Invariants Discretization for Individuality Representation in Handwritten Authorship

Azah Kamilah Muda[1], Siti Mariyam Shamsuddin[1], and Maslina Darus[2]

[1] Faculty of Computer Science and Information Systems
Universiti Teknologi Malaysia, Malaysia
[2] Faculty of Sciences and Technology
Universiti Kebangsaan Malaysia, Malaysia
{azah@utem.edu.my, mariyam@utm.my, maslina}@pkrisc.cc.ukm.my

Abstract. Writer identification is one of the areas in pattern recognition that have created a center of attention by many researchers to work in. Its focal point is in forensics and biometric application as such the writing style can be used as biometric features for authenticating a writer. Handwriting style is a personal to individual and it is implicitly represented by unique features that are hidden in individual's handwriting. These unique features can be used to identify the handwritten authorship accordingly. Many researches have been done to develop algorithms for extracting good features that can reflect the authorship with good performance. However, this paper investigates the individuality representation of individual features through discretization technique. Discretization is a procedure to explore the partition of attributes into intervals and to unify the values for each interval. It illustrates the pattern of data systematically which improved the identification accuracy. An experiment has been conducted using IAM database with 3520 training data and 880 testing data (70% training data and 30% testing data) and 2639 training data and 1760 testing data (60% training data and 40% testing data). The results reveal that with invariants discretization, the accuracy of handwritten identification is improved significantly with the classification accuracy of 99.90% compared to undiscretized data.

Keywords: Writer Identification, Authorship Invarianceness, Invariants Discretization.

1 Introduction

Pattern recognition is imperative in various engineering and scientific disciplines such as computer vision, marketing, biology, psychology, medicine, artificial intelligence, remote sensing and etc. One of the areas in pattern recognition is handwriting analysis. Handwriting analysis is important in forensic application such as Writer Identification (WI). Writer identification (WI) can be considered as a particular kind of dynamic biometric since the shape and style of writing can be used as biometric features for authenticating an identity [1-4], similar to signature, fingerprint, iris or face identification. Frequently, writer identification performed on legal papers by a way of signature. However, there is also exist a scenario where to identify a handwritten

S.N. Srihari and K. Franke (Eds.): IWCF 2008, LNCS 5158, pp. 218–228, 2008.

document without a signature such as in a threaten letter, authorship determination of old or historical manuscript, film script (to identify the original idea) and others. In this work, the shape of cursive word is employed and extracted to obtain the features with a proposed descritized process prior to identification task.

Handwriting is individualistic where consistent individual's features are hidden in the shape and writing style. The writing styles are different from one to another, but it is personal to individual. Any written word by the same author must have the same characteristic features, despite of the word shape or writing style. The main issue in writer identification is to acquire the features that reflect the author for varieties of handwriting [3, 5-9] and more important is the unique individual features of handwriting. Previous works have developed new approach or technique for better feature extraction and to proof the individuality concept in handwriting. However, from the literature we found that most of the works are focus on how to extract the individual features and not on illustrating the individual characteristic of handwriting with systematic representation.

The performance of pattern recognition largely depends on the feature extraction and classification/learning scheme [10 - 11]. These two tasks are vital to achieve a good performance in identifying handwritten authorship. Extracting and selecting the meaningful features are a crucial task in the process of pattern recognition prior to classification task, where the extracted features will be classified into categories. Low performance in terms of accuracy is due to various features are representing the same author. It makes the identification process become intricate and complex. The same characteristics are easily identified if all of different features values for same author are having a standard representation for the generalized unique features or individual features. It can make the identification process simpler. Therefore, illustration of individuality features is required to portray the individual's unique features in a systematic representation. This can be achieved by executing the discretization process to demonstrate the pattern of individual features thoroughly.

