, resulting in three datasets of 1200 fingerprint images each (one dataset per sensor). Some example fingerprints from this database are shown in Fig. 4. We consider the different fingers as different users enrolled into the system. The following comparisons are performed

for each fingerprint matcher and for each sensor: i) *genuine matchings:* each fingerprint image is considered as an enrollment fingerprint which is compared to the remaining images of the same finger, avoiding symmetric matches, resulting in $100 \times 12 \times 11/2 = 6,600$ genuine scores per matcher and per sensor; and ii) *impostor matchings:* the second fingerprint image of each finger is compared with three images of the remaining fingers, resulting in $100 \times 99 \times 3 = 29,700$ impostor scores per matcher and per sensor.

C. Results

In Fig. 5 we can see the quality distribution of the datasets used for the experiments provided by the NFIS2 software (see Sect. IV-A). The NFIS2 software uses the extracted minutiae to compute the quality of a given fingerprint [16], [18]. We observe that the dataset acquired with the thermal sweeping sensor has better quality than the datasets acquired with the two optical sensors, although it is known that sweeping sensors have to reconstruct the fingerprint image from slices, which usually results in spurious artifacts [21].

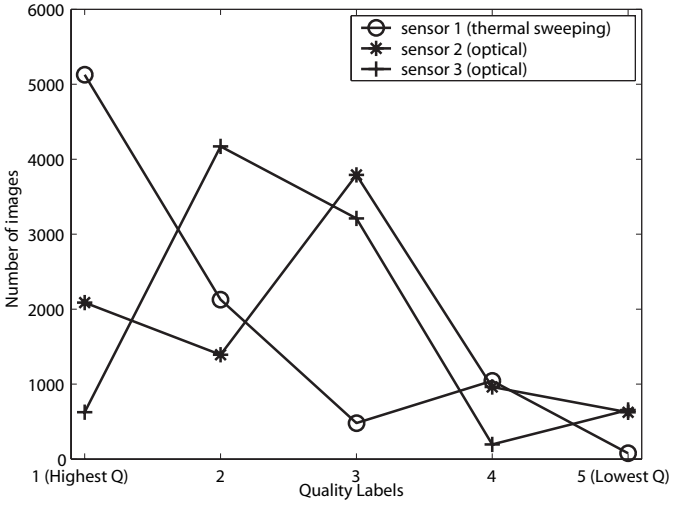

Fig. 5. Quality distribution of the datasets used for the experiments provided by the NFIS2 software.

Individual sensors. In Fig. 6 we plot the verification performance of the two matchers on the three different datasets according to the experimental protocol defined in Sect. IV-B. We observe that the minutiae-based matcher performs better than the ridge-based matcher. It is known that minutiae are more discriminative than other fingerprint features [21]. Interestingly, the performance on *sensor 1* (thermal sweeping) is better than the performance on *sensor 2* (optical) for the minutiae-based matcher, although sweeping sensors may result in errors and spurious artifacts (due to the reconstruction process that they perform) [21].

Also worth noting, the minutiae-based matcher results in the best performance on the $sensor3$ (Polaroid optical), whereas the ridge-based matcher performs best on the $sensor1$ (Atmel thermal). Due to the acquisition process of a sweeping sensor, there is practically no rotation [21]. Since the alignment performed in our ridge-based matcher only accounts for translation [19], this should be the reason of the improved performance observed with respect to the optical sensors.

