

Vehicle License Plate Segmentation In Natural Images

Javier Cano and Juan-Carlos Pérez-Cortés

Instituto Tecnológico de Informática, Universidad Politécnica de Valencia,
Camino de Vera, s/n 46071 Valencia (SPAIN)
{jcano,jcperez}@iti.upv.es *

Abstract. A robust method for plate segmentation in a License Plate Recognition (LPR) system is presented, designed to work in a wide range of acquisition conditions, including unrestricted scene environments, light, perspective and camera-to-car distance. Although this novel text-region segmentation technique has been applied to a very specific problem, it is extensible to more general contexts, like difficult text segmentation tasks dealing with natural images. Extensive experimentation has been performed in order to estimate the best parameters for the task at hand, and the results obtained are presented.

1 Introduction

Text-region segmentation has been largely studied over the last years, [9], [8], [5], [2], [4], however, even today it remains an open field of work, interesting for many different applications in which complex images are to be processed. Reasonable advances have been actually achieved in the task of extracting text from some kind of restricted images, as in the case of scanned documents, artificially edited video, electronic boards, synthetic images, etc. In all of them, the text included in the image has a number of "a priori" defined properties (localisation, intensity, homogeneity) that makes possible to tackle the segmentation task using filters, morphology or connectivity based approximations.

Historically, the methods devised to solve the text segmentation problem fall into one of two different branches: a morphology and/or connectivity approach, most useful for dealing with the kind of images previously described, and a textural (statistical) approach that has been successfully used to find text regions over non-restricted natural images. This is the problem that arises in the segmentation phase of an LPR system, where images are composed of a great variety of objects and affected by illumination and perspective variations. All these variable environment conditions result in a complex scene, where text regions are embedded within the scene and nearly impossible to identify by the methods employed in the morphology approximation.

Thus, the task of text (license plate) segmentation in a LPR system is included in the second category. Moreover, due the nature of images (as we will

* Work partially supported by the Spanish CICYT under grant TIC2000-1703-CO3-01

see in Section 2), it is also desirable to use a segmentation method capable of generating various hypothesis for each image in order to prevent the loss of any possible license plate region. In this way, it is possible to design a subsequent recognition phase that filters the final results without discarding beforehand any reasonable segmentation hypothesis.

The segmentation method proposed can be also useful for detecting any kind of text regions in natural and complex images. However, since we are concerned with a very particular task, all the parameters have been specifically adapted to improve the detection of text regions which match the constraints imposed by the shape and content of a typical vehicle license plate.

As it will be shown in the experiments section, very promising results have been achieved for the segmentation phase, therefore the next step in the design of a complete license plate recognition system requires further work on the design of a complementary recognition phase able to take advantage of the multiple data (multiple hypothesis) provided by this segmentation.

The rest of the paper is organized as follows: Section 3 describes the data and their acquisition conditions. In Section 2, the proposed methodology is presented. Extensive experimentation and results are reported in Section 4, and finally, in Section 5, some conclusions are given and future work is proposed.

2 Corpus

A number of experiments have been performed in order to evaluate the performance of the novel segmentation technique. In other application areas, there are typically one or more standard databases which are commonly used to test different approaches to solve a specific task, and it is possible to compare results among them. This is not the case for our application, as far as we know, perhaps because license plate segmentation in non-restricted images is a fairly recent topic of interest in the pattern recognition community.

Therefore, we have used a locally acquired database. It is composed of 1307 color images of 640×480 pixels randomly divided into a test set of 131 and a training set of 1176 images. The experiments were carried out using only the gray-level information.

The scenes have been freely captured without any distance, perspective, illumination, position, background or framing constraints, except that the plate number has to be reasonably legible for a human observer. Several examples of images in the database are shown in Figure 1. In applications such as parking time control or police surveillance, the camera can be located in a vehicle and the images captured may be similar to the ones in this database. In other applications, such as access control or traffic surveillance, cameras are typically fixed in a place and thus the scene features (perspective, distance, background, etc.) are easily predictable.

