Computer Vision and Image Understanding 114 (2010) 66–84

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Content-based retrieval of logo and trademarks in unconstrained color image databases using Color Edge Gradient Co-occurrence Histograms

Raymond Phan *, Dimitrios Androutsos

Department of Electrical and Computer Engineering, Ryerson University, 350 Victoria St., Toronto, ON, Canada M5B 2K3

ARTICLE INFO

Article history:
Received 22 December 2008
Accepted 26 July 2009
Available online 3 August 2009

Keywords:
Color Edge Gradient Co-occurrence Histogram (CEGCH)
Logo and trademark retrieval
Color edge detection
HSV color quantization
Pattern recognition

ABSTRACT

In this paper, we present an algorithm that extends the Color Edge Co-occurrence Histogram (CECH) object detection scheme on compound color objects, for the retrieval of logos and trademarks in unconstrained color image databases. We introduce more accurate information to the CECH, by virtue of incorporating color edge detection using vector order statistics. This produces a more accurate representation of edges in color images, as compared to the simple color pixel difference classification of edges seen with the CECH. Our proposed method is thus reliant on edge gradient information, and so we call it the Color Edge Gradient Co-occurrence Histogram (CEGCH). We also introduce a color quantization method based in the hue–saturation–value (HSV) color space, illustrating that it is a more suitable scheme of quantization for image retrieval, compared to the color quantization scheme introduced with the CECH. Experimental results demonstrate that the CEGCH and the HSV color quantization scheme is insensitive to scaling, rotation, and partial deformations, and outperforms the use of the CECH in image retrieval, with higher precision and recall. We also perform experiments on a closely related algorithm based on the Color Co-occurrence Histogram (CCH) and demonstrate that our algorithm is also superior in comparison, with higher precision and recall.

1. Introduction

Content Based Image Retrieval (CBIR) is an area of signal and multimedia processing with many promising applications. CBIR is an active area of research with many years of attention, producing a high volume of research output, causing many subsets of CBIR to exist. Because of this, in this study, we concentrate on one subset of CBIR, which is on logo and trademark retrieval. Logo and trademark retrieval has received much attention in literature, due to its many promising commercial applications. It is seen as an increasingly vital tool for industry, commerce and trademark registration [1,2], and for sports entertainment [3,4].

For any image retrieval system, a suitable representation or feature for representing these objects should be chosen to facilitate a meaningful comparison of the input query to the images in the database. To choose this feature, the amount of distortion expected from the object is taken into account. As there are many different features to use when performing logo and trademark retrieval, one may suggest that a hybrid approach is necessary: combining more than one feature into a single feature descriptor. However, for the feature representation in logo and trademark retrieval, different feature representations ultimately have different goals in mind, and is fairly difficult to create a hybrid approach in combining more than one feature into a single descriptor. However, for selecting a feature representation of interest, there are a wide variety of approaches to solve this problem, each having their own assumptions on the properties of the logos and trademarks to detect or retrieve.

1.1. Related work

Relevant prior research in logo and trademark detection and retrieval include the following, each concentrating on a particular feature or approach to solve the problem.

1.1.1. Shape

Many approaches to solve the problem of retrieving logos and trademarks use shape descriptors of varying complexities as indices for detection or retrieval. Kim and Kim [5] employ the use of Zernike and pseudo-Zernike moments to build the feature set used for the retrieval of trademarks in trademark databases. Eakins et al. [2] designed their trademark retrieval system such that it segments all trademarks in the database into components, and derives all features from these component boundaries. Features such as perimeter, aspect ratio, edges, area, and straightness were employed.

* Corresponding author.
E-mail addresses: rphan@ee.ryerson.ca (R. Phan), dimitri@ee.ryerson.ca (D. Androutsos).

1077-3142/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2009.07.004

1.1.2. Neural networks

Logo and trademark retrieval can be treated as a pattern classification problem. Neural networks require no mathematical model to determine whether a particular logo or trademark is the one of interest. Instead, they behave as model-free estimators and the network learns by training with input samples, or in this case, the logos and trademarks already present in a database. The work seen in Zhao et al. [6], and Alwis and Austin [1,7] concentrate on logo and trademark retrieval in this viewpoint.

1.1.3. Contour-based (query by sketch)

There have been approaches where the user submits a hand drawn sketch using rectangles, polygons, ellipses and B-spline curves, as opposed to an already prepared logo or trademark to submit to the system. Once these are drawn and submitted as an input query, contour information of these hand drawn curves are used and features are extracted, serving as the input query to the logo and trademark database. Folkers and Samet [8] use Fourier Descriptors to describe the contours. Leung and Chen [9] extract the contour of a hand drawn sketch by thinning it, segmenting it into regions and extracting the contour by edge detection.

