



Fusion of Digital Fingerprint Quality Assessment Metrics

Christophe Rosenberger, Christophe Charrier

► To cite this version:

Christophe Rosenberger, Christophe Charrier. Fusion of Digital Fingerprint Quality Assessment Metrics. 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX), May 2020, Athlone, Ireland. hal-02541010

HAL Id: hal-02541010

<https://hal.science/hal-02541010>

Submitted on 12 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fusion of Digital Fingerprint Quality Assessment Metrics

Christophe Rosenberger and Christophe Charrier

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

christophe.rosenberger@ensicaen.fr, christophe.charrier@unicaen.fr

Abstract—The quality assessment of biometric samples is a crucial issue in biometrics, indeed, many studies showed its significant impact on the subsequent performance of the biometric system. Many metrics have been proposed and studied in the literature in order to quantify their usefulness. In this paper, we propose to merge different metrics in order to improve the utility estimation of the quality assessment. We use the enrollment selection validation approach in order to compute the utility estimation of the fused metrics. We show the efficiency of the proposed approach comparing with 7 well known metrics on the 12 FVC datasets and 5 synthesized SFinGE-based databases with two matching algorithms. Experimental results show a good improvement on the fused metric to better qualify the quality of digital fingerprints. Those results demonstrate the effectiveness of the approach.

Index Terms—Fingerprint, quality assessment, fusion

I. INTRODUCTION

The aim of biometrics is to identify or verify the identity of human being based on morphological characteristics or behavior. Identification and/or verification is a mature technology used in many government and civilian applications such as e-passports, ID cards, and border control. To this end, the biometric system exploits a biometric reference template of the individual constructed from one or more captures of the biometric data. During the past several years, biometric sample quality assessment became a significant issue because of biometric systems' poor performance on degraded samples. Studies and benchmarks have shown that biometric sample quality have a direct influence on the overall performance of a biometric recognition system [3], [9]. Indeed, using a poor quality biometric sample (fingerprint, face, iris, etc.) in the enrollment phase of the subject, the recognition of the person cannot be ensured with a high level of accuracy. In order to guarantee the best performance of the biometric system, it is generally attempted to collect during the enrollment step the best possible quality of the biometric data. Thus, when acquiring a fingerprint for passport, a metric is involved to ensure a sufficient quality level [19].

The quality of a biometric data depends on several factors, including the acquisition conditions (environment, good use of the sensor, ...), the biometric sensor (problem of adjustment, ...) or the biometric data (damaged finger, part of the hidden face, ...). Figure 1 shows some examples of fingerprints with different quality. It is commonly assumed that the performance



Fig. 1. Examples of fingerprints of different quality (normal, dry finger, wet finger, wrinkles)

of a fingerprint comparison algorithm will be different in all of these cases.

In the last decade, several research works have been focused on the definition of biometric quality metrics for face [13], [20], venous network [15] and especially fingerprint [12], [19], [22]. Yet, several problems remain whatever the considered biometric modality. The first one concerns the relevance of quality metrics. Since we address at least two aspects of quality (image and biometrics data), how the relevance of biometric data quality metric can be formulated and measured? The second one concerns the ranking of measures. How to objectively demonstrate the superiority of a quality metric over another?

The first contribution of this paper is to define a methodological framework to quantify the relevance of a quality metric of a biometric data. The main advantage of the proposed approach is to provide a generic framework that can be applied to any biometric modality. The other advantage of the method is its ability to objectively compare two quality metrics. This method also makes it possible to estimate how close a metric is to an optimal judgment. The computed relevance indicator is therefore limited. The second contribution is the application of this methodology to compare main state-of-the-art Fingerprint Quality Assessment (FQA) metrics for 16 datasets.

The paper is organized as follows. Section 2 contains related works concerning the definition of FQA metrics. We present in section 3 the Enrollment Selection (ES) validation approach. Section 4 is dedicated to the comparative study of FQA metrics by applying the ES validation approach. We finally conclude and give some perspectives.

