MVA '96 IAPR Workshop on Machine Vision Applications, November. 12-14, 1996, Tokyo, Japan

Cooperation between Patterns Recognition and Contextual Analysis : Application to the Reading of Printed Documents.

HENRY Jean-Luc, LEBOURGEOIS Frank
Laboratoire de Reconnaissance de Formes et Vision
I.N.S.A. de LYON - Bât 403
20, Avenue A. Einstein 69621 Villeurbanne Cedex FRANCE
Tél : (+33) 72 43 80 93 Fax : (+33) 72 43 80 97
E-mail : flebourg@rfv.insa-lyon.fr

Abstract

This paper presents a cooperation between the recognition stage and the contextual analysis stage for printed documents recognition which makes contextual correction easier and the continuous training of the recognition stage possible. In the first part we recall the particular organization of our implementation which is based on a preclassification of identical characters patterns called prototypes. The second part describes the information provided by the recognition stage to the contextual stage in order to improve the syntactic correction. This information is represented by an « initial recognition table » which gives all the hypotheses about the recognition of characters associated with a confidence rate computed by the recognition stage. We also demonstrate that the utilization of this table is more efficient for a syntactic stage using the Viterbi algorithm than the substitution table generally computed in advance on a large collection of characters. The third part describes a backloop from the contextual analysis stage towards the recognition stage. This will allow the modification of the decision rules and the correction of the interpretation. It constitutes the last link of the information chain that forms the cooperation between the context and the recognition.

1 Introduction

It is generally agreed that the pattern recognition is a complex process which used graphic and contextual information. But most of documents reading machines make a separation between the optical recognition of characters and the contextual processing. A classic process of recognition of characters chains sequentially operations of isolated characters segmentation, features extraction, comparison to models patterns, reconstruction of words and finally syntactic word correction. For few years, we have proposed to change this organization such as to improve the performances by introducing backtrack loops between the stages of the reading system. The idea is to make a cooperation between different stages of the reading system and to integrate diverse sources of knowledge during each step of the process. Some works already deal with the combination of information ; for examples between the recognition and the segmentation stages in order to separate composite characters patterns. In this paper, we present a cooperation between the recognition stage and the contextual stage limited to a syntactic word correction. We have study the links which are of interest between these two stages.

2 Preclassification of identical patterns

Context cannot be reduced to grammar or syntactic rules, we define the context by all the clues which can help the recognition task. For examples, during natural reading , people generally memorize styles and deformations of all characters patterns. This training helps us to recognize degraded characters which cannot be recognized alone. For handwritten text, this example is especially appropriate when one recognizes some bad scripted characters because the pattern has been recognized by deduction from some words at different locations in the text.

We have already tested this natural human task in early works [1] by implementing a preclassification stage which identifies all identical patterns in the text. In images of printed document, the characters patterns are redundant and it is useless to recognize all characters designed by identical styles and fonts as well as to develop the same recognition process on image of similar patterns. The preclassification of identical patterns of characters use a method introduced by R.G. Casey [2][3]. We call prototypes the different patterns of printed characters found in a document. This step has three advantages :

- It reduces the number of patterns to be recognized. Generally we compute less than few hundred of prototypes for several thousand of characters during the reading of the first page. The number of prototypes is stationary after the reading of few pages of the same document. The number of prototypes depends on the quality of the document and on the number of fonts used.

- It gives the capacity to learn all fonts and their declinations (body, bold and placement in italic) and also the deformations caused by noises, allowing the processing of various documents.

- It helps the contextual processing by providing relevant information about patterns redundancies and their positions in the text. We have proposed a specific contextual processing [4] based on the syntactic analysis of all the words in the text which contain a particular character prototype. As a character prototype can be found in several words, we improve the correction rate of the syntactic stage and we can also use a limited dictionary for a large variety of documents. We have shown that the number of words found in the dictionary which contain the same identified prototype is more efficient than a word by word validation.

This previous works show that we can use more contextual information in different steps of a reading system.

In spite of these improvements, we have noticed that the recognition stage makes the same errors on same patterns and that the contextual stage supports all the computational effort to correct always the same errors. These errors are explained by the difficulty for a recognition stage to modelize all possible deformations of characters patterns (numerous styles and fonts exist, bad quality of the print...).

The proposed idea consist to use the errors detected by the contextual stage to retrain the recognition stage. Moreover we can also improve the information provided by the recognition stage to the contextual stage in order to reduce the complexity of syntactic correction.

3 Information from the recognition stage to the contextual analysis stage

The contextual analysis stage used a classic syntactic correction module based on the Dictionary Viterbi Algorithm (D.V.A) [5][6]. This algorithm mainly use the Viterbi algorithm associated with a dictionary which reduces the search in the Viterbi net for valid words. But the Viterbi Algorithm use transition probabilities between characters computed from the dictionary. The V.A. also needs characters substitution probabilities generally computed from the substitution matrix. For each character, one limits the number of hypotheses by tresholding the substitution matrix. The treshold strategy defines the number of assumptions to check in the Viterbi net, the correction rate, and also computational complexity.

