LOGO APPEARANCE DETECTION AND CLASSIFICATION IN A SPORT VIDEO

B. Kovář
Dept. of Applied Mathematics.
Faculty of Transportation Sciences - CTU
Prague, Czech Republic

A. Hanjalic
Dept. of Mediamatics.
Faculty of Information Technology
TU Delft - The Netherlands

Abstract - This paper presents a novel approach to detecting and classifying a trademark logo in frames of a sport video. In view of the fact that we attempt to detect and recognize a logo in a natural scene, the algorithm developed in this paper differs from traditional techniques for logo detection and classification that are applicable either to well-structured general text documents (e.g. invoices, memos, bank cheques) or to specialized trademark logo databases where logos appear isolated on a clear background and where their detection and classification is not disturbed by the surrounding visual detail. Although the development of our algorithm is still in its starting phase, experimental results performed so far on a set of soccer TV broadcasts are very encouraging.

INTRODUCTION

Measuring the statistics of the appearance of a trademark logo in a video is of great importance in the marketing and sponsoring sector. This statistics takes into account the number N of appearances of a certain logo along a video, positions L of a logo in a frame and the duration T of each logo appearance. Based on these three parameters one can compute the likelihood that a logo was noticed by a wide TV audience at home. As such, the parameters N, L and T can be used to determine the optimal position to place a logo on a scene that is to be broadcasted and to determine the advertising price per position. Finally, these parameters can be used as a feedback for a sponsor to check if the visibility of a logo along a video justifies its sponsorship engagement: the visibility of a logo can be seen as an indication of the expected revenue for the sponsor.

UNDERLYING PRINCIPLES

Having an arbitrary natural scene as shown in any of the frames in Figure 1, it is, on the first sight, highly unlikely that a logo can successfully be separated from its surrounding. A logo is only one of many objects displayed in a frame and an attempt to localize, isolate and properly classify that logo seems to be analogous to the attempt to localize, isolate and classify any arbitrary object in a natural scene.

The latter is, however, not viable for the general case, given the current state-of-the-art in the area of image interpretation. Additional difficulties are that the actual shape and texture of the object (due to 3-D deformation and occlusion), its color composition (due to varying lighting conditions) and position in the frame (due to camera and object motion) are not known a priori. Nevertheless, we believe that there might be a possibility to solve this problem specifically for the case of logo objects. This possibility exists due to some specific visual characteristics of logos compared to other objects, and in view of practical constraints that are related to the applications addressed in the previous section.

(a) (b) (c)

Figure 1: Examples of frames taken from a soccer video and containing logos that are to be detected and classified: (a) Champion League logo in the middle of the stadium, (b) ABN-AMRO bank logo on player's shirt, (c) EuroCard, Amstel Beer and Canon logos on the boards surrounding the playing field.

First, a good logo is designed precisely to be easily distinguishable from the rest of the scene, in other words, to act as "eye-catcher". We can translate this property into visual characteristics like:

- very sharp inner and outer edges,
- high contrast of neighboring logo regions,
- unique shapes of logo regions,
- unique color composition.

TECHNIQUE DESCRIPTION

Figure 2 shows the flow diagram of our proposed logo extraction and classification algorithm. The problem we address by this algorithm can be formulated as follows: *"In a given video frame find all appearances of a logo that was specified a priori!"*. As can be seen in Figure 2, our algorithm consists of three major phases. We will explain each of them separately in the following subsections.

Preprocessing

The activities in this phase serve to provide a reliable base for subsequent logo recognition and classification steps, the first of which is edge detection. We first blur the Y frame to reduce false or featureless edges. Gaussian smoothing is followed by the edge detection process for which we choose the Canny detector. The edge map is then processed by morphological filters for dilation and erosion. This is done in order to remove small discontinuities between neighboring edge lines that most likely belong to the same object. Close edges in the binary edge map are grouped by dilation using eight-connected structuring elements. Then, small connected components in the dilated image are filtered out using erosion. In this step, insignificant edges without connection to the neighbor area are removed. The output is a binary image that contains only most significant connected edges.

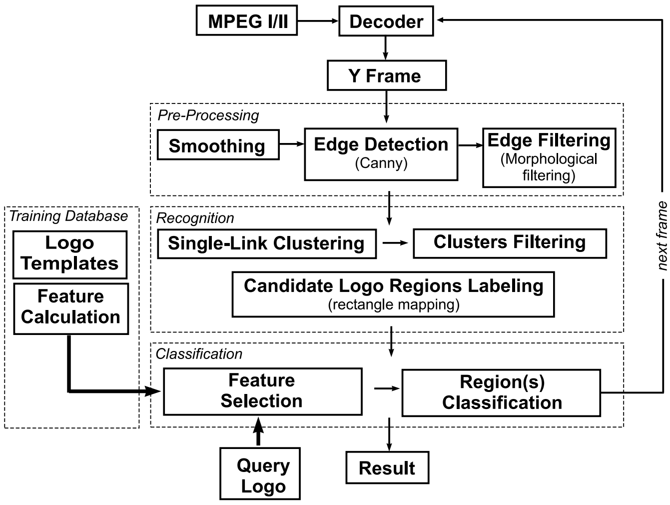


Figure 2: Flow diagram of the proposed logo detection and classification algorithm

Detection

In this phase, the set of most significant edges is refined in order to obtain edges that mark candidate logo regions. This is done first by applying single-link hierarchical clustering on the set of most significant edges resulted from the previous phase. Here, the cluster-merging criterion is the shortest distance between the terminal points (beginning and end) of two neighboring edge lines. The merging process is stopped once the distances between each two edge clusters are larger than the prespecified threshold d.

