SPATIAL CONNECTED COMPONENT PRE-LOCATING ALGORITHM FOR RAPID LOGO DETECTION

Yuan Zhang, Shuwu Zhang, Wei Liang, Hai Wang
Hi-tech Innovation Center, Institute of Automation, Chinese Academy of Sciences, Beijing, China
{yzhang, swzhang, wliang, haiwang}@hitic.ia.ac.cn

ABSTRACT

This paper introduces a novel pre-locating algorithm for rapid logo detection in unconstrained color images. This work is distinguished by two major contributions. The first is a new method of representation for logo called "spatial connected component descriptor" (SCCD) containing connected component (CC) prediction model and effective-CC pixel distribution histogram. The former represents combinations between CCs based on color and spatial relationships of CCs. While the latter describes the pixel distribution information of effective CCs. The two parts capture the layout of logos from different points. The second is a logo pre-locating algorithm by the means of SCCD to search for logo prediction regions in test images, on which some content-based features are used for logo matching. Experimental results illustrate that our pre-locating algorithm speeds up logo detection to a great extent and shows precise location compared to previous systems.

***Index Terms*—**spatial connected component descriptor, logo detection, content-based features, object detection and recognition

1. INTRODUCTION

Logo detection is an important image understanding task in Content-based Image Retrieval (CBIR), which has won much attention in literature due to its various applications [1]. Color and shape are two features that are mostly important for logos, with which human beings can distinguish various kinds of logo. Therefore, they have been embraced by many different approaches for logo retrieval and detection.

Color histogram [2] has been researched thoroughly for retrieval because of its simplicity and ease of implementation. But it only relies on the color without considering shape and texture. Chang and Krumm [3] proposed the Color Co-occurrence Histogram (CCH) to capture both color and spatial content of an image. Luo and Crandall [4] presented an improvement to CCH by only considering edge pixels, named Color Edge Co-occurrence Histogram (CECH), which performed better than CCH. Phan et. al. [5] proposed a new algorithm that introduced more accurate information to the CECH by virtue of incorporating color edge detection using vector order statistics and a HSV color quantization method, called Color Edge Gradient Co-occurrence Histogram (CEGCH). Hesson introduced Color Wavelet Co-occurrence Histograms by adding wavelet transforms into CCH [6]. Kleban proposed a spatial pyramid mining scheme to identify frequent spatial configurations of local features [7]. Penfei Xu et. al. adopted the integration model of spatial feature correlations [8].

Despite inspiring results of these researches in logo detection, they only laid much stress on studying features while overlooking searching strategy, which gives rise to two defects. Firstly, some methods subsample the database images to a low resolution in order to reduce computation. However, logos in the images may become too small to be detected. Secondly, most of the works still resort to overlapping sliding windows (OSW) to search for the input query at different scales, which brings about much unnecessary computation. Furthermore, the search window is not perfect enough to locate logos with arbitrary size precisely.

Thus, in this study, we focus on developing a novel pre-locating algorithm to offset the two flaws in order to realize rapid logo detection in arbitrary size images, guaranteeing both high precision and efficiency. Logos are compound color objects which have a specific set of multiple colors that are arranged in a unique spatial layout [4]. Hence they can be taken as varied spatial combinations of some CCs gotten by color segmentation. Meanwhile, in unconstrained images [5], logos are a tempting target for detection as they often appear in high contrast regions with distinctive shape and edge information, making it convenient to get CCs of logos. Therefore, as the pre-locating feature, CCs are distinctive enough to obtain the probable region for logo matching.

The major contributions of our work include:

- introducing spatial CC descriptor (SCCD) of logos for pre-locating logos in unconstrained images, which captures color and spatial information of CCs. SCCD is compact and robust across some degree of rotation, self-occlusion and spatial distortion, difficult illumination, scaling, etc.
- proposing a pre-locating algorithm based on CC matching for logo detection, which can generate as few and precise prediction rectangles as possible.

The rest of our paper is organized as follows. Section 2 describes our pre-locating algorithm, including a detailed description of SCCD for logo and CC descriptor (CCD) for unconstrained color images, and a fast pre-locating algorithm. Section 3 gives the logo matching procedures. The experimental results are presented in section 4 and final conclusions are given in section 5.

2. LOGO PRE-LOCATING

The pre-locating process contains two parts. One is to get SCCD for logos and CCD for unconstrained color images respectively. The other is to gain logo prediction rectangles (LPR) based on our pre-locating strategy.

2.1. Spatial Connected component descriptor

For each logo, we use a model image as the query logo image. Fig.1 illustrates the procedure of generating SCCD, which is composed of two main parts: HSV space CC segmentation and establishment of SCCD.