We have noticed that the treshold is difficult to define without more specific information about the real substitution rate computed from the document itself. The problem is that we cannot evaluate the substitution rate before the contextual processing. So we generally compute a substitution matrix on a large collection of characters from different fonts and styles. But the substitution matrix results are not specific to the patterns of characters which must be recognized from the input document. For example, generally confused characters (**'l','1','l'**) can easily be recognized for particular fonts. Then it is not necessary for these fonts to consider all these substitution cases. In the other hand, easy recognizable characters (**'g','8','a'**)(**'h','b'**) can be confused because of the font or the quality of the print.

We propose to substitute an initial recognition table provided by the recognition stage for the confusion matrix. This table gives for each prototype, the list of possible interpretations associated with recognition scores which are specific to the patterns of characters found in the document. These recognition scores give better substitution probabilities for the search in the Viterbi net. As the number of possible interpretations for each prototype is reduced to the real number of substitution cases, we do not need any more a treshold strategy.

Figure 2 illustrates the advantage to use more information on the characters interpretation from the recognition stage instead of using a fix substitution table computed statistically. The first table gives the possible interpretations for the right recognition with scores (Figure 1). We can notice the effect of the identical pattern detection which avoids to present again the same identical printed characters (**'a','l'**) to the recognition stage to correct the word '**allocate**'. This example shows that the hypotheses about the possible identification are more precise than the confusion risks computed from a fix tresholded substitution matrix.

a		l		o		c		t		e	
Prototype 6		Prototype 11		Prototype 23		Prototype 42		Prototype 31		Prototype 81	
identif	Score	identif	Score	identif	Score	identif	Score	identif	Score	identif	Score
a	100%	I l t	40% 35% 25%	o Q	70% 30%	c o	72% 28%	t	100%	e	100%

Figure 1. Initial recognition table provided by the recognition stage to the contextual stage

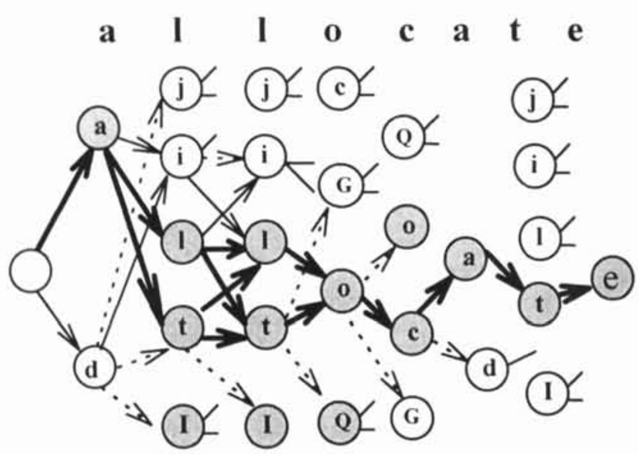

Figure 2. Reduction search in the Viterbi net with the initial recognition table provided by the recognition stage to the contextual stage

Path not explored because the partial word is not valid in the dictionary (DVA)

Path explored by the DVA with fix assumptions given by the substitution matrix

Path explored by the DVA with the initial recognition table provided by the recognition stage

Possible substitution proposed by the recognition stage

Possible substitution proposed by the treshold of the substitution matrix

This approach can be applied to different recognition method which can give a list of possible interpretation associated with a recognition score.

4 Information from the contextual analysis stage to the recognition stage

In order to obtain what is called as the cooperation between the last two stages, we have to construct a feedback loop from the contextual stage to the recognition stage. The contextual analysis stage corrects most characters which have been substituted or rejected by the recognition stage. We use these contextual corrections in order to modify the behaviour of the recognition stage by adjusting the internal representation of character models. This cooperation permits the realization of continuous learning on different character patterns used in various documents.

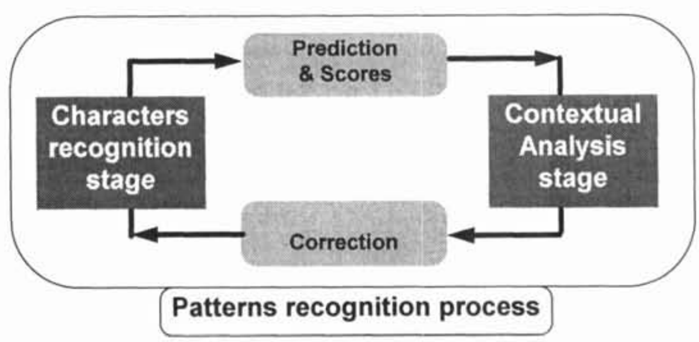

Figure 3. Cooperation between the recognition stage and the contextual stage for a continuous training.

5 Continuous training

In Pattern Recognition field, the learning is essentially done on a limited basis of representative models. The maximal performances of the pattern recognition system depend on the choice of the models. Continuous learning solves the problems of training set limitation and allows the system to reach a sufficient level of knowledge for an omnifont recognition. However, it raises other problems like the ability to learn new observations, the risk of overtraining, the choice of information to preserve, the coherence of representation or the necessary degree of sensitivity to modify the decisions.