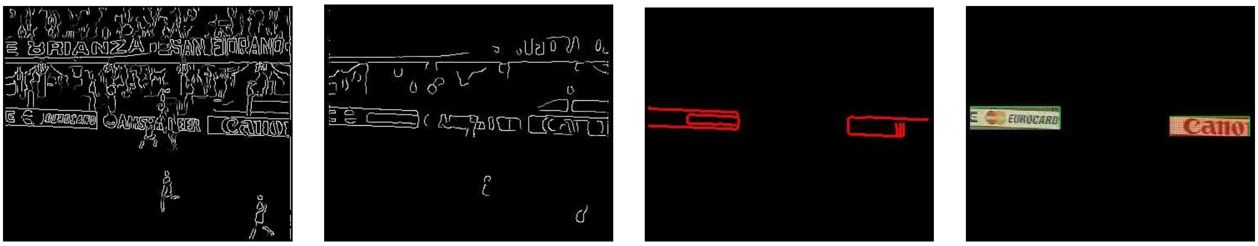
Not all the clusters obtained through the above procedure are relevant for further analysis. This is mainly the case for small isolated clusters. First, these clusters are unlikely to belong to a logo: such a small edge concentration is unrealistic in view of desirable ”eye-catching” properties of a logo. Second, even if there was a logo on the place of such a small cluster, that logo can be assumed rather small and invisible, and therefore not interesting for the applications described in the introduction to this paper. For these reasons we remove all clusters that contain less than allowed number of edge points. Each of the remaining edge clusters is considered as a marker of one candidate logo region, and is labeled by an identification number and position in the frame.

Classification

The classification step is based on matching the candidate logo regions with logo templates that are collected a priori in a large database. For each logo that is to be traced in a video a large number of templates are collected. These templates are extracted from video frames of many hours of sport videos simply by grabbing the image region with the logo. In this way, the templates are taken under similar lighting conditions, with similar types of occlusions and with similar 3-D deformations as to be found in logos in the video under consideration. For each logo class, the feature vector F is extracted. In the composition of the above feature vector shape descriptors dominate. In the final classification stage, simple nearest-neighbor classifier is used.

EXPERIMENTAL RESULTS

The proposed algorithm was evaluated using a number of MPEG video sequences taken from the Champion League soccer video. Here we present results for one of them, namely the Canon-MasterCard. In the Canon-MasterCard sequence (Figure 3) too many most significant edges were filtered out in the image region around the Amstel Beer logo. We explain this by the fact that due to occlusion the edge map in this region was split into four small clusters, that were then improperly classified as irrelevant. Good classification results for the other two logos (MasterCard and Canon) show that we might have obtained proper classification also for the Amstel logo, if there was no occlusion.



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

Figure 3: (a) Edge map obtained by a Canny edge detector with $t_{min} = 150$, $t_{max} = 255$, after smoothing with $\sigma = 2.0$, (b) the corresponding set of most significant connected edges as a result of morphological operations, (c) the result of single-linkage clustering: edge clusters marking candidate logo regions, (d) candidate logo regions marked by a rectangle spanned over each edge cluster and after removing all small clusters.

DISCUSSION

We like to start the discussion about the material presented in this paper by emphasizing that the algorithm for logo detection and classification proposed here is still in the infancy of its development. This becomes visible from many problems that we came across in describing the algorithm in previous sections. The most important problem is the one of defining a suitable parameter set that would work for a vast scope of logo categories. If no constant parameter set can be found, mechanisms should be known for adapting this set per logo class. Nevertheless, two important issues were shown in this paper. Detecting and classifying a logo in a natural scene is possible by using the edge and shape information only. Good detection and classification results can be obtained also in cases of an occlusion and 3-D logo deformation.

References

[1] Anil K. Jain, Aditya Vailaya: *Shape Based Retrieval: A Case Study with Trademark Image Database*, Department of Computer Science, Michigan State University

[2] Ciocca G., Schettini R.: *Similarity retrieval of trademark images*, Proceedings. International Conference on Image Analysis and Processing, Page(s): 915-920, 1999

[3] Eakins J.P., Boardman J.M., Graham M.E.: *Similarity retrieval of trademark images*, IEEE Multimedia , Volume: 5 Issue: 2 , Page(s):53 -63, April-June 1998