PAPER • **OPEN ACCESS**

Enhancement of Individuality Representation for Multi-Biometric Identification

To cite this article: W Y Leng *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **884** 012061

View the article online for updates and enhancements.

You may also like

- Biometric Blockchain: A Better Solution for the Security and Trust of Food Logistics
Bing Xu, Tobechukwu Agbele and Richard Jiang
- Review of modern biometric user authentication and their development prospects
Z V Boriev, S S Sokolov and A P Nyrkov
- Study for Integration of Multi Modal Biometric Personal Identification Using Heart Rate Variability (HRV) Parameter
Priatna Ahmad Budiman, Teni Tresnawati, Ahmad Tossin Alamsyah et al.

This content was downloaded from IP address 161.139.222.41 on 03/11/2021 at 01:10

Enhancement of Individuality Representation for Multi-Biometric Identification

W Y Leng[1,3], **S M Shamsuddin**[2] **and S Sulaiman**[2,3]
[1] School of Computing, Faculty Engineering, Universiti Teknologi Malaysia, 81300 Skudai JB, Malaysia
[2] UTM Big Data Centre, Universiti Teknologi Malaysia, 81300 Skudai JB, Malaysia
[3] Soft Computing Research Group (SCRG), Universiti Teknologi Malaysia, 81300 Skudai JB, Malaysia

E-mail: nureiliyah@utm.my

Abstract. Personal identification is one of the areas in pattern recognition that has created a center of attention by many researchers to work in. Recently, its focal point is in forensic investigation and biometric identification as such the physical (i.e., iris, fingerprint) and behavioural (i.e., signature) style can be used as biometric features for authenticating an individual. In this study, an improved approach of presenting biometric features of true individual from multi-form of biometric images is presented. The discriminability of the features is proposed by discretizing the extracted features of each person using improved Biometric Feature Discretization (BFD). BFD is introduced for features perseverance to obtain better individual representations and discriminations without the use of normalization. Our experiments have revealed that by using the proposed improved BFD in Multi-Biometric System, the individual identification is significantly increased with an average identification rate of 98%.

1. Introduction

Due to increasing level of fraud and security threats, the requirement for more secure personal identification technologies is becoming apparent. In recent years, Uni-biometric identification has seen considerable improvements in reliability and accuracy, with some of the traits offering good performance. However, even the best biometric traits till date are facing numerous problems; some of them are inherent to the technology itself. Generally, a Uni-biometric system is composed of four main elements namely sensor, feature extraction, matching and decision making elements. The Uni-biometric identification systems highly depend on the sensor used and the features extracted from the biometric traits. If the sensed signal from biometric features is being destructed (for instance, fingerprint with scar or ink and voice with background noise), the decision made by decision making element based on matching score may not be accurate. Furthermore, the insufficient performance of Uni-biometric system is also due to non-universal biometric traits, easy susceptibility to biometric spoofing (Choras, 2019). One solution to cope with the issues is the use of more than one biometric trait. A system that uses more than one biometric trait (physical and behaviour) is referred to as Multimodal biometric system. It combines multiple sources from different biometric traits, enabling a user who does not possess a particular biometric identifier to still enrol and authenticate using other traits, thus eliminating the enrolment problems and making it universal.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

