A New Chain-code Quantization Approach Enabling High Performance Handwriting Recognition based on Multi-Classifier Schemes

S. Hoque K. Sirlantzis M. C. Fairhurst

Department of Electronics, University of Kent, Canterbury, Kent, United Kingdom.
E-mail: {S.Hoque,K.Sirlantzis,M.C.Fairhurst}@kent.ac.uk

Abstract

In this paper initially we propose a novel approach to classify handwritten characters based on a directional decomposition of the corresponding chain-code representation. This is alternative to previous transformations of the chain-codes proposed by the authors, namely the ordered and random decomposition of the bit-planes resulting from the binary representation of the chain-codes. Subsequently we utilize the power of the recently developed multiple classifier schemes using sntuple classifiers to integrate the complimentary information encapsulated in all three transformations into a more powerful and robust character recognition system. The results obtained through a series of cross-validation experiments show that the proposed fusion scheme not only outperforms its constituent parts and a number of other successful classifiers, but also enables significant savings in memory requirements compared to the original sntuple-based recognition system.

1. Introduction

Recognition of printed and handwritten documents is still one of the most challenging areas in pattern recognition with profound implications for the machine vision field. Although many different methods have been reported and some have shown very high performance, none has been able to achieve the accuracy and speed of human readers, which is the ultimate target. So there is ample scope for improvement in this well-researched problem. This paper focuses on how a simple high performance recognition scheme, the sn-tuple classifier, can be further improved both in performance and in computational resources' requirements by the use of the recently introduced technology of multiple classifier systems. A complete recognition system involves modules such as image capture, pre-processing, character segmentation, feature extraction, post-processing, etc. The experiments reported in this paper involve only the recognition phase of the pre-segmented characters.

The sn-tuple classifier trained on a chain-code representation of the character images has been shown to achieve very high recognition rates at a reasonably high speed [11, 12]. Unfortunately there exists a trade-off between high performance and increased memory requirements. Motivated by this considerations we have previously proposed two methods [6, 7] to effectively reduce the memory needed while keeping the corresponding performance at a high level. These proposals were based on either ordered or random sampling of the bit planes resulting from the binary representation of the contour chain code. The binary strings obtained formed a less complex, transformed feature space on which sn-tuple classifiers can be trained. The idea of this type of decomposition is based on an approach proposed by Schwarz [14] as a means for compressing data. The realization and success of these proposals were made possible by exploiting the power of multiple classifier systems in which the output of simple and possibly not very accurate classifiers can be combined to provide classification decisions which have been shown to be highly successful in a variety of applications.

Although our previous proposals have been shown to be effective, it should be recognized that every transformation of the initial feature space can cause some loss of information; a fact reflected in the final performance. The information loss due to one transformation can be potentially counter-acted by complimentary information encapsulated through a different technique. So, in this work, we propose a transformation of the chain code significantly different from the previous two. This new approach is concerned with extracting and decomposing the directional information of the original chain code. Utilizing the flexibility offered by multi-classifier schemes we can subsequently exploit the potential complementarity of the different transformations of the original feature space while at the same time decompose it to simpler input spaces thus reducing the resources' requirements of the sntuple classifiers used. So the character recognition system proposed at the end consists of such groups of classifiers trained on the input spaces obtained from the three transformation methods and their

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 **IEEE**

decisions are fused according to a number of combination schemes chosen to be simple enough so that they will not add excessive computational load to the system.

The following sections of the paper first briefly describe the sn-tuple classifier and introduce the decomposition methods used. The paper then proposes a multiple classifier scheme to combine the resulting classifiers, presenting results obtained over a series of cross-validating experiments using off-line handwriting images. The paper ends with a concluding discussion.

