

A method of TV Logo Recognition based on SIFT

Da Pan¹ , Ping Shi²

Abstract TV logo is an important symbol of TV station. Automatic identification of TV logo can be used in television monitoring and media assets management. By analyzing the characteristics of TV logos, a new algorithm for TV logo recognition is proposed in this paper. First of all, the TV logo information is obtained by applying a special mask image to the captured video frames. Then the coarse matching is performed based on color tone. Further more, the region of TV logo is separated into the master region and sub region and the SIFT keypoints are obtained based on this separation. Finally, the TV logo is recognized by comparing the number of keypoints with a threshold. Experiments results show that the algorithm proposed has a good performance in TV logo recognition, especially for transparent logo.

Keywords: TV logo, SIFT descriptor, Recognition, keypoint

1 Introduction

TV logo implies semantics about TV channel name and program tendency and is one of the important semantic sources for video analysis, comprehension and retrieval. It is widely used in television monitoring and media assets management.

Generally, TV logos can be divided into two main types: static logos and dynamic logos. Most TV logos are static. In addition, according to different transparency, TV logos can also be divided into opaque and transparent. For static images, Shi Yingchun[9] proposed a method of Space color histogram describing the characteristic of TV logos, and combined knowledge based on histogram statistics to go through all the elements. But its computation efficiency is low with the increase of the knowledge and it doesn't work effectively for transparent TV logos. Yang Qiang[10]presented an algorithm which extracts the

¹ Da Pan (✉)

School of information engineering, Communication University of China
e-mail: 823454879@qq.com

² Ping Shi

School of information engineering, Communication University of China

wavelet-character of TV Logo and realizes the TV symbol recognition. But its accuracy is not high enough. Jin Yang [11] utilized a method based on binary image corner matching by three—step process. But it is only suitable for long video sequences. Liang Zhang[12] presented a method of Hollow TV logo detection which has a high recognition rate. But this method is only for hollow TV logo.

Scale Invariant Feature Transform (SIFT) was introduced by Lowe[6-7] and is invariant to rotation, translation and scale variation between images and partially invariant to affine distortion, illumination variance and noise. Apostolos Psyllos[15] presented an interesting approach based on an enhanced Scale Invariant Feature Transform (Merge-SIFT or M-SIFT) for vehicle Logo Recognition. In paper[13] , SIFT has already been used for recognizing faces in controlled situations. In paper[14] , a new vehicle tracking algorithm based on UKF and SIFT is proposed.

In this paper, a method of TV logo Recognition based on SIFT is proposed which uses the structure and color features of TV logo and combines the region separation and SIFT matching. Experiments results show that this method has a good performance in TV logo recognition, especially for transparent logo.

2 The Algorithm

In most TV programs, the backgrounds of TV logo have the following features: the background is simple and there is no complicated interference; complicated background of TV program surrounds the region of TV logo but do not cover the region; TV logo is located in the complicated background, and the hue of background differs from that of TV logo; TV logo is located in the complicated background and the hue of background is similar to that of TV logo.

2.1 A diagram of the algorithm

The flow diagram of algorithm is shown in figure 1. This algorithm can be divided into three steps:

(1)Pre-judgment of TV logo. The TV logo information is obtained by applying a special mask image to the captured video frames. Then the coarse matching is performed on the HSV color space to select the similar TV logo templates.

(2)Separation of TV logo region. According to the information given by TV logo template, the logo region is separated into the master region and sub region.

(3)SIFT matching. The keypoints obtained from the master region are used to identify which categories the logo belongs to and the keypoints from the sub

region are used to make a final judgment according to the number of matching points.

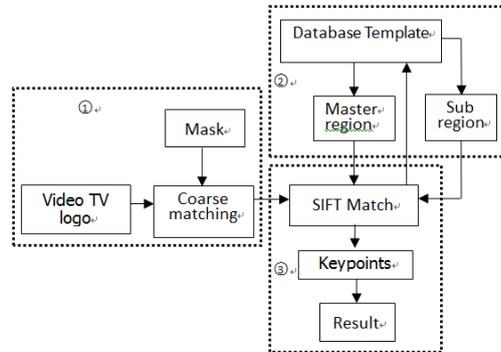


Fig. 1 The flow diagram of algorithm

2.2 Pre-judgment of TV logo

SIFT has high recognition rate, but huge computation and complex algorithm are the limitations of it. Hue histogram matching needs small amount of calculation. A coarse match based on hue histogram matching is used to remove some TV logo templates which have dissimilar tone.

Figure 2 shows an original video frame and its corresponding mask and mapping image. TV logo mask shown in figure 2(b) is used to extract information from the original TV logo. Here we call the result of extraction mapping of TV logo, as shown in figure 2(c). The RGB information in the region of TV logo is obtained by mask image and then converted to the HSV color space. The H is compressed from 0 ~ 360 to 0 ~ 36 to generate the hue histogram. Then Euclidean distance method is used to find the similarity between the mapping of TV logo and logo template. The similarity distance is smaller and the two logos are more similar. The top 15 logos are chosen as SIFT matching templates.

