Utility Validation of a New Fingerprint Quality Metric

Zhigang Yao, Christophe Charrier, Christophe Rosenberger

Utility Validation of a New Fingerprint Quality Metric

Z. Yao, C. Charrier and C. Rosenberger
Université de Caen Basse-Normandie, UMR 6072 GREYC, F-14032 Caen, France
ENSICAEN, UMR 6072 GREYC, F-14050 Caen, France
CNRS, UMR 6072 GREYC, F-14032 Caen, France
Email: christophe.rosenberger@ensicaen.fr

I. Introduction

Fingerprint somehow can be regarded as a relatively full-fledged application in biometrics. The use of this biometric modality is not limited to traditional public security area, but spread into the daily life, smart phone authentication control and e-payment, for instance. However, quality control of biometric sample is still a necessary task due in order to optimize the operational performance. Research works had shown that biometric systems performance could be greatly depressed for those uncontrolled conditions [2]. For fingerprint, the quality can be intuitively described as the clarity of its ridge and valley pattern, noise condition, and the feasibility of feature extraction such as minutiae points. However, intuitively bad quality fingerprint sample might generate high matching result for some cases and vice versa [10]. In this case, many previous research works contributed to validate fingerprint quality metrics in terms of the relation between biometric sample quality and system performance. Chen et al. [7] proposed quality metrics in both frequency domain and spatial domain, and discussed several criteria to evaluate quality metric, such as quality index predicting matching performance. Fernandez et al. [1] compared correlations between several previously proposed quality metrics. Grother et al. [10] discussed quality evaluation methods in terms of detection error trade-off characteristics (DET), the effect of low quality samples rejection rates on improving performance, Kolmogorov Smirnov (KS) statistic, and etc. Mohamad et al. [8] calculated the KS statistics for image-based quality metric of multimodal biometric data.

The study of this paper makes an evaluation of a proposed quality metric generated by using a utility-based multi-features fusion approach [9]. The quality metric involves in several quality criteria, including no-reference image quality metric (NR-IQA) [24], measures derived from texture features, and measure based on the triplet representation of fingerprint minutiae.

The paper is organized as follows. Section 2 presents a state of the art on fingerprint quality. In section 3, we describe the proposed quality metric called Q. Section 4 details experimental results. Conclusion and future works are given in section 5.

II. State of the art on Fingerprint Quality

Quality assessment of fingerprint samples is mostly studied considering feature extraction. Earlier studies considered feature extraction in both at a local and global level [4], [22], [3], [25], [16], [6]. Most of these studies qualifying fingerprint quality in terms of ridge-valley structure clarity, orientation variations, and so on. Studies in that time considered fingerprint quality from image features and gray level contrast, loosely speaking. Details of fingerprint such as minutiae characteristics and singularity points are rarely considered due to the development of this issue in that time. In addition, performance evaluation was taken into account in terms of the comparison between quality metric and manual quality assessment result. Later, NBIS quality algorithm [26] had been proposed, in which fingerprint quality has been defined as a predictor of the further matching performance. The evaluation of fingerprint quality metric had been considered by the interaction between sample quality and biometric system performance [15], [7]. ISO/INCITS then established a standard for biometric quality, among which the quality of a biometric sample is considered in 3 different aspects [1]:

(1) Character, referring to the quality of subject physical features;
(2) Fidelity, which means that how similar is a biometric sample to its reference sample;
(3) Utility, indicating the impact of biometric sample quality on the overall performance of a biometric system.

Mohamad et al. [9] proposed an utility-based quality assessment approach, in which several criteria of quality were linear combined by using an optimization method [14], genetic algorithm (GA). The quality metric proposed in that study shows a relatively ideal result for fingerprint samples with several kinds of alterations. The quality metric carried out in this study is an extension of the one that had just mentioned, and mainly focuses on the original database of fingerprint, and the purpose of evaluating its performance was achieved.

