A MOVING WINDOW CLASSIFIER FOR OFF-LINE CHARACTER RECOGNITION

M. S. HOQUE, M. C. FAIRHURST
Electronic Engineering Laboratory, University of Kent, Canterbury, Kent CT2 7NT, United Kingdom.
E-mail: {msh4,mcf}@ukc.ac.uk

A new classification scheme, primarily aimed at applications in document image processing, is presented. Features are extracted from a partial image and a sub-classifier generates scores based on the likelihood of the sub-image belonging to the candidate classes. This partial classification is carried out for several overlapping image segments and scores are combined to make the final classification. The scheme shows promising results in OCR applications where high processing speeds are achievable with minimal compromise in the recognition accuracy.

1 Introduction

Automatic recognition of document images has witnessed significant progress in the last few decades, and industry is employing more and more automated computer vision techniques to solve real life problems in commercial environments. Among the stringent requirements usually associated with such systems, high classification accuracy and high speed processing are of particular significance. High reliability is achievable but usually at the expense of extensive and complex computations which in turn makes the processing slow. Thus, it is extremely important to devise a system that is highly accurate but, simultaneously, the processing demands must be restrained to make the adoption of the system economically viable. This paper focuses on the off-line recognition of pre-segmented characters and proposes a classification scheme that is capable of producing high recognition rates at high speed.

2 Proposed Methodology

The proposed Moving Window Classifier(MWC) scheme is illustrated in Figure 1. First a window smaller than the image is defined. Only a portion of the image is visible through this window. Features are extracted from this part image and a Sub-Classifier assigns scores corresponding to the likelihood of the pattern viewed belonging to the individual classes. Then the window is shifted and classification is carried out for the new part image visible. Thus, the window is moved left to right and top to bottom in single pixel displacement steps until the entire image is covered and part classification is carried out for all different window positions. A decision fusion stage then combines

In: L.R.B. Schomaker and L.G. Vuurpijl (Eds.), Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition, September 11-13 2000, Amsterdam, ISBN 90-76942-01-3, Nijmegen: International Unipen Foundation, pp 595-600

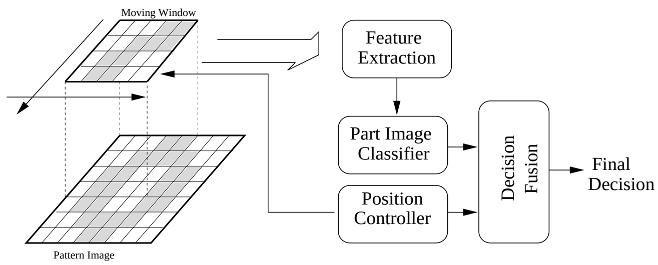

Figure 1: The Moving Window Classifier (MWC).

these partial classification scores and, accordingly, assigns a class label to the test image as a whole.

The MWC uses the n-tuple memory network classifier [1,2] at its core. In an n-tuple classifier, n-tuples are formed by selecting multiple sets of n distinct pixels from a pattern space. During the training phase, the system keeps count of the different pixel-patterns each individual n-tuple encounters. During the testing phase, the sum of the counts for the pixel-patterns in the test image is computed and the image is assigned to the class generating the highest score.

3 Experimental Results

To study the characteristics of the MWC, experiments were conducted for the classification of pre-segmented characters consisting of digits and uppercase letters only. No distinction is made between '0'/'O' and '1'/'I' character pairs. Two character image databases were used. Database A consists of machine printed character images (extracted from post codes on envelopes in the UK mail [3]) and contains 300 images per character class each of resolution 24×16 pixels. The database is divided randomly into two disjoint sets of 200 and 100 images for training and testing respectively. Database B is the standard NIST database [4] consisting of handwritten character images. There are 30899 images for training and 20667 for testing, each of resolution 32×32 pixels.

Individual n-tuples were randomly selected. Performance varied significantly with the chosen mapping, and selection of train/test divisions also had a considerable effect on the recognition performance. Therefore, all experiments were carried out with several different mappings and train/test divisions and the reported results correspond to arithmetic means of the individual performances. To keep computation to a minimum, all pattern normalization measures such as slant correction, broken-image and stroke thickness mending, etc. were avoided. The recognition performances of the MWC scheme on

Table 1: Performance of the Proposed MWC Classifier.