II. RELATED WORKS

Phillips et al. [14] highlighted the importance of evaluating biometric systems. However, some previous studies on biometric data quality assessment do not provide fully open protocols [9]. This widely used method (in particular for the

standardization of quality metrics) is mainly based on the study of the statistical correlation of legitimate scores and the value of the metrics of the samples. This type of approach requires a very large biometric database (several million samples) that is difficult to access for researchers.

Other studies evaluate the benefit of a quality metric with respect to subjective evaluation elements [7]. Ratha and Bolle also suggested to measure the relevance of a biometric quality metric by statistical comparisons (with the Pearson correlation factor) with the values of other quality metrics [16]. Similarly, Shen et al. [17] proposed to assign the fingerprints of a database according to their quality evaluated by a metric and to compare these classes with others generated by other quality metrics. These approaches do not quantify the relationship between the quality metric and the corresponding performance. Moreover, these attempts are more or less linked to subjective observations when validating quality metrics. However, this type of operation could be easily estimated using a synthetic fingerprint generator such as SFinGe [6] that allows us to generate a huge number of synthesized fingerprint images.

Tabassi et al. [18] defined the quality of the biometric sample as a predictor of the recognition performance by observing that good quality biometric samples produce relatively high legitimate scores and well separated from impostor scores. However, the prediction depends entirely on the performance of the used comparison algorithm. Grother et al. [9] have proposed several evaluation approaches associated with the decision threshold levels and corresponding quality, including the Detection Error Tradeoff (DET) curve and the Kolmogorov Smirnov (KS) test approach. These latter approaches are interesting but difficult to interpret (because based on a hypothesis test), not really adapted to quantitatively compare two quality metrics.

The next section introduces a new method addressing these two limitations. This method allow us to estimate the optimal relevance to be reached for a biometric database and a comparison algorithm.

III. PERFORMANCE OF A QUALITY METRIC

The principle of the proposed method for validating a biometric data quality metric is to evaluate its relevance for the enrollment process (which is the most important use of quality metrics). This makes sense as the principal use of a quality metric to optimize the enrollment process. We assume to have a dataset of biometric samples for many users and an algorithm for comparing these samples.

For the performance evaluation of a biometric system on a dataset, one heuristic must be applied to select the sample as reference template for each individual. Once this choice is made, we can calculate the set of legitimate scores by comparing the samples with the reference template from the same user. Similarly, impostor scores are obtained by comparing a reference template with the samples from all other individuals in the dataset, as illustrated in Fig. 2. Thus, if we have a dataset composed of N individuals with M samples

per individual, this process generates $N \times (M - 1)$ legitimate scores and $N \times (M - 1) \times (N - 1)$ impostor scores.

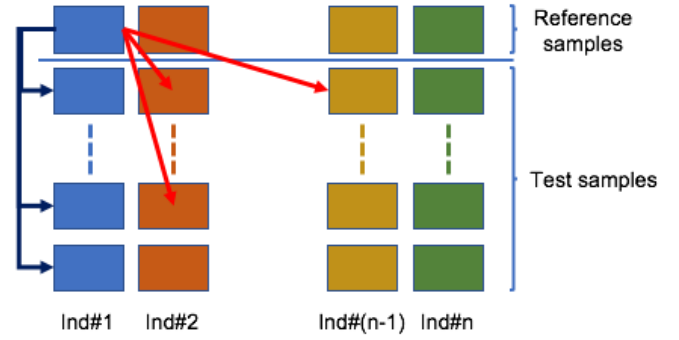


Fig. 2. illustration of the computation of genuine scores (blue solid lines) and impostor scores (red solid lines).

These scores allow us to calculate the False Rejection Rate (FRR) and False Acceptance Rate (FAR) for different values of the decision threshold (from which the proof of verification is considered as sufficient). This yields us to compute different classical measurements in biometrics: 1) the DET curve (evolution of the False Rejection as a function of the False Acceptance Rate), 2) the Equal Error Rate (EER) and 3) the Area Under the DET Curve (AUC).

