Distinguishing between Handwritten and Machine Printed Text in Bank Cheque Images

José Eduardo Bastos Dos Santos[1,2], Bernard Dubuisson[1], and Flávio Bortolozzi[2]

Heudiasyc – Université de Technologie de Compiègne(UTC)
BP 20529 – 60205 Compiegne cedex France
Tel. 33 3 44 23 44 23 – Fax. 33 3 44 23 44 77
{jose-eduardo.santos,bernard.dubuisson}@hds.utc.fr

[2]LUCIA –Pontifícia Universidade Católica do Paraná (PUCPR)
Rua Imaculada Conceição, 1155
80.215-901 Curitiba – Brasil
Tel. 55 41 330-1543 – Fax. 55 41 330-1392
{jesantos,fborto}@ppgia.pucpr.br

Abstract. In the current literature about textual element identification in bank cheque images, many strategies put forward are strongly dependent on document layout. This means searching and employing contextual information as a pointer to a search region on the image. However human handwriting, as well as machine printed characters, are not dependent on the document in which they are inserted. Components of handwritten and machine printed behavior can be maintained in a generic and independent way. Based on these observations this paper presents a new approach to identifying textual elements from a set of local features enabling the category of a textual element to be established, without needing to observe its environment. The use of local features might allow a more generic and reach classificatory process, enabling it in some cases to be used over different sorts of documents. Based on this assumption, in our tests we used bank cheque images from Brazil, USA, Canada and France. The preliminary results show the efficiency and the potential of this approach.

1 Introduction

The automatic treatment of bank cheques is a task that has been receiving increasing attention from scientists over past years. It remains a complex and challenging task where automatic detection of filled-in information is one of the main causes of difficulties. Up until to now, different solutions have been proposed to resolve these technical problems. Many of these are based on assumptions such as the spatial position of the information sought after on the image, the location of baselines or other kinds of contextual information. In this paper we present a new vision of how to treat the handwritten text identification of bank cheque images based on some local features. The main idea of this methodology centres on the fact that both machine printed and handwritten texts retain some characteristic features irrespective of the environment in which they are printed, i.e. independent of the document type. This allows us to look at these elements in such a way as to design a model able to characterize them without needing to look at other elements on the image besides

D. Lopresti, J. Hu, and R. Kashi (Eds.): DAS 2002, LNCS 2423, pp. 58–61, 2002.

textual objects. This means that we believe it is possible to determine a set of features devoted to textual elements identification and discrimination, with a strong degree of independence of the documents involved. Once we have defined this set, a classification process based on this feature group, can be applied in a generic way over an wide variety of documents.

This paper describes our experiences of obtaining a condensed yet sufficient group of features applied to handwritten text extraction. The documents used to test the selected features are bank cheque images presenting a great variety of different backgrounds, fonts and handwritten styles.

Preliminary results show the efficiency of our approach since cheque images from different banks and different countries were tested, allowing a more extensive view of features performance.

2 How to Observe Local Features

As mentioned in the previous section we observe some features locally in textual elements. These elements are isolated from the background by a morphological tophat that aims to suppress slow trends while enhancing the contrast of some elements on the image.

In order to observe selected features we have to divide the image into small portions containing only individual parts of textual elements. Firstly the entire image is equally shared between small frames of 11 x 11 pixels each. This can cause some bad apportionment since frames containing minute portions of elements or more than one object can be obtained. In this latter case we can isolate each of the elements on the frame image and treat them as individual elements. Where the portions of an element in a single frame are too small, we carry out an aggregation phase responsible for re-organizing portion sub-division. This gives a better representation of isolated objects. These small images would normally contain a portion of a handwritten or pre-printed character, since having portions of two different objects in the same frame can cause problems during the classification phase, since this is also based on some geometrical features. A well-divided image avoids the need for an eventual verification of some parts that have been poorly classified due to an inadequate representation on the frame (too few pixels).

Since we have an image that is well divided up into small elements, we will extract from these, some features representing their shape and their contents. Statistical features normally provide enough information about content, which due to different printing processes for handwritten and machine printed text, should provide a reliable and important discrimination factor. Mean and standard deviation are classical examples of content representation, since the ink distribution changes from printed to handwritten characters.

Even if we are not considering the characters on the image in its integrality, shape is another source of information that should not be forgotten when characterizing the two different types of text. Eccentricity and symmetry (related to the convex area) are examples of shape descriptors extracted from the objects.