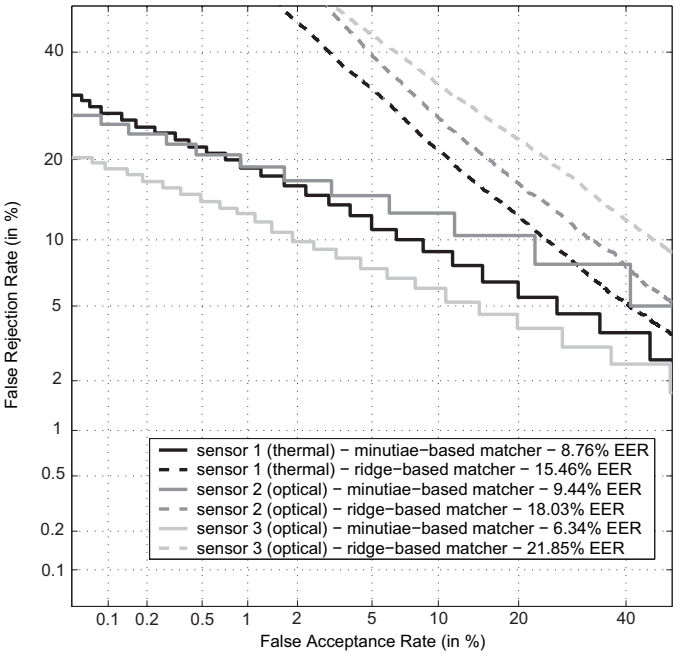

Fig. 6. Verification performance of the two matchers.

Sensor Interoperability Experiments. We study the effects of sensor interoperability by following the experimental protocol of Sect. IV-B for the individual sensors but considering different sensors for enrolment and testing. Verification performance results are given in Table I. It can be observed that when matching images from different sensors, the performance drops dramatically for both the minutiae- and ridge-based matchers. The best performance is obtained when matching images from sensors of the same technology (i.e. *sensor*2 and *sensor*3). However, in all cases the performance is insufficient for practical applications (EER higher than 40%).

To evaluate the effects of the fingerprint quality in the interoperability of sensors, we have next considered only users with medium to high quality genuine fingerprint samples (i.e. quality label of 3, 2 or 1, according to the labeling assessed by the NFIS2 software, see Sect. IV-A), as in the MINEX evaluation [8]. Verification performance results considering only images of good quality are given in Table II. Also in this case, the performance dramatically decreases for both matchers. In our experiments, we observe that image quality does not play a primary role in the drop of performance found when matching images from different sensors, both in the minutiae- and the ridge-based matcher.

Sensor Fusion Experiments. We compare fusion of different sensors with fusion of different instances of each sensor, in

EER %	testing					
	*s*1 (thermal)		*s*2 (optical)		*s*3 (optical)	
enrolm.	minut.	ridge	minut.	ridge	minut.	ridge
*s*1	**8.76%**	**15.46%**	47.71%	55.19%	45.61%	54.05%
*s*2	47.66%	50.89%	**9.44%**	**18.03%**	40.55%	43.65%
*s*3	44.61%	52.07%	39.26%	46.53%	**6.34%**	**21.85%**

TABLE I

ERROR RATES OF THE INDIVIDUAL MATCHERS (MINUTIAE- AND RIDGE-BASED) IN TERMS OF EER FOR THE EXPERIMENTS EVALUATING INTEROPERABILITY OF SENSORS. *s*1, *s*2 AND *s*3 STAND FOR *sensor*1 (THERMAL), *sensor*2 (OPTICAL) AND *sensor*3 (OPTICAL), RESPECTIVELY.

EER %	testing					
	*s*1 (thermal)		*s*2 (optical)		*s*3 (optical)	
enrolm.	minut.	ridge	minut.	ridge	minut.	ridge
s1	**6.25%**	**15.62%**	54.47%	61.89%	47.35%	57.99%
s2	47.95%	56.93%	**3.99%**	**17.05%**	37.58%	47.59%
s3	43.66%	59.18%	37.35%	50.11%	**2.86%**	**22.63%**

TABLE II

ERROR RATES OF THE INDIVIDUAL MATCHERS (MINUTIAE- AND RIDGE-BASED) IN TERMS OF EER FOR THE EXPERIMENTS EVALUATING INTEROPERABILITY OF SENSORS CONSIDERING ONLY GOOD QUALITY IMAGES. *s1*, *s2* AND *s3* STAND FOR *sensor1* (THERMAL), *sensor2* (OPTICAL) AND *sensor3* (OPTICAL), RESPECTIVELY.

order to reveal the real benefits of considering information provided from different sensors [10], [14]. In this work we have used a simple fusion approach at match-score level based on the mean rule. The use of this simple fusion rule is motivated by the fact that complex trained fusion approaches do not clearly outperform simple fusion approaches, e.g. see [3].