III. Proposed Fingerprint Quality Metric

The quality metric defined in [9] is in the form of

$$Q = \frac{1}{A}\sum_{i=1}^{N}\alpha_i C_i \ , \tag{1}$$

where N is the number of quality criteria C_i $(i = 1, , N)$, α_i is the linear coefficients, and A is a normalization constant. . In this study, the feature involved in generating quality metric includes 3 classes, one is features proposed in [9], the second class is consisted of several texture features extracted directly from fingerprint image, and the third type of feature is derived from the triplet representation of minutiae-based template, $m_i = \{x,\ y,\ \theta\}$, (x, y) is the location of minutia point and θ is the orientation of the minutia point. The framework of the quality metric in this study can be illustrated by figure 1.

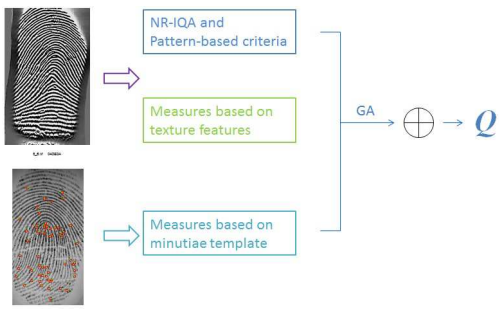

Fig. 1. Illustration of the calculation of quality metric.

A. Features in Previous Works

In the previous work, there were 11 features contributing to obtain the quality metric, including one derived from a NR-IQA algorithm and several image-based features.

1) NR-IQA [24] is an algorithm to evaluate the quality of an image, and it is classified into non-distortion specific approaches. This quality metric generally involves 3 kinds of DCT-based feature, DCT-based contrast feature v_1, DCT-based structure feature v_2, and DCT-based anisotropy orientation features v_3 and v_4;
v_1 is an average value of the contrast of kth DCT path of an image,
v_2 is the a global image kurtosis based on the kurtosis value of each DCT patch,
v_3 and v_4 are variance and max value of the mean value of each DCT patchs Renyi entropy.
Then, a global quality score was calculated by using a multi-scale approach as:

$$BLIINDS = \prod_{i}^{L} v_1^{\alpha_1^i} v_2^{\alpha_2^i} v_3^{\alpha_3^i} v_4^{\alpha_4^i} \ , \tag{2}$$

where α_j^i are calculated by using the correlation of v_i with the subjective notes given by human observers.
2) Salient features based on features extracted by using Scale Invariant Feature Transform (SIFT) operator, which include the number NB of detected keypoints in the image, DC coefficient of the descriptor matrix, mean and standard deviation of the scales related to the keypoints.
3) Patched features, in addition to the blocks number of the image, were computed from the root mean square (RMS) values of each block of the image, including mean value of blocks RMS values, standard deviation of RMSs, median value of RMSs, skewness of RMSs and kurtosis value of them.

B. Texture Features

Texture features are widely used for image classification and retrieval applications. But it is not sure that whether some of them are able to contribute distinctive result for quality assessment of fingerprint image. In this study, 4 classes, 11 texture features were selected as the components for generating quality metric, given in table I.

TABLE I
List of texture features.

Feature	Format	NO.
LBP	256-level LBP histogram vector	1-C1
Four-patch LBP	Descriptor code vector	2-C1
Completed LBP	512-bit 3D joint histogram vector	3-C1
GLCM measures	8-bit GLCM vector	-C2
LBP histogram FT	LBP histogram Fourier transform vector	5-C1
2-S 16-O[1] Gabor	64-bit Gabor response vector	6-C3
4-S 16-O Gabor	128-bit Gabor response vector	7-C3
8-S 16-O Gabor	256-bit Gabor response vector	8-C3
16-S 16-O Gabor	512-bit Gabor response vector	9-C3
LRS	81-bit LRS motif histogram vector	10-C4
Median LBP	256-level MBP histogram	11-C1

[1]. 'S' and 'O': abbreviations of scale and orientation.

1) The first class is local binary pattern (LBP) feature and its extensions or transforms. The LBP feature is proposed by Ojala et al [19], using for image classification. This feature is simple yet efficient so that it is widely used for texture analysis. The idea of LBP operator was that the two-dimensional surface textures can be described by two complementary measures: local spatial pattern and gray scale contrast [21]. Basic LBP operator generates a binary string by thresholding each 3-by-3 neighborhood of every pixel of the image. Figure 2 illustrates the procedure of extracting LBP histogram from fingerprint image.

