Symbol Recognition by Multi-class Blurred Shape Models

Alicia Fornés, Sergio Escalera, Josep LLadós, and Gemma Sánchez

Computer Vision Center, Dept. of Computer Science, Universitat Autònoma de Barcelona, 08193, Bellaterra, Spain

1 Introduction

One of the main difficulties in the document analysis field is the recognition of handwriten documents. High variability among symbols because of different writer styles, different sizes, shape deformations, noise, or intensity changes are just a few problems. In handwriting recognition language models can be used to assist the recognition process. However, in Graphics Recognition hand drawn symbols use to be classified without context information. Here, two major focus of interest can be stated: firstly, the definition of expressive and compact shape description signatures (Zhang [1] reviews the main techniques used) that guarantee inter-class compaction and intra-class separability, even when describing noisy and distorted shapes. It has been proved that some descriptors, robust with some affine transformations and occlusions in printed symbols, are not efficient enough for handwritten symbols. Secondly, the formulation of robust classification methods according to such descriptors is required. One of the most well-known techniques in this domain is the Adaboost algorithm due to its ability for feature selection, detection and classification problems [2]. To design a single multi-classifier is a difficult problem, and it is common to conceive just binary classifiers and combine them in some way. One-versus-one voting scheme or one-versus-all strategies are the schemes most frequently applied. In this topic, the Error Correcting Output Codes framework has shown to efficiently combine binary classifiers to address the multi-class problem [3].

In this paper, we present an approach to model and classify both printed and handwritten symbols with elastic deformations. The descriptor, classification, and experimental results are explained in the next sections.

2 Overview of the system

The steps of the system are shown in Fig.1: As a preprocessing step, contours or skeletons are obtained from the input image in order to have thickness invariant features. The use of skeletons or contours is decided depending on the shape database (see Fig.2). Thus, we prefer skeletons for line-based structures, and contours for silhouette-based shapes. Secondly, the Hotelling transform based on principal components [5] is applied to find the main axis of the object so the alignment can be performed. Third, the method defines a blurred region of the shape that makes the technique robust against elastic deformations. Afterwards, Adaboost is applied to each pair of classes to train relevant features that split better object classes. And finally, the set of binary classifiers is embedded in the framework of Error Correcting Output Codes (ECOC) to improve final classification.

Fig. 1. Process scheme.

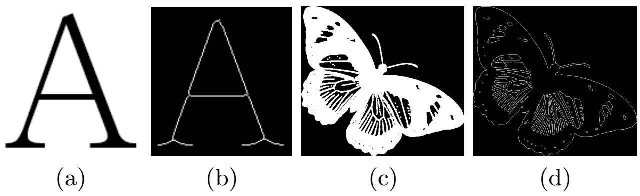
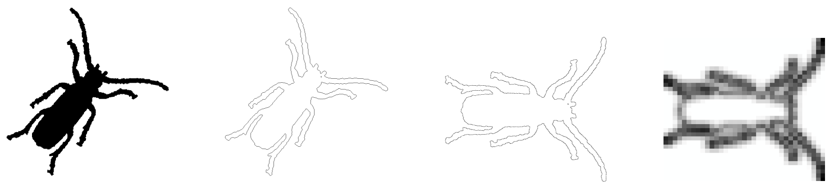


Fig. 2. (a)(b) Shape by skeleton. (c)(d) Shape by contour map.

3 Blurred Shape Descriptor

Referring to the present descriptor, the symbol is described by a probability density function of blurred shape model (BSM) that encodes the probability of pixel densities of image regions: The image is divided in a grid of $n \times n$ equal-sized subregions, and each bin receives votes from the shape points in it and also from the shape points in the neighboring bins. Thus, each shape point contributes to a density measure of its bin and its neighboring ones. The output descriptor is a vector histogram where each position corresponds to the amount of shape points in the context of the sub-region. The resulting vector histogram, obtained by processing all feature points, is normalized in the range [0..1] to obtain the probability density function (pdf) of $n \times n$ bins. In this way, the output descriptor represents a distribution of probabilities of the object shape considering spatial distortions. In Fig. 3(a) an input shape is processed. The symbol is filtered to obtain the contour map (Fig. 3(b)), and the sequent figures correspond to the aligned shape (Fig. 3(c)) and the 30×30 BSM (Fig. 3(d)), respectively. For further details, see [4].



(a) (b) (c) (d)

Fig. 3. (a) Input image. (b) Contour map. (c) Aligned shape. (d) 30×30 BSM.

4 Classification

Concerning the classification step, the Adaboost algorithm is proposed to learn the descriptor features that best split classes, training the classifier from Blurred Shape Model descriptors. The BSM has a probabilistic parametrization on the object shape considering its possible shape distortions. Due to the fact that different types of objects may share local features, Adaboost has been chosen to boost the BSM model in order to define a classifier based on the features that best discriminate one classes against the others. In particular, we use the Discrete Adaboost version [2] with 50 iterations of decision stumps. To outperform the Adaboost behavior, we embed the Adaboost binary classifiers in the framework of Error Correcting Output Codes [3].

5 Results

The evaluation of the method has been performed on the GREC2005 database. Our initial tests are applied on the first level of distortions. We have generated 140 artificial images per model (thus, for each of the 25 classes) applying different distortions such as morphological operations, noise addition, and partial occlusions. In this way, the ECOC Adaboost is able to learn a high space of transformations for each class. The BSM descriptor uses a grid of 30×30 bins. In this sense, 900 features are extracted from every image, from which Adaboost selects a maximum of 50. Comparing our technique with the kernel density matching method (KDM) proposed by Zhang [6], our first results (table 5) are very promising.

Method	Distortion Level 1	Distortion Level 2	Distortion Level 3	Distortion Level 4	Distortion Level 5	Distortion Level 6
KDM	100	100	100	96	88	76
BSM	100	100	100	100	96	92

Table 1. Descriptors classification accuracy increasing the distortion level of GREC05 database using 25 models and 50 test images.

6 Conclusions and future work

In summary, our proposed methodology shows promising results in comparison to the state-of-the-art descriptors, being robust against rotation, noise, scale and elastic deformations.

Nowadays, we are extending the experiments on different printed databases, such as the MPEG7, GREC03 or GREC05, increasing the set of distortions to evaluate the robustness of the present methodology against to a wide set of image transformations. Besides, we also take use of several handwritten symbol databases to show the suitability of our scheme dealing with elastic deformations and different writing styles.

References

1. D. Zhang and G. Lu, "Review of shape representation and description techniques", *Pattern Recognition*, vol. 37, pp. 1-19, 2004.
2. J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting", *The Annals of Statistics*, vol. 8, issue 2, pp. 337-374, 1998.
3. T. Dietterich and G. Bakiri, "Solving multiclass learning problems via error-correcting output codes", *Artificial Intelligence Research*, vol. 2, pp. 263-286, 1995.
4. A. Fornés, S. Escalera, J. LLadós, G. Sánchez, P. Radeva and O. Pujol, "Handwritten symbol recognition by a boosted blurred shape model with error correction", Iberian Conference on Pattern Recognition and Image Analysis (IBPRIA), Part I, LNCS 4477, Girona, Spain, pp. 13-21, June 2007.
5. George H. Dunteman, "Principal components analysis (Quantitative Applications in the Social Sciences)", Sage Publications, May 1989.
6. W. Zhang, L. Wenyin, K. Zhang, "Symbol recognition with kernel density matching", IEEE Trans. Pattern Anal. Mach. Intell. vol. 28, issue 12, pp. 2020-2024, 2006.