• HSV space CC segmentation

We firstly refer to HSV quantization implementation [5]. However, in order to overcome lighting and over-segmentation, we partition the HSV color space into a set of 8 colors: *S*= {*black, white, red, green, yellow, blue, orange, purple*}. Then in color smoothing, for a noise color (e.g. pixel number < 20) pixel, we set the main color that appears most frequently surrounding it as its color. We obtain effective CCs reliant on color regions except for background (e.g. white) and number all CCs in descending order of area. In order to reduce computation, the top five CCs are marked as anchor CCs representing the logo, which are used for CC matching in the process of getting LPRs. If the number of CCs is less than five, all CCs are marked as anchor CCs.

• Establishment of SCCD

After getting effective CCs, it is time to set up SCCD which is made up of two main parts, illustrating in Fig.2.

The first is CC prediction model containing anchor-CC-combination and spatial unit of effective CCs. Three logos in Fig.2a~c are arranged to show the process. The representation of anchor-CC-combination is color-combinations (<=3). In Fig.2a, the representation is {red-[(1)], blue-[(2)]}, while in Fig.2c it's {red-[(1), (2), (3), (4), (1 2), (1 3), (1 4), (2 3), (2 4), (3 4), (1 2 3), (1 2 4), (1 3 4), (2 3 4)]}. In each combination, we keep a record of its color, circumscribed rectangle, area and edge point-centroid distance-angle histogram (e.g. 10×12). When the logo in a test image is so small that the same color CCs cannot be divided precisely, our anchor-CC-combination can deal with this situation. For spatial unit of effective CCs, we employ circumscribed rectangle-CCs to describe it. If the circumscribed rectangle of a CC is covered by another, we add the CC to the bigger circumscribed rectangle. If not, we take down this circumscribed rectangle-CC. In Fig.2b the circumscribed rectangle of CC-2 is enclosed by that of CC-1.

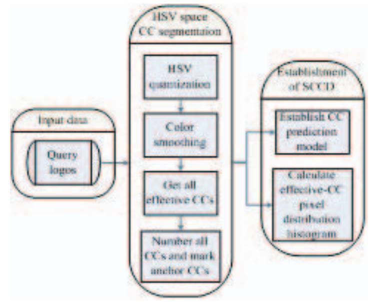
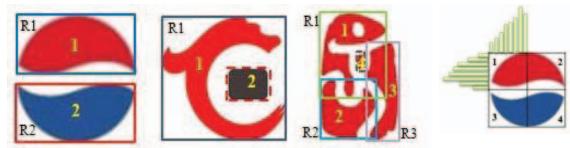


Fig. 1 Process of getting SCCD



a b c d

Fig. 2 Establishment of SCCD

So its spatial units is {R1-(1 2)}. Similarly, spatial unit of Fig.2a and c are {R1-(1), R2-(2)} and {R1-(1 4), R2-(2), R3-(3)} respectively.

The second is effective-CC pixel distribution histogram, the function of which is to prescreen LPRs. In Fig.2d, a logo is divided into four equal blocks. We calculate horizontal and vertical 8-bin effective-CC pixel distribution histograms located in each block. Then each histogram is normalized by the block area. Therefore, for each logo there are eight histograms containing space information of the logo to some extent.

2.2. Connected component descriptor

Due to the complicated background of logos in unconstrained images, we add further segmentation in gray space to obtain as accurate CCs of logos as possible. The procedure of generating CCD includes three main parts: HSV space CC segmentation, gray space CC segmentation and establishment of CCD.

• HSV space CC segmentation

Test images are first processed in the very way that logos have adopted in HSV quantization and color smoothing. Then we get primary CCs based on all color regions.

• Gray space CC segmentation

When the area of an original CC is above a threshold (e.g. 100) and its gray-level variance is above a threshold (e.g. 50), the original CC is separated under Otsu's algorithm [9] into two parts. If the difference between gray-level averages of the two parts is above a threshold (e.g. 50), the original CC is replaced by the sub-CCs obtained based on the two parts. If not, the original CC is kept.

Each CC is divided in gray-level as described above until each of the CCs could not be divided into any smaller parts. Then CCs with the pixel number below a threshold (e.g. 25) are abandoned. The left are effective CCs.

• Establishment of CCD

Fig. 3 Getting a LPR in the test image for Pepsi

There is no need for unconstrained images to get spatial information of CCs, thus compared to SCCD, CCD is much simpler. For each effective CC, we note down its color, area, circumscribed rectangle and edge point-centroid distance-angle histogram.

2.3. Logo prediction rectangle

The procedure of logo pre-locating comprises two main steps: search for LPRs based on prediction model and prescreen LPRs based on pixel distribution histogram.