A specific preprocessing step has to be performed prior to the training and test phases. This preprocessing task consists of a manual labelling, where each



Fig. 1. Four real example images from the test set. Different acquisition conditions are shown, as illumination, perspective, distance, background, etc.

license plate is located in the image and a four-sided polygon corresponding to the minimum inclusion box of the plate is defined and associated to that image.

3 Methodology

The aim of the segmentation phase is to obtain a rectangular window of a test image that should include the license plate of a vehicle present in a given scene. The task of detecting the skew and accurately finding the borders of the plate is left for the next phase, as well as the recognition proper, which is beyond the scope of this paper.

The method proposed for the automatic location of the license plate is based on a supervised classifier trained on the features of the plates in the training set. To reduce the computational load, a preselection of the candidate points that are more likely to belong to the plate is performed. The original image is subject to three operations. First, an histogram equalization is carried out to normalize the illumination. Next, a Sobel filter is applied to the whole image to highlight non-homogeneous areas. Finally, a simple threshold and a sub-sampling are applied to select the subset of points of interest. The complete procedure is depicted in Figure 2.

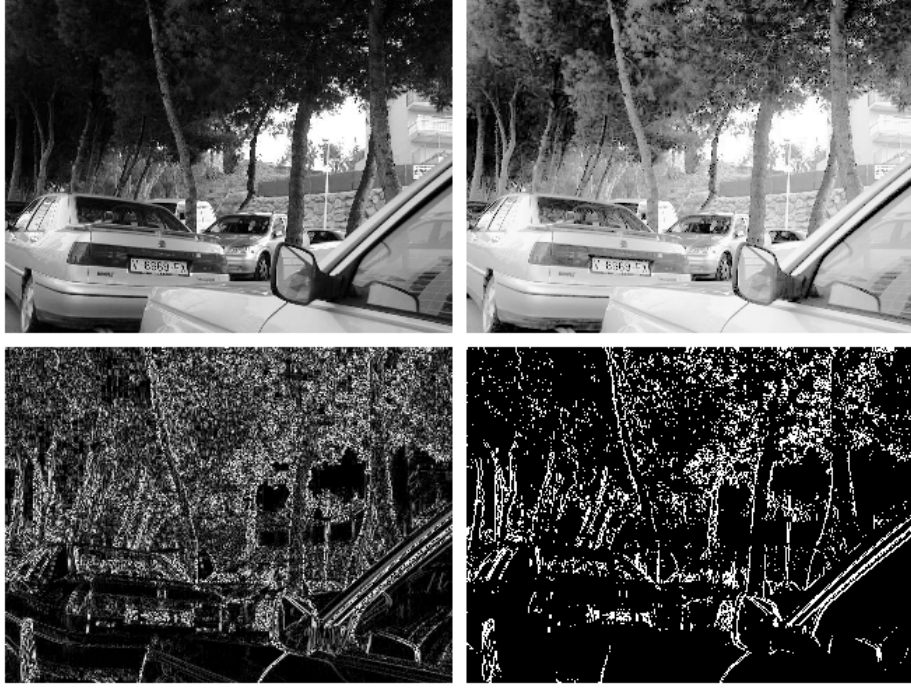


Fig. 2. Test image preprocess example. Upper-left: Original image. Upper-right: Equalization. Lower-left: Horizontal Sobel filter and Lower-right: Threshold binarization.

3.1 Multi-hypothesis scheme

Ideally, one segmentation hypothesis per image should be enough to detect a single vehicle plate, but because of the unrestricted nature of the images, it is possible that false positives appear when particular areas have features typically found in a license plate, like signs, advertisements and many other similarly textured regions. Therefore, it is important to save every hypothesis that can represent a plate region and leave the decision of discarding wrong hypotheses for the recognition phase, where all the details about the task are taken into account.

