Fuzzy Rule Based Quality Measures for Adaptive Multimodal Biometric Fusion at Operation Time

Madeena Sultana[1], Marina Gavrilova[1] and Svetlana Yanushkevich[2]
[1]*Dept. of Computer Science, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada*
[2]*Dept. of Electrical and Computer Engineering, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada*

Keywords: Adaptive Multimodal Fusion, Fuzzy Quality Measure, Non-intrusive Biometrics, Quality Score Fusion.

Abstract: Sample quality variation at operation time is one of the major concerns of real time biometric authentication and surveillance systems. Quality deviations of samples affect the performance of many benchmark biometric trait recognition systems. Moreover, large variation between enrolled and probe samples is very uncertain since it may arise at operation time for various reasons. In this paper, a novel adaptive multimodal biometric system is presented that can adapt the uncertainty of the quality degradation during operation. Fuzzy rule based method is applied for the first time to calculate the quality score of template-probe pairs dynamically. Feature extraction is accomplished using a novel shift invariant multi-resolution fusion approach. Finally, face and ear modalities are fused adaptively at rank level based on the quality scores. Proposed method relies more on good quality samples and disregards misclassification of poor quality samples. Experimental results demonstrate significant performance improvement of the proposed adaptive multimodal approach over baseline i.e. non-adaptive method.

1 INTRODUCTION

Person identification or authentication is a basic requirement of preventing the adverse effects of growing security threats all over the world. From smart phone to immigration system, person is needed to be identified to get the right access of the right information or service. Although passwords remain the most common mechanism of person authentication, reports on security of traditional password based systems point out how easy it is nowadays to break majority of "strong" passwords (Monwar and Gavrilova, 2009). Moreover, passwords or tokens could be either forgotten, stolen, or lost. To overcome these drawbacks, biometric-based authentication systems have an increasing demand for many security applications (Jain and Kumar, 2012). Scientific research revealed that multimodal biometric provides a higher recognition accuracy over single biometrics (Ross et al., 2006; Bhanu and Govindaraju, 2011).

However, biometric trait recognition is still a challenging problem due to large variations between enrolled and probe samples (Yampolskiy and Gavrilova, 2012; Sultana and Gavrilova, 2013). Sample variations mostly occur during the acquisition time for several reasons such as lightning variation, camera movement, pose variation, mood of the subject, clothing and accessories, human-sensor interaction, multiple acquisition devices, image capturing distance, quality of the sensor or acquisition equipment etc. Performance of a biometric system may compromise significantly for the presence of large variation between enrolled and probe images (Poh et al., 2012; Sultana et al., 2014). Most importantly, the aforementioned factors are very uncertain and may arise any time after deploying the biometric authentication system. For example, large variations in lightning at different time of the day may be observed if a camera is placed in a glass surrounded room or corridor. In such case, biometric samples such as face, ear etc. may suffer a significant amount of lightning varioations if acquired at different day times or from different viewpoints. Another very common problem of surveillance systems is low resolution images are captured from long distances (Marciniak, 2013). A novel multimodal biometric system capable of adapting uncertain illumination and resolution degradation is presented in this paper to address these issues. Our fuzzy inference system can measure the degree of quality deviation of samples

Sultana M., Gavrilova M. and Yanushkevich S..
Fuzzy Rule Based Quality Measures for Adaptive Multimodal Biometric Fusion at Operation Time.
DOI: 10.5220/0005126301460152
In *Proceedings of the International Conference on Fuzzy Computation Theory and Applications* (FCTA-2014), pages 146-152
ISBN: 978-989-758-053-6

during operation. Proposed fuzzy quality scores are fused adaptively in a multimodal system to enhance the confidence of the classifier of good sample. For example, if a face biometric sample receives bad quality score and ear sample obtains good quality score then the person would be identified mostly relying on ear rather than both. In this way, the overall recognition performance would be maximized and the proposed multimodal biometric system would be fully adaptive to resolution and illumination changes at operational time.

2 RELATED WORK

Multimodal biometric traits maximize the recognition performance over single biometrics by reinforcing one another's confidence level (Ross et al., 2006). It also extends users acceptance allowing alternate traits for being recognized. A number of state-of-the-art multimodal architectures, sources, learning and fusion strategies, and novel research directions have been summarized by Gavrilova and Monwar (2013). Multimodality remained a hot topic of research over last couple of years due to growing demand of security applications and surveillance systems. However, performance of the conventional multimodal systems may compromise significantly if the quality of acquired biometric samples degrades at operation time.