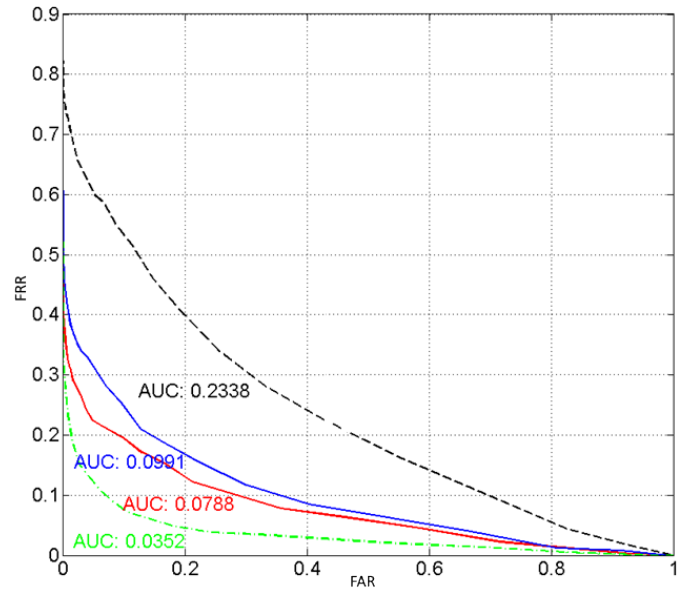


Fig. 3. Representation of performance according to choice of reference template: worst choice (black), best choice (green), choice by metric 1 (blue) and choice by metric 2 (red).

Let us returning now to the choice of the sample as a reference template for one user. In this paper, we mainly consider (but not only) the AUC value since it is a global measure of performance to be minimized. Several possibilities exist:

- Choice of the first sample as the reference of the individual. It is a common practice considered as a default choice,
- Choice of reference considering a quality metric. In this case, the samples are evaluated by the considered metric and the "best" sample (based of this metric) is selected as the user's reference template,
- Choice of the best sample: In general, a biometric database contains a finite and small number of M samples of each individual. One can test all possibilities, *i.e.*, to estimate the value of the AUC obtained for each sample possibility as a reference. It is therefore possible to determine the optimal choice of the reference sample considering the performance of the biometric system (minimum value of the AUC),
- Selection of the worst sample: the previous procedure is repeated by taking the sample leading to the highest value of the AUC.

By applying those different heuristics, one can measure the performance of the biometric system for a biometric dataset and a comparison algorithm. We may calculate AUC_{best} the value of the AUC with the optimal choice of the reference, AUC_{worst} the value of the AUC with the choice of the worst reference and AUC_{metric} the value of the AUC with the choice of reference template guided by a quality metric. We propose to calculate the relevance P of a metric as follows:

$$P = 1 - \frac{(AUC_{metric} - AUC_{best})}{(AUC_{worst} - AUC_{best})} \quad (1)$$

Figure 5 shows the performance obtained on a biometric dataset for different heuristics of choice of the reference. In black, we represented the DET curve taking the worst sample as reference and in green the best choice. This figure shows that the choice of the reference will lead to a performance of the system between an AUC of 0.0352 to 0.2338. Using two metrics, one achieves a performance of 0.0991 (blue) and the other of 0.0788 (red). Two remarks can be formulated:

- 1) metric 1 (blue curve) is less efficient than metric 2 (red curve). Indeed, metric 2 allows a better performance of the biometric system.
- 2) metric 2 has a relevance of $P = 78\%$ against $P = 67\%$ for the other metric. This means that there are still possibilities for improvement to make the best choice of the reference.

A. Fusion of quality metrics

All existing biometrics data quality assessment algorithms fail for at least one to correctly predict the quality of data when degradation, such as rotation, cropping, zoom out, stripes, and so on, is not considered. Degradation failure is not necessarily the same for the trial algorithms. Taking into account those failure cases, a fusion algorithm is proposed to improve the final quality scoring. The used fusion process is a naive one, based on a fusion scoring process. Let $\mathcal{F} = \{f_1, f_2, \dots, f_N\}$ a set of N biometrics data quality assessment algorithms. Let $S(\cdot)$ the scoring process for any biometrics data quality

assessment algorithm. The final score $S(\text{FUSION})$ is then computed as follows:

$$S(\text{FUSION}) = \sum_{i=1}^N S(f_i). \quad (2)$$

IV. EXPERIMENTAL PROTOCOL

In this section, the proposed framework is applied within the fingerprint context, since many quality measures exist in the literature to score the quality of fingerprints. A comparative study of Fingerprint Quality Assessment (FQA) metrics is investigated applying the alluded above methodology.

