Trademark logo recognition: Preliminary results on a comparative between Haar-like features and Local Binary Patterns

Miguel Angel Medina Nieto, Jorge de la Calleja, Ma. Auxilio Medina, Antonio Benitez Ruiz

Cuerpo Académico de Sistemas y Cómputo Inteligente, Universidad Politécnica de Puebla, Tercer Carril del Ejido Serrano S/N, San Mateo Cuanalá, Juan C. Bonilla, Puebla, CP 72640.

E-mail: miguel.medina@uppuebla.edu.mx

Abstract

In this paper, a comparative study between two computer vision methods for recognizing trademark logos in images, is presented. Trademark logo recognition has different real-world applications, for example the identification of the source of a store. Logos in images present different conditions that affect their automatic recognition, such as shape, scale, location, illumination and perspective, among others. We present a training and detection phase to recognize logos using Haar-like features and Local Binary Patterns, two well-known methods in the computer vision community that obtain high accuracy for object recognition. Our preliminary results, using an image database of logos, show that the method of Haar-like features obtains a 20% better performance than Local Binary Patterns. Also, the reported results indicate a fast processing time, making it suitable for real-time applications such as augmented reality marketing applications.

Keywords: Computer vision, Object recognition, Haar-life features, Locally Binary Patterns.

1 Introduction

Object recognition is a fundamental problem in computer vision, however this is not a straightforward task due to several factors such as pose, rotation, illumination, scale, low image resolution, among others. In order to solve this problem, the first step is to extract the main characteristics of the objects. Then, these ones are used to perform matching over unknown images that may contain desired objects [1].

According to literature, detection, classification and logo recognition applications have been carried out mainly for document analysis. However, trademark logo recognition from a sequence of images is a few explored filed.

Trademark logos can be categorized in different ways. A logo can be made of words, images or word and images. Those containing only words or phrases are called word logo. The ones that consist of only images are named device logo and, the logos made of words and images are called device-and-word logo [2]. This work is related to the last one: device-and-word logo recognition approach.

Xia [3] states that logo recognition has a high interest by industrial companies, mainly due to the development of mobile applications. Besides that, trademark logo recognition provides enterprises profitable information to build immersive systems. Capturing logos by using mobile phones is promising, because it is possible to retrieve semantic information about the owners of the logos. Nevertheless, to build automatic systems to robustly recognize logos in an efficient way are required to establish appropriate methods.

In scenarios where users of mobile apps walk through a mall, the logos can be captured by the device's camera. Then, these images generally are influenced by changes in illumination, and deformed by the perspective of the user, for instance skewness, rotation and size affect the images.

The main challenge in object recognition systems is to look for increasing the accuracy and efficiency on the recognition task. At present time, the most common approach is to use image descriptors like Histogram of Oriented Gradients (HOG) and Scale Invariant Feature Transform (SIFT). These are combined with trainable classifiers such as Support Vector Machines (SVM), boosted classifiers, among others. [3]. Nevertheless, other computer vision techniques have been proposed that obtain high accuracy in real-world applications, such as Haar-like features and Locally Binary Patterns.

In this paper we present preliminary results to define a method for logo recognition in real-world domains. Our approach is based on two computer vision techniques: Haar-like features and Local Binary Patterns (LBP). Experimental results were carried out using an image database with several images for recognizing four trademark logos. The remainder of the paper is organized as follows: In Section 2 we present the methodology; in Section 3 the experimental results are presented; and finally conclusions and directions for future work are described in Section 4.

2 Methodology

In order to perform logo recognition, the first step is to select trademark logos. This selection is based on some of the most popular trademarks according to [4]. Thus, for this work, we only select four logos, which are shown in Figure 1.

Then we perform a training phase with the goal of generate a classifier that learns to recognize logos among the different ones. In next subsections we give descriptions of this stage and a brief explanation of the methods.

Figure 1. Selected trademark logos. Top left, Ferrioni. Top right, Sanborns. Bottom left, Tous. Bottom right, Puma.