Quality degradation have always been considered as an image restoration problem and attempted to be solved by various preprocessing techniques. For example, the most common approach of illumination quality enhancement is preprocessing all enrolled and probe images using a blind normalization. However, blind normalization of biometric samples may degrade the quality of good samples and not all normalization methods perform equally well at different degree of illumination change (Sellahewa and Jassim, 2010). Therefore, conventional pre-processing approaches only deal with predefined problems and unable to adapt uncertain degradation of samples during operation time. The concept of quality adaptive biometrics has been emerged lately to address such uncertain issues. The aim of adaptive biometric systems is to adapt the variations in samples observed during operation without compromising the overall performance of the system (Poh, and J. Kittler, 2012, Fernandez et al., 2010). However, Jain and Kumar (2012) mentioned in their book chapter, "It is not easy to design adaptive multimodal biometrics systems that are flexible enough to consider user preference for biometric modalities, user constraints, and/or varying biometric image quality." Moreover, quality fusion in multimodal system itself is very challenging because of the multi-faceted data, different quality measures, system dependency, different application scenarios etc. (Nandakumar et al., 2006; Abaza and Ross, 2009). Most of the existing quality score based biometric systems are designed for fingerprint and iris. Dong et al. (Dong et al., 2009) reported that quality adaptive approach for selecting the decision threshold of iris recognition system improves recognition performance. Performance improvement of score level and rank level quality score fusion of fingerprints have been studied by Nandakumar et al. (2006) and Abaza and Ross (2009), respectively.

Nowadays, non-intrusive biometric traits such as face, ear etc. have increasing demand to enhance the security in public sectors (Sultana et al, 20104, Kumar and Wu, 2012). Illumination variation and low resolution at operation time are two of the most common causes of quality deviation of non-intrusive biometrics (Marciniak, 2013; Sellahewa and Jassim, 2010). Therefore, in this paper we are proposing a novel non-intrusive multimodal system using face and ear that is capable of adapting uncertain lightning and resolution change. Proposed fuzzy rule based inference system overcomes the problems of multi-faceted and multi ranged data of different quality measures. In addition, it makes our system extendable for more quality measures being integrated.

3 METHODOLOGY

An adaptive multimodal biometric system is proposed in this article, which is capable of adapting quality degradation of samples at operation time. In other words, the performance of the proposed biometric recognition system will not be compromised due to quality degradation of the acquired samples up to a certain level. We developed a novel fuzzy multi-modal inference system is to assign quality scores on acquired (probe) samples according to the degree of deviation. Initially, face and ear biometrics are identified separately as unimodal systems. Our newly developed shift invariant multi-resolution Fusion (MRF) approach is applied for feature extraction from each modality. Finally, ranks of each modality are fused adaptively based on the quality scores of the samples instead of using predefined weights. Proposed adaptive multimodal fusion improves the confidence of a good quality sample and degrades the confidence of the bad quality sample. The former improves the genuine recognition rate while the latter degrades the false acceptance rate of the system. The proposed method has three important phases: fuzzy quality measure, unimodal trait identification, and adaptive multimodal fusion. The detailed descriptions of the three major stages are presented hereafter.

3.1 Fuzzy Quality Measure

In real scenarios of non-intrusive biometric recognition systems, enrolled or template images are mostly obtained under uniform lightning conditions with good resolution. Image quality drastically varies at operation time due to uncertain factors. Therefore, in the proposed method, quality of the probe image is measured with respect to the average quality of enrolled samples. In this work, we are interested in measuring illumination and resolution quality of the probe image. The Illumination Quality (IQ) of the probe image is measured as luminance distortion in comparison to a reference image. The average of all enrolled images is considered as the reference image. Luminance distortion is calculated using the following equation of Wang and Bovik's universal quality measure (Wang and Bovik, 2002).

$$IQ = \frac{2\overline{x}\,\overline{y}}{(\overline{x})^2 + (\overline{y})^2}, \tag{1}$$

where $x=\{x_i \mid i=1,2,\ldots,N\}$ is the probe image and $y=\{y_i \mid i=1,2,\ldots,N\}$ is the reference image. $\overline{x}$ and $\overline{y}$ are the average intensity of probe image and reference image.