There is an additional important reason to adopt a multi-hypothesis scheme. Images have been acquired at different distances from the camera to the vehicle and, as a result, different sizes of plates can be seen in the images. This variability can be overcome using size-invariant features, including in the training set features from images of various sizes or using a multi-resolution scheme producing additional hypotheses. Informal tests have been performed that suggest that the first two options give rise to less accurate models of the “license plate texture” and thus lead to more false positives. For this reason, the third option has been



Fig. 3. Different hypothesis in a multi-resolution segmentation scheme. The brighter points indicate pixels classified as “license plate”.

chosen. In Figure 3, an example of this multi-hypothesis detection procedure is shown.

3.2 Feature vectors

A feature extraction technique that has proven its success in other image recognition tasks [3], [6] has been used in this case. It consists on using the gray values of a small local window centered on each pixel and applying a PCA transformation to reduce its dimensionality.

Each feature vector of the training set is labelled as belonging to one of two classes: positive (pixel in a license plate region), or negative (any other region). Obviously this gives rise to a huge set of negative samples, compared to the relatively small set of vectors of the “plate” class. Many of the negative samples can be very similar and add very little value to the “non-plate” class representation if they come from common background areas such as car bodies, buildings, etc.

To reduce the negative set, editing and condensing procedures can be probably used with good results, but we have applied a simpler and more efficient method that can be regarded as a bootstrapping technique. The procedure starts

up with no negative samples in the training set and then proceeds by iteratively adding those training samples that are misclassified in each iteration. In the first iteration, since the train set it is only composed by positive samples, the classification relies on a threshold on the average distance of the k -nearest neighbours.

We have found that a more compact and accurate training set is built if another distance threshold is used to limit the number of misclassified samples included at each iteration.

3.3 Classification

A conventional statistical classifier based on the *k nearest neighbours rule* is used to classify every pixel of a test image to obtain a pixel map where well-differentiated groups of positive samples probably indicate the location of a license plate.

In order to achieve a reasonable speed, a combination of a “*kd-tree*” data structure and an “*approximate nearest neighbour*” search technique have been used. This data structure and search algorithm combination has been successfully used in other pattern recognition tasks, as in [7] and [1]. Moreover, the “*approximate*” search algorithm provides us with a simple way to control the tradeoff between speed and precision.

4 Experiments

The results of the proposed segmentation technique are highly dependent on the classifier performance, which in turn depends on the use of a complete and accurate training set. Several parameters in this regard have been varied in initial tests.

In Figure 4, the segmentation results are shown for four iterations of the bootstrap process. A clear improvement is found in the 3 first iterations, but after that the results do not improve significantly. In this experiment, a window size of 40×8 pixels and the training images scaled so as the plate has a similar size, have been tested. Larger values of the window size proved to add little classification improvement.

However, the most promising results have been obtained for a window size of 40×8 pixels and the training images scaled so as the plate is around three times as large in each dimension, as suggest the results in Figure 5. In the experiments reported in that figure, as receiver operating curves, the size of the window is fixed to 40×8 pixels, while the normalized plate size ranges from 40×8 to 160×40 .

The best tradeoff between segmentation accuracy and cost is probably for a plate size of 100×25 pixels. Only slightly better results are found for higher plate sizes.

All the results are given at the pixel level. A simple post-processing procedure that isolates areas with a large number of pixels labelled as “license plate” with the correct size has to be applied before the plate recognitions phase. The shape

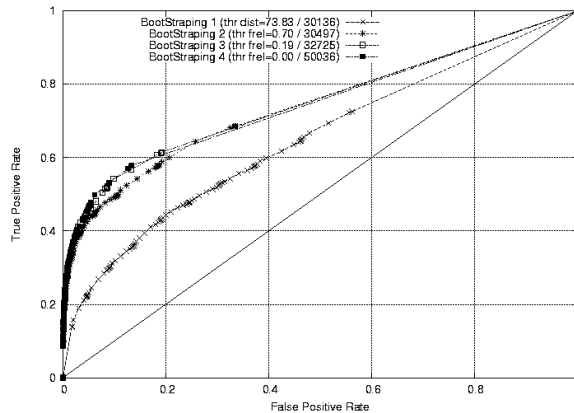


Fig. 4. Improvement of the classification performance in four bootstrap iterations.

of that area must also be taken into consideration to minimize the number of false positives at the plate segmentation level.