Criterion	Numerals only		Alpha-Numerics	
	Database A	Database B	Database A	Database B
Accuracy	99.8%	89.4%	99.3%	75.5%
Throughput	450cps	150cps	210cps	66cps

Table 2: Effect of the Window Size on MWC Recognition Rate (n=12).

		Row pixels							
		32	31	30	29	28	27	26	25
column pixels	32	85.08	86.40	86.89	87.02	86.73	86.35	85.96	85.56
	31	86.25	87.57	87.76	87.67	87.41	86.85	86.49	86.16
	30	87.04	87.99	88.33	88.13	87.97	87.37	87.03	86.56
	29	87.67	88.37	88.62	88.48	88.14	87.62	87.26	86.86
	28	87.74	88.66	88.94	88.62	88.32	87.83	87.36	87.12
	27	88.02	88.66	89.36	88.72	88.42	87.82	87.54	87.26
	26	88.07	88.77	88.96	88.82	88.33	87.96	87.49	87.20

the two data sets are shown in Table 1. These performances were achieved with n=12 using 21×13 window for Database A and 30×27 window for Database B. There are several important parameters which determine the performance of the proposed MWC scheme, and the following sections describe an experimental investigation of the principal factors determining attainable performance.

3.1 Effect of Window size

The size of the moving window is an important criterion in determining classifier performance. If the window size equals the image size, the scheme becomes the standard n-tuple scheme. Table 2 illustrates the variation of recognition rates under different window resolutions as tested on the NIST data set (Numerals only). The peak performance is 89.4% at the window resolution of 30×27. At the same resolution, the accuracy is 75.5% for the alphanumeric character set. For Database A, the peak performances are 99.8% and 99.3% respectively at a window resolution of 21×13.

3.2 Throughput

Speed (Throughput) is the performance criterion often overlooked when reporting high performance OCR results. The proposed MWC classifier aims at achieving high speed without unduly compromising the accuracy achievable. The MWC scheme was implemented on a SUN Ultra 10 workstation

Table 3: Effect of Feature Dimension Reduction.

k	Numerals only				Alpha-Numerics			
	Database A		Database B		Database A		Database B	
	Accuracy %	Speed cps	Accuracy %	Speed cps	Accuracy %	Speed cps	Accuracy %	Speed cps
0	99.70	450	89.36	150	99.20	210	75.47	66
20	99.71	620	89.20	186	99.15	260	75.25	82
40	99.68	770	89.11	255	99.12	370	75.02	113
60	99.62	1150	88.85	376	99.01	480	74.68	170
75	99.54	2140	88.12	630	98.65	1000	73.81	286

with 640MB real memory running under the Solaris 7 operating system. The speed of classification depends on the window sizes used as well as on the resolution of the character images. The smaller the window chosen, the slower it becomes, since a larger scan-time is required for the entire image. A 21×13 window with database A generates optimal accuracy with a throughput of 450 characters per second(cps) for the numeral classes and 210 cps for the alphanumeric classes. For the NIST data set (database B), the throughput figures at maximum accuracy are 150cps and 66cps respectively. The reason for lower throughput is that the NIST image resolution is larger (by a factor of 2.67) and correspondingly more n-tuples need to be processed.

Improving Processing Speed

Two principal factors, other than the image and the moving window resolution, significantly affect the speed of the proposed MWC scheme. Firstly, the dimension of the feature vector (here, the number of n-tuples per window) and the number of part classifications. Both of these parameters were varied to evaluate their effect on the overall performance.

Reducing the Feature Dimension : Initially, all available pixels in the window were organized into disjoint n-tuples. In order to improve the speed, k% of the available pixels were arbitrarily masked, thereby reducing the feature dimension by k%. The resulting recognition rates and corresponding throughput are presented in Table 3, for a window size of 21×13 pixels. It is evident that a significant gain in speed is achieved at the cost of only a small loss in recognition accuracy. For machine-printed characters, 250% throughput gain is achieved (at k=60) for about 0.20% loss in the recognition accuracy. For the hand-printed characters, similar throughput gains are achieved at a somewhat larger ($\approx$ 0.5% for numerals and $\approx$0.8% for alphanumeric) loss in accuracy.