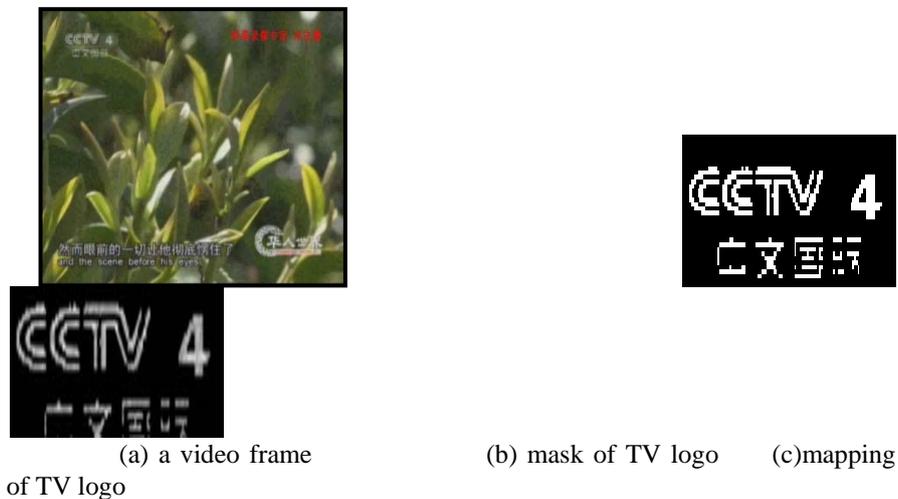


Fig. 2 mask and mapping of TV logo

2.3 Separation of TV logo region

Separation of TV logo region plays an important part in this method. The master region usually represents the TV station and sub region represents different channels belonging to the TV station. In other words, for a certain TV station, almost all channels have the same master region. Figure 3 gives an example of the separation of TV logo region. The region of TV logo in figure 3 is separated into two parts. The part of “BTV” represents Beijing TV station and is the master region. The part of “财经” represents Finance and Economics Channel and is sub region. In order to introduce less background interference and noise, the edge of the region should be as far as possible close to the logo.



Fig. 3 TV logo separation

3 SIFT matching

Scale-invariant feature transform (SIFT) feature has been widely accepted as an effective local keypoint descriptor for its invariance to rotation, scale, and lighting changes in images. There are four stages of computation used to generate the set of image features:

- 1) Constructing the scale space of an image, detecting the scale-space local extrema.

To efficiently detect stable keypoints locations in scale space, Lowe[6] has proposed the scale space, the difference-of-Gaussian (DoG) function convolves with the image:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) \times I(x, y) = L(x, y, k) \quad (1)$$

- 2) Accurate keypoints localization.

At each candidate location, a detailed model is fit to determine location and scale. Boundary influence is reduced using trilinear interpolation to distribute the value of each gradient sample into adjacent histogram bins. The figure 4 gives the accurate keypoints in TV logo.

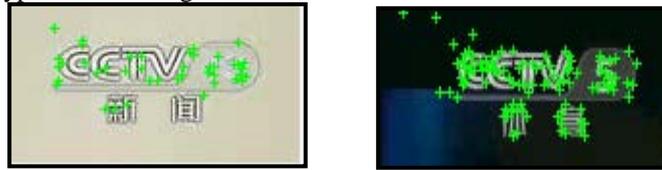


Fig. 4 Keypoints of SIFT

- 3) Assigning a consistent orientation to each feature point.

In this stage, the SIFT algorithm assigns orientations to each keypoint location base on local image gradient directions

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (2)$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (3)$$

$m(x, y)$ is the gradient magnitude, and $\theta(x, y)$ is the orientation. L is the convolution of a Gaussian mask to image in the first stage. From the gradient orientations of the sample points within a region around the feature point, the orientation histogram, which has 36 bins covering the 360 degree range of orientations can be computed. With the gradient magnitude and a Gaussian-weighted circular window each sample is weighted before inserting them into the histogram.

- 4) Creating the feature point descriptors -- SIFT feature vectors.

Here we used 16 orientation histogram aligned in a 4x4 grid. Each histogram has 8 orientation bins. This 4x4x8 elements is combined as the SIFT feature vector. After being constructed, this 128-element vector is then normalized to unit length. Finally, a normalized 128-element float type vector is created as the SIFT

keypoint descriptor.

5) Feature matching

Euclidean distance is used as similarity measure function. Generate sets of keypoint descriptor for template image and identifying image .TV logo is identified through two sets of keypoint descriptor. Steps are as follows:

step 1、Template image descriptors:

$$R_i(r_{i1}, r_{i2}, r_{i3}, \dots, r_{i128})$$

step 2、Identifying image descriptors:

$$R_i(s_{i1}, s_{i2}, s_{i3}, \dots, s_{i128})$$

step 3、A similarity measure between any two descriptors:

$$d(R_i, S_i) = \sqrt{\sum_{j=1}^{128} (r_{ij} - s_{ij})^2}$$

(4)

step 4、 $J/P < TH$, J is the nearest point in identifying image to S_j , P is the second nearest point in identifying image to R_i , TH is threshold value.