This paper focal point is to investigate the invariant discretization process of features in order to represent the individual features of writers and significantly illustrates related features in systematic way. In return, it is easily classified and performed better identification result. The paper is systematized as follows. Individuality of handwriting is explained in Section 2. Followed with the authorship invarianceness of moment in Section 3. Section 4 describes the proposed approach of invariants descritization process in this work. The experiment and results is discussed in Section 5. And finally, the conclusion is drawn in Section 6.

2 Individuality of Handwriting

Handwriting has long been considered individualistic and writer individuality rests on the hypothesis that each individual has consistent handwriting [12 -16]. The relation of character, words, shape or style of writing is different from one person to another. Even for one person, they are different in times. However, there are still unique features for each person. These unique features can be generalized as individual's handwriting even though one person has many styles of writing. Fig. 1 is example of words by different authors. Each person's handwriting is seen as having a specific texture [4]. The shape is slightly different for the same author and quite difference for different authors. It shows that each person has its individual style in handwriting. Intra-class measurement is exhibited for features of the same author, and inter-class for different authors. To benchmark these measurements, similarity error is computed for both inter-class and intra-class where the similarity error for intra-class must be lowers than inter-class. This reflects the individuality concept in handwriting. This is called as authorship invarianceness in this work due to the concept of moment function. Moment function is used to extract the features in this work.

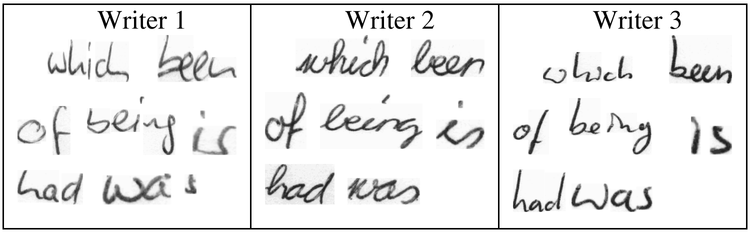


Fig. 1. Various words for different writer

Each image of word is performed the feature extraction task to obtain the features of the image. In this work, the images are extracted to obtain the invariant features using the proposed moment function of integrated Aspect Scale Invariant (ASI) into United Moment Invariant (UMI). It is based on the original United Moment Invariant function. Detail procedure on proposed moment function of integrated ASI into UMI can be referred in [17]. Example of extracted features is shown in Fig 2. Further stage, the extracted features are performed authorship invarianceness analysis to evaluate the individuality concept of handwriting in WI.

3 Authorship Invarianceness

An invarianceness in the context of moment functions can be defined as the *perseverance of the images regardless of its transformations.* In this work, the invarianceness of authorship in WI is given as small similarity error for intra-class (same writer) and large similarity error for inter-class (different writers) of words and regardless of word shape. This is due to the uniqueness features of person in handwriting that called as individuality of handwriting concept in handwriting analysis. The main process of identification in WI is to look for similar characteristic of handwriting based on the nearest unknown handwriting in the database. This can be solved by implementing the individuality of handwriting concept. To achieve this, intra-class and inter-class measurement are implemented to find the nearest characteristic using word shape with the lowest Mean Absolute Error (MAE) value in order to obtain authorship invarianceness. Intra-class should give smaller MAE value compared to inter-class, regardless of any types of word. The range of deviation between intra-class MAE value and inter-class MAE value is not a concern. This is due to the characteristic of Moment

Function where the intra-class value must be lower than inter-class value confirm it can be classified as authorship invarianceness. The MAE function is given by Equation (1):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|(x_i - r_i)\right| \tag{1}$$

where :

n is number of image.
x_i is the current image.
r_i is the reference image.
i is the feature's column of image.

The result in Table 1 and Table 2 show that the proposed technique of ASI into UMI is worth for further exploration in WI domain. The initial result of similarity error shows that invarianceness of authorship for intra-class (same writer) is smaller compared to inter-class (different writers) for same word and different words, respectively. It is proof the individuality handwriting concept in WI, where MAE value for intra-class (same writer) is smaller value compared to inter-class (different writers) for the same or different words, regardless of short or long word such as the word of "To" or "Being". This is due to the capability of Moment Function in extracting object shape without any constrain in terms of length. Thus, this authorship invarianceness analysis confirms the integrated ASI into UMI techniques can be used to extract features for WI domain.

