UNCONSTRAINED LOGO AND TRADEMARK RETRIEVAL IN GENERAL COLOR IMAGE DATABASES USING COLOR EDGE GRADIENT CO-OCCURRENCE HISTOGRAMS

Raymond Phan *, *John Chia* †, *and Dimitrios Androutsos* ‡

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

ABSTRACT

In this paper, we present a logo and trademark retrieval system for general, unconstrained, color image databases, extending the Color Edge Co-occurrence Histogram (CECH) object detection scheme. 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, in comparison to the simple color pixel difference classification of edges as seen in the CECH. Our proposed method is thus reliant on edge gradient information, thus we call it the Color Edge Gradient Co-occurrence Histogram (CEGCH). We also introduce a novel color quantization scheme based in the Hue-Saturation-Value (HSV) color space, illustrating that it is more suitable for image retrieval in comparison to the color quantization scheme introduced with the CECH. Results illustrate that our retrieval system retrieves logos and trademarks with good accuracy, outperforming the use of the CECH in image retrieval with higher precision and recall.

***Index Terms*—** Color Edge Gradient Co-occurrence Histogram (CEGCH), logo and trademark retrieval, color edge detection, HSV quantization, pattern recognition

1. INTRODUCTION

Content Based Image Retrieval (CBIR) is an area of multimedia processing with many promising applications. Much research has been devoted to CBIR, and due to its many subsets, in this paper, we focus solely on the logo and trademark retrieval viewpoint. It is seen as an increasingly vital tool for industry, commerce, and trademark registration [1] and also for sports entertainment [2]. Before research in logo and trademark retrieval began, Swain and Ballard [3] proposed a method for color image retrieval involving color histograms. Though their illustrated results were very promising, the color histogram only captures the color content of an image, resulting in the misclassification of retrieved images. These may have similar color distributions but different spatial statistics in compared to the query. Therefore, to resolve this problem, methods were proposed to add additional information to color histograms.

Chang and Krumm [4] proposed the **Color Co-occurrence Histogram (CCH)** for use in their object detection scheme, being a subset of image retrieval. This modification to the color histogram captures both the color and spatial relationships between color image pixels, effectively modelling texture. The CCH is a three-dimensional histogram containing the number of color pixel pairs, c_1 and c_2, that occur at different spatial distances, d. They used CCHs as the main tool for their object detection scheme, and proved that it was quite an effective method. However, Luo and Crandall [5] 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. As such, they 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 was the tool used in their object detection scheme, which also focused on *unconstrained* color images.

In Luo and Crandall's work, the construction of the CECH was to first quantize the query and search images using a two-stage color quantization process. The first stage involves quantizing an image using a set of 267 colors defined by the Inter-Society Color Council and the National Bureau of Standards (ISCC-NBS). The second stage performs further quantization down to a set of 11 colors: S = $\{$*black, white, gray, brown, yellow, orange, pink, red, purple, blue, green*$\}$. 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. Their definition of an edge pixel is any pixel having a different color from any of its 8 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 their work were very good, we feel that a more appropriate definition of what an edge pixel is in a color image should be employed, as opposed to their simple color pixel difference for edge classification. To this end, we present a paper that extends their work of using CECHs in object detection for use in logo and trademark retrieval in unconstrained color image databases. We introduce *color edge detection* to the CECH, producing an edge map determining valid edge points in color images with greater accuracy. This edge map is used in conjunction with the CECH to perform logo and trademark retrieval in unconstrained color images. The inclusion of this information to the CCH thus makes it dependent on the edge gradient, and so we name this the **Color Edge Gradient Co-occurrence Histogram (CEGCH)**. We also introduce a novel color quantization method based in the Hue-Saturation-Value (HSV) color space that is more suitable for image retrieval in unconstrained color images in comparison to the quantization scheme seen in the CECH.

The rest of this paper presents our retrieval system and experimental findings. Section 2 describes our methodology, including

*rphan@ee.ryerson.ca
†jchia@ee.ryerson.ca
‡dimitri@ee.ryerson.ca

the color edge detection algorithm used, the HSV color quantization method proposed, and the mechanics used for logo and trademark retrieval in unconstrained color image databases. Section 3 illustrates retrieval results using the proposed framework applied to an unconstrained color image database, and the evaluation of our CEGCH in comparison to the CECH as seen in Luo and Crandall's work. Finally, Section 4 will present our conclusions and future research in this area.