2. The Scanning n-tuple Classifier (sn-tuple)

The Scanning n-tuple Classifier is an n-tuple based classifier. It has been introduced by Lucas *et al.*[11] and is shown to have achieved very high recognition rates while retaining many of the benefits of the simple n-tuple algorithm. In an sn-tuple system, each sn-tuple defines a set of relative offsets between its input points which then scans over a 1-D representation of the character image. This unidimensional model of the character image is obtained by tracing the contour edges of the image and representing the path by Freeman chain-codes [3]. In the case of multiple contours, all strings are mapped to a single string by concatenation after discarding the positional information. As different characters produce contour strings of widely varying lengths, all these chains are proportionately expanded to a predefined fixed length. Details of the sn-tuple classification algorithm (including pseudo-code) can be found in [12].

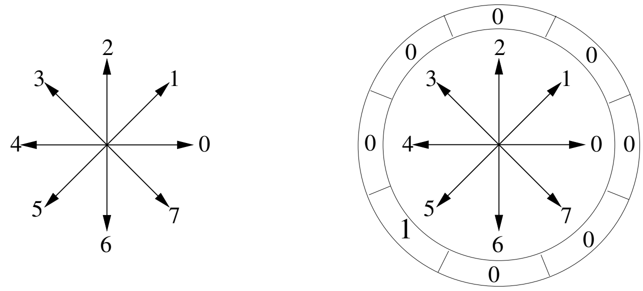
3. Transformation of the Contour Chain-code String

In the following we give a brief description of the three transformation strategies to be used in our proposed multiple classifier system.

3.1. Ordered and Random Decomposition Technique

These two transformation strategies have already been reported by the authors [6, 7, 16, 5]. For the sake of completeness, these are described here again very briefly.

In both the cases, the Freeman direction codes are represented in binary. (It is also possible to use other forms of binary notation, for example, Gray coding, etc., to express the direction codes prior to decomposition and classifier performance is sometimes dependent on this choice. See [5] for details). Since there are 8 possible distinct direction codes, 3-bit binary numbers are sufficient to represent them. For Ordered decomposition (also known as *bit plane decomposition*, the chain-code string is decomposed into 3 separate



Freeman code (a) | Directional Filtering (b)

...56656656671070122233223455565556...

(c) an arbitrary chain code segment

...10010010000000000000000011101110...

(d) Directional layer-5 of the above

Figure 1. Directional Quantization of Contour Chain-code String (example for direction/layer 5)

strings (called *Layers*) such that *Layer*-'i' is composed only of the 'i'th bits of the corresponding direction code.

On the contrary, in the Random Decomposition technique (which is based on Random Subspace Method [4]), bits for decomposed layers are chosen arbitrarily from the Freeman direction-codes. Since the same bits must always be chosen from a given contour position, an array of randomly selected numbers from the set {0,1,2} is generated identifying the bit to be sampled from the corresponding chain element. An arbitrary number of templates can be generated, hence the random transformation approach can create many different binary layers, while for the ordered scheme only as many layers as the number of bits representing each symbol in original chain code can be generated.

3.2. Proposed Directional Filtering Transformation

The proposed Directional Quantization approach in fact is a filter detecting the presence (as well as the location) of a particular direction code in a given contour chain. Since there are 8 possible Freeman directions, only 8 directional layers can be created. If C_f denote a contour chain using Freeman codes d_k such that,

$$C_f = d_1 d_2 \ldots d_k \ldots d_N$$

where $d_k \in \{0, 1, \ldots, 7\}$, and N is the length of chain.

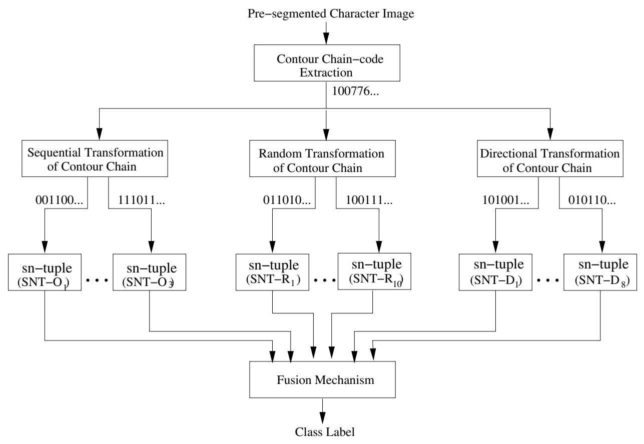


Figure 2. A Schematic of the Proposed System.