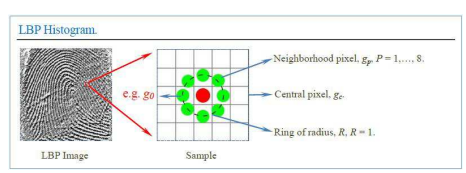

Fig. 2. Illustration of general LBP operator.

The basic operator also extended by many improvements such as uniformed LBP. The transforms of LBP involved in this study include four-patch LBP (FLBP), completed LBP (CLBP), LBP histogram Fourier transform (LBPHFT [18], and median LBP (MLBP) [12]. The

original LBP is a gray-scale invariant operator and its general form (with radius) is rotation invariant. Thus, the general form of basic LBP is robust to changes of both luminance and rotation. Because of this, the general form LBP has been widely adopted in the applications of face modality. CLBP modeling the general form of LBP operator and splits it into 3 components, in which sign components is said to be most informative, and they proposed two forms of combination of the components to obtain more effective performance. FLBP is a patch-based descriptor based on central symmetric patches in different scales around a pixel. This operator considers the similarities between neighboring patches of pixels, and is a complementary of pixel based descriptors. LBPHFT is a rotation invariant image descriptor based on the Fourier transform of ULBP histogram of the original image. This operator considers complex conjugate of the ULBP histogram Fourier transform to deal with the cyclic shift in ULBP histogram caused by image rotation. MLBP is and transform of LBP operator and invariant to monotonic change of gray-scale. Instead of using central pixel, it uses the median value of the local block as the threshold to obtain LBP string.

2) The second class is Haralick features or gray level co-occurrence matrix (GLCM) [13]. Figure 3 presents the calculation of several statistic measures generated from the GLCM matrix which involves in 4 directions combination of neighbor pixels.

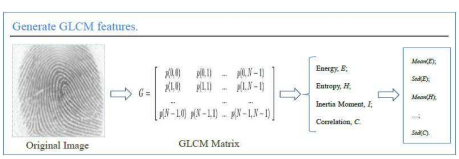

Fig. 3. Illustration of Haralick features.

3) The 2D Gabor function is a sinusoidal function modulated by a Gaussian window. In this case, the basis of Gabor function is complete but not orthogonal. In the last few decades, it has been widely applied to fingerprint image and other biometric data, such as classification and segmentation tasks. Shen et al. [25] proposed using Gabor response to qualify fingerprint image, in which it is said that one or several Gabor features of 8-direction Gabor response are larger than that of the others. Olsen et al. [20] proposed a quality index based on 4-direction Gabor response and it is said that 4-direction is sufficient to qualify fingerprint. However, in this study, it is observed that 2-scale 4-direction Gabor filters do not bring out distinctive regularity for a specified database.

4) The last one is local relational string (LRS) [11] which is an illumination invariant operator, and it reflects variation of local gray level of the image. The operator is based on the local pixels relation in a specified scale, and it uses 3 relations to generate local relation motif histogram for measuring local spatial variations of the image.

C. Measures based Minutiae Template

It can be observed that minutiae are greatly associated with minutiae distributions, locations, and their orientation through the past studies of fingerprint minutiae [5]. In this study, several simple measures based on minutiae locations and orientations were calculated to generate quality metric. Some of these measures were inspired by features for matching operation, and some others were derived from studies of classification. This is due to the consideration of utility of image quality, i.e. features carried out for matching and classification somehow might be associated with fingerprint image quality and one or several distinctive properties of the image. In addition, the qualities of fingerprint images decided by human vision do not fully accord with the practical matching performance. The first kind of measures are based on minutiae numbers and DFT of three components of minutia template. In addition to minutiae number-based features, several statistic-based rotation and translation invariant measures based on minutiae orientations and neighboring minutiae are also calculated [23]. These features were proposed to distinguish fingerprint images of different classes in terms of the relations of both minutiae orientations and locations.

IV. Experimental Results

In order to validate the behavior of the quality metric of this study, a comparative study have been made between the proposed quality metric and a representative quality assessment tool, NFIQ [26].

A. Protocol and Databases

In this study, three FVC databases [17] have been used for experiments. Details on databases are given in table II.

TABLE II
Details of 3 experiment databases.