A. FQA metrics

Fernandez et al. [4] proposed a comparative study of the FQA prior to 2006, in which they categorized FQA algorithms into several classes known as local feature-based approaches, global feature-based methods and solutions with classifiers. The assessing approaches reviewed can be simply summarized in several points: quality metrics based on the orientation of fingerprint pattern; algorithms depending on the variation of Gabor responses; approaches in frequency domain; measurements that are based on pixel information and quality indexes rely on classification with multi-features. In addition, this study also analyzed quality metrics in terms of linear correlation between the metrics.

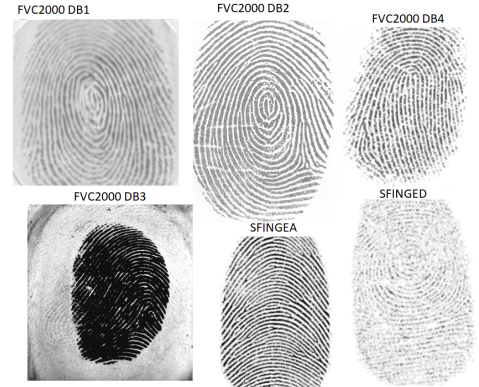


Fig. 4. Examples of fingerprint images from both FVC and SFINGE datasets

In this study, seven state-of-the-art FQA metrics are considered:

- 1) **NFIQ**: This metric involves a 11-dimension feature to estimate a matching score and classify a fingerprint image to five quality levels using neural network [19]. Since 15 years, this metric has been used as the standard FQA metric embedded in all commercial biometric systems API,
- 2) **NFIQ 2.0**: Olsen et al. [2] trained a two-layer self-organizing map (SOM neural network) to obtain a histogram of SOM unit activation with an intensity vector of image block. The histogram is the frequency of the occurrence of the best-matching unit (with respect to the competitive layer) assigned to each block. The

Dataset	NFIQ	NFIQ2	OCL	QMF	NBIS	MSEG	MQF
FVC2000DB1	2	65	0.73	83.81	14.16	0.44	59802
FVC2000DB3	4	40	0.71	28.06	15.11	0.18	29804
SFINGEA	1	69	0.90	76.09	57.46	0.83	55720
SFINGED	3	28	0.47	91.19	10	0.006	43546

TABLE I
VALUE OF THE DIFFERENT FQA METRICS FOR THE FIRST SAMPLE OF DIFFERENT DATASETS

trained feature is then threw to a Random Forest (RF) to estimate the binned genuine matching scores (GMS). This is the first study of FQA to generate a learning-based feature by using unsupervised approach and quite large dataset. However, the quality feature is also a regularity of the proposed histogram, and the RF is used to classify samples in terms of a prior-knowledge of matching score. So far, there is no study which may perform a perfect matching algorithm. Indeed the matching scores between two bad quality genuine or impostor samples are somehow unforeseeable [9]. NFIQ 2.0 is the new ISO standard for FQA.

- 3) **OCL**: Lim et al. [11] proposed a quality metric through weighted combination of local and global quality scores which are estimated in terms of several features such as orientation certainty level (OCL) and so on. Their quality metric also involves in several thresholds to classify the local blocks into variant levels.
- 4) **QMF**: This metric is computed by considering several different aspects 1) the fingerprint image itself (through a blind quality image assessment and textures features) and 2) the associated minutiae template [22]. The quality metric is hence implemented via a linear combination of these quality features.
- 5) **NBIS**: The quality index is a simple FQA metric based on the minutiae quality extracted by the NIST NBIS software [10]. This metric corresponds to the average value of minutiae quality in the template.
- 6) **MSEG**: This metric computation is based upon trimming foreground pixels of bad quality image as much as possible [22]. It generates a quality metric with multiple coarse segmentation. This framework is almost a two-step (or more) work which firstly performs one coarse segmentation to the fingerprint image and followed by another segmentation-like operation for a further pixel-removing. Finally, each of the segmentation results is simply used as a feature, which allows to merge features in segmentation phase.
- 7) **MQF**: It is an original metric as it computes a quality score from a minutiae template associated to the fingerprint [21]. In another word, it could be viewed as a quality metric for assessing the global quality of a minutiae template. The benefit of having this type of metric could be related to embedded biometric systems in smart cards or smart objects where only the minutiae template is available due to computational and storage constraints.