2. METHODOLOGY

As with Luo and Crandall's work, the input to the retrieval system is the logo or trademark of interest to be retrieved, and its size is the smallest possible representation of the logo or trademark to detect. The dimensions of this logo or trademark are defined as m_i x n_i. The following is an overview of the algorithm:

Processing of the Input Image: The input image has its edge map calculated by performing color edge detection on it. After, the image is quantized using our HSV color quantization algorithm, and the CEGCH of the input query is built and stored using the aforesaid information.

Processing the database images: To reduce computation time, each database image is subsampled to a resolution of 256 x 384 pixels, as done in Luo and Crandall's work [5] before further processing. For detecting multiple sizes of the input query, multiple scale factors of the input query must be considered. As such, we search for the input query at multiple sizes by further subsampling the database image in question at multiple scale factors. We consider scale factors between 1.0 and γ_H. γ_H is defined in Equation 1 as:

$$\gamma_H = \min\left(\frac{m_d}{m_i}, \frac{n_d}{n_i}\right), \tag{1}$$

where m_d x n_d are the dimensions of the database image of interest. This restriction is in place, for instances of the input to detect cannot exist beyond the dimensions of the database image. Choosing scale factors that are uniformly spaced from one another is one possible option. However, proper size matching is more important at smaller scale factors than at larger ones. As such, we choose our set of scale factors, $\gamma_S = \{\gamma_0, \gamma_1, \gamma_2, \ldots, \gamma_H\}$, such that $\gamma_0 = 1.0$, and $\gamma_{j+1} = \alpha\gamma_j$, where $\alpha > 1.0$ is a constant, as seen in [5]. Color edge detection is performed on each subsampled database image to produce their corresponding edge maps, and color quantization is performed after using our HSV color quantization method.

Image Retrieval: Finally, for a quantized database image at a scale factor, overlapping search windows of size m_i x n_i are used to search for the input query. The CEGCHs of every search window at a scale factor are computed and are compared to the input query CEGCH. Once searching has completed at a scale factor, the next subsampled quantize database image is used, and overlapping search windows of the aforesaid size are used to search for the query, and is repeated for the rest of the scale factors for the database image. The search window at any scale factor within this image, producing the greatest similarity value, is used as the overall similarity for this image. Also, the co-ordinates of this search window at this scale factor are recorded, and a bounding box is drawn around it, illustrating that the system found the logo or trademark to be within this search window. This aforesaid procedure is repeated for all database images. Once overall similarity values are generated for all database images, they are used to rank the images. It should be noted that our current algorithm does not have the ability to detect multiple logos or trademarks, or those that are subject to large amounts of deformation. Also, in terms analyzing computational complexity, its illustration is not the objective of this paper; rather, this paper is to demonstrate that the CEGCH is a viable mechanism to facilitate logo and trademark retrieval in unconstrained color images. However, these are areas of research we are currently investigating.

2.1. Color Edge Detection

The color edge detection approach that we take is in the vector space domain, using vector order statistics, proposed by Trahanias and Venetsanopoulos [6][7]. This family of color edge detectors proves to have extremely good results by retaining edges, and is robust against the presence of noise [7]. In the vector space, color pixels are treated as three-dimensional vectors. Pixels in a neighborhood (e.g. 3 x 3, 5 x 5, ...) are ordered based on some criteria. We employ *reduced ordering* to order these color pixels. Each color pixel in the neighborhood is reduced to a scalar value, determined by a distance measure based on color pixel values, known as the *reduction function*. Once each color pixel is reduced to a scalar value, ordering of the color pixels is based on these scalar values. As such, let ***X*** = $\{X_1, X_2, X_3, \ldots, X_n\}$ represent the color pixels within a pixel neighborhood. The reduction function that we used is based on the *aggregate* distance metric. It is defined as the cumulative distance from the i^{th} color pixel to every pixel within this neighborhood. Or in other words, as seen in Equation 2:

$$d_i = \sum_{k=1}^{n} ||X_i - X_k|| \;,\; i = 1, 2, \cdots, n \;, \tag{2}$$

where $||\cdot||$ represents the Euclidean distance, or L_2 norm. The color edge detector we chose from the vector order statistics family was the *minimum vector dispersion* edge detector (MVD). As such, let **Z** = $\{X^{(1)}, X^{(2)}, X^{(3)}, \ldots, X^{(n)}\}$ represent the sorted color pixels within a pixel neighborhood in ascending order after applying the reduction function. The output of the MVD edge detector defined at the center of a pixel neighborhood is:

$$MVD = \min_j \left(\|X^{(n-j+1)} - \sum_{i=1}^{m} \frac{X^{(i)}}{m}\| \right), \tag{3}$$

where $j = 1, 2, \ldots, k$ and $k, m < n$.

Thresholding these values creates an edge map, and there is no need to filter any of the images for noise, for this edge detector automatically performs noise filtering [6][7]. By consulting Equation 3, averaging the lowest ranked m color pixels results in reducing short-tailed, or Gaussian noise. Choosing the minimum value of the distances between the k highest ranked pixels with the average color pixel eliminates impulsive noise, simultaneously performing two different methods of noise suppression in a single edge detector, making its selection a robust choice. Also, it more accurately represents what is considered an edge in an image in comparison to the edge map produced with the CECH. The parameters *k*, *m*, and the size of the pixel neighborhood for the MVD are tunable to produce the best possible edge detection results for different images in applications where accuracy needs to be controlled. By increasing k, the likelihood of impulsive noise appearing as an edge in the output is decreased, but will introduce possible short-tailed noise at the output. Increasing m will increase the amount of averaging, reducing the amount of short-tailed noise that will be seen at the output, but will create more inaccurate edge points due to blurring.

Color	Lower Threshold	Upper Threshold
Orange	25^o	46^o
Yellow	46^o	70^o
Green	70^o	165^o
Blue	165^o	270^o
Purple	270^o	340^o
Red	340^o	25^o

Table 1. Range of hues for color mapping. Note, $360^o = 0^o$.

2.2. HSV Color Quantization Method

The HSV color space was chosen as the basis for our quantization scheme, for its parameters naturally mimic our perception of color, thus it is a huge advantage for color quantization. It simplifies the task for a human to identify the problematic areas of an algorithm and fine tune them to produce better results [8]. Starting with an adaptation of nearest neighbor for HSV, we introduced thresholds to compensate for the non-linearity of the color space and finally adjusted some thresholds to be non-linear themselves. We partition the HSV color space into 11 regions which correspond to 11 colors that correspond to the same basic set of 11 colors seen in [5]. We first perform transformation of a given color in RGB to the HSV color space using the standard conversion formulae [8]. We then determine if the color is chromatic or achromatic by examining the given color's saturation in relation to its value. With respect to its value, if a color is achromatic, we treat it as either white ($V > \frac{2}{3}V_{max}$), grey ($\frac{1}{3}V_{max} \leq V \leq \frac{2}{3}V_{max}$) or black ($V < \frac{1}{3}V_{max}$). Chromatic colors are treated similarly, except there are six bins (red, blue, purple, green, yellow and orange) instead of three bins, and the bins are unequal in size. The range of hues which map to each color is given in Table 1. To assign the remaining two colors (pink and brown), if the color is assigned to the red bin, we perform a simple nearest neighbor search on the color set *C* = {*red, pink, brown*}. By examining a standard CIE color chooser, we found that the chromatic-achromatic border resembled a parabolic curve in the Saturation-Value plane, forming the basis of our method to determine if a given color is achromatic. We treat a color as chromatic or achromatic by the position of its saturation and value with respect to the chromatic-achromatic border that we identified in the CIE color chooser. We modelled the border by trial and error to be roughly: $S = 2(V-1)^4 + 0.15$ (where S and V are normalized). Therefore, the chromatic region is when $S > 2(V-1)^4 + 0.15$ and the achromatic region is when $S \leq 2(V-1)^4 + 0.15$.

Figure 1 illustrates a color quantization comparison between the method devised by Luo and Crandall with our method. Compared to the method in [5], our method produces subjectively better results. Also, for the color quantization method devised in [5], the ISCC-NBS color set was modified each time for the object to detect. The color quantization result using the method in [5] was generated without any modification to the color set. As such, our quantization method is more generalized; it does not require tuning for the desired object to detect.