DB	Sensor Type	Resolution	Image Dim	DB Size
2002DB3A	Capacitive	500dpi	296×560	100×8
2004DB1A	Optical	500dpi	480×640	100×8
2004DB3A	Thermal Sweeping	512dpi	300×480	100×8

Each of these 3 databases involves 100 fingertips, and 8 samples for each fingertip. In this case, the matching scores involved in the experiment have been calculated by using NBIS tool [27], Bozorth3. The intra-class scores contain $1\times7\times100 = 100$ genuine scores, and the inter-class scores are consisted of $1\times7\times99\times100 = 69300$ impostor scores for the whole database. Minutiae template used in the experiment was also extracted by using NBIS tool, MINDTCT. This software generates a quadruple representation of minutia point, $m_i = \{x, y, o, q\}$, where (x, y) is the location of minutia point, o indicates orientation, and q is a quality score of minutia point. In the experiment, just the first 3 components have been used for calculating minutiae-based measures.

B. Evaluation

In the experiments, the EER values of the 3 databases have been calculated using the FVC principles [17], i.e. the first sample of each fingerprint is used as the reference while other seven samples are used as the authentication samples. The results are shown by ROCs in figure 4, in which the EER values of FVC2002DB3A, FVC2004DB1A and FVC2004DB3A are 11.6 %, 12.9%, and 8.5%, respectively.

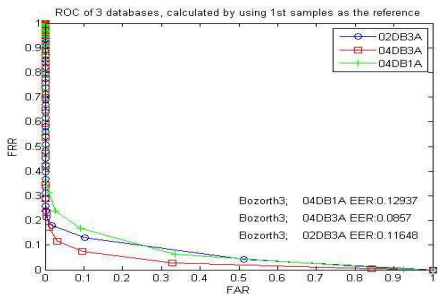

Fig. 4. ROCs of 3 databases.

Second, we computed the EER values of 3 databases by choosing the best samples as the reference (considering the resulting EER values). This is due to the consideration that good quality references should decrease matching error rate. In this case, two kinds of EER values are computed based on NFIQ and proposed approach, respectively. Here, the graphic results of FVC2004DB1A are presented only because the Pearson correlation value of this database is the lowest one among the 3 databases, illustrated by figure 5. The EER value based on NFIQ is 13.2%, and 10.6% corresponds to the proposed approach. For FVC2004DB1A and FVC2004DB3A, the EER values are 14.7% (NFIQ), 14% (proposed), 8.3% (NFIQ), and 6.7% (proposed). Third, the correlation between genuine matching score and the corresponding quality index of each database was calculated, as given in table III.

TABLE III
CORRELATION ANALYSIS BETWEEN SCORES AND QUALITY METRICS

Quality Metric \ Database	02DB3A	04DB1A	04DB3A
NFIQ	-0.2690	-0.2067	-0.2458
Q	0.4600	0.3164	0.5351

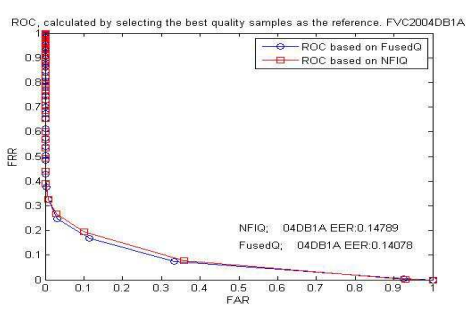

Fig. 5. ROCs and EERs of FVC2004DB1A, calculated by using the best samples as the reference.

It can be observed that NFIQ does not show high numbers for the correlation values. This fact have been mentioned in [1]. In addition, this experiment shows a similar result given by [1], i.e. it is relatively difficult to qualify fingerprint images captured by optical sensors. The proposed metric give much better results (but not perfect).

We used another validation approach defined in [7] to test the proposed quality metric. In order to have a comparative study with the reference algorithm, the discrete quality values of each database have been divided into 5 isometric bins which correspond to the 5 quality labels of NFIQ. Then, the EER values of the divided bins have been calculated. The result is given in figure 6. For NFIQ-based quality values, it is easier to calculate the EER values of the 5 label bins, as depicted in figure 7.