2.1 Training phase

The goal of this phase is to train a classifier for trademark logo recognition. As we mention in previous section, Haar-like features and Locally Binary Patters were used.

First, two set of samples must be generated for training: negative samples and positive samples

Negative samples are images in which there is no one logo to detect. Positive samples are images that contain at least one logo of our interest. For this work, 500 images were used to create the negative sample set. These images are color images with a resolution of 160x200 pixels. In Figure 2 a sample of these images is presented.

In the case of the positive sample set, it was built using 200 images for each logo. These images were generated applying different image processing transformations to the original images. These transformations were rotation and skewness. The resolution of the resulting images was of 50x50 pixels. In Figure 3 an example of these generated images is presented. It is important to note that all positive samples contain only one logo within the image.

Figure 2. Example images for negative samples.

2.2 Recognition phase

In order to produce a stable classifier it is needed to obtain abstract details of the object to recognize within the image. These details are known as features. Next we give a brief description of the methods used for feature extraction.

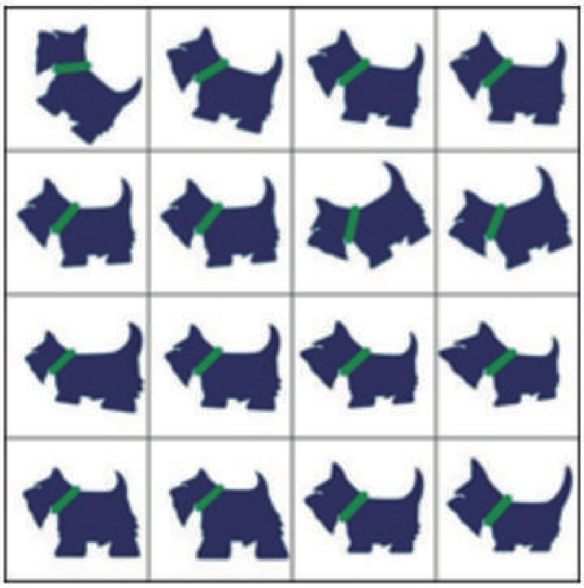

Figure 3. Example images for positive samples, in this case for the Ferrioni logo.

Haar-like features

The research of Viola and Jones has shown that AdaBoost can be used to select features and train classifiers [5]. Haar-like features are rectangular areas composed by white and black regions. Each feature type can indicate if a certain characteristic is present or not in image, such as an edge, a line and so on. Figure 4 shows examples of Haar-like features for detecting edges and lines.

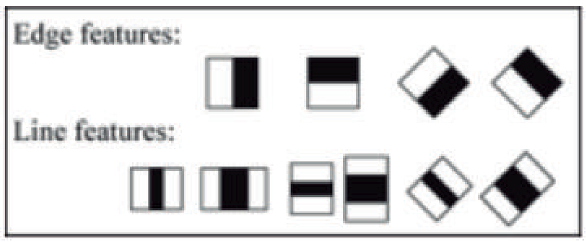

Figure 4. Examples of Haar-like features.

The value of a feature is calculated using an integral image, that is the sum of all pixels to the left and above of the original image for each pixel.

The object detection is realized by using weak classifiers. Each classifier evaluates a single feature, then it determines an optimal threshold classification function. In the end, several weak classifiers will be combined to create a strong classifier in order to obtain better accuracy that the yield for a single one.

A cascade of classifiers is a degenerated decision tree where at each stage a classifier is trained to detect the objects of interest while rejecting certain fraction of the non-object patterns. The purpose of the cascade is to achieve specified hit rate and false alarm rate through each stage. Increasing the number of stage, increases the number of weak classifiers. In Figure 5 a cascade of classifiers is shown

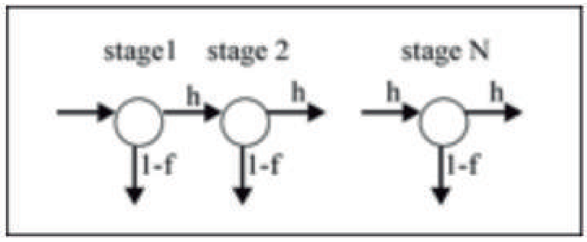

Figure 5. Cascade of Classifiers.