Some work on multimodal-biometric identification has already been reported beginning as early as 1997 and 1998, especially focuses on fusing at decision level by combining the weak classifiers in order to increase the overall performance of traditional biometric system (Wang et al., 2014). Damousis and Argyropoulos (2012) used two biometric traits; face and voice as a form of user's identity for recognition process. The authors examined the efficiency of identification by using four machine learning algorithms namely Gaussian Mixture Models (GMMs), Artificial Neural Networks (ANNs), Fuzzy Expert Systems (FESs), and Support Vector Machines (SVMs). Final comparison showed that the classification performance of the algorithms provided higher recognition than Uni-biometric. Besides on multi-classifiers of biometric system, much effort has also been carried out to fuse multiple modalities at different levels. For instance, score fusion at decision level by Turki in 2016 integrated face and palm print for biometric identification. They compared different levels of fusion schemes and the best results was obtained with AND Rule of 91 at 0.01% FAR. At maching score level, Parkavi et al., (2017) combined finger print and iris of a person. The proposed technique has been evaluated and accuracy has been increased by minimizing the FAR (False Acceptance Rate) and FRR (False Rejection Rate). Meanwhile, multimodal biometric fusion at feature level proposed by Xin et al., (2018) which is based on secondary calculation of the Fisher vector and uses three biometric modalities: face, fingerprint, and finger vein. It was reported that their fused extractor strategy performed better than a single biometric extractor. However, fusion at feature level is not an easy task because each biometric system contains its own unique processing techniques (i.e., feature extraction and matching process), thus fusing their scores require an additional step such as score normalisation and a complex fusion approach. Commonly, normalisation is applied to overcome the salient weaknesses of feature representation due to high dimensional, heterogeneous data in extracted features. Careful observation and experimental analysis need to be performed in order to improve the performance of identification. Too much normalisation will diminish the original characteristics of an individual from the multiple biometric images (Schlapbach and Bunke, 2005). This is the first issue that needs to be coped. In relation to this, adding another phase of feature filtering or enhancement procedure on the already available systems prior to the identification decision, in which, each layer puts emphasis on each other would increase the system's effectiveness. Therefore, in this paper, the improved Discretization so-called Biometric Feature Discretization (BFD) is proposed to represent the distinctiveness and uniqueness of features granularity for personal identification. The proposed BFD here is the improved version of Azah's Discretization (Azah et al., 2008). Work on Discretization has been explored widely in the past since 1991 by Catlet, followed by Holte (1993); Dougherty, Kohavi, & Sahami (1995); Zighed &Rakotomalala (2000); Shamsuddin, S.M et al., (2004) on real values of mathematic symbol recognition; Azah et al., (2008) in English handwriting identification, Chin et al., (2011) on multimodal biometric for template protection, Bayan and Shamsuddin (2012) on twins identification, and recently on finger vein biometric identification based on Discretization by Yahaya et al., (2019).

2. Biometric Feature Discretization for Individuality Representation

The main contribution of this study is the utilization of descriptive feature representation based on improved Discretization on multiple biometric images to improve the performance of identification. The improved Discretisation so-called Biometric Feature Discretisation (BFD) has the capability of standardising the extracted features, as well as reducing the dimensionality and complication of feature sets where they extract only significant information from multiple images to represent a person. In this work, the feature extraction module first processes the biometric images by extracting them into an appropriate feature set. Let the raw input signature images of i^{th} individual be denoted by $S_{i1}, S_{i2}, S_{i3}, \ldots, S_{iM}$, whereas, input fingerprint images be $F_{i1}, F_{i2}, F_{i3}, \ldots, F_{iM}$, and input of iris images represented by $I_{i1}, I_{i2}, I_{i3}, \ldots, I_{iM}$, where M represents the total number of images for each individual. In this study, M is four. After feature extraction, $actFeature$ set is produced, representing the original extracted features of signatures, fingerprints and irises. Here, let the original

extracted signature features of i^{th} individual be denoted by $fS_{i1}, fS_{i2}, fS_{i3},...., fS_{iM}$ meanwhile, fingerprint images be $fF_{i1}, fF_{i2}, fF_{i3},...., fF_{iM}$, and the original extracted iris images denoted by $fI_{i1}, fI_{i2}, fI_{i3},...., fI_{iM}$. These features do not reflect the statistical distinctiveness between individuals, and therefore lack the capability of preserving the discriminative power in multi-biometric samples of individuals. In this paper, improved Discretization is proposed to the identification in such a way that it represents a compact information of each biometric modalities called $disFeature$ set to assist the fusion or classification process. Let the discriminatory $disFeature$ feature discretized from original extracted signatures of i^{th} individual $\{fS_{i1}, fS_{i2}, fS_{i3},...., fS_{iM}\}$ be $Dsign_i$ meanwhile, the $disFeature$ feature discretized from fingerprint features of i^{th} individual $\{fF_{i1}, fF_{i2}, fF_{i3},...., fF_{iM}\}$ be $Dfin_i$, whereas, $disFeature$ feature discretized from irises of i^{th} individual $\{fI_{i1}, fI_{i2}, fI_{i3},...., fI_{iM}\}$ be $Diris_i$. The process of proposed improved Discretization is represented in Figure 1. These features which are discriminability features for each individual are then fed to matching and fusion or classification for identification process.