However, it is difficult to modify the decision rules for a lot of recognition methods. The continuous training of a recognition stage tends to increase particular cases and this stage risks to loose its ability to recognize general cases. In the same way, it is difficult to work out continuous learning on pattern recognition system based on neural networks, decision trees or chains comparison. From existing methods, it is necessary to choose those that are able to tolerate a progressive and sequential reorganization of their representation or their decision rules.

We have proposed [7] a modified k-nearest neighbours classifier specially designed for the continuous training. For each model, we introduce a weight that corresponds to the importance of its contribution to the decision. The models that contribute to successive failures must observe a diminution of their importance by a reduction of their weight to the benefit of models that contribute to improving recognition. The modification of the weights will be based on the confrontation of the results, between the recognition stage and the contextual stage.

But models representing unusual or rare patterns clutter the representative space unnecessarily and complicate the recognition process. The system select automatically the models according to performance and efficiency criteria. In this way, we can maintain a set of models that participate actively in accurate recognition.

Moreover, models that are rarely used by the system during the recognition process were introduced for the reading of specific texts presenting very particular deformations. These little-used models are also removed automatically in order to keep only efficient and active models.

During the reading of a same document, the results have shown a rise and a convergence of the recognition rate with the training on character patterns corrected by the contextual stage. The absolute limit of the recognition rate after adaptation could be explained by the choice of the features or by the limitation of the syntactic corrections for difficult documents.

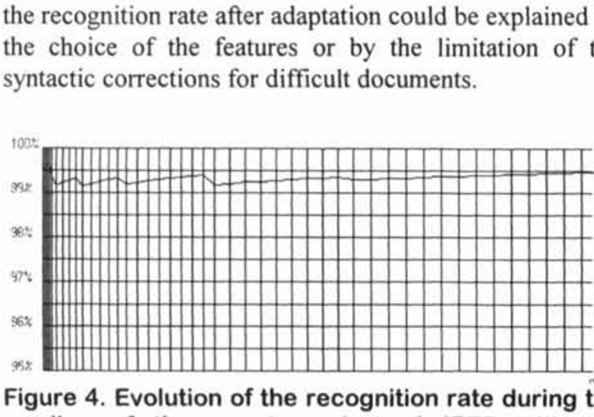

Figure 4. Evolution of the recognition rate during the reading of the spectrum journal IEEE with the continuous training

But if the recognition rate improves continually on average, its value tends to decrease occasionally with texts in different styles until the system adapts once again. These performance jumps are characteristical in our method based on the selection of models according to their efficiency. The mechanism for the destruction of inefficient or little-used models, which economize the memory, leads the system to forgetting its former knowledge.

6. Conclusion

We have proposed a cooperation loop between the contextual analysis stage and the recognition stage. The recognition stage give hypotheses that the contextual stage will check. These information also serve to optimize the syntactic correction by guiding the heuristic search from a real confusion risks linked to the patterns read from the document. The analysis stage can also provide the necessary corrections for retraining the recognition stage. The problem of continuous training has been presented. The solution given in this paper about this difficult problem, show an overtraining of our classifier. The recognition rate increase during the reading of the same document. But we have also observed major drops in recognition performance when we read different documents until the system adapts once again.

References

[1] **LEBOURGEOIS, F.** and **HENRY, J.L.** *An OCR System for Printed Documents.* The Proceedings of the IAPR Workshop on Machine Vision Applications, Tokyo, December 7-9,1992.p.83-86.

[2] **CASEY, R.G.** and **JIH, C.R.** A processor-Based OCR System. *IBM. J. Res. Dev., Vol. 27, N° 4, July,* 1983, p. 386-399.

[3] **BRICKMAN, N.F.** and **ROSENBAUM, W.S.** Word Autocorrelation Redundancy Match (WARM) Technology. *IBM. J. Res. Dev.*, 1982, Vol. 26, N° 6, p. 681-686.

[4] **LEBOURGEOIS, F.** and **HENRY, J.L.** *A Contextual Processing for an OCR System, Based on Pattern Learning.* In : The Proceedings of the 2nd International Conference on Document Analysis, Tsukuba Science City, October 20-22, 1993. p. 862-865.

[5] **HULL, J.J.**, **SRIHARI, S.N.** and **CHOUDHARI, R.** An Integrated Algorithm for Text Recognition : Comparison with a Cascaded Algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1983, Vol. 5, No. 4, p. 384-395.

[6] **SINHA, K.**, **BIRENDA**, and **PRASADA**, Visual text recognition through contextual processing *pattern Recognition,* 1988, Vol 21,N°5, p463-479

[7] **LEBOURGEOIS, F.** and **HENRY, J.L.** A*n Evolutive OCR system Based on Continuous Learning* Third IEEE Workshop on Applications of Computer Vision, Sarasota Florida, December 2-4 1996