Where hit rate is equal to h^N and the false alarms correspond to f^N.

AdaBoost is a powerful machine learning algorithm used to find the best weak classifiers. This algorithm maximizes the margin between positive and negative samples, and also performs several trials with each of the weak classifiers to produce the smallest misclassification error.

Local Binary Patterns

Local Binary Patterns (LBP) is a simple but efficient method for extracting features, and it is rotation invariant for gray scale images. This method has been applied in areas such as face recognition, texture recognition, background modeling, facial expression analysis, among others [6].
LBP are computed by comparing a pixel with its neighbors according to the equation 2.

$$LBP_{P,R} = \sum_{p=0}^{p-1} s(g_p - g_c)2^p, \qquad \text{Eq. (2)}$$

$$s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$$

In equation 2, represents the center pixel, $g_p(p = 0, .. p-1)$ denotes the neighborhood of the center pixel with a radius *R*, and *P* is the total numbers of neighbors. The function $s(x)$ generates a new image based on a specific threshold. As an example in Figure 6, the pixels are mapped for a threshold value of 120.

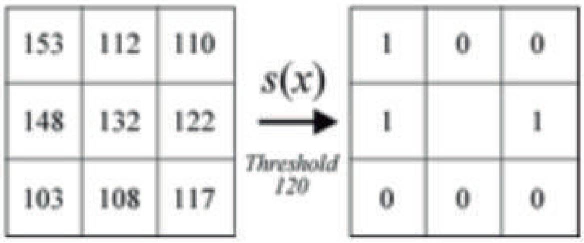

Figure 6. LBP calculation with a threshold of 120.

The idea of using LBP to logo recognition is motivated by the fact that this method is robust to illumination and rotation conditions, which are present when logos have to be recognized.

3 Experimental results

In order to test the effectiveness of our classifiers, a new set of positive images was established. These images contain logos based on real scenarios. The sets are composed by 250 images of size of 640 x 480 pixels for each logo. In Figure 7 a subset of these images is presented.

Figure 7. Subset of positive images for testing our proposed method. As we can observe, logos appear in several parts of the images and with different illumination and pose conditions.

Another set of 250 negative examples, with the same size, was used to measure the accuracy of the classifiers. The accuracy metric is defined according to Equation 1.

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn} \qquad \text{Eq. (1)}$$

Where TP corresponds to the true positives detections, FP are false positives, FN are the false negatives and, TN are the true negatives.

The evaluation of the classifiers was made by establishing the same parameters for all of them. The scale factor for the Haar-like classifier was set to a 20 percent, while the value of the minimum neighbors was 3 for LBP.

Table 1 shows the results of the accuracy for each logo detection. As we can observe, the recognition accuracy is over .92 for the Adidas, Ferrioni and Tous logos, however, the Puma logo was only recognized with a .5 accuracy. Also, we can see that Haar obtains the best results results for all cases.

Table 1. Comparative accuracy results.

Logo	Haar	LBP
Adidas	0.948	0.718
Ferrioni	0.926	0.732
Puma	0.564	0.500
Tous	0.948	0.748

In table 2 the average of time detection per image over positive samples is shown. The experiments were executed using a PC with an i5 Intel processor and 4Gb of RAM. Now, we can note that LBP obtains the lowest execution times for detections in all cases.

Table 2. Positive samples time detection in milliseconds.

Logo	Haar	LBP
Adidas	11.979	8.146
Ferrioni	11.792	8.265
Puma	12.681	10.258
Tous	12.475	10.551

Results from Table 1 reflect that Haar-like features classifier obtains better accuracy than LBP. But it is also clear, according to Table 2, that LBP is a faster classifier.