2.3. Similarity Measure

The similarity measure used to compare CEGCHs is by means of histogram intersection, introduced in Swain and Ballard's work [3] for their image retrieval scheme. For CEGCHs, histogram intersection is defined as:

$$s(h_s, h_i) = \frac{\sum_{i=1}^{c}\sum_{j=1}^{c}\sum_{k=1}^{d}\min(h_s[i][j][k], h_i[i][j][k])}{\sum_{i=1}^{c}\sum_{j=1}^{c}\sum_{k=1}^{d} h_i[i][j][k]}, \quad (4)$$

where h_s is defined as the CEGCH within a search window of a database image at a scale factor, and h_i represents the CEGCH of an input image. c represents the total number of colors (in our case, 11) and d represents the maximum distance two color pixels can be separated over a spatial distance. The spatial distance metric we used was the quantized Euclidean distance, as seen in [5]. This is defined as the mathematical floor of the Euclidean distance between two color pixels in an image. We chose histogram intersection as the comparison metric for it is a very simple similarity measure to implement and it provides a very quick way in determining whether our modification to the CECH yields acceptable retrieval results.

3. EXPERIMENTAL RESULTS

To test the feasibility of our retrieval scheme, retrieval trials were performed using two different logos: The Ferrari, and Lufthansa logos. These are shown in Figure 2. Also, the dimensions of the logos are 37 x 40, and 28 x 31 pixels respectively.

A retrieval test for each logo was performed using our proposed CEGCH and HSV color quantization scheme, and for comparative purposes, we performed the same tests using the CECH and its color quantization scheme. As with Figure 1, we did not modify the original ISCC-NBS color set to ensure unbiased results. Our image database consists of a 5,000 unconstrained color image database we compiled, where 100 images contained the Ferrari logo, and 100 images contained the Lufthansa logo. The parameters for the MVD edge detector were $k = 8$, $m = 12$, and the pixel neighborhood was 5 x 5, as seen in [6]. These parameters were chosen based on their results, for they generated excellent edge maps under the influence of different combinations of short-tailed and impulsive noise. Also, the gradient threshold for the MVD edge detector was set to 48, and $\alpha = 1.1$, as seen in Luo and Crandall's work [5]. The maximum spatial distance that two color pixels can be separated in an image to calculate co-occurrences was 2, corresponding to a 5 x 5 window. Also, to decrease computational time, γ_H was empirically found to be 3.0, which suited the detection of the input queries of interest with all images containing the input logos in order to generate faster retrieval results. To ensure uniformity for comparison, histogram intersection was the similarity measure used for both tests, and for both methods. The best measure to evaluate the performance of a retrieval system would be through Precision-Recall graphs. Figure 3 represent the Precision-Recall graphs for each test, demonstrating the performance of using our proposed CEGCH in comparison to

Fig. 1. Original Image (left) and comparison of result of Luo and Crandall's quantization method (middle) versus HSV quantization (right)

Fig. 2. Original input Ferrari (left), and Lufthansa (right) logos

Luo's CECH with the Ferrari and Lufthansa logos as the input query images.

Results in Figure 3 demonstrate that the proposed CEGCH shows higher precision and recall than the CECH. The CECH demonstrates low precision and recall in comparison to the proposed CEGCH, meaning that images containing the logo of interest using the CECH are ranked much lower in comparison to the CEGCH. Also, this states that the CEGCH requires fewer images to retrieve the entire logo data set in the database in comparison to the CECH. The low precision and recall using the CECH can be attributed to many factors; most notably, its edge map definition misclassifies edges. Pixels can be included in the calculation of the co-occurrence histogram which should not originally be there, rendering the CECH to be susceptible to false positives. This would rank images that did not contain the logo of interest to be higher than what they should be. To attest to this statement, Figure 4 illustrates some sample retrieval results indicating where the system located the Ferrari logo for two unconstrained color images using the CECH and color quantization method in [5] and with our proposed system. Using the CECH fails to correctly delineate where the logo is in the images, where our system locates them properly. Also, 23 of the 100 Ferrari logo images and 4 of the 100 Lufthansa logo images were incorrectly delineated using the CECH. Our method correctly delineates where the logo is for all of the logo images. This illustrates that our method has higher precision and recall in comparison to the CECH, and our method accurately locates the logos in spite of rotation and partial deformation.