For the fusion experiments, we have considered all the available scores resulting from the experimental protocol defined in Sect. IV-B. To perform the fusion of different instances from the same sensor, we make groups of consecutive scores having the same fingerprint for enrolment. This results in 3,000 genuine scores and 9,900 impostor scores when fusing two instances; and 1,800 genuine scores and 9,900 impostor scores when fusing three instances from the same sensor. To perform the fusion of different sensors, we fuse all the available scores from each sensor, resulting in 6,600 genuine scores and 29,700 impostor scores.

Verification performance results are given in Table III. We observe that fusing scores from different sensors is better than fusing different instances from the same sensor, for both matchers. This reveals that the complementarity between different sensors provides capability to recover fingerprints wrongly recognized by the individual sensors [6]. This behavior has been also observed in other biometric traits [10]. Moreover, the best EER value and the best relative improvement is obtained in most cases when fusing scores from sensors with different technology, i.e. *sensor*1 (thermal) with *sensor*2 or *sensor*3 (both optical), revealing another complementarity

		minutiae-based	ridge-based
Individual	*s1*	8.76%	15.46%
	s2	9.44%	18.03%
	s3	6.34%	21.85%
Multi-instance	*s1-s1*	6.12% (-30.14%)	13.33% (-13.78%)
	s2-s2	6.85% (-27.44%)	15.04% (-16.58%)
	s3-s3	4.52% (-28.71%)	17.67% (-19.13%)
Multi-sensor	*s1-s2*	3.26% (-62.79%)	10.14% (-34.41%)
	s1-s3	2.75% (-56.62%)	13.08% (-15.46%)
	s2-s3	3.71% (-41.48%)	14.72% (-18.36%)
Multi-instance	*s1-s1-s1*	5.02% (-42.69%)	12.56% (-18.76%)
	s2-s2-s2	5.64% (-40.25%)	13.63% (-24.40%)
	s3-s3-s3	3.94% (-37.85%)	17.66% (-19.17%)
Multi-sensor	*s1-s2-s3*	1.93% (-69.56%)	9.53% (-38.36%)

TABLE III

ERROR RATES OF THE INDIVIDUAL MATCHERS TESTED (MINUTIAE- AND RIDGE-BASED) IN TERMS OF EER FOR THE EXPERIMENTS EVALUATING FUSION OF SENSORS. *s*1, *s*2 AND *s*3 STAND FOR *sensor*1 (THERMAL), *sensor*2 (OPTICAL) AND *sensor*3 (OPTICAL), RESPECTIVELY. THE RELATIVE PERFORMANCE GAIN COMPARED TO THE BEST INDIVIDUAL MATCHER INVOLVED IS ALSO GIVEN.

,

based on the technology. Also worth noting, the minutiae-based matcher obtains higher relative EER improvements than the ridge-based matcher in all cases.

V. CONCLUSIONS

Sensor interoperability and sensor fusion have been studied using a minutiae- and a ridge-based fingerprint matchers. Experiments are reported using a database acquired with three different fingerprint sensors, one with sweeping thermal and two with optical technology. We have also used an automatic quality assessment software which computes the quality of a given fingerprint based on their extracted minutiae. We have observed that the overall quality of the dataset acquired with the thermal sweeping sensor is higher than the quality of the datasets acquired with the two optical sensors, although it is known that sweeping sensors usually produces errors and spurious artifacts due to its acquisition process [21].

The minutiae- matcher performs better than the ridge-based matcher for all the datasets. Sensor interoperability experiments show that when matching images from different sensors, the performance drops dramatically for both matchers. This problem outlines the importance of system development and benchmarking using different and heterogeneous data.

Regarding sensor fusion, we have observed for both matchers that fusing scores from different sensors results in better performance than fusing different instances from the same sensor, revealing the complementarity between different sensors. Moreover, the best relative improvement is obtained when fusing scores from sensors with different technology, revealing another source of complementarity. Also worth noting, the highest relative improvements are always obtained with the minutiae-based matcher. This should be because minutiae-based matchers are strongly dependent on image morphology and quality thus more complementarity information is provided by different sensors.

ACKNOWLEDGMENTS

This work has been supported by BioSecure NoE and the TIC2003-08382-C05-01 project of the Spanish Ministry of Science and Technology. F. A.-F. and J. F.-A. are supported by a FPI scholarship from Comunidad de Madrid. Authors want to thank to L.-M. Muñoz-Serrano for valuable system development.

REFERENCES

[1] A.K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4–20, January 2004.

