International Journal of Emerging Trend in Engineering and Basic Sciences (IJEEBS)
ISSN (Online) 2349-6967
Volume 2 , Issue 1(Jan-Feb 2015), PP128-135

Improved Approach for Detection and Recognition of Logo from Videos

Deepak Gupta[1], Prajakta Prakash Musale[2]
[1,2] *Computer Science Department, Siddhant Collage of Engineering, Pune.*

Abstract:- Logo matching and recognition is important for brand advertising and surveillance applications and it discovers either improper or non-authorized use of logos. An effective logo matching and recognition method for detect logos in a high motion videos. The central issues of this technology are fast localization and accurate matching and unveil the malicious use of logos that have small variation with respect to the original. In this paper, an implementation system for logo detection and recognition in video stream which contains one or more than one logos. A novel solution for logo matching and recognition based on Context Dependent Similarity (CDS) kernel is proposed and it's able to match and recognize multiple instances of multiple reference logos in video. Context is a collection of interest points and Context Dependent Similarity Matrix is created to find interest point correspondences between two images in order to tackle logo detection. Feature set using SIFT (Scale Invariant Feature Transform) key-points is created for each query logo and the video frame being processed. Also novel threshold scheme is used to finally make decision regarding presence of logo. Finally boundary is detected around the matched logo. The system provides various kinds of statistics like frequency of recognized logos, their visibility time and their locations in the video.

Keywords:- Context-dependent kernel, CDS, logo detection, logo recognition, SIFT

I. INTRODUCTION

Logos are commonly used in business and government documents as a declaration of document source and ownership. The problem of logo detection and recognition is of great interest in the document domain as it enables us to identify the source of documents based on the organization where a document originates. Automatic image and video annotation has received an increasing attention from the research and industrial community in the recent years. This is mainly due to the growing request for content based search and retrieval of interesting visual elements, resulting from the exponential growth of multimedia sharing systems such as Flickr and YouTube. In particular, a really challenging task is the detection and recognition of advertising trademarks/logos, which are of great interest for several real world applications. In fact, logos are key elements for companies and play essential role in industry and commerce; they also recall the expectations associated with a particular product or service. Logo analysis in videos involves 3 main steps: (i) Detecting the probable logo from Video frames extraction,(ii) Context Formation

(iii)Matching. Logos are graphic productions that either recall some real world objects, or emphasize a name, or simply display some abstract signs that have strong perceptual appeal [see Fig. 1.1(a)]. Color may have some relevance to assess the logo identity. But the distinctiveness of logos is more often given by a few details carefully studied by graphic designers, sociologists' and experts of social communication. The graphic layout is equally important to attract the attention of the customer and convey the message appropriately and permanently. Different logos may have similar layout with slightly different spatial disposition of the graphic elements, localized differences in the orientation, size and shape or in the case of malicious tampering differ by the presence/absence of one or few traits [see Fig. 1.1(b)].

Logos however often appear in images/videos of real world indoor or outdoor scenes superimposed on objects of any geometry, shirts of persons or jerseys of players, boards of shops or billboards and posters in sports playfields. In most of the cases they are subjected to perspective transformations and deformations, often corrupted by noise or lighting effects, or partially occluded. Such images – and logos thereafter – have often relatively low resolution and quality. Regions that include logos might be small and contain few information [see Fig. 1.1(c)].

Logo detection and recognition in these scenarios has become important for a number of applications. Such as the automatic identification of products on the web to improve commercial search-engines, the verification of the visibility of advertising logos in sports events, the detection of near-duplicate logos and

ISSN (Online) 2349-6967
Volume 2 , Issue 1(Jan-Feb 2015), PP128-135

unauthorized uses. Special applications of social utility have also been reported such as the recognition of groceries in stores for assisting the blind.