B. Fingerprint datasets

In order to evaluate the performance of the seven trial algorithms and the FUSION one, 17 fingerprint datasets are considered:

- 12 are from the FVC competition [1]:
 - FVC 2000: DB1,DB2,DB3 and DB4
 - FVC 2002: DB1,DB2,DB3 and DB4
 - FVC 2004: DB1,DB2,DB3 and DB4
- Five datasets have been generated using SFINGE [6]:
 - SFINGE0 with different quality images randomly ranging from low quality to very high quality,
 - SFINGEA where images are of very high quality,
 - SFINGEB with high quality images,
 - SFINGEC where images are of medium quality,
 - SFINGED containing low quality images

Figure 4 displays some digital fingerprints from the above mentioned datasets.

C. Minutiae extractor and comparison

In this paper, we used the NIST [10] Mindtct extractor. Two fingerprint comparison algorithms have also been used. The first one is the Bozorth3 algorithm proposed by the NIST [8]. The MCC algorithm [5] is the second comparison one.

V. EXPERIMENTAL RESULTS

Table I shows the quality score obtained from the trial FQA metrics computed on the first sample of four datasets. Note that a low value of NFIQ means a high quality of the fingerprint, whereas for all other metrics, it has to be maximized. We can note that the evaluation results seem efficient, e.g. the best quality fingerprint image for nearly all metrics is obtained from the SFINGEA database which is a very high quality database. One observes that there are some important scores differences from FQA metrics especially between the fingerprint image from the FVC2002DB3 database and the one for SFINGED. Since FQA metrics have values in very different intervals, it is not so easy to compare them, in the following, we use the enrollment selection approach to compare their efficiency.

We computed the P value (Eq. (1)) for all datasets and all trial FQA metrics. Figure 5 presents the DET curves associated to the choice of the biometric sample as reference template for each trial FQA metric. The EER value obtained by using each selection strategy of the reference template is presented. Some differences occur between FQA metrics and none of them are close to the optimal curve meaning there is room for improvement.