Table 1. Invarianceness of Authorship for Same Word

Word	Intra-class (1 writer)	Inter-class (10 writers)	Inter-class (30 writers)	Inter-class (60 writers)
To	1.08086	1.10181	1.21423	1.2927
He	0.486922	0.865588	0.721937	0.737597
Of	0.486201	0.702867	0.691087	0.754485
Is	0.489428	0.599104	0.684848	0.779217
Had	0.454727	0.566663	0.670911	0.675404
And	0.564578	0.856195	0.797005	0.782162
The	0.39991	0.718456	0.643291	0.611504
Was	0.736664	0.951713	1.0253	0.955763
Been	1.02514	1.35783	1.28346	1.27161
That	0.677631	1.0147	0.847687	0.768499
With	0.394996	0.706262	0.739905	0.718119
Which	0.335732	0.491985	0.556506	0.599928
Being	0.291463	0.557977	0.581267	0.552889

Table 2. Invarianceness of Authorship for Various Word

Various words	Intra-class (1 writer)	Inter-class (10 writers)	Inter-class (30 writers)	Inter-class (60 writers)
60 words	0.733659	1.11315	0.931423	0.882049
90 words	0.693564	1.03028	0.94499	0.924337
120 words	0.852839	0.975387	0.939999	0.936329

The uniqueness or individual characteristic for each writer in handwriting describes the above result. Similarity error for inter-class (different writers) should be higher than intra-class (same writer) in authorship invarianceness concept. It has been proven in Table 1 and Table 2. For further exploration, these similarity errors can be associated into discretization technique in order to illustrate the data by discerning the individual features into category. The idea is to acquire objects, attributes, decision values, and generate rules for lower, upper and boundary approximations of the set. With these rules, a new object can easily be classified into one of the region or interval which is called as discretization process.

4 Discretization

Discretization is a process of dividing the range of continuous attributes into disjoint regions (interval) which labels can then be used to replace the actual data values [18]. It engages searching for "cuts" that determine intervals and unifying the values over each interval. All values that lie within each interval are mapped to the same value, in effect converting numerical attributes that can be treated as being symbolic [19]. Empirical results show the superiority of classification methods depends on the discretization algorithm used in preprocessing process. There are abundant of discretization algorithms exist based on three basic perspectives. They are supervised versus unsupervised, global versus local and dynamic versus static [20]. Supervised method considers class information is on hand and no classification information available for unsupervised. Another perspective of global versus local describes global method discritized entire data before classification while local method discretized specific amount of defined data. Furthermore, static versus dynamic perspective explains static method discretized each attribute independently without consider interaction between attributes. Meanwhile, dynamic method considered attributes interdependencies while discretization process.

Proposed discretization method is resemblance with the simplest unsupervised methods of Equal Width Binning. However, proposed method is categorized in supervised method because it needs class information to perform discretization process. It globally process for all integrated invariants feature vector for all writers with dynamic characteristic of features in WI domain. The continuous values of invariant features are discretized to obtain the detachment of authors' individuality for better data representation. In this work, invariant features are in real value format, extracted using integrated ASI into UMI technique. Discretization of real value attributes is an essential task in data mining, predominantly the classification problem. Our results disclose that the performance of the classification on writers' handwriting is much improved with discretized data of proposed Invariant Discretization algorithm.