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \quad \overline{y} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2}$$

The value of luminance distortion ranges from 0 to 1. The authors of (Marciniak, 2013) demonstrated that low resolution has serious impact on the performance of biometric recognition. Therefore, we also measured the ratio of resolution degradation of probe image compared to enrolled images. For this purpose, the average resolution of enrolled images is considered as reference. Resolution degradation is calculated as the ratio of the size of probe image and the reference. If the resolution of the probe image is better than the reference then the ratio is considered as 1.

A fuzzy inference system is developed to measure the degree of quality deviation of the probe image based on the quality scores of illumination and resolution. We defined four fuzzy linguistic variables, extremely bad, bad, moderate, and good for illumination quality (IQ) as follows:

$$\begin{cases} good, & IQ \ge 0.85 \\ moderate, & 0.85 > IQ > .6 \\ bad, & 0.6 \ge IQ > 0.4 \\ extremely\_bad, & IQ \le 0.4 \end{cases} \tag{3}$$

Four fuzzy linguistic variables for Resolution Quality (RQ) are defined as follows:

$$\begin{cases} high, & RQ > 0.6 \\ medium, & 0.6 \ge RQ > .4 \\ low, & 0.4 \ge RQ \ge 0.1 \\ extremely\_low, & RQ < 0.1 \end{cases} \tag{4}$$

We defined eight fuzzy rules to combine the illumination quality (IQ) and the resolution quality into a single Fuzzy Quality (FQ) score. The four linguistic variables of our final fuzzy quality score are extremely poor, poor, average, and excellent. Fig. 1 shows the rules of our fuzzy inference system. The flow diagram of the proposed quality measure method is depicted in Fig. 2.

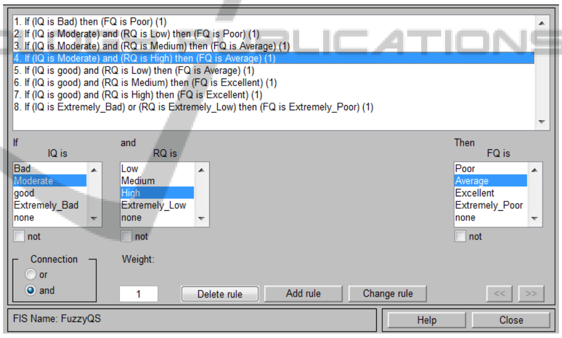


Figure 1: Fuzzy rules to measure the quality of biometric samples.

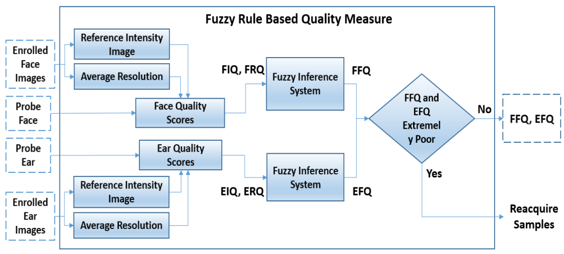


Figure 2: Flow diagram of the proposed quality measure.

At first, the RQ score is calculated for each probe image. Then the probe image is resized to the size of reference image and IQ score is measured. Next, IQ and RQ scores of face and ear are computed and fed to the fuzzy inference system. Finally, fuzzy quality scores of face (FFQ) and ear (EFQ) are computed based on fuzzy rules. The membership function of FQ assigns a singleton score zero if the quality of any modality is extremely poor. Therefore, if any of the modalities (face or ear) has extremely poor quality, recognition would entirely be determined based on the other modality. If both of the modalities have extremely poor quality, both would be rejected and needed to be reacquired. Thus, the proposed method enhances user acceptability by reducing the number of reacquisition of biometric samples as long as the quality of one modality remains within an acceptable range.

3.2 Unimodal Trait Identification

In the proposed method, face and ear traits are identified independently. Fig. 3 shows the block diagram of unimodal face and ear enrolment and recognition process. We introduce feature extraction method called multi-resolution fusion (MRF) to extract features from face and ear. Shift invariance and computational efficiency are the two major advantages of MRF feature extraction method.

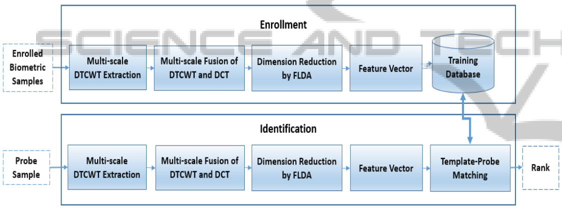


Figure 3: Flow diagram of unimodal trait (face/ear) enrollment and identification.