(a) BTV-1

(b) CCTV-7



(c) BTV-3 and BTV-2

(d) CCTV-1 and CCTV-2

Fig. 5 The result of SIFT matching

In figure 5(a), the total number of keypoints in TV logo template matched with identifying image is 33. In figure 5(b), the total number of feature points in CCTV7 TV logo template matched with identifying image is 12. In figure 5(c), "BTV-3" template matches with "BTV-2", keypoints only locate in the region of "BTV" block, and there is not any feature points in the region of "科教" or "文艺"; In figure 5(d), keypoints also only locate in the region of "CCTV" block.

According to the master region classification, the first step is to match the master region template, and then match the corresponding sub region. The image in which the number of keypoints is larger than threshold Q_0 is the best matching image.

4 Experimental Results

The experiments are performed with Intel(R)Core(TM)i7-2600CPU@3.40GHz and the software platform is Windows7 System, VC6.0, Opencv1.0. The video sequences are captured from the broadcasting TV programs in China. The size of image is 640*475 and the number of TV Logo models in data base is 53.

Table 1 Result of TV logos recognition

index	TV logo	frames	Correct identification	recognition rate/(%)
1	CCTV1	300	284	94.7
2	CCTV2	300	276	92
3	CCTV3	300	281	93.7
4	CCTV5	300	283	94.3
5	BTV-1	300	287	95.7
6	BTV-2	300	292	97.3
7	BTV-3	300	280	93.3
8	HNTV	300	300	100
9	PhoenixTV	300	300	100
10	SBN	300	296	98.7
11	GDTV	300	300	100
12	GXTV	300	300	100

Table 1 gives the recognition rate of 12 different TV logos. In this paper, we select CCTV and some provincial TV station logo as experiment sequences. There are total 3900 frames and the threshold of the number of keypoints for SIFT is 3. We can see from table 1 that the recognition rate can achieve approximately 95%. In addition, this method is also effective for transparent logo. In table 1, the logos of CCTV1, CCTV2, CCTV3, CCTV5, BTV-1, BTV-2 and BTV-3 are all transparent and the recognition rate for these logos are also high.

5 Conclusions

The method of TV logo recognition based on SIFT is proposed in this paper. It successfully suppresses the noises and complex background. Moreover, associated with tone matching and blocks, it can narrow down search of SIFT and reduce calculation amount. In the experiment, the method adapt to changing background of TV logo, with advantages of rapid and accuracy. In the future, we will focus on

the development of rough matching and try to look for a method to minimize the search of SIFT.

Acknowledgments This work is supported by "863" national project, No. 2012AA01172

References

1. Wan-Lei Zhao , Chong-Wah Ngo . Flip-Invariant SIFT for Copy and Object Detection. Image Processing, IEEE Transactions on. 980–991. 2013.
2. Kumar, N.A.M.;Sathidevi,P.S.Image match using wavelet — Colour SIFT features. Industrial and Information Systems (ICIIS), 2012 7th IEEE International Conference on.1–6.2012.
3. Tan Chunlin , Wang Hongqiao.PEI Deli SWF-SIFT Approach for Infrared Face Recognition[J]. Tsinghua Science and Technology. 2010.6:357 – 362.
4. Dandan Li,Yongzhen Ke,Guiling Zhang. A SIFT descriptor with local kernel color histogramsMechanic Automation and Control Engineering.2011. 992 – 995.
5. P. Nieto, et al., “A TV-logo Classification and Learning System,” in Proc. ICIP, pp.2548-2551, 2008.
6. D. Lowe. Object Recognition from Local Scale-Invariant Features. International Conference On Computer Vision, Corfu, Greece, pp.1150-1157, 1999.
7. D. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60(2):91-110, 2004.
8. Jalilvand,A. Boroujeni,H.S. Charkari,N.M.CH-SIFT: A local kernel color histogram SIFT based descriptor.Multimedia Technology. 2011.:6269-6272.
9. Shi Yingchun ,Zhou Xian,Wang tao.Weighted Hu moment invariants based on HSV color space of the sticker to identify [J]. 2005, 29 (3): 364-367.
10. Yang Qiang, Cheng Yukun ,Ma Sen.An algorithm of TV symbol recognition based on wavelet transform [J].Computer Application and Software.2009,26(3):63-65.
11. Jin Yang ,Cheng Jianhua, Ren Tong, Ku Xishu.TV Logos Recognition Method Base on Binary Image Coner Matching.Digital Video, 2012,36(17):14-19.
12. Liang Zhang,Tian Xia,Yongdong Zhang,Jintao Li.Hollow TV logo detection .IEEE conference publications. 2011,3581-3584.
13. C. Geng ,X. Jiang,“Face recognition using sift features”, Proc.ICIP 2009, pp.3313-3316.
14. Jiang qingbing, Wang Hao, Zhao Li. Research on Vehicle Tracking Algorithm Based on UKF and SIFT. Computer Measurement & Control .2012.20(10):2824-2826.
15. Apostolos Psyllos, Christos-Nikolaos Anagnostopoulos, Eleftherios Kayafas, M-SIFT: A new method for Vehicle Logo Recognition .IEEE conference publication. 2012,261- 266.