4.1 Proposed Invariant Discretization Algorithm

Discretization is important in this work because it leads to the better accuracy in classification phase compared to undiscretize data. Proposed discretization algorithm is applied where class information is given for the each image to represent the writer. In the process of discretization, it will search the suitable set of cuts to represent the real data for each writer. It divides the range of minimum to maximum data of each writer with the equal size of interval or cuts. Lower and upper approximation is given to the each cut. Number of cuts is defined based on number of feature vector for the each word image, i.e, eight feature vector values of ASI into UMI are used to represent a pattern of image. This is to keep the original number of invariant vector in moment invariant function that has been applied. Each cuts will represented with one defined representation value. Feature's values that fall within the same cut will have the same representation. The proposed discretization algorithm is given below :

Algorithm of Proposed Discretization

```
For each writer {
   Min = min feature;  Max = max feature;
   No_bin = no_feature_invariant;
   Interval = (Max - Min)/ No_bin;

   For each bin {
      Find lower and upper value of interval;
      RepValue = (upper -lower)/2;
   }

   For (1 to no_feature_invariant) {
      For each bin {
         If (feature in range of interval)
            Dis_Feature = RepValue;
      }
   }
}
```

Process to calculate the interval and representation value for the each cut is done based on writer classes. This is due to the concept in WI domain where each person has their own style of writing or individuality in handwriting. To make sure the uniqueness or individuality characteristic is preserved, the interval and representation value is calculated based on each writer. If there are two different writers that have closed or same invariant feature, there will be the same or quite similar interval or cuts for these two classes. Therefore, the representation value of each cut will be same or quite similar. Thus, this proposed algorithm is not changed the information gather or characteristic of writers. It just represents the real invariant data into better data representation. Discretization process is implemented to illustrate the features clearly and not to change the characteristics of features. Therefore, the proposed discretization of each writer's class approach is seen as acceptable and match with the individuality concept in WI.

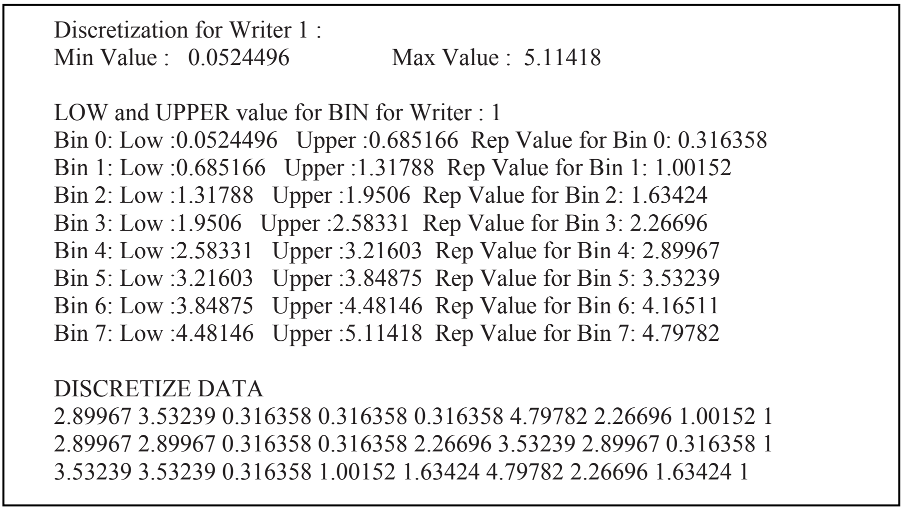
Example of transformation of feature invariant vector to discretize feature vector is ilustrated in Fig. 2 through Fig. 4 below. Fig. 2 shows the example of data before discretization process for various images of writers. There are eight columns of extracted invariant feature vectors and the last column is the label of author's class. Eight invariant vectors of feature in one row represent one word image for the writer in the last column.