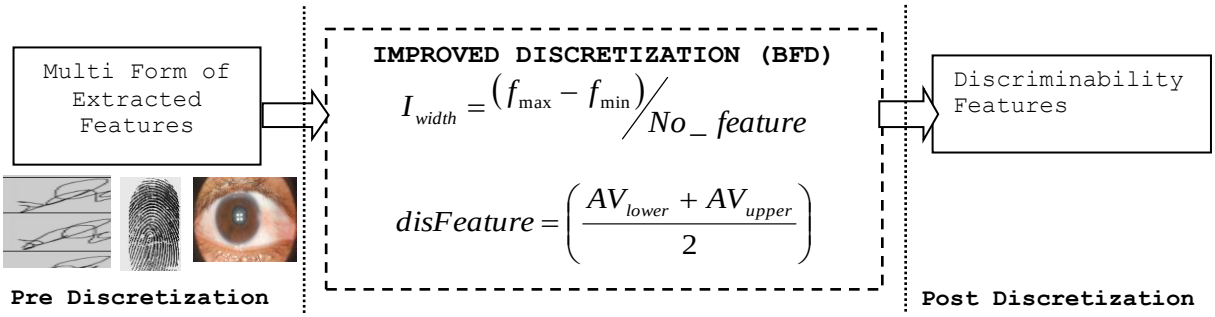

Figure 1. Before and after the process of Biometric Feature Discretization (BFD).

The description in two main steps below explains how the actual feature sets of an individual are discretized in multi-biometric for personal identification.

Step 1: Compute the size of interval, I_{width}. Given a set of features, the discretization algorithm first computes the size of interval, I_{width} as defined in Equation (1), i.e., it determines its upper and lower bounds. The range is divided by the number of features which then gives each interval upper approximation, AV_{upper} and lower approximation, AV_{lower} as demonstrated in Figure 2.

$$I_{width} = (f_{max} - f_{min}) / No\_feature \quad (1)$$

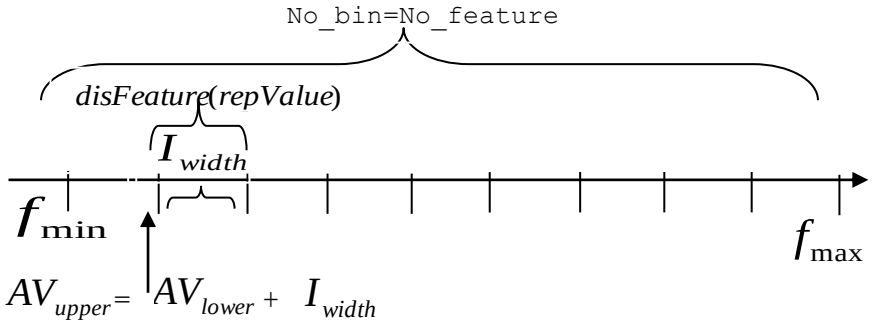

Figure 2. The improved Discretization; Biometric Feature Discretization (BFD).

The number of intervals generated is equal to the dimensionality of the feature vectors. In this research, there are nine actual features for a signature, thus nine bins are created; eight actual features for an iris, thus eight bins are created. In other words, the number of intervals created corresponds to the number of features of each modality. This is to maintain the original number of extracted features from multiple extraction methods.

Step 2: Compute a single discriminatory value, $disFeature$ of each interval. Instead of taking the range between the interval, the $disFeature$ here is considered by taking the *midpoint of the* AV_{upper} *and* AV_{lower} *interval* as defined in Equation (2) (improved from Azah's Invariant Discretization).

$$disFeature = \left(\frac{AV_{lower} + AV_{upper}}{2} \right) \tag{2}$$

With this improved procedure in computing intervals, the estimated representation feature values are more close to the actual multi-biometric features distribution (true feature values). This preserves the discriminative power of the original multi-biometric features and enhances the statistical distinctiveness between individuals. Therefore, here, the Discretization scheme is said to be robust and efficient enough to handle the issues in feature distribution of Multi-biometric identification. Unlike other discretization approaches, the main motivation behind this scheme is to maximise inter-class distances for all biometric samples that do not belong to the same individual class. By representing the features into a set of intervals, the issues of dimensionality caused by overlapping features can be avoided. This makes the identification process easier and faster due to the easier clarification by the classification and decision tasks.

3. Experimental and Main Results

This section investigates the improvement of identification performance using the proposed Biometric Feature Discretization by utilising the different types of Discretization methods, determined on a variety of classification methods. The experiment is tested on fifty subjects, where each subject contributes four samples of different fingerprint impressions, four samples of handwritten signatures and four samples of irises. These raw biometric modalities are extracted by appropriate feature extraction to produce feature samples. In this experiment, six most pertinent discretization methods are implemented for the comparison purpose.