4. CONCLUSIONS AND FUTURE WORK

We presented an extension of the color object detection method devised by Luo and Crandall that was intended for use in logo and trademark retrieval in unconstrained color image databases. Luo and Crandall's work erroneously classifies edges in their edge map; therefore, we generated with a more reliable edge map with color edge detection using vector order statistics. We also proposed a novel color quantization scheme based in the HSV color space, and demonstrated that it is a more suitable mechanism for image retrieval in comparison to scheme proposed with the CECH. The retrieval results illustrate that our system can accurately determine the location of the logo or trademark of interest in spite of partial deformation, and produces higher precision and recall in comparison to the CECH. However, there are several aspects of our system that could be improved upon. Better retrieval results can be achieved by creating a more robust similarity measure, as opposed to histogram intersection. We are investigating pre-screening approaches, where images in the database not containing the correct spatial-color statistics to that of the input logo are not retrieved, ultimately reducing computation time by effectively narrowing the search to images possibly containing the logo of interest. Finally, we are investigating the detection of logos and trademarks under significant deformations, and detecting multiple instances of a logo or trademark within a database image.

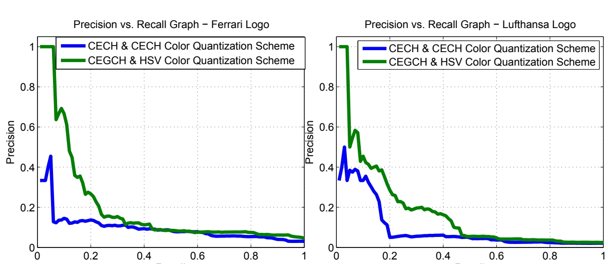

Fig. 3. Precision-Recall graphs for the Ferrari (left) and Lufthansa (right) logos

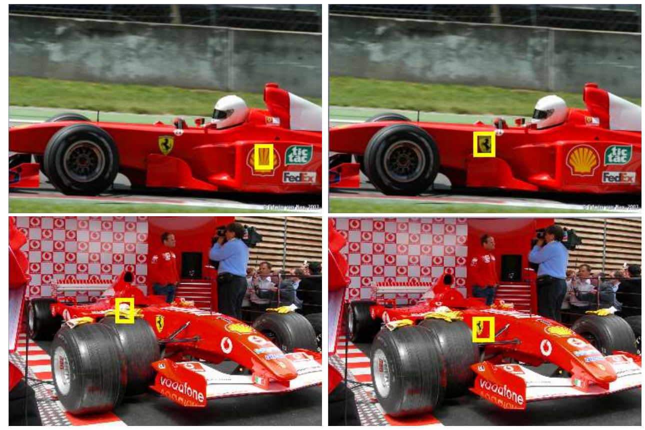

Fig. 4. Sample detection results: The CECH (left column) and the CEGCH with HSV color quantization (right column)

5. REFERENCES

[1] J. P. Eakins, J. M. Boardman, and M. E. Graham, "Similarity retrieval of trademark images," *IEEE Multimedia*, vol. 5, no. 2, pp. 53–63, 1998.

[2] H. Pan, B. Li, and M. I. Sezan, "Automatic detection of replay segments in broadcast sports programs by detection of logos in scene transitions," in *Proc. IEEE ICASSP*, 2002, vol. 4, pp. 13–17.

[3] M. J. Swain and D. H. Ballard, "Color indexing," *Intl. Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.

[4] P. Chang and J. Krumm, "Object recognition with color cooccurrence histograms," in *IEEE Trans. on CVPR*, 1999, vol. 2, pp. 23–25.

[5] J. Luo and D. Crandall, "Color object detection using spatial-color joint probability functions," *IEEE Trans. on Image Processing*, vol. 15, no. 6, pp. 1443–1453, 2006.

[6] P. Trahanias and A. N. Venetsanopoulos, "Color edge detection using vector-order statistics," *IEEE Trans. on Image Processing*, vol. 2, no. 2, pp. 259–264, 1993.

[7] K. N. Plataniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*, Berlin, Germany: Springer-Verlag, 2000.

[8] A. Watt, *3D Computer Graphics*, Addison Wesley: Boston, MA, USA, 1989.