• Search for LPRs

When it comes to CC in the test image and each combination with the same color in the logo anchor-CC-combination, we first compute the similarity of their circumscribed rectangles, given by

$$S_1(R_{tc}, R_{lc}) = W_{tc}/H_{tc} - W_{lc}/H_{lc} > 0.8 \qquad (1)$$

$$S_2(R_{tc}, R_{lc}) = W_{tc}/W_{lc} - H_{tc}/H_{lc} > 0.5 \qquad (2)$$

where *tc*, *lc*, *R*, *W* and *H* are the CC in the test image, the combination in the logo, the circumscribed rectangle, width and height of the circumscribed rectangle. If S_1 and S_2 satisfy (1) and (2) respectively, we reckon the similarity of their edge point-centroid distance-angle histogram by means of histogram intersection [5], defined as

$$S_3(h_{tc}, h_{lc}) = \frac{\sum_{i=1}^{a}\sum_{j=1}^{d} min\,(h_{tc}(i,j), h_{lc}(i,j))}{\sum_{i=1}^{d}\sum_{j=1}^{a}\; h_{lc}(i,j)} > 0.6 \quad (3)$$

where *h*, *a* and *d* denote histogram, dimension of angle and distance respectively. If S_3 corresponds with (3), the *tc* and *lc* are regarded as a CC matching pair. Then a LPR which *tc* belongs to is delineated based on the scale of r_{tc} relative to r_{lc} and the relative-location of r_{lc} in the logo circumscribed rectangle. Each CC in test images is dealt with as the above process, getting a LPR. One example is illustrated in Fig. 3.

• Prescreen LPRs

Firstly, LPRs are filtered based on overlapping ratio of area. Each LPR, considered as a target LPR, is replaced by the combination of it and all other LPRs that satisfy

$$r_t = S_o/S_t > 0.9 \;\&\&\; r_a = S_o/S_a > 0.9 \qquad (4)$$

where r_t and r_a are ratio of overlapping area over target LPR and another LPR. S_o, S_t and S_a are area of overlapping region, target LPR and another LPR. After the combination, one of the LPRs with the same location is kept.

Secondly, the remaining LPRs are prescreened based on effective-CC pixel distribution histogram. We process all the LPRs as described below. All rectangles in spatial unit of the logo are projected into each LPR based on scale of the LPR relative to the logo circumscribed rectangle and relative-locations of these rectangles in the logo circumscribed rectangle. Any CC with its 80% area located in one of these rectangles of the LPR is considered as effective CC. After acquiring all effective CCs for the LPR, we compute effective-CC pixel distribution histogram of the LPR and compare it with that of the logo. When the similarity is above a threshold (e.g. 0.35), the LPR is kept.

3. LOGO MATCHING

After getting LPRs through logo pre-locating, we embrace some content-based features for logo matching.

In our method, we use block color histograms [2]. We calculate 8-color block color histogram in HSV quantization space for the query logo and each LPR, and adopt histogram intersection (3) as the similarity measure for color histogram. Then we apply CEGCH to test the query and the LPR. It is important to note that color edge detection is just processed in the local region of the LPR instead of the whole test image as in [5]. We employ the similarity measure presented in [4] for CEGCH because of its better performance than histogram intersection. Finally, overall similarity between the query and LPR is give by

$$S(Q,L) = k_1 S(C_Q, C_L) + (1-k_1) S(CH_Q, CH_L) \qquad (5)$$

where $0<k_1<1.0$ is a constant (e.g. k_1=0.5). Each of the LPRs with its similarity above a threshold (e.g. 0.6) is declared to be the target logo. Moreover, the LPR with the best score is regarded as the most likely logo region, and its score is marked as the score of the test image.

4. EXPERIMENTS

We evaluate the pre-locating method in logo detection on an unconstrained color image [5] database, which consists of 2000 unrelated images and 480 logo images containing 6 logo categories: *Starbucks (S), Pepsi (P), Baidu (B), Mengniu (M), HSBC (H) and China Netcom (C) logos* collected from the web. This is a challenging database because of the resolutions of images ranging from 116×48 to 1536×2048, the logo in images extending from 18×20 to 300×300, and different photography appearances. All experiments were conducted on a standard desktop PC (Intel Core 2 Duo CPU 3.16GHz E8500) running Windows XP compiled in Microsoft Visual Studio 2008.