1.1.4. Objects at multiple scales

Even though much research has been put forth on logo and trademark detection and retrieval, the above work already consists of logos or trademarks themselves in the database, and all of the images are of the same scale. This viewpoint of retrieval is most commonly encountered in copyright protection and trademark registration applications, as seen in Eakins et al.'s work [2]. However, the previous attempts do not address the case of unconstrained logo and trademark retrieval, and the acquisition of the images that populate the databases are under controllable conditions. There have been attempts for retrieving objects at multiple scales, not necessarily limiting the scope to logos and trademarks. Sebe et al. [10] created a system where they partition an unconstrained image into sub-images, and retrieve an entire image using a single partition. Ke et al. [11] created a system for "near-duplicate" detection of copyrights in images for the purpose of detecting forged images and copyright violations.

However, with the introduction of the Internet as an effective method for communication, the low cost of scanners and storage devices, digital images are a vital tool for the dissemination of pictorial information, introducing the problem of *unconstrained* images. Unconstrained images are those that have been captured by an acquisition device in natural environments where there are factors beyond one's control. For these images, the objects of interest can be subject to many uncontrollable factors. High amounts of deformation, and large variations of color depending on illumination conditions are inevitable in unconstrained images. The objects to detect can also be of an arbitrary size, and can be in any location when the image is captured. These last two aforementioned attempts only consider a subset of the unconstrained domain, and do not consider situations of heavy deformation. Also, most of the aforementioned work does not consider the use of color information, and the logos are either monochromatic, or binary. Color is one feature that is most important, for it is considered as a powerful visual cue, and is one of the most discriminatory and commonly used features in image retrieval [12]. Human beings can discern many different kinds of colors, but only a few shades of gray intensities.

With regards to color, much work has been put forth by considering color as the primary feature. The first example of this approach was by Swain and Ballard [12], using color histograms for building indices and retrieving images based on these indices. Retrieval based on histograms has been researched thoroughly, due to their simplicity and ease of implementation. Though the results in their work were very promising, the color histogram only captures color content of an image, creating misclassifications for retrieved images. Images can have similar color distributions, but different spatial statistics in comparison to the input query. As such, to resolve this problem, methods were proposed to add additional information to color histograms.

Chang and Krumm [13] proposed the *Color Co-occurrence Histogram (CCH)* for use in the recognition of color objects in images. In fact, techniques based on co-occurrence histograms, or co-occurrence matrices, have been used in the past, especially for texture analysis. The earliest published work in this area was authored by Haralick et al. in [14]. This modification to the color histogram captures both the color and spatial relationships between color pixels, effectively modeling color patterns. As stated earlier, though it is difficult to design a feature descriptor that combines more than one feature into a single descriptor, the CCH is a descriptor that ultimately combines color and texture together, thus making the descriptor more attractive to use. The CCH is a three-dimensional histogram, where each bin contains the number of color pixel pairs occurring at different spatial distances. Chang and Krumm used CCHs as the main tool for their object detection scheme, and proved that it was quite an effective method. However, their results illustrate the detection of the *same* rigid object amidst cluttered scenes, and experiments were performed under controlled conditions. The illumination, size, and orientation of the objects were kept *constant*, which is not true in general, especially for unconstrained images.

Luo and Crandall [15] discovered a fundamental flaw with the CCH. Regions in the image of a uniform color contribute a disproportionate amount of energy to the CCH, overwhelming the comparison metrics. Therefore, the CCH would demonstrate similarity between images that contain similar solid regions of color, but arranged quite differently in terms of their spatial layout. Luo and Crandall then proposed a modification to the CCH named the *Color Edge Co-occurrence Histogram (CECH)*, where the objective was to eliminate the solid color contribution problem to the CCH. The CECH only captures the separation of pairs of color pixels at different spatial distances when these pixels lie in edge neighborhoods, alleviating the disproportionate energy contributions a single color would have on the CCH. The CECH is the tool used in the object detection scheme proposed by Luo and Crandall, which concentrated an unconstrained detection of compound color objects.

A compound color object is one that has a set of multiple, particular colors arranged in a particular layout. These include (and are not limited to) flags, logos, cartoon characters and people in uniformed clothing. In Luo and Crandall's work, the construction of the CECH is performed by first quantizing the query and search images using a two-stage color quantization process. After, CECHs are built using the information from the quantized input query and search images, and the edge map information originating from the quantized images as well. Luo and Crandall's definition of an edge pixel is any pixel having a different color from any of its eight connected neighbors. This implies that only pixels that are on, or very close to, an edge are included in the CECH, and these tend to be pixels containing the most important spatial–color information. Though the object detection results in Luo and Crandall's work were very good, we feel that a more appropriate definition of what defines an edge pixel in a color image should be employed, as opposed to the simple color pixel difference for edge classification as introduced in their work.

To this end, we present work that extends the CECH object detection method for use in logo and trademark retrieval in unconstrained color image databases. We introduce color edge detection to the CECH using vector order statistics, which produces an edge map determining valid edge points in color images with greater accuracy. This edge map is used in conjunction with CECHs to per-