If i be the direction of interest, then after filtering,

$$C_d^i = b_1 b_2 \dots b_k \dots b_N$$

where

$$b_k = \begin{cases} 1, & \text{if } d_k = i \\ 0, & \text{otherwise.} \end{cases}$$

Figure 1 illustrates this transformation.

4. Proposed Multi-Classifier Architecture

The proposed recognition system is illustrated in Figure 2. It is a parallel combination of many sn-tuple classifier implementations. Unlike the conventional sn-tuple classifier [12], where the classifier is trained on the Freeman chain-code representation of the character contours directly, the sntuple classifiers used in the proposed system are trained on the binary transformed strings of the Freeman chain-code as described in Section 3. Finally, the decision fusion stage combines the outputs of these individual classifiers and generates a final class label for the test image.

5. Experiments and Discussion of Results

A database of pre-segmented handwritten character images has been used for the experiments [9]. This database consists only of digits and uppercase letters with no distinctions made between '0'/'O' and '1'/'I' character pairs. There are 300 binary images for each character class, each of resolution 24×16 pixels. This data set is randomly partitioned into two disjoint sets for training and testing. For the experiments reported here, 150 images per character class were used for the training. Experiments were conducted in two task domains, the first involving only the digits and the second involving all 34 alphanumeric characters. For multi-classifier combination, five fusion schemes were investigated. These are *sum, median, min, max, and simple majority voting*. Details of these rules can be found in [10]. All the tables presented forthwith include performance statistics averaged over the 10-fold cross validation experiments performed on the unseen test sets from the above database.

Tables 1 and 2 show the mean error rates achieved by the individual classifiers trained on the different input data (layers) produced by the decomposition process for each one of the three transformation methods we described. There are two important observations to be made here about the directional filtering approach we proposed here. First, both in the 10 class (digits) and the 34-class (alphanumeric) cases, significant performance differences can be observed among the classifiers trained on different layers. It seems that layers 1,3,5,7 corresponding to the main directions (0,2,4,6 in Figure 1) are more informative and hence more successful in discriminating the shapes of different characters. Second, it is not difficult to observe that directional filtering is the worst performing among the three decomposition schemes as far as individual layer performances are concerned. However, if we move to Table 3, we can easily observe that , at least for three of the 5 fusion rules tested (sum, median, and majority vote), the combination of the classifiers trained on the layers resulting from directional filtering exhibit significantly better performance. In fact, directionally filtered layering proved to be the best of the three transformation methods in both the digit and alphanumeric examples.

Table 4 presents error rates corresponding to combinations of the classifiers, based on the sum-rule, trained on the layers obtained by the proposed directional decomposition (indicated as SNT-D in the table) with either those trained on the ordered layers (SNT-O) or those trained on the random layers (SNT-R), and finally with both these groups (last row in the table). For ease of comparison, we also included the original sntuple results in the first row. Before proceeding to discuss further the results in this table, it is important to make the following observation. Since, as shown in Tables 1 and 2 the individual classifiers of the directional layering are the worst performers, one should reasonably expect that if they do not encapsulate information complimentary to that of the classifiers obtained form the other two decomposition schemes, their combination should not result in any performance improvement. However, all combinations presented in Table 4 exhibit significant improvements not only with respect to individual classifiers but also with respect to the combinations in Table 3 and the original sntuple. This observation, we believe, supports the existence of complimentarity in the information captured by the three decomposition methods presented.