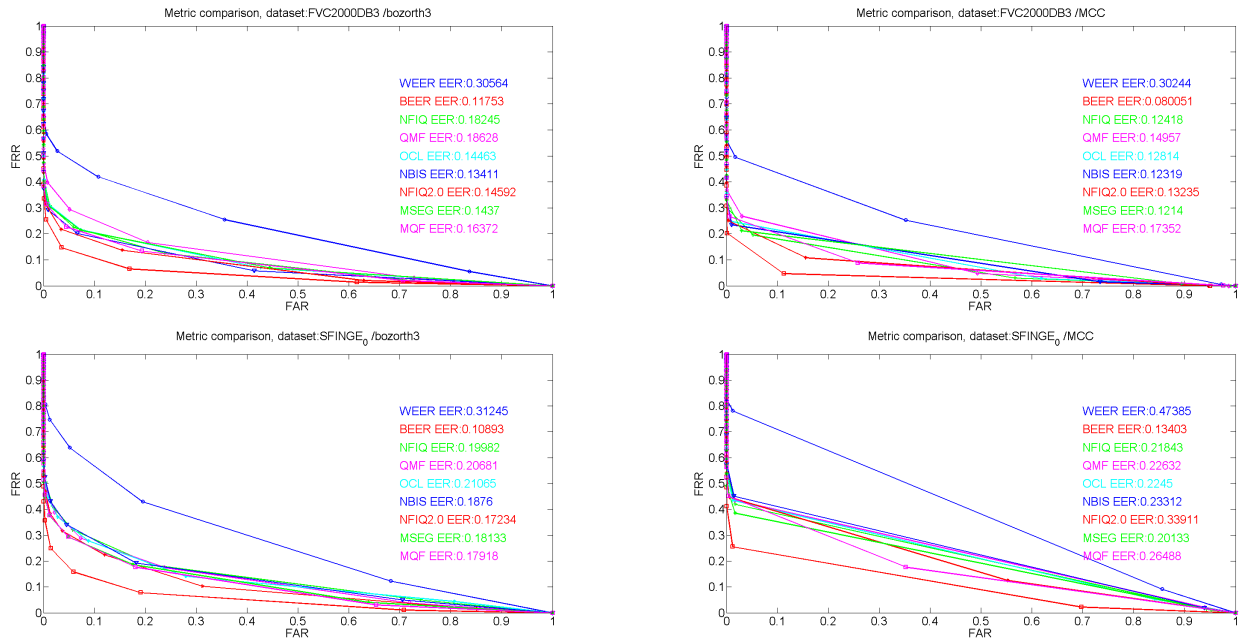


Fig. 5. DET curves for the different FQA metrics for two datasets: FVC2000DB3 and SFINGE0 (left: bozorth3, right: MCC)

Dataset	NFIQ	NFIQ2	OCL	QMF	NBIS	MSEG	MQF	FUSION
FVC2000DB1	71.7%	70.4%	79.4%	72.8%	73.7%	76.5%	71.5%	93.5%
FVC2000DB2	74.8%	82.9%	63.3%	69.4%	80.0%	79.5%	68.9%	87.6%
FVC2000DB3	74.5%	82.8%	72.1 %	61.3%	78.6%	71.4%	77.0%	89.3%
FVC2000DB4	63.7%	69.1%	68.8%	61.5%	56.4%	69.3%	69.0%	100%
FVC2002DB1	82.6%	73%	72.3%	68.2%	72.2%	79.7%	80.2%	83.3%
FVC2002DB2	82.2%	85.4%	85.3%	79.3%	81.4%	91.5%	85.4%	93.7%
FVC2002DB3	82.2%	85.4%	85.2%	79.3%	81.4%	91.4%	85.3%	93.7%
FVC2002DB4	82.6%	73%	72.3%	68.2%	72.2%	79.7%	80.2%	86.9%
FVC2004DB1	82.2%	85.4%	85.3%	79.3%	81.4%	91.5%	85.4%	93.8%
FVC2004DB2	79%	71.9%	71.2%	68.6%	71.6%	77%	69.2%	80.2%
FVC2004DB3	70.9%	87.5%	80.4%	55.2%	76.6%	77.1%	76.1%	90.7%
FVC2004DB4	75.9%	66.5%	68.3%	62.3%	63.5%	64.1%	71.4%	77.1%
SFINGE0	67.0%	67.0%	64.5%	68.0%	66.1%	74.7%	77.0%	77.0%
SFINGEA	87.2%	14.5%	97.3%	63.1%	78.1%	24%	89.4%	100%
SFINGEB	95.2%	68.2%	55.3%	55.3%	63.5%	55.1%	80.4%	95.2%
SFINGEC	85.3%	92.7%	91.6%	100%	76.3%	92.2%	87.8%	100%
SFINGED	68.8%	75.3%	94.6%	71.4%	71.3%	61.2%	70.9%	98.2%

TABLE II

VALUE OF THE P PERFORMANCE OF EACH METRIC ON THE 17 TRIAL BIOMETRIC DATASETS WHEN THE MATCHING BOZORTH3 ALGORITHM IS CONSIDERED.

Table II presents the results obtained using the Bozorth3 matcher. It is noted that the best FQA metric (bold values) is the FUSION scheme. This means that applying a combination of FQA methods, existing individual failure for each FQA scheme can be mitigated. The performance gain is not negligible. The proposed approach reaches a P value higher than 90% for ten databases. The second best FQA algorithm reaching a P value higher than 90% is MSEG (only for four databases). This illustrates the superiority of the proposed approach and validates the concept of fusion of FQA techniques.