The results of the experiments for three multi-biometrics using six Discretization and four Classification methods are summarised and reported into a single Table 1. Based on Table 1, the best Discretization results turned out to be the proposed improved Discretization scheme in both training

and testing dataset of irises, fingerprints and signatures exhibiting better performance than other six Discretization methods with K-NN classification accuracy rate of more than 95%. Also, it can be seen that the proposed BFD statically outperforms the other CAIM, CACC, ChiM, Chi2, ExtChi2, and Khiops Discretization. It yields the accuracy of 92.116% for testing dataset and 98.681% for the combination of three biometric training datasets. The six other Discretization turn out to be unfeasible in both training and testing dataset of C-45 classification since almost all performance measures yields the average accuracy rate of 76% for training, whereas average accuracy rate of 66% for testing datasets. Based on Table 1, again, interestingly, the proposed BFD statically outperforms the other CAIM, CACC, ChiM, Chi2, ExtChi2, and Khiops Discretization in both training and testing datasets of NB classification. It yields the accuracy of 94.415% for testing dataset and 96.301% for the combination of three biometric training dataset. The six other Discretization turn out to be unfeasible in both training and testing datasets of NB classification since almost all performance measures show the degrading accuracy rate of about 24% for training, and about 29% for testing dataset.

The same discussions hold for SVM classification, where the best Discretization results turn out to be the Proposed BFD in both, training and testing datasets of irises, fingerprints, and signatures revealing more superior performance than CAIM, CACC, ChiM, Chi2, ExtChi2, and Khiops Discretization. The proposed improved Discretization when integrates with SVM classification yields the average accuracy rate of 98.462% for training and 91.670% for testing dataset of the tree biometric modalities. Finally, the results of the experiments for three multi-biometrics using six Discretization and four classification methods are summarised and reported into a single Table 1 presented below. The table consists of Discretization performance in row and the average accuracy of four classifications in columns.

Table 1. Performance of four classification methods on various types of Discretization algorithms.

Classification /	***K-NN***		***C45***		***NB***		***SVM***	
Discretization	Train	Test	Train	Test	Train	Test	Train	Test
CAIM	67.581	68.625	85.099	75.495	74.426	67.546	71.750	67.476
CACC	47.655	46.364	70.012	63.941	60.274	55.772	58.218	56.642
ChiM	60.592	59.358	74.789	65.903	74.319	66.794	71.384	66.660
Chi2	65.988	67.344	78.588	66.826	77.479	66.821	70.828	67.816
ExtChi2	67.581	68.625	76.644	61.519	81.171	69.181	71.350	68.223
Khiops	59.328	57.734	71.870	65.282	68.122	64.357	70.843	67.920
BFD (Improved Discretization)	**97.016**	**97.623**	**98.681**	**92.116**	**96.301**	**94.415**	**98.462**	**91.670**

Overall, as it can be seen here, the combination of proposed BFD scheme with K-NN, C-45, NB and SVM classification on training datasets successfully achieved the best performance with the average accuracy rate of 97.016%, 98.681%, 96.301%, and 98.462% respectively. Whereas, for testing dataset, the performance of K-NN, C-45, NB and SVM classification also yields a higher performance with the average accuracy of 97.623%, 92.116%, 94.415%, and 91.670% after applying the proposed BFD on multiple modalities. Meanwhile, the second best on the combination of CAIM and four classification methods on biometric datasets, while the worst for the CACC method.

4. Conclusion and Future Work

The effort towards the development of Multi-biometrics is an exciting research in the area of forensic security and pattern recognition. This research presents other alternatives to improve the performance of individual identification by using standardised and informative feature set for multiple biometric modalities. The research outcomes are directed towards the understanding of the complexity of multiple trait features and fusion methods for accurate and reliable personal information. Thus, the success of this research could be treated as a benchmark for researchers to conduct more robust, effective and accurate biometric system in the near future. A key to successful Multi-biometric identification is an effective methodology organisation and fusion process, capable to integrate and handle important information such as distinctive characteristics of an individual. Individual's distinctive characteristics are unique to biometric identification especially in the field of forensic. The principle contribution of this study is the utilisation of feature representation based on discretisation and fusion process, based on single-matching fusion scheme on multiple biometric images to improve the performance of identification; aiming to be treated as Multi-biometric based forensic authentication; a promising model and good practice to assist forensic investigation that could be best serve for both academia and industry. In the near future, we expect to see a higher adoption rate and better technology for authentication and personalization system. This includes potential applications, from security, forensics, financial activities to archaeology (i.e., to identify writers and validate the owners of ancient documents, law enforcement agencies, judicial systems and border control system).