5 Conclusions and further work

A robust text segmentation technique has been presented. This technique seems to be able to cope with highly variable acquisition conditions (background, illumination, perspective, camera-to-car distance, etc.) in a License Plate Recognition task.

From the experiments performed, it can be concluded that a good tradeoff between segmentation accuracy and computation cost can be obtained for a plate normalization size of 100x25 pixels and a local window of 40x8 pixels for the feature vectors. In this conditions, a ratio of 0% False Positive Rate against a 40% True Positive Rate can be obtained with the most restrictive confidence threshold, that is, a 100% of classification reliability at the pixel level.

According to visual inspection of the whole set of 131 test images, the segmentation system has correctly located all the plates but two. Due to the unrestricted nature of the test set this can be considered a very promising result.

The computational resource demand of this segmentation technique is currently the main drawback, taking an average of 34 seconds the processing of a single 640x480 image on a AMD Athlon PC, at 1.2GHz in the conditions of the experiments reported. With some code and parameter optimizations, however, much shorter times, of only a few seconds are being already obtained in our laboratory.

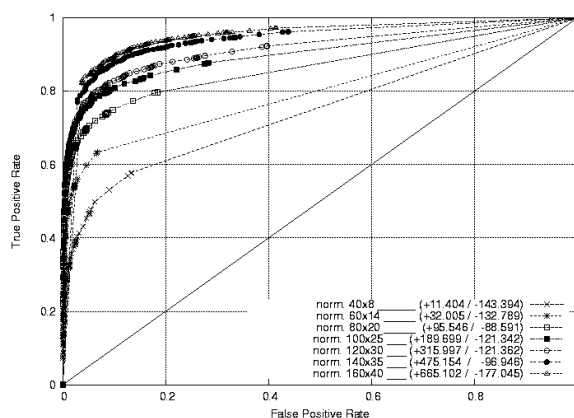


Fig. 5. Classification performance for a fixed local window size and a range of plate normalizations. Only the results of the last bootstrap iteration are presented.

References

1. J. Cano, J.C. Perez-Cortes, J. Arlandis, and R. Llobet. Trainig set expansion in handwritten character recognition. In *Workshop on Statistical Pattern Recognition SPR-2002*, Windsor (Canada), 2002.
2. Paul Clark and Majid Mirmehdi. Finding text regions using localised measures. In Majid Mirmehdi and Barry Thomas, editors, *Proceedings of the 11th British Machine Vision Conference*, pages 675–684. BMVA Press, 2000.
3. H. Ney D. Keyzers, R. Paredes and E. Vidal. Combination of tangent vectors and local representations for handwritten digit recognition. In *Workshop on Statistical Pattern Recognition SPR-2002*, Windsor (Canada), 2002.
4. A. Jain and S. Bhattacharjee. Text segmentation using gabor filters for automatic document processing. *Machine Vision and Applications*, 5:169–184, 1992.
5. A. Jain and B. Yu. Automatic text location in images and video frames. In *Proceedings of ICPR*, pages 1497–1499, 1998.
6. R. Paredes, J. Perez-Cortes, A. Juan, and E. Vidal. Face recognition using local representations and a direct voting scheme. In *Proc. of the IX Spanish Symposium on Pattern Recognition and Image Analysis*, volume I, pages 249–254, Benicassim (Spain), May 2001.
7. J.C. Perez-Cortes, J. Arlandis, and R. Llobet. Fast and accurate handwritten character recognition using approximate nearest neighbours search on large databases. In *Workshop on Statistical Pattern Recognition SPR-2000*, Alicante (Spain), 2000.
8. V. Wu and E. Riseman. Textfinder: An automatic system to detect and recognize text in images. *IEEE Transactions on pattern analysis and machine intelligence*, 21(11), 1999.
9. Y. Zhong, K. Karu, and A. Jain. Locating text in complex color images. *Pattern Recognition*, 28(10):1523–1236, 1995.