Reducing Part Classifications : Alternatively, the number of part classi-

Table 4: Effect of Reduced Feature Dimension with scan step=2

k	Numerals only				Alpha-Numerics			
	Database A		Database B		Database A		Database B	
	Accuracy %	Speed cps	Accuracy %	Speed cps	Accuracy %	Speed cps	Accuracy %	Speed cps
0	99.58	1690	88.61	387	98.69	690	74.53	159
20	99.60	2085	88.47	485	98.60	860	74.30	198
40	99.59	2890	88.31	661	98.46	1215	73.94	274
60	99.40	4250	88.00	984	98.22	1790	73.39	415

fications carried out can be reduced by increasing the step size by which the moving window is shifted. By simply moving the window in 2-pixel steps, the number of computations can be reduced by a factor of 4. This leads to some loss in accuracy of the system because of the sacrifice of some discriminating information that could have been derived from those skipped positions. Using 2-pixel steps and a 21×13 window with database A, the throughput increases to 1690 cps (from 450 cps) for numerals and 690 cps for alphanumerics. The recognition accuracies are shown in the first row of Table 4, which also demonstrates the effect of combining both speed enhancement schemes.

3.3 Performance under Forced Rejection

In many practical applications, it is essential to maintain a very low error rate. This is often achieved by rejecting the classification of certain test patterns when the confidence is low. The use of a 'reject class' can help reduce the misclassification rate in tasks where exceptional handling (for example, by another automatic expert or a human) of particularly ambiguous cases is feasible. In this study, patterns were rejected when the highest combined score and the succeeding score differ by less than a threshold 'θ'. Error rates can now be arbitrarily lowered by increasing θ, but with a consequent increase in the rejection rate. Table 5 illustrates the effect of forced rejection on error rates for the proposed MWC classification algorithms as tested on dataset B.

4 Conclusion

A new classification algorithm which is particularly suited to OCR applications for document processing has been presented. Although it may be possible to identify techniques which can outperform the MWC scheme with respect either to classification accuracy or pattern throughput alone, the trade-off achievable by MWC between these two factors shows very favourable characteristics, since

Table 5: Error rates under forced rejection.

Rejection rate	Database A			Rejection rate	Database B	
	C=10	C=34			C=10	C=34
0%	0.2%	0.74%		0%	10.88%	25.52%
0.5%	-	0.47%		5%	8.38%	22.29%
1%	0.1%	0.31%		10%	5.79%	19.24%
2.5%	-	0.15%		20%	2.51%	13.72%
5%	0%	0.09%		50%	0.15%	3.50%

significant gains in throughput can be demonstrated at minimal cost in accuracy. MWC performance has been compared against that achievable with a range of other classification algorithms implemented in the same processing environment. The MWC outperformed most of these schemes when both accuracy and speed are considered simultaneously.[5]

Although the technique is aimed particularly at OCR applications, it is clearly applicable to a range of other task domains and, indeed, is likely to be particularly suited to problems which can currently be addressed using systems based on n-tuple processing.[6] Preliminary work on face recognition using the MWC approach has already shown very promising results.[7]

References

1. S. Lucas and A. Amiri. Statistical syntactic methods for high-performance ocr. *IEE Proc on Vision, Image and Signal Processing*, 143(1):23–30, 1996.
2. M. C. Fairhurst and T. J. Stonham. A classification system for alpha-numeric characters based on learning network techniques. *Digital Processes*, 2:321–339, 1976.
3. Image Processing & Computer Vision, Electronic Engineering, University of Kent, Canterbury, CT2 7NT, UK.
4. NIST Special Databases 1-3, 6-8, 19, 20, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA.
5. M. C. Fairhurst and M. S. Hoque. Moving window classifier: a new approach to off-line image recognition. *Electronics Letters*, 36(7):628–630, 2000.
6. M. Morciniec and R. Rohwer. The n-tuple classifier: Too good to ignore. Technical Report NCRG/95/013, Aston University, UK, 1995.
7. M. S. Hoque and M. C. Fairhurst. Face recognition using the moving window classifier. In *Proc of 11th British Machine Vision Conference (BMVC2000)*, Bristol, UK, 2000. (*to appear*).