We found that taking the logos of Figure 1 and applying the detection on images in Figure 7 was a challenging task. The logos in real images are not always flat, that is we can see that they appear in 3D figures. In addition, illumination or transparency behind them affect directly the detection accuracy.

Some of these 3D figures, such as Puma, have sources of light that causes that images have a glow effect making the classifier to fail in its detection task. However, the obtained results guide us on a future work: and augmented reality application with an immersive environment that uses these classifiers to detect logo trademark and present adds to the user (See Figure 8).

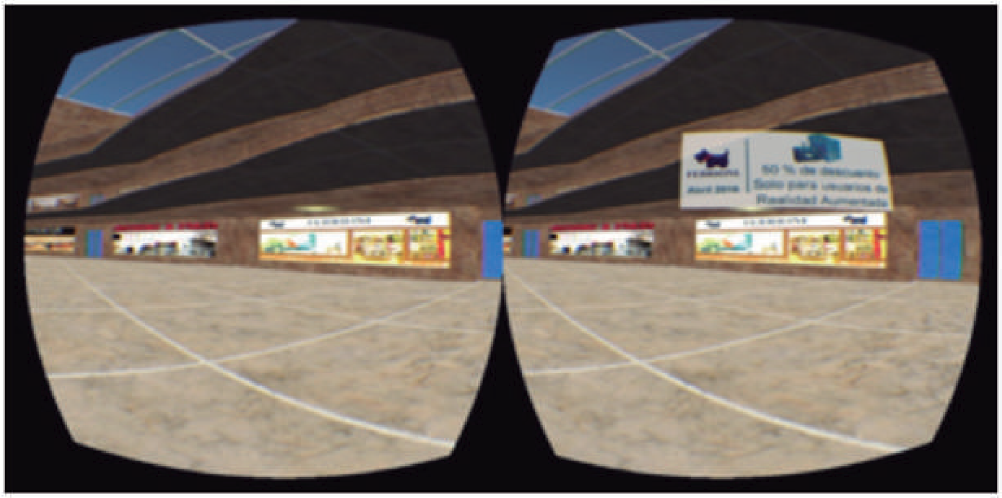

Figure 8. Immersive environment showing information about particular stores, which have been recognized by their logos.

4 Conclusions

In this paper we have presented preliminary results using two methods to recognize and classify trademark logos within low resolution images. The whole framework was proved to be effective in real images. Also, we have tested two different classifiers applied to the same kind of images. The experimental results show that Haar-like features obtains the best average recognition performance, yielding 94.8% recognition accuracy. We consider that Haar classifiers have a great application prospect in real-time logo detection applications, particularly for marketing in real-time video streams.

Acknowledgments

The first author wants to thank to INAOE (National Institute of Astrophysics, Optics and Electronics) for the facilities granted to finish this work.

References

[1] Psyllos, A., Christos-Nikolaos, A., & Eleftherios, K. (2010). Vehicle Logo Recognition Using a SIFT-Based Enhanced Matching Scheme. IEEE Transactions on Intelligent Transportation Systems, 11(2), 322-328.

[2] Xia L., Qi F., Zhou Q., (2008). A Learning-based Logo Recognition Algorithm Using SIFT and Efficient Correspondence Matching. International Conference on Information and Automation, 2008. Changsha: IEEE, 1767-1772.

[3] Wei, C.-H., Li, Y., Chau, W.-Y., & Li, C.-T. (2009). Trademark image retrieval using synthetic features for describing global shape and interior structure. Pattern Recognition, 42(3), 386-394.

[4] Relaciones, T. (s.f.). Los Centros y Plazas Comerciales más Valorados de México. Recovered on January 10th, 2016, from tnrelaciones.com: http://www.tnrelaciones.com/cm/preguntas_y_respuestas/content/391/4681/es/los-centros-y-plazas-comerciales-mas-valorados-de-mexico.html.

[5] Viola P, Jones M. J. (2004). Robust Real-Time Face Detection. International Journal of Computer Vision, 2004, 57(2), 137-154.