Table 1: Average detection time for different logos

Logo / T(s)	S	P	B	M	H	C
SCCD	0.6	0.4	0.3	0.2	0.4	0.3
SCCD+CEGCH	**3.5**	**1.1**	**2.5**	**0.8**	**0.9**	**0.8**
OSW+CEGCH	32.3	19.7	26.1	17.7	20.2	18.4

We compare SCCD with OSW on a resolution of 256×384 database, subsampled from our own [5]. Table 1 summarizes average detection time of SCCD pre-locating, CEGCH with SCCD and OSW, which demonstrates that our pre-locating algorithm speeds up the detection time to a large extent and accounts for a minor part of the whole detection time. Moreover, we present the detection time-area (resolution) curve of pre-locating and overall detection

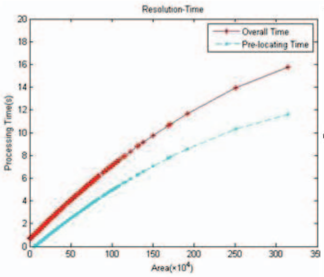


Fig. 4 The curve of detection time-area

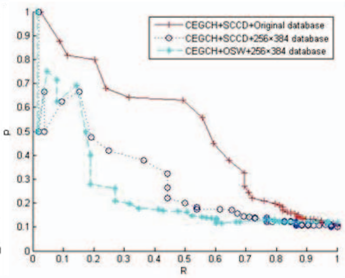


Fig. 5 Precision-Recall curves

in our original image database in Fig. 4, which indicates that with the growing resolution, both pre-locating and overall detection time of our method grow steadily. SCCD describes color and spatial features of logos in a simple manner without much computation and complexity. Meanwhile our pre-locating procedure is reliant on color indexing and structure feature of logos, both of which are easy to carry out. Therefore, our method can handle arbitrary size images in a short time.

In addition, Fig. 5 shows the precision & recall curves for CEGCH with SCCD and OSW on the subsampled database, and CEGCH with SCCD on our original database. It illustrates that our proposed approach outperforms OSW in all precision and recall settlements with both resolution databases, due to the precise LPRs.

Table 2: Proper and missed detections with SCCD pre-locating and OSW in the subsampled database (Ss &Os), and SCCD in original image database (So).

Category	Starbucks			Pepsi			Baidu		
	So	Ss	Os	So	Ss	Os	So	Ss	Os
Proper	**65**	44	35	**77**	66	56	**71**	61	45
Missed	**15**	36	45	**3**	14	24	**9**	19	35
Category	Mengniu			HSBC			China Netcom		
	So	Ss	Os	So	Ss	Os	So	Ss	Os
Proper	**73**	36	27	**74**	48	37	**72**	45	34
Missed	**7**	44	53	**6**	32	43	**8**	35	46

Finally, we present precision of the bounding box [5] in Table 2, which illustrates that our pre-locating method on original image database performs much better because of our accurate depiction of layout in SCCD. Fig. 6 indicates the reasonable detection results of the proposed method, which demonstrate that the proposed approach is robust to handle scaling, occlusions, illumination change, etc. Besides, extremely small logos (e.g. 18×20) and multiple logos per image can be detected very quickly via our method.

5. CONCLUSION

This paper presents a new and general pre-locating algorithm based on SCCD for rapid logo detection in unconstrained images. By employing SCCD for CC matching, we get as few and exact LPRs as possible for logo detection. Results show that our method can speed up logo detection considerably while ensuring precise detection. In the future work, we would further integrate our pre-locating algorithm for more application scenarios such as rapid object recognition and retrieval in the data deluge. In addition, we are interested in developing more robust features for logo matching. We believe our pre-locating research has far reaching implications for many applications in computer vision.

Fig. 6 Some results on our own dataset.

6. ACKNOWLEDGEMENTS

The wok is supported by National Key Technology R&D Program of China under Grant No.2009BAH48B02, 2009BAH43B04, 2011BAH16B01 and 2011BAH16B02.

7. REFERENCES

[1] S.Y. Arafat, S.A. Husain, I.A. Niaz and M. Saleem, "Logo Detection and Recognition in Video Stream," *ICDIM*, pp. 163-168, 2010.

[2] M.J. Swain and D.H. Ballard, "Color indexing," *International Journal of Computer Vision 7 (1)*, pp. 11–32, 1991.

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

[4] J. Luo and D. Crandall, "Color object detection using spatial–color joint probability functions," *IEEE Transactions on Image Processing 15 (6),* pp. 1443–1453, 2006.

[5] R. Phan and D. Androutsos, "Content-based retrieval of logo and trademarks in unconstrained color image databases using Color Edge Gradient Co-occurrence Histograms," *Computer Vision and Image Understanding (114)*, pp.66–84, 2010.

[6] A. Hesson and D. Androutsos, "Logo and trademark detection in images using color wavelet co-occurrence histograms," *ICASSP*, pp.1233 – 1236, 2008.

[7] J. Kleban, X. Xie, and W. Ma, "Spatial Pyramid Mining for Logo Detection in Natural Scenes," *ICME*, pp.1077-1080, 2008.

[8] Pengfei Xu, Hongxun Yao and Rongrong Ji, " SIGMA: Spatial integrated matching association algorithm for logo detection," *ICASSP*, pp. 1086 – 1089, 2010.

[9] N. Otsu, "A threshold selection method from gray level histograms," *IEEE Transactions on Systems Man and Cybernetics (9)*, pp.62–66, 1979.