PERGAMON

Pattern Recognition 34 (2001) 1639–1655

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Content-based similarity retrieval of trademarks using relevance feedback

G. Ciocca, R. Schettini*

Istituto Tecnologie Informatiche Multimediali, Consiglio Nazionale delle Ricerche, Via Ampere 56, 20131 Milano, Italy

Received 11 March 1999; received in revised form 14 March 2000; accepted 14 March 2000

Abstract

This paper addresses the problem of how to efficiently and effectively retrieve images similar to a query from a trademark database purely on the basis of low-level feature analysis. It investigates the hypothesis that the low-level image features used to index the trademark images can be correlated with image contents by applying a relevance feedback mechanism that evaluates the feature distributions of the images the user has judged relevant, or not relevant and dynamically updates both the similarity measure and query in order to better represent the user's particular information needs. Experimental results on a database of 1100 trademarks are reported and commented.

Keywords: Content-based retrieval; Relevance feedback; Image similarity; Trademarks

1. Introduction

The content-based retrieval of trademarks is "*extremely challenging and instructive to study*", due to the high complexity and diversity of the data involved, also often composed of several distinct components [1]. Our study has addressed the problem of how to efficiently and effectively retrieve images similar to a query from a trademark database purely on the basis of low-level feature analysis. As already pointed out by several authors [1,2], perceptually similar images are not necessarily similar in terms of low-level features. We have investigated the hypothesis that the low-level features used to index the images can be correlated with their semantic contents by applying a relevance feedback mechanism.

A few applications designed specifically for the registration of trademarks are available. Wu et al. have developed a prototype system, STAR, using their content-based retrieval engine for multimedia information systems [2–4]. Eakins et al. have developed a prototype system (ARTISAN) for the UK Patent Office Trade Marks Registry, to retrieve trademarks when these consist of abstract geometric designs [5]. Another system, called TRADEMARK, operates on the trademark database of the Patent Office of Japan [6]. A detailed analysis of the problems involved in trademark registration can be found in a recent paper by Jain and Vailaya [1]. These authors propose a computational strategy in which multiple feature description schemes of the same visual cue (shape) are used to improve retrieval accuracy without significantly increasing computational costs. At the first stage of processing (pruning) Jain and Vailaya represent the trademark images in terms of invariant moments and the histogram of the edge directions, integrating the dissimilarity of these features by a weighted mean. A small set of plausible candidates is then presented to a detail matcher based on deformable templates to eliminate false matches. This second phase makes it possible to eliminate the false matches, but cannot cope with trademarks that have not been retrieved in the first stage, although actually perceptually similar to the query. We have attempted to improve the effectiveness of the first stage of retrieval by relevance feedback, i.e. by allowing the user to progressively refine the system's response to

* Corresponding author. Tel.: + 39-2-706-43288; fax: + 39-2-706-43292.

E-mail addresses: ciocca@itim.mi.cnr.it (G. Ciocca), centaura@itim.mi.cnr.it (R. Schettini).

0031-3203/01/$20.00
PII: S0031-3203(00)00055-8

his query. The key concept of the relevance feedback we propose is the statistical analysis of the feature distributions of the retrieved images the user has judged relevant, or not relevant, in order to understand what features he has taken into account (and to what extent) in formulating this judgment, so that we can then accentuate their influence in the overall evaluation of image similarity, as well as in the formulation of a new query iteration.

Section 2 of the paper briefly describes the feature sets used to index the images. Section 3 presents the relevance feedback mechanism implemented. Experimental results are reported in Section 4, followed by our conclusions.

2. Image indexing

The features used for indexing have been selected for three basic properties [7]:

(i) perceptual similarity (the feature distance between two images is large only if the images are not "similar"),
(ii) efficiency (the features can be rapidly computed), and
(iii) economy (small dimensions that do not affect retrieval efficiency).

Loncavic [8] has recently published a review of shape analysis, while Mehthre et al. [9] and Scasselati et al. [10] have reported experimental comparisons of methods of image retrieval based on shape similarity. Jain and Vailaya have experimented different feature sets for trademark indexing, finding that invariant moments and the histogram of the edge directions are the most effective [1]. These features are applied here, together with the mean and variance of the absolute values of the coefficients of the sub-images of the first three levels of the multiresolution wavelet transform of the image. We have used this somewhat redundant image description as none of the features can univocally identify an image: completely different images may yield similar feature values. The very different natures of the indices chosen here should limit the possibility of different trademarks corresponding to very close points in the feature space.

2.1. Moments

Moments, in general, describe numerical quantities at some distance from a reference point or axis [11,12]. Their use in image analysis is straightforward if we consider the image a two-dimensional density distribution function. However, characterizing all the information contained in an image would require an infinite number of moment values. The challenge, therefore, is to select a meaningful subset of moment values that contains sufficient information to describe image appearance. For an image $f(x, y)$ the central moments are given by

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y), \tag{1}$$

where $\bar{x}$ and $\bar{y}$ are the coordinates of the center of mass. From the following combinations of second and third moments:

$$\begin{aligned}
M_1 &= (\mu_{20} + \mu_{02}),\\
M_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2,\\
M_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2,\\
M_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2,\\
M_5 &= (\mu_{30} + \mu_{12})(\mu_{30} - 3\mu_{12})[(\mu_{30} + \mu_{12})^2\\
&\quad - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(\mu_{21} + 3\mu_{03})\\
&\quad \times [3(\mu_{03} + \mu_{21})^2 - (\mu_{21} - \mu_{03})^2],\\
M_6 &= (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]\\
&\quad + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}),\\
M_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2\\
&\quad - 3(\mu_{21} + \mu_{03})^2] + (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})\\
&\quad \times [3(\mu_{03} + \mu_{21})^2 - (\mu_{21} - \mu_{03})^2],
\end{aligned} \tag{2}$$

a set of invariant moments which have the useful properties of being invariant to the object's scale, rotation, and position has been derived [11]:

$$\begin{aligned}
&M'_1 = M_1/\mu_{00}, \quad M'_2 = M_2/r^4, \quad M'_3 = M_3/r^6,\\
&M'_4 = M_4/r^6, \quad M'_5 = M_5/r^{12}, \quad M'_6 = M_6/r^8,\\
&M'_7 = M_7/r^{12},
\end{aligned} \tag{3}$$

where $r = (\mu_{20} + \mu_{02})^{1/2}$ is the radius of gyration of the object.

Moments, however, do not suffice to completely describe the perceptual appearance of trademarks. The two trademarks shown in Fig. 1, for example, are actually similar from a perceptual point of view, although they have quite different moment values.

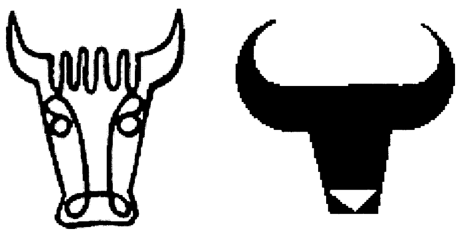

Fig. 1. Two similar trademarks having very different descriptions in terms of invariant moments.