In order to evaluate the pertinence of the features set chosen, we used a multi-layer perceptron based classifier. The input data set was composed of selected features extracted from the sample images. Different configurations were tested and the results proved not to be very dependent on network architecture. The diversity of the database needs to be underlined, since it was composed of cheques from more than 30 different banks based in four different countries. The features employed were selected using a neural network based method which considers how the algorithm and the training sets interact. Nine mainly geometrical and statistical features make up the resulting features sub-set.

Results obtained point to a significant improvement with regard to previous experiments, where frames were considered since their initial division and features arrangements were not taken into account. The images are represented by 256 gray levels and a 300 dpi resolution. More than 6800 sample images, equally divided between handwritten and machine printed text, were used in the classification process. The input data are reduced and centred on the mean after being used. The goal is to distinguish between handwritten or machine printed objects. A tax of nearly 90% of well-classified images was achieved in the test phase. The misclassified group of sample images is well composed of the two represented classes. Ambiguous samples were mainly composed of printed elements. A great diversity of writing styles were tested since the whole database contains more than 700 images from several different banks.

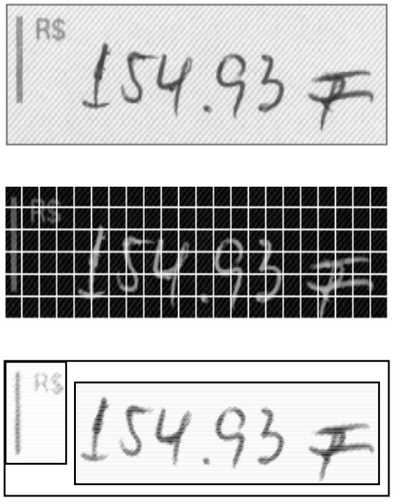

Fig. 1. Extract of a Brazilian bank cheque, its inital partition and final classification result.

Ambiguous samples were treated in order to increase the number of well-classified images. In this case, observing its neighbourhood gave us the probability of each ambiguous sample belonging to the handwritten or the machine printed class. The results were improved. One such example is shown in image 1.

3 Concluding Remarks

In this paper we presented a local feature based textual classification process. The idea of using local features is linked to the assumption that handwritten and machine printed text preserve their general typographical features independent of the context in which they are inserted. Defining an adequate set of features related to textual elements' shape and content allows the identification of these elements in a more generic and context free way.

Our main objective in this paper is to demonstrate the effectiveness of handwriting identification when performed through modelling textual elements. This modelling is carried out by observing some local features in a way that is totally independent of context and which can allow for the use of the current methodology with similar tasks for different kinds of documents.

Using content and shape as discriminative elements for the two kinds of text analyzed, we assume that is possible to verify directly the behaviour of a textual element - handwritten or not - over its component pixels, without needing to observe the environment in which it is inserted. This assumption can be easily verified for example when writing one's name on a form or on another document type. Most people tend to do this in the same manner for every kind of document.

Initial results prove that this methodology contributes a new perspective to the textual element identification of documents which is broad reaching and generic.

We are currently working to improve our features set in order to ameliorate our results. The use of other faster classifiers is also under study.

References

1. John D. Hobby. Using shape and layout information to find signatures, text and graphics. *Computer Vision and Image Understanding*, 80(1): 88–110, October, 2000.
2. J. E. B. Santos, B. Dubuisson and F. Bortolozzi. *Handwritten Text Extraction from Bank Cheque Images by a Multivariate Classification Process*. 6th World Multi Conference on Systemics, Cybernetics and Informatics – SCI'02, Orlando – USA, 2002.
3. Nikolay Gorski, Valery Anisimov, Emmanuel Augustin, Olivier Baret, Sergey Maximov. *Industrial bank check processing: the a2ia check reader*. International Journal on Document Analysis and Recognition, 3(4) :196–206, May, 2001.
4. P. Clark and M. Mirhehdi. *Combining statistical measures to find image text regions*. In ICPR'00, pages 450 – 453, Barcelona – España, 2000.
5. Xiangyun Ye, Mohamed Cheriet and Ching Y. Suen. *A generic system to extract and clean handwritten data from business forms*. In Seventh International Workshop on Frontiers in Handwriting Recognition, pages 63–72, Amsterdam, 2000.