2.59224	3.23024	0.332166	0.672428	0.617473	4.56811	2.55781	1.02415	1
3.61109	3.62337	0.0471209	0.10731	3.39726	3.82502	3. 51606	0.366274	4
2.91782	3.11856	0.0524496	0.204262	2.40792	3.42825	2.9143	0.43011	1
3.34655	3.40755	0.284003	1.13843	1.57912	5.11418	2.26912	1.43088	1
2.74886	2.75738	0.0650583	0.31401	2.29621	3.20228	2.44336	0.512494	2
3.18126	3.18186	0.229476	0.475357	2.24635	4.11646	2.7065	0.752626	8
3.54961	3.74973	0.180705	1.65463	1.33345	5.76589	2.0951	1.8395	8
3.05499	4.58202	0.163657	0.588422	0.612222	5.49814	3.9936	0.801845	2
3.18019	3.49778	0.0694599	0.81009	1.9136	4.44707	2.68769	0.925833	2
3.36354	3.67488	0.115037	0.471541	2.81074	3.91654	3.20334	0.553371	2
3.24221	3.526	0.0506261	0.928334	2.13133	4.35333	2.59767	0.937993	3
3.39974	3.40077	0.0320461	0.249581	3.08504	3.71462	3.15119	0.453545	4
3.50443	5.83822	0.0726602	0.182035	0.843275	6.16572	5.65619	0.422188	4
4.19887	5.14676	0.31637	0.30295	4.18666	4.2111	5.44971	0.364019	6
3.51602	3.57551	0.0456017	0.187287	3.36045	3.67173	3.38822	0.397325	7
2.68472	2.68664	0.357513	0.104451	3.29337	2.07696	2.58219	0.452733	6
3.66434	3.77518	0.176983	0.50243	2.6971	4.63167	3.27275	0.736009	8
3.58092	4.42651	0.139926	0.501636	2.51353	4.64841	3.92488	0.565854	8
3.55531	3.84184	0.177988	0.32283	3.94758	3.16315	4.16467	0.260595	8

Fig. 2. Real data of invariant feature vector

Data in Fig. 2 is continued to perform discretization process as shown in Fig. 3. It is an example to discretize data for writer 1. Discretized feature data of discretization process is shown in Fig. 4 for all data in Fig. 2.

From the discretized feature data in Fig. 4, it shows that each writer has its own representation data which illustrates the characteristic of each writer. It represents the individuality concept of handwriting in WI domain where each person has its own style of handwriting. These discretized features data then undergo identification process in order to analyze the performance of identification.

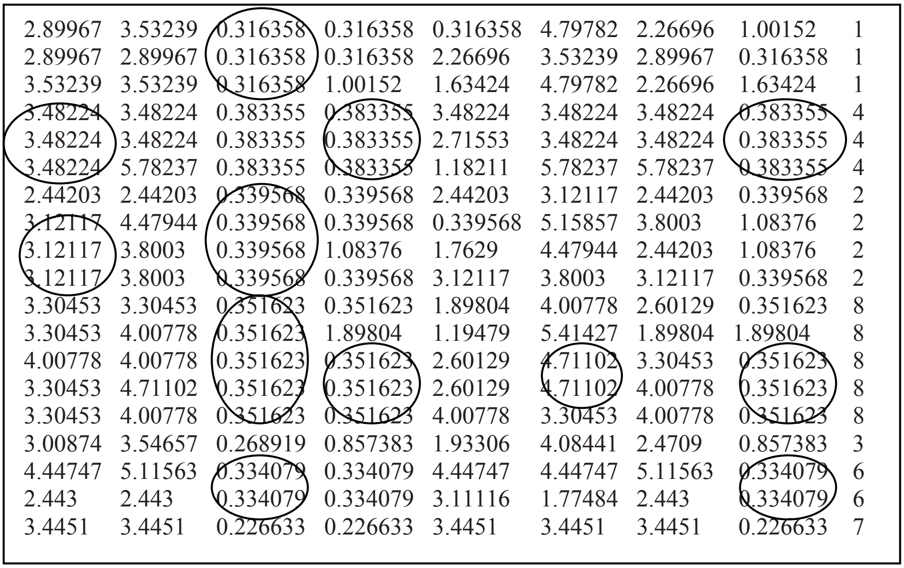


```
Discretization for Writer 1 :
Min Value :   0.0524496              Max Value :  5.11418

LOW and UPPER value for BIN for Writer : 1
Bin 0: Low :0.0524496   Upper :0.685166  Rep Value for Bin 0: 0.316358
Bin 1: Low :0.685166   Upper :1.31788  Rep Value for Bin 1: 1.00152
Bin 2: Low :1.31788   Upper :1.9506  Rep Value for Bin 2: 1.63424
Bin 3: Low :1.9506   Upper :2.58331  Rep Value for Bin 3: 2.26696
Bin 4: Low :2.58331   Upper :3.21603  Rep Value for Bin 4: 2.89967
Bin 5: Low :3.21603   Upper :3.84875  Rep Value for Bin 5: 3.53239
Bin 6: Low :3.84875   Upper :4.48146  Rep Value for Bin 6: 4.16511
Bin 7: Low :4.48146   Upper :5.11418  Rep Value for Bin 7: 4.79782

DISCRETIZE DATA
2.89967 3.53239 0.316358 0.316358 0.316358 4.79782 2.26696 1.00152 1
2.89967 2.89967 0.316358 0.316358 2.26696 3.53239 2.89967 0.316358 1
3.53239 3.53239 0.316358 1.00152 1.63424 4.79782 2.26696 1.63424 1
```