In MRF, the multi-scale property of 2D Dual-Tree Wavelet Transform (DTCWT) (Kingsbury, 1998) has been utilized to enhance the accuracy and robustness of biometric recognition system. The first step of MRF method is to apply 2D DTCWT on face/ear image until 4^{th} level, which generates a series of different scale subband images. The magnitude subimages of DTCWT provide accurate measures of energy and are approximately shift invariant. Therefore, magnitude images of 24 complex bandpass subimages are computed as multi-scale feature set. The lowpass real image from 4^{th} level is also included in the feature set since it contains large amount of information of the input image. However, the main obstacle of using the multi-scale property is the high dimensional feature set. DTCWT of an image of size 128×128 until 4^{th} scale produces a feature vector of size 32640, which is computationally very expensive. This problem is overcome by fusing 2D Discrete Cosine Transform (DCT) to obtain decorrelated features from each scale. Therefore, the second step of MRF method is to compute 2D DCT coefficients from 24 multi-scale bandpass and one lowpass subimages. Then 8×8 DCT coefficients from the upper left corner of each coefficient matrix of every subband are extracted. Each matrix is converted to a row vector of size 64 and all are concatenated to form a feature descriptor of each image. This process allows extracting the most informative coefficients from each subband. Lastly, Fisher's Linear Discriminant (FLD) is applied to the above feature vector to reduce redundancy and extract a small sized feature descriptor containing the most discriminative biometric information. Therefore, finally, a shift invariant, non-redundant, and computationally efficient feature vector is formed by applying MRF method on each modality. The feature vectors of enrolled images are stored in training database. Face and ear modality of each person is then separately matched against the enrolled face and ear vectors. Euclidian distance is used to compute the ranked similarity score of each modality. Ranks of each modality accompanying with corresponding fuzzy quality scores are then fused adaptively to obtain the final recognition result.

3.3 Adaptive Multimodal Fusion

Final decision is obtained by adaptively fusing face and ear biometrics along with corresponding quality scores at rank level. In the proposed system, a variant of the Borda count method (Ross et al., 2006) has been exploited to derive the final ranks. Traditional Borda count method calculates final ranks by summing up all ranks from independent matchers and assumes that all matchers perform equally. In the proposed method, we computed the weighted sum from face and ear matchers in lieu of adding ranks of each modality. The weights are the fuzzy quality (FQ) scores calculated by the proposed fuzzy inference system. Fig. 4 shows an example of the proposed adaptive multimodal fusion of ranks and FQ scores. In this example, the ranks of person three are 1 and 2 from face and ear matcher, respectively.

The quality of face sample (FFQ) is high whereas the quality of ear sample (EFQ) is moderate. Therefore, the face matcher should produce more reliable result than the ear matcher should. In the proposed adaptive multimodal fusion, ranks of face matcher are weighted with high score and ranks of ear matcher are weighted with moderate score. Thus, the final rank of person 3 using adaptively weighted Borda count method is 1. In this way, the proposed adaptive multimodal system can obtain maximum recognition rate with low quality samples at operation time.

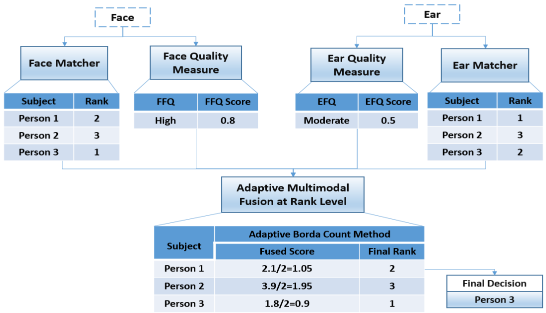


Figure 4: Adaptive multimodal fusion at rank level.