Table 5 presents a comparative study of error rates from a diverse set of classification schemes. In addition to the er-

Table 1. Individual performance for Numeral Data Set (error rates)

Decomposition Method	Layer No. 1	2	3	4	5	6	7	8	9	10
Ordered (O)	23.89	21.28	22.45	–	–	–	–	–	–	–
Random (R)	25.31	27.39	27.09	27.79	26.21	25.97	27.23	27.60	26.21	27.44
Directional (D)	30.48	39.81	27.68	49.95	27.47	43.49	23.92	45.15	–	–

Table 2. Individual performance for Alphanumeric Data Set (error rates)

Decomposition Method	Layer No. 1	2	3	4	5	6	7	8	9	10
Ordered (O)	44.92	40.68	42.09	–	–	–	–	–	–	–
Random (R)	43.05	43.54	42.33	44.29	41.91	41.93	44.32	45.12	43.07	45.04
Directional (D)	56.71	67.90	46.50	68.70	51.95	66.58	41.80	64.64	–	–

Table 3. Effect of fusion of the sn-tuple classifier trained on the three transformed feature spaces

Fusion rules applied	Classification Error Rates					
	Numerals			Alphanumerics		
	Ordered	Random	Directional	Ordered	Random	Directional
Sum	11.41%	17.28%	10.93%	23.69%	30.45%	22.10%
Median	13.07%	17.44%	12.43%	28.30%	31.04%	26.89%
Min	14.03%	22.43%	23.39%	30.55%	37.80%	43.83%
Max	19.65%	21.57%	24.80%	40.96%	40.58%	52.19%
Majority Vote	15.73%	17.09%	13.95%	34.97%	31.47%	31.21%

Table 4. Mean error rates of the proposed multi-classifier system

Layers Fused	Classification Error Rates	
	Numerals	Alphanumerics
Original Sntuple	4.59%	12.42%
SNT-D + SNT-R	4.08%	10.96%
SNT-D + SNT-O	3.28%	9.93%
SNT-D + SNT-O + SNT-R	2.96%	8.96%

Table 5. Comparison with Different Classifiers

Classifiers	Classification Error Rates	
	Numerals	Alphanumerics
Proposed system	**2.96%**	**8.96%**
FWS	10.00%	22.28%
sn-tuple	4.59%	12.42%
MPC	15.00%	21.29%
MLP	7.20%	18.22%
MWC	5.84%	14.59%
GA	3.40%	8.62%

ror rates of the proposed architecture, error rates achieved by six other classifiers are shown. The conventional sn-tuple scheme has already been described in this paper. The FWS [2] is a frequency-weighted n-tuple classifier using 10 bit tuples. The MLP is the standard multi-layer perceptron [8] trained using Back-propagation learning. The reported MLP had 40 hidden nodes and used zonal pixel density as the input feature. The MPC is a Maximum-likelihood-based [13] statistical classifier which explores possible cluster formation with respect to a distance measure. The particular implementation reported here used the Mahalanobis distance metric calculated on geometric moments (up to 7th order) of the binary image as features. The MWC [1] is an n-tuple based system where features are extracted from a sub-image isolated by a scanning window. The reported results are based on an MWC using a 21×13 pixel window and 12-tuples. The GA is a parallel multiple classifier system optimized using genetic algorithm techniques, originally introduced in [15]. It is readily evident from Table 5 that the proposed layout is capable of producing very low error classification decisions. When compared to the GA which uses a trainable fusion scheme, the proposed system performed very favourably even though it is using a very simple fixed (i.e., non-trainable) fusion mechanism. The proposed architecture outperformed the GA system in numeral classification and error rates are very comparable in the 34-class task.

6. Conclusion

This paper proposes a novel off-line handwritten character recognition system consisting of a number of sntuple classifiers using a simple decision fusion scheme. All these sntuple classifiers are trained on a transformed feature space. For this, in addition to the two previously reported techniques, a third strategy named *Directional filtering* has been introduced. All of these transformation schemes involve very simple operations and as such, the computational overhead is very limited. The recognition accuracy of the proposed scheme not only outperformed the sntuple classifier trained on the un-decomposed Freeman chain-code but also outperformed a number of other popular high performing classification algorithms.

The principal merit of the proposed architecture is in the savings of physical memory space requirement. Where the conventional sntuple implementation demands the presence of a large physical memory (in the order of 8^n, where n is the tuple size). Whereas, for sntuple classifiers using the transformed feature space this requirement is in the order of 2^n. Thus, in spite of using 21 distinct sntuple classifiers, the total memory requirement is slightly more than 2% of the physical memory needed by a single sntuple system using the original chain code.