Fig. 3. Example of Discretization Process for Writer 1



2.89967	3.53239	0.316358	0.316358	0.316358	4.79782	2.26696	1.00152	1
2.89967	2.89967	0.316358	0.316358	2.26696	3.53239	2.89967	0.316358	1
3.53239	3.53239	0.316358	1.00152	1.63424	4.79782	2.26696	1.63424	1
3.48224	3.48224	0.383355	0.383355	3.48224	3.48224	3.48224	0.383355	4
3.48224	3.48224	0.383355	0.383355	2.71553	3.48224	3.48224	0.383355	4
3.48224	5.78237	0.383355	0.383355	1.18211	5.78237	5.78237	0.383355	4
2.44203	2.44203	0.339568	0.339568	2.44203	3.12117	2.44203	0.339568	2
3.12117	4.47944	0.339568	0.339568	0.339568	5.15857	3.8003	1.08376	2
3.12117	3.8003	0.339568	1.08376	1.7629	4.47944	2.44203	1.08376	2
3.12117	3.8003	0.339568	0.339568	3.12117	3.8003	3.12117	0.339568	2
3.30453	3.30453	0.351623	0.351623	1.89804	4.00778	2.60129	0.351623	8
3.30453	4.00778	0.351623	1.89804	1.19479	5.41427	1.89804	1.89804	8
4.00778	4.00778	0.351623	0.351623	2.60129	4.71102	3.30453	0.351623	8
3.30453	4.71102	0.351623	0.351623	2.60129	4.71102	4.00778	0.351623	8
3.30453	4.00778	0.351623	0.351623	4.00778	3.30453	4.00778	0.351623	8
3.00874	3.54657	0.268919	0.857383	1.93306	4.08441	2.4709	0.857383	3
4.44747	5.11563	0.334079	0.334079	4.44747	4.44747	5.11563	0.334079	6
2.443	2.443	0.334079	0.334079	3.11116	1.77484	2.443	0.334079	6
3.4451	3.4451	0.226633	0.226633	3.4451	3.4451	3.4451	0.226633	7

Fig. 4. Example of Descritized Feature Data

5 Experiment Result and Discussion

Experiment is conducted to proof the discretization process can improve the performance of identification in WI domain. The comparisons of identification accuracy (%) for discretized data with un-discretize data are shown in Table 3 and

Table 4. Two techniques have been used to extract the features from the various written words, which are original UMI and proposed ASI into UMI. Identification accuracy is compared for these two techniques. For identification task, discretized data and un-discretized data are run with Johnson Algorithm and 1R Algorithm, which are the techniques that embedded in Rosetta toolkit [21]. Meanwhile R-Chunk is the pattern matching that applied in Modified Negative Selection Algorithm (MNSA) classifier [22]. Un-discretized data is the original extracted features meanwhile discretized data is the extracted features that performed discretized process using proposed invariant discretization algorithm. IAM database [23] with 60 writers from the various types of word images were used to run this experiment. Table 3 is for 3520 training data and 880 testing data (70% training data and 30% testing data) and Table 4 is for 2639 training data and 1760 testing data (60% training data and 40% testing data).