4 EXPERIMENTAL RESULTS

Experimentation is conducted to evaluate the performance of the proposed adaptive multimodal biometric recognition system. Due to lack of existing publicly available multimodal dataset containing face and ear, we have created our own multimodal virtual database to conduct the experiments. We have used FERET (Phillips et al., 2000) and USTB-II (USTB Ear Database, 2014), two most widely used publicly available biometric databases, to create virtual multimodal database. FERET consists of 14051 eight-bit grayscale face images of resolution 256×384 of 1199 individuals. Images samples have different facial expressions, poses, elevations, and illumination conditions. In this study, frontal faces with different expressions and illuminations are considered. USTB database II contains total 308 ear images of size 300×400 pixels of 77 subjects. The first and the fourth samples having different illuminations are used in our study. The goal of our experimentation is to demonstrate the adaptiveness of the proposed method under uncertain resolution and illumination distortions. Therefore, a training database is created using mostly good and uniform samples whereas four test databases are created containing large variations of size and illumination of samples to provide reasonable amount of quality difference during operation.

Training DB: The first profile images of 77 subjects are selected from USTB II. Corresponding face images of 77 subjects are randomly picked up from FERET database having uniform illumination. The two reference images are created by averaging the intensities of all face and ear images in training database. The reference resolution of face and ear samples are 128×128 and 128×192, respectively.

***Good-FaceDB*:** Single samples under different lightning conditions of the 77 subjects are selected from FERET database. Illumination qualities of the images mostly fall into 'moderate' category and resolution is same as of the training samples.

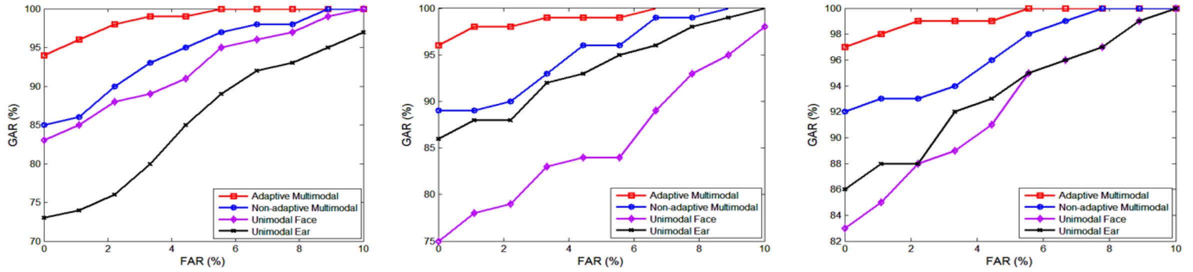
***Poor-FaceDB*:** This database is created by randomly downsampling all the samples of Good-FaceDB four to ten times. Therefore, it comprises of samples having both illumination and resolution distortions.

***Good-EarDB*:** The fourth image of each of the 77 subjects from USTB-II database is selected for this database.

***Poor-EarDB*:** For this database, all images of Good-EarDB are accumulated and downsampled four to ten times randomly.

Three sets of experiments have been conducted to evaluate the recognition performance of unimodal and multimodal systems. Each set of experiment has been conducted five times using different set of 77 face samples from FERET database. All experiments were carried out on Windows 7 operating system, 2.7 GHz Quad-Core Intel Core i7 processor with 16GB RAM using Matlab R2013a. Experimental results are represented by plotting the Receiver Operating Characteristic (ROC) curves of unimodal face, unimodal ear, non-adaptive multimodal, and the proposed adaptive multimodal methods. During the first experiment, the test set comprises of 'good' faces and 'poor' ears. The ROC curves of this experiment are plotted in Fig. 5 (a). Fig. 5 (a) shows that 'poor' quality of ear samples could not degrade the performance of the proposed adaptive approach. The proposed method obtained 94% Genuine Acceptance Rate (GAR) at 0.1% False Acceptance Rate (FAR) whereas unimodal face, ear, and non-adaptive multimodal systems have only 83%, 73%, and 85% GAR, respectively.

For the second experiment, the test set is created using 'poor' faces and 'good' ears. ROC curves of this experiment are plotted in Fig 5 (b). In Fig. 5 (b), similar performance improvement of the proposed method is observed for 'poor' face and 'good' ear quality over unimodal and non-adaptive multimodal systems. The third experiment is conducted using test set containing 'good' faces and 'good' ears. ROC curves of this experiment are shown in Fig 5 (c). A performance boost of the proposed method over unimodal and non-adaptive multimodal systems is observed even for good face and ear samples. This is because of the proposed system is adaptively weighting the corresponding modality based on its illumination condition. However, quality deviation of one modality significantly affects the performance of unimodal and non-adaptive multimodal biometric systems. Results also show that the



(a) Good face and poor ear (b) Poor face and good ear (c) Good face and good ear

Figure 5: ROC curves of different biometric systems for a) GOOD face and POOR ear recognition, b) POOR face and GOOD ear recognition, and c) GOOD face and GOOD ear recognition.

proposed method is capable of adapting uncertain quality deviation.