[2] A. Ross, A.K. Jain, and J. Reisman, "A hybrid fingerprint matcher," *Pattern Recognition*, vol. 36, no. 7, pp. 1661–1673, July 2003.

[3] J. Fierrez-Aguilar, L. Nanni, J. Ortega-Garcia, R. Capelli, and D. Maltoni, "Combining multiple matchers for fingerprint verification: A case study in FVC2004," *Proc. ICIAP, Springer LNCS 3617*, pp. 1035–1042, 2005.

[4] G.L. Marcialis and F. Roli, "Fusion of multiple fingerprint matchers by single-layer perceptron with class-separation loss function," *Pattern Recognition Letters*, vol. 26, pp. 1830–1839, 2005.

[5] J. Fierrez-Aguilar, Y. Chen, J. Ortega-Garcia, and A.K. Jain, "Incorporating image quality in multi-algorithm fingerprint verification," *Proc. IAPR Intl. Conf. on Biometrics, ICB*, vol. Springer LNCS-3832, pp. 213–220, 2006.

[6] G.L. Marcialis and F. Roli, "Fingerprint verification by fusion of optical and capacitive sensors," *Pattern Recognition Letters*, vol. 25, pp. 1315–1322, 2004.

[7] A. Ross and A.K. Jain, "Biometric sensor interoperability: A case study in fingerprints," *Proc. Biometric Authentication: ECCV 2004 International Workshop, BioAW 2004 - LNCS 3087*, vol. 3087, pp. 134–145, May 2004.

[8] P. Grother and et al., "Minex - performance and interoperability of the INCITS 378 fingerprint template," *NISTIR 7296 - http://fingerprint.nist.gov/minex*, 2005.

[9] A. Martin, M. Przybocki, G. Doddington, and D. Reynolds, "The NIST speaker recognition evaluation - overview, methodology, systems, results, perspectives," *Speech Communications*, p. 225254, 2000.

[10] F. Alonso-Fernandez, J. Fierrez-Aguilar, and J. Ortega-Garcia, "Sensor interoperability and fusion in signature verification: a case study using tablet pc," *Proc. IWBRS 2005, Springer LNCS-3781*, pp. 180–187, 2005.

[11] "ANSI-INCITS 378, fingerprint minutiae format for data interchange," *American National Standard*, 2004.

[12] A.K. Jain and A. Ross, "Multibiometric systems," *Communications of the ACM, Special Issue on Multimodal Interfaces*, vol. 47, no. 1, pp. 34–40, January 2004.

[13] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On combining classifiers," *IEEE Trans on PAMI*, vol. 20, no. 3, pp. 226–239, March 1998.

[14] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, and J. Bigun, "Discriminative multimodal biometric authentication based on quality measures," *Pattern Recognition*, vol. 38, no. 5, pp. 777–779, 2005.

[15] K.I. Chang, K.W. Bowyer, and P.J. Flynn, "An evaluation of multimodal 2D+3D face biometrics," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 619–624, April 2005.

[16] C.I. Watson, M.D. Garris, E. Tabassi, C.L. Wilson, R.M. McCabe, and S. Janet, *User's Guide to Fingerprint Image Software 2 - NFIS2 (http://fingerprint.nist.gov/NFIS)*, NIST, 2004.

[17] J. Fierrez-Aguilar, L.M. Munoz-Serrano, F. Alonso-Fernandez, and J. Ortega-Garcia, "On the effects of image quality degradation on minutiae- and ridge-based automatic fingerprint recognition," *Proc. IEEE ICCST*, pp. 79–82, 2005.

[18] E. Tabassi and C.L. Wilson, "A novel approach to fingerprint image quality," *Proc. IEEE Intl. Conf. on Image Processing, ICIP*, vol. 2, pp. 37–40, 2005.

[19] A. Ross, K. Reisman, and A.K. Jain, "Fingerprint matching using feature space correlation," *Proc. BioAW, Springer LNCS*, vol. 2359, pp. 48–57, 2002.

[20] L. Hong, Y. Wan, and A.K. Jain, "Fingerprint imagen enhancement: Algorithm and performance evaluation," *IEEE Trans. on PAMI*, vol. 20, no. 8, pp. 777–789, August 1998.

[21] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*, Springer, New York, 2003.