Acknowledgments

This research is sponsored partly by the School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia. Authors would especially like to thank UTM Big Data Centre and Soft Computing Research Group (SCRG), UTM for their excellent cooperation and contributions to improve this paper.

References

[1] Choras, R.S., 2019. Multimodal Biometrics for Person Authentication. In *Digital Identity*. IntechOpen.

[2] Damousis, G. and Argyropoulos, S. Four Machine Learning Algorithms for Biometrics Fusion: A Comparative Study, *Applied Computational Intelligence and Soft Computing*, vol. 2012, Article ID 242401, 7 pages.

[3] He, Y. L., Wang, R., Kwong, S., & Wang, X. Z. Bayesian classifiers based on probability density estimation and their applications to simultaneous fault diagnosis. *Information Sciences*, 259 (2014), 252-268.

[4] Turki, A. A. Evaluation of Multimodal Biometrics at Different Levels of Face and Palm Print Fusion Schemes. *Asian Journal of Applied Sciences, 9 (2016) 126-130.*

[5] Parkavi, K.R., Babu C. and Kumar, J.A. Multimodal Biometrics for user authentication. *11th International Conference on Intelligent Systems and Control (ISCO)*, Coimbatore (2017), 501-505.

[6] Xin, Y. X., Kong, L., Liu, Z., Wang, C., Zhu, H., Gao, M., Zhao, C., and Xu, X. Multimodal Feature-Level Fusion for Biometrics Identification System on IoMT Platform, in *IEEE Access*, vol. 6 (2018) pp. 21418-21426.

[7] Schlapbach, A., & Bunke, H. Writer identification using an HMM-based handwriting recognition system: To normalize the input or not. In *Proc. 12th Conf. of the Int. Graphonomics Society* (2005), 138-142.

[8] Muda, A. K., Shamsuddin, S. M., and Darus, M. Invariants Discretization for Individuality Representation in Handwritten Authorship. *Lecture Notes in Computer Science Series on Computational Forensics*, Springer-Verlag,(2008), 218-228.

[9] Catlett, J. Onchanging continuous attributes into ordered discrete attributes. In *Proceedings of the European Working Session on Learning,* Springer-Verlag (1991), 87–102.

[10] Holte, R. C. Very simple classification rules perform well on most commonly used datasets. *Machine Learning, 11* (1993), 63–90.

[11] Dougherty, J.,Kohavi, R.,&Sahami, M. Supervised and unsupervised discretization of continuous features. In *Proceedings of the 12th International Conference on Machine Learning,* San Francisco,

CA: Morgan Kaufmann (1995),194–202.
[12] Zighed, D. A., & Rakotomalala, R. *Graphes d'induction,* HERMES Science Publications,(2000), 327–359.
[13] Ahmad, R., Darus, M., Shamsuddin, S. M., and Bakar, A. A. Pendiskretan Set Kasar Menggunakan Ta'akulan Boolean Terhadap Pencaman Simbol Matematik. *Journal of Information Technology & Multimedia,* UKM, **1** (2004),15-26.
[14] Mohammed, B. O., and Shamsuddin, S. M. Feature Discretization for Individuality Representation in Twins Handwritten Identification. *Journal of Computer Science*, **7** (7) (2011)1080-1087.
[15] Chin, Y. J., Ong, T. S., Teoh, A.B.J., and Goh, M. K. O. Multimodal biometrics based bit extraction method for template security, *2011 6th IEEE Conference on Industrial Electronics and Applications*, Beijing, (2011),1971-1976.
[16] Yahaya, Y.H., Shamsuddin, S. M. and Leng, Wong, Y.L. Finger Vein Identification Based on Maximum Curvature Directional Feature Extraction, *Journal of Creative Practices in Language Learning and Teaching (CPLT) Special Issue*: *Generating New Knowledge through Best Practices in Computing and Mathematical Sciences.*, vol 7(1),(2019).