5 CONCLUSIONS

This paper proposes to consider biometric sample quality variation during operation time to improve biometric authentication. To accomplish this goal, a novel adaptive multimodal biometric system using fuzzy quality scores is presented. Proposed adaptive fusion scheme strengthens the confidence of good samples and reduces misclassification due to poor samples. Therefore, the issue of performance degradation for poor samples during operation has been overcome. Experimental results show that significant quality distortion of one modality has no impact on the overall performance of our system. Comparative analysis to non-adaptive multimodal and unimodal approaches demonstrates the superiority of the proposed method with poor quality samples. Future research will look into incorporating more quality factors and higher granularity quality classification to improve the recognition rate further.

ACKNOWLEDGEMENTS

Authors would like to thank NSERC, NSERC Vanier CGS, and URGC Seed grant for partial support of this project.

REFERENCES

Abaza, A., Ross, A., 2009. Quality based rank-level fusion in multibiometric systems. *BTAS*, 1-6.

Bhanu, B., V. Govindaraju, V. (Eds), 2011. *Multibiometrics for Human Identification*. Cambridge University Press.

Dong, W., Sun, Z., Tan, T., Wei, Z., 2009. Quality-based dynamic threshold for iris matching. *ICIP*, 1949-1952.

Fernandez, F. A., Fierrez, J., Ramos, D., Rodriguez, J. G., 2010. Quality-based conditional processing in multi-biometrics: application to sensor interoperability. *IEEE SMCA,* 40(6), 1168-1179.

Gavrilova, M. L., Monwar, M., 2013. *Multimodal biometrics and intelligent image processing for security systems*. IGI Global.

Jain, A. K., Kumar, A., 2012. Biometric recognition: An overview. *Second Generation Biometrics: The Ethical, Legal and Social Context*, 49-79. Springer.

Kingsbury, N. G., 1998. The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement. *EUSIPCO*, 98, 319-322.

Kumar, A., Wu, C., 2012. Automated human identification using ear imaging. *Pattern Recognit.*, 45(3), 956-968.

Marciniak, T., Chmielewska, A. Weychan, R., Parzych, M., Dabrowski, A., 2013. Influence of low resolution of images on reliability of face detection and recognition. *Multimedia Tools and Applications*, 1-21.

Monwar, Md. M., Gavrilova, M. L., 2009. Multimodal biometric system using rank-level fusion approach. *IEEE Trans. SMCB*, 39(4), 867–878.

Nandakumar, K., Chen, Y., Jain, A. K., Dass, S. C., 2006. Quality-based score level fusion in multibiometric systems. *ICPR*, 4, 473-476.

Phillips, P. J., Moon, H., Rauss, P. J., Rizvi, S., 2000. The FERET evaluation methodology for face recognition algorithms. *IEEE TPAMI*, 22(10), 1090–1104.

Poh, N., Kittler, J., 2010. A unified framework for biometric expert fusion incorporating quality measures. *IEEE TPAMI*, 34(1), 3-18.

Poh, N., Rattani, A., Roli, F., 2012. Critical analysis of adaptive biometric systems. Biometrics, *IET*, 1(4), 179-187.

Ross, A., Nandakumar, K., Jain, A. K., 2006. *Handbook of Multibiometrics*, New York: Springer-Verlag.

Sellahewa, H, Jassim, S. A., 2010. Image-quality-based adaptive face recognition. *IEEE TIM*, 59(4), 805-813.

Sultana, M., Gavrilova, M., 2013. A content based feature combination method for face recognition. *CORES*, 197-206.

Sultana, M., Gavrilova, M., Yanushkevich, S., 2014. Expression, pose, and illumination invariant face

recognition using lower order pseudo Zernike moments, *VISAPP*, 1, 216-221.
USTB Ear Database, available at http://www.ustb.edu.cn/resb/ Accessed on August 13, 2014.
Wang, Z., Bovik, A. C., 2002. A universal image quality index. *IEEE Signal Processing Letter*, 9(3), 81–84.
Yampolskiy, R., Gavrilova, M., 2012. Artimetrics: Biometrics for artificial entities. *IEEE Robot. Autom. Mag.*, 19(4), 48-58.