The third point of concern may be from the classification speed viewpoint. The major computational load in an sntuple classifier comes from the extraction of the contour chain code from the off-line image and the computational load for the actual classification task is very small. In our proposed model, the chain-code needs to be extracted only once (the same as in conventional sntuple classification). The additional computation for feature transformation and classification by the individual sntuple classifiers are very fast and the overall computational speed remains nearly the same to that of conventional sntuple classifier.

Acknowledgements

The authors gratefully acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC).

References

[1] M. C. Fairhurst and M. S. Hoque. Moving window classifier: Approach to off-line image recognition. *Electronics Letters*, 36(7):628–630, March 2000.

[2] M. C. Fairhurst and H. M. S. A. Whab. An approach to performance optimisation in frequency-weighted memory network pattern classifiers. *Electronics Letters*, 23:1116–1118, 1987.

[3] H. Freeman. Computer processing of line-drawing images. *ACM Computing Surveys*, 6(1):57–98, March 1974.

[4] T. K. Ho. The random subspace method for constructing decision forests. *IEEE Trans Pattern Analysis and Machine Intelligence*, 20(8):832–844, 1998.

[5] M. S. Hoque and M. C. Fairhurst. Face recognition using the moving window classifier. In *Proceedings of the 11th British Machine Vision Conference (BMVC2000)*, volume 1, pages 312–321, Bristol, UK., September 2000.

[6] S. Hoque, K. Sirlantzis, and M. C. Fairhurst. Bit plane decomposition and the scanning n-tuple classifier. In *Proceedings of 8th International Workshop on Frontiers in Handwriting Recognition (IWFHR-8)*, pages 207–211, Niagara-on-the-Lake, Ontario, Canada, 6-8 August 2002.

[7] S. Hoque, K. Sirlantzis, and M. C. Fairhurst. Intelligent chain-code quantization for multiple classifier-based shape recognition. In J. A. Bullinaria, editor, *Proceedings of the 2002 U.K. Workshop on Computational Intelligence (UKCI-02)*, pages 61–67, Birmingham, United Kingdom., 2-4 September 2002.

[8] A. K. Jain, J. Mao, and K. M. Mohiuddin. Artificial neural networks: A tutorial. *Computer*, pages 31–44, March 1996.

[9] Digital Systems Research Group (DSRG), Department of Electronics, University of Kent, Canterbury CT2 7NT, United Kingdom.

[10] J. Kittler, M. Hatef, R. P. Duin, and J. Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, March 1998.

[11] S. Lucas and A. Amiri. Recognition of chain-coded handwritten character images with scanning n-tuple method. *Electronic Letters*, 31(24):2088–2089, November 1995.

[12] S. Lucas and A. Amiri. Statistical syntactic methods for high performance ocr. *IEE Proceedings Vision, Image and Signal Processing*, 143(1):23–30, February 1996.

[13] R. Schalkoff. *Pattern Recognition: Statistical, Structural and Neural Approaches*. John Wiley and Sons, New York, 1992.

[14] J. W. Schwarz and R. C. Barker. Bit-plane encoding: A technique for source encoding. *IEEE Transaction on Aerospace and Electronic Systems*, 2(4):385–392, 1966.

[15] K. Sirlantzis, M. C. Fairhurst, and M. S. Hoque. Genetic algorithms for multi-classifier system configuration: A case study in character recognition. In J. Kittler and F. Roli, editors, *Multiple Clasifier Systems (Second International Workshop, MCS 2001, Cambridge, UK)*, volume 2096 of *Lecture Notes in Computer Science*, pages 99–108. Springer, July 2001.

[16] K. Sirlantzis, S. Hoque, and M. C. Fairhurst. Classifier diversity estimation in a multiclassifier recognition system based on binary feature quantization. In *Proceedings of 4th International Conference on Recent Advances in Soft Computing (RASC 2002)*, pages 84–89, Nottingham, UK., 12-13 December 2002.