We are able to observe that the matching performances on FVC2002DB3A and FVC2004DB3A were monotonically increased by pruning bad quality samples gradually. NFIQ generated quality levels from 1 to 4 for FVC2002DB3A, and no samples of level 5 were figured out for this database. This might be due to the minutiae points detected on the images of this database, because NFIQ algorithm involves in minutia quality of the fingerprint image. This situation was observed when calculated the correlation between 14 minutiae measure and genuine matching scores in the experiment of this study. It shows a relatively higher correlation on FVC2002DB3A, while the values of two other databases are relatively lower. For FVC2004DB1, both the proposed quality metric and the reference algorithm showed certain difficulties. This is consistent with the past result in [1] and the previous result in this study. Here, just graphic results on FVC2002DB3A are presented only, while the 5 bins' EER values based on proposed approach and NFIQ of FVC2004DB1A and FVC2004DB3A are given in table IV.

V. CONCLUSION

This study evaluated the performance of a multi-feature fusion-based quality metric for fingerprint samples. In the study, the proposed quality metric was evaluated on 3 different FVC databases, FVC2002 DB3 A, FVC2004 DB1 A, and FVC2004 DB3 A. Among the evaluation result, it can be observed that database consisted of images of capacitive sensor is relatively easier to be qualified. This is due to several factors impacted on image quality and matching performance. In addition to external factors such as sensor type and environment, it might be involved in image factors, such as contrast, image size, pixel density, foreground and background area; and correspondingly the factors caused by minutiae template, such as minutiae location, minutiae reliability, and other minutiae properties if they are considered.

Future works of this study will focus on improving the current quality metric and feature processing for the quality metric.

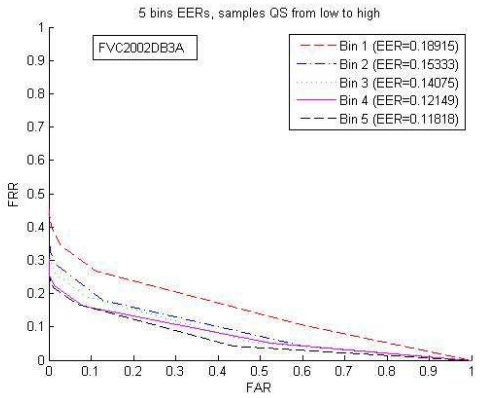

Fig. 6. Monotonic increasing matching perfromance validation of FVC2002DB3A for proposed quality metric, calculated by dividing quality values into 5 isometric bins.

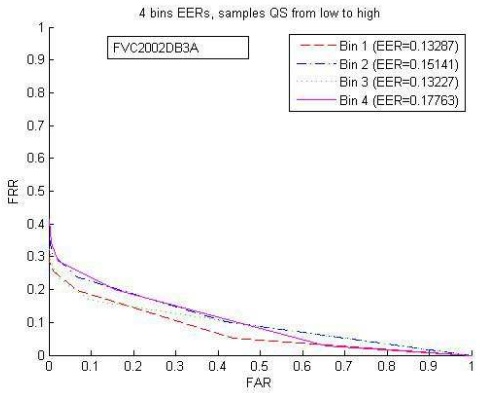

Fig. 7. Monotonic increasing matching perfromance validation of FVC2002DB3A for NFIQ, calculated by dividing quality values into 5 isometric bins.

TABLE IV
5 BINS EER VALUES BASED ON PROPOSED APPROACH AND NFIQ OF FVC2004DB3A.

Bin No.	B1	B2	B3	B4	B5
Q (04DB1)	22.2%	16.6%	17.2%	17.8%	13.3%
NFIQ (04DB1)	15.8%	18.1%	17.7%	23.2%	26.5%
Q (04DB3)	14.2%	8.9%	7.4%	5.8%	4.2%
NFIQ (04DB3)	7.5%	8.1%	13.4%	12.9%	29.8%

REFERENCES

[1] 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," *Information Forensics and Security, IEEE Transactions on*, vol. 2, no. 4, pp. 734–743, 2007.

[2] F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia, "Quality measures in biometric systems," 2011.

[3] A. M. Bazen and S. H. Gerez, "Segmentation of fingerprint images," in *Proc. Workshop on Circuits Systems and Signal Processing (ProRISC 2001)*. Citeseer, 2001, pp. 276–280.