Table 3. Comparison of Identification Accuracy for 3520 Training Data and 880 Testing Data

Technique	Original UMI	ASI into UMI	Data
Johnson Algorithm	33.56	35.34	Un_Dis
	99.09	99.55	Dis
1R Algorithma	33.67	35.34	Un_Dis
	99.90	99.90	Dis
R-Chunck Algorithm	45.80	46.68	Un_Dis
	95.34	99.88	Dis

Table 4. Comparison of Identification Accuracy for 2639 Training Data and 1760 Testing Data

Technique	Original UMI	ASI into UMI	Data
Johnson Algorithm	29.92	31.70	Un_Dis
	97.95	98.75	Dis
1R Algorithma	30.03	31.70	Un_Dis
	99.90	99.90	Dis
R-Chunck Algorithm	37.54	38.63	Un_Dis
	95.52	99.89	Dis

Discretized data gives higher accuracy for both feature extraction techniques and all classifiers tested in the experiment. Discretization is performed to represent the features of data systematically. Thus, the individuality of handwriting is clearly illustrated in discretized data. The same characteristics are easily identified if all of different features values for each author are having a standard represented value for the generalized unique features or individual features. Therefore increased the performance of identification compared to un-discretized data. The focus in this paper is to show that discretized data performed much better in identifying author compared to un-discretized data. Both tables show that discretized data give much better performance in identification and it is proven in the experiment result.

6 Conclusion

This paper proposed an approach of invariants discretization to represent the individual features systematically. Discrete features extracted from the various words undergo discretization process for granular mining of writer authorship. Similarity errors are reduced between these data, thus handwritten authorship can be defined easily. It is experimentally evaluated that discretized data give much better performance of identification compared to un-discretized data in WI domain. Our experiments have revealed better results with various identification techniques in classification process of Rosetta toolkit [21] and MNSA classifier[22].

Acknowledgments

This work was supported by Ministry of Higher of Education (MOHE) under Fundamental Research Grant Scheme (Vot 78182). Authors would like to thank Research Management Center (RMC), Universiti Teknologi Malaysia for the research activities and *Soft Computing Research Group* (SCRG) for their support in making this research success.

References

[1] Srihari, S.N., Huang, C., Srinivasan, H., Shah, V.A.: Biometric and forensic aspects of digital document processing. In: Chaudhuri, B.B. (ed.) Digital Document Processing. Springer, Heidelberg (2006)

[2] Tapiador, M., Sigüenza, J.A.: Writer identification method based on forensic knowledge, Biometric Authentication. In: Zhang, D., Jain, A.K. (eds.) ICBA 2004. LNCS, vol. 3072, pp. 555–561. Springer, Heidelberg (2004)

[3] Kun, Y., Yunhong, W., Tieniu, T.: Writer identification using dynamic features, Biometric Authentication. In: Zhang, D., Jain, A.K. (eds.) ICBA 2004. LNCS, vol. 3072, pp. 512–518. Springer, Heidelberg (2004)

[4] Zhu, Y., Tan, T., Wang, Y.: Biometric personal identification based on handwriting, Pattern Recognition. In: 15th International Conference, September 3-7, vol. 2, pp. 797–800 (2000)

[5] Bensefia, A., Paquet, T., Heutte, L.: A writer identification and verification system. Pattern Recognition Letters 26(10), 2080–2092 (2005)

[6] Schlapbach, A., Bunke, H.: Off-line handwriting identification using HMM based recognizers. In: 17th Int. Conf. on Pattern Recognition, Cambridge, UK, August 23-26, pp. 654–658 (2004)