Table III the results obtained using the MCC matcher. One can observe a similar behavior of FQA metric than

the one noticed when the Bozorth3 matcher is used. Once, the FUSION algorithm outperforms the trail FQA methods whatever the database is. Furthermore, a P value higher than 90 is obtained for 14 databases whereas the second best FQA method (namely MSEG) reaches such a value for only four databases. This is a relevant property as the intrinsic reliability of the FQA metric to assess the quality of a sample.

We also note that in some cases, the performance of a FQA metric may be poor (such as NFIQ2 for the SFINGEA dataset). It is surprising as these synthetic fingerprints have a high quality. If we consider the average value of the FQA metrics, NFIQ2 provides the best result for real digital fingerprints

Dataset	NFIQ	NFIQ2	OCL	QMF	NBIS	MSEG	MQF	FUSION
FVC2000DB1	79.2%	78.3%	82.6%	81.6%	68.4%	76.5%	82.3%	96.7%
FVC2000DB2	68.7%	79.1%	81.3%	75.5%	72.8%	92.6%	78.6%	96.7%
FVC2000DB3	75.8%	79.3%	67.7%	66.1%	65.6%	61.1%	75.1%	91.8%
FVC2000DB4	71.5%	76.4%	72.6%	67.6%	16.1%	74.8%	69.6%	92.6%
FVC2002DB1	89.5%	78.1%	77.2%	77.7%	72.2%	88.5%	89.5%	88.2%
FVC2002DB2	68.7%	79%	81.3%	75.5%	72.8%	92.6%	78.6%	96.7%
FVC2002DB3	79.2%	78.3%	82.6%	81.6%	68.4%	76.5%	82.3%	92.2%
FVC2002DB4	89.5%	78.1%	77.2%	77.7%	72.3%	88.5%	89.5%	94.9%
FVC2004DB1	79.2%	78.3%	82.6%	81.6%	68.4%	76.5%	82.3%	92.9%
FVC2004DB2	68.7%	79%	81.3%	75.5%	72.8%	92.6%	78.6%	92.6%
FVC2004DB3	89.5%	78.1%	77.2%	77.7%	72.3%	88.5%	89.5%	95.3%
FVC2004DB4	58.4%	56%	55.8%	56.2%	48.6%	53%	64.2%	66.7%
SFINGE0	89.5%	78.1%	77.2%	77.7%	72.3%	88.5%	89.5%	94.2%
SFINGEA	91.5%	0%	42.6%	74.1%	34.5%	58.5%	81.5%	100%
SFINGEB	89.5%	78.1%	77.2%	77.7%	72.3%	88.5%	89.5%	100%
SFINGEC	89.5%	78.1%	77.2%	77.7%	72.3%	88.5%	89.5%	100%
SFINGED	68.7%	59.4%	82.1	72.7%	61.1%	64.7%	67.2%	84.7%

TABLE III

VALUE OF THE P PERFORMANCE OF EACH METRIC ON THE 17 TRIAL BIOMETRIC DATASETS WHEN THE MATCHING MCC ALGORITHM IS CONSIDERED.

(from FVC databases). It is obvious to identify NFIQ2 as a good quality FQA metric (recent standard replacing NFIQ). It could be interesting to test another minutiae extractor in order to consider the behavior of NFIQ2 on this dataset. Nevertheless, considering all the datasets and excluding the proposed approach, MQF provides the best result. It is also surprising as this metric only considers the minutiae template.

VI. CONCLUSION AND PERSPECTIVES

This paper provides a framework to compare seven state-of-the-art FQA metrics from the literature, evaluating their relevance. We use the enrollment selection approach to estimate the reliability of a metric. Results on 17 datasets illustrated the good behavior of two metrics, NFIQ2 on real digital fingerprints and QMF on all datasets. The fusion of FQA metrics has been hypothesized, tested and validated using the proposed framework. The fusion scheme is based on a naive process (a fusion scoring process). The obtained FUSION FQA scheme outperforms all the seven trial FQA algorithms that demonstrate its efficiency. Perspectives of this study concern the fusion process. A weighted combination of all FQA metrics considering the P value can be investigated. It has been shown that, even if the seven trial FQA algorithms provide high values of P , they present high variability of P depending on considered degradation and database. Thus selecting features from each FQA schemes providing highest P values, one should design one algorithm that provides higher P value than any usual FQA schemes.