[4] R. M. Bolle, S. U. Pankanti, and Y.-S. Yao, "System and method for determining the quality of fingerprint images," Oct. 5 1999, uS Patent 5,963,656.

[5] J. Chen and Y.-S. Moon, "The statistical modelling of fingerprint minutiae distribution with implications for fingerprint individuality studies," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–7.

[6] T. Chen, X. Jiang, and W. Yau, "Fingerprint image quality analysis," in *Image Processing, 2004. ICIP '04. 2004 International Conference on*, vol. 2, 2004, pp. 1253–1256 Vol.2.

[7] Y. Chen, S. C. Dass, and A. K. Jain, "Fingerprint quality indices for predicting authentication performance," in *Audio-and Video-Based Biometric Person Authentication*. Springer, 2005, pp. 160–170.

[8] M. El-Abed, C. Charrier, and C. Rosenberger, "Quality assessment of image-based biometric information."

[9] M. El Abed, A. Ninassi, C. Charrier, and C. Rosenberger, "Fingerprint quality assessment using a no-reference image quality metric," in *European Signal Processing Conference (EUSIPCO)*, 2013, p. 6.

[10] P. Grother and E. Tabassi, "Performance of biometric quality measures," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 4, pp. 531–543, 2007.

[11] A. Hafiane and B. Zavidovique, "Local relational string for textures classification," in *Image Processing, 2006 IEEE International Conference on*, 2006, pp. 2157–2160.

[12] A. Hafiane, G. Seetharaman, and B. Zavidovique, "Median binary pattern for textures classification," in *Image Analysis and Recognition*. Springer, 2007, pp. 387–398.

[13] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *Systems, Man and Cybernetics, IEEE Transactions on*, no. 6, pp. 610–621, 1973.

[14] C. R. Houck, J. A. Joines, and M. G. Kay, "A genetic algorithm for function optimization: a matlab implementation," *NCSU-IE TR*, vol. 95, no. 09, 1995.

[15] B. Lee, J. Moon, and H. Kim, "A novel measure of fingerprint image quality using the Fourier spectrum," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, A. K. Jain and N. K. Ratha, Eds., vol. 5779, Mar. 2005, pp. 105–112.

[16] E. Lim, X. Jiang, and W. Yau, "Fingerprint quality and validity analysis," in *Image Processing. 2002. Proceedings. 2002 International Conference on*, vol. 1, 2002, pp. I–469–I–472 vol.1.

[17] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain, "Fvc2004: Third fingerprint verification competition," in *Biometric Authentication*. Springer, 2004, pp. 1–7.

[18] L. Nanni, A. Lumini, and S. Brahnam, "Survey on lbp based texture descriptors for image classification," *Expert Systems with Applications*, vol. 39, no. 3, pp. 3634 – 3641, 2012.

[19] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971–987, 2002.

[20] M. Olsen, H. Xu, and C. Busch, "Gabor filters as candidate quality measure for nfiq 2.0," in *Biometrics (ICB), 2012 5th IAPR International Conference on*, 2012, pp. 158–163.

[21] M. Pietikäinen, *Computer vision using local binary patterns*. Springer, 2011, vol. 40.

[22] N. K. Ratha and R. Bolle, *Fingerprint image quality estimation*. IBM TJ Watson Research Center, 1999.

[23] A. A. Ross, J. Shah, and A. K. Jain, "Toward reconstructing fingerprints from minutiae points," in *Defense and Security*. International Society for Optics and Photonics, 2005, pp. 68–80.

[24] M. A. Saad, A. C. Bovik, and C. Charrier, "A dct statistics-based blind image quality index," *Signal Processing Letters, IEEE*, vol. 17, no. 6, pp. 583–586, 2010.

[25] L. Shen, A. Kot, and W. Koo, "Quality measures of fingerprint images," in *IN: PROC. AVBPA, SPRINGER LNCS-2091*, 2001, pp. 266–271.

[26] E. Tabassi, C. Wilson, and C. Watson, "Nist fingerprint image quality," *NIST Res. Rep. NISTIR7151*, 2004.

[27] C. I. Watson, M. D. Garris, E. Tabassi, C. L. Wilson, R. M. Mccabe, S. Janet, and K. Ko, "User's guide to nist biometric image software (nbis)," 2007.