[7] He, Z.Y., Tang, Y.Y.: Chinese handwriting-based writer identification by texture analysis, Machine Learning and Cybernetics. In: 2004 International Conference, August 26-29, vol. 6, pp. 3488–3491 (2004)

[8] Srihari, S.N., Cha, S.-H., Arora, H., Lee, S.: Individuality of handwriting. Journal of Forensic Sciences 47(4), 1–17 (2002)

[9] Shen, C., Ruan, X.-G., Mao, T.-L.: Writer identification using Gabor wavelet, Intelligent Control and Automation. In: 4th World Congress, June 10-14, vol. 3, pp. 2061–2064 (2002)

[10] Liu, C.-L., Nakashima, K., Sako, H., Fujisawa, H.: Handwritten digit recognition: benchmarking of state-of-the-art techniques. Pattern Recognition 36(10), 2271–2285 (2003)

[11] Liu, C.-L., Nakashima, K., Sako, H., Fujisawa, H.: Handwritten digit recognition: investiga-tion of normalization and feature extraction techniques. Pattern Recognition 37(2), 265–279 (2004)

[12] Srihari, S.N., Huang, C., Srinivasan, H., Shah, V.A.: Biometric and forensic aspects of digital document processing. In: Chaudhuri, B.B. (ed.) Digital Document Processing. Springer, Heidelberg (2006)

[13] Zhang, B., Srihari, S.N.: Analysis of handwriting individuality using word features, Document Analysis and Recognition. In: Seventh International Conference, August 3-6, pp. 1142–1146 (2003)

[14] Srihari, S.N., Cha, S.-H., Arora, H., Lee, S.: Individuality of handwriting. Journal of Forensic Sciences 47(4), 1–17 (2002)

[15] Srihari, S.N., Cha, S.-H., Lee, S.: Establishing handwriting individuality using pattern recog-nition techniques, Document Analysis and Recognition. In: Sixth International Conference on Document Analysis and Recognition (ICDAR 2001), Seattle, September 10-13, pp. 1195–1204 (2001)

[16] Zhu, Y., Tan, T., Wang, Y.: Biometric personal identification based on handwriting, Pattern Recognition. In: 15th International Conference, September 3-7, vol. 2, pp. 797–800 (2000)

[17] Muda, A.K., Shamsuddin, S.M., Darus, M.: Embedded scale united moment invariant for identification of handwriting individuality. In: Gervasi, O., Gavrilova, M.L. (eds.) ICCSA 2007, Part I. LNCS, vol. 4705, pp. 385–396. Springer, Heidelberg (2007)

[18] Agre, G., Peev, S.: On supervised and unsupervised discretization. Cybernetics And Information Technologies 2(2), 43–57 (2002)

[19] Nguyen, H.S.: Discretization problems for rough set methods. In: Polkowski, L., Skowron, A. (eds.) RSCTC 1998. LNCS (LNAI), vol. 1424, pp. 545–552. Springer, Heidelberg (1998)

[20] Dougherty, J., Kohavi, R., Sahami, M.: Supervised and unsupervised discretization of con-tinuous features. In: Twelfth International Conference on Machine Learning, pp. 194–202. Morgan Kaufmann, Los Altos (1995)

[21] Øhrn, A., Komorowski, J.: ROSETTA: A rough set toolkit for analysis of data. In: Wang, P.P. (ed.) Proc. Third International Joint Conference on Information Sciences, Durham, NC, vol. 3, pp. 403–407 (March 1997)

[22] Muda, A.K., Shamsuddin, S.M., Darus, M.: Bio-inspired generalized global shape approach for writer identification. Transaction on Engineering, Computing and Technology 16, 55–59 (2006)

[23] Marti, U.-V., Bunke, H.: The IAM-database: an english sentence database for off-line handwriting recognition. Int. Journal on Document Analysis and Recognition 5, 39–46 (2002)