REFERENCES

- [1] Fingerprint verification competition databases.
- [2] M. Aastrup Olsen, E. Tabassi, A. Makarov, and C. Busch. Self-organizing maps for fingerprint image quality assessment. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 138–145, 2013.
- [3] F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia. Quality measures in biometric systems. *IEEE Security & Privacy*, 10(6):52–62, 2012.
- [4] F. Alonso-Fernandez, J. Fierrez, J. Ortega-Garcia, J. Gonzalez-Rodriguez, H. Fronthaler, K. Kollreider, and J. Bigun. A comparative study of fingerprint image-quality estimation methods. *IEEE Transactions on Information Forensics and Security*, 2(4):734–743, 2007.
- [5] R. Cappelli, M. Ferrara, D. Maltoni, and M. Tistarelli. Mcc: A baseline algorithm for fingerprint verification in fvc-ongoing. In *Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on*, pages 19–23. IEEE, 2010.
- [6] R. Cappelli, D. Maio, and D. Maltoni. Sfinge: an approach to synthetic fingerprint generation. In *International Workshop on Biometric Technologies (BT2004)*, pages 147–154, 2004.
- [7] T. P. Chen, X. Jiang, and W.-Y. Yau. Fingerprint image quality analysis. In *Image Processing, 2004. ICIP'04. 2004 International Conference on*, volume 2, pages 1253–1256. IEEE, 2004.
- [8] P. A. Flanagan. Nist biometric image software (nbis). 2010.
- [9] P. Grother and E. Tabassi. Performance of biometric quality measures. *IEEE transactions on pattern analysis and machine intelligence*, 29(4), 2007.
- [10] K. Ko. User's guide to nist biometric image software (nbis). *NIST Interagency/Internal Report (NISTIR)-7392*, 2007.
- [11] E. Lim, X. Jiang, and W. Yau. Fingerprint quality and validity analysis. In *Image Processing, 2002. Proceedings. 2002 International Conference on*, volume 1, pages I–I. IEEE, 2002.
- [12] X. Liu, M. Pedersen, C. Charrier, P. Bours, and C. Busch. The influence of fingerprint image degradations on the performance of biometric system and quality assessment. In *Biometrics Special Interest Group (BIOSIG), 2016 International Conference of the*, pages 1–6. IEEE, 2016.
- [13] K. Nasrollahi and T. B. Moeslund. Face quality assessment system in video sequences. In *European Workshop on Biometrics and Identity Management*, pages 10–18. Springer, 2008.
- [14] P. J. Phillips, A. Martin, C. L. Wilson, and M. Przybocki. An introduction evaluating biometric systems. *Computer*, 33(2):56–63, 2000.
- [15] H. Qin and M. A. El Yacoubi. Deep representation for finger-vein image quality assessment. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017.
- [16] N. K. Ratha and R. Bolle. *Fingerprint image quality estimation*. IBM TJ Watson Research Center, 1999.
- [17] L. Shen, A. Kot, and W. M. Koo. Quality measures of fingerprint images. In *AVBPA*, pages 266–271. Springer, 2001.
- [18] E. Tabassi and P. Grother. Fingerprint image quality. In *Encyclopedia of Biometrics*, pages 482–490. Springer, 2009.
- [19] E. Tabassi, C. Wilson, and C. Watson. Fingerprint image quality (nfiq). nistir 7151. august 2004, 2011.
- [20] P. Wasiuk, K. B. Raja, R. Ramachandra, and C. Busch. Assessing face image quality for smartphone based face recognition system. In *Biometrics and Forensics (IWFBI), 2017 5th International Workshop on*, pages 1–6. IEEE, 2017.
- [21] Z. Yao, C. Charrier, C. Rosenberger, et al. Quality assessment of fingerprints with minutiae delaunay triangulation. In *Information Systems Security and Privacy (ICISSP), 2015 International Conference on*, pages 315–321. IEEE, 2015.
- [22] Z. Yao, J.-M. Le Bars, C. Charrier, and C. Rosenberger. Fingerprint quality assessment combining blind image quality, texture and minutiae features. In *ICISSP 2015*, 2015.