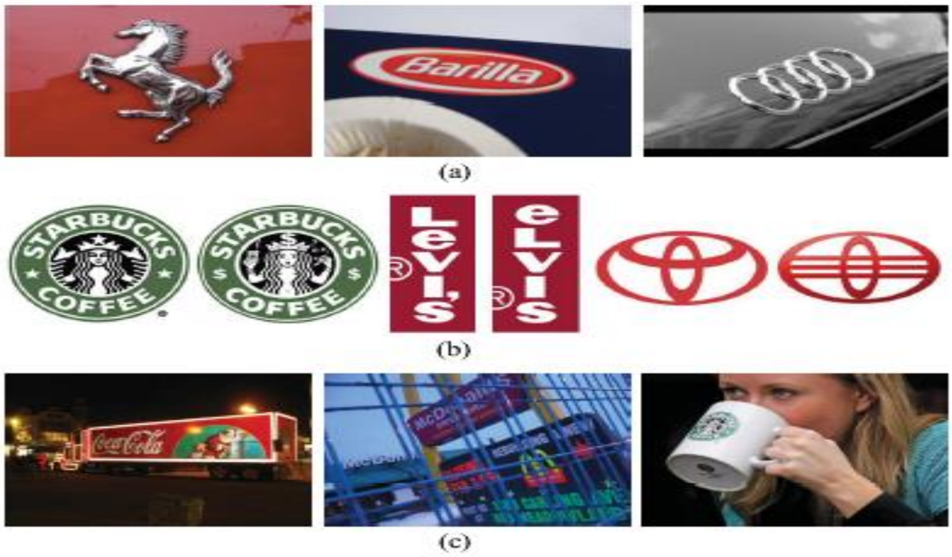


Fig. 1.1. (a) Examples of popular logos depicting real world objects, text, graphic signs, and complex layouts with graphic details. (b) Pairs of logos with malicious small changes in details or spatial arrangements. (c) Examples of logos displayed in real world images in bad light conditions, with partial occlusions and deformations.

II. PROBLEM DESCRIPTION & REQUIREMENT SPECIFICATIONS

2.1 Problem Statement

Logo is a key visual feature for readers to distinguish the origin or ownership of a document along with other features such as title and seal. In the applications of automatic document image processing, the main focus of logo detection is to find and extract logos with high speed and reliability from videos. The existing algorithm matches the one logo from video with dataset of logos.

2.2 Proposed System

The main aim of this paper is video frames will be generated and by taking that array of frame, each image will be checked simultaneously one by one also to present a highly effective and scalable framework for matching and recognizing logos from real environment. Given a query image from video and a large logo database, the goal is to recognize the logo contained in the query, if any, in many cases one video contains more than one logo. Previously efficient method presented which outperform the existing method in terms of FRR and FPR. In this project we are extending the same method for improved scalability of logo detection and recognition. The recent method of logo detection and recognition which is based on the definition of a “Context-Dependent Similarity” (CDS) kernel that directly incorporates the spatial context of local features is under investigation. Formally, the CDS function is defined as the fixed-point of three terms: (i) an energy function which balances a fidelity term; (ii) a context criterion; (iii) an entropy term. In this project we are extending this method further for scalability as well as other rigid and non-rigid logo transformations. During the simulation we will first do comparative analysis proposed CDS matching and detection procedure against nearest-neighbor SIFT matching and nearest-neighbor matching with RANSAC verification so that we can claim the proposed method is best as compared existing once. Second we will evaluate the performance of proposed by considering the scalability factor and compute its precision and recall rate.

III. PROPOSED APPROACH FRAMEWORK AND DESIGN

3.1 Problem Definition

Logo recognition in these scenarios and for a number of key presentations Beach Them, some self-acting as demonstrations publications, has been reported in financial search engine on the world wide web to, the verification of the sporting events, explore near-duplicate logo advocated improving visibility and identification of unauthorized uses of products. Special presentations of social utility in addition to the recognition of such as groceries in stores for aiding the unseeing have been described and recognition in images system to a normal logo detection took real world environment must obey with diverging needs a hand, Geometric and photometric transformations is of invariance to a large variety with all the conditions of the image video/notes need to obey. Since in genuine world images logos are not apprehended in isolation, logo detection and recognition should furthermore be robust to partial occlusions. At the identical time, especially if we desire to find out malicious tampering or get logos with some localized peculiarities, we should further more require that the small differences in the localized organizations area apprehended in the localized descriptor and are adequately differentiating for recognition. Existing methods assume that a logo picture is fully visible in the image, is not corrupted by noise and is not subjected to transformations. According to this, they cannot be applied to real world image which are may be corrupted by noise. Logo and acknowledgement which of course incorporates is the spatial context of the localized features a "depending on context similarity" (CDS) based on the definition of the latest kernel under interrogation process. Formally, the CDs function as three periods of real estate is characterized by: an energy function which is the remaining one term of allegiance, a standard reference period of entropy. Video contains many times more than one logos, we done matching each query logo with reference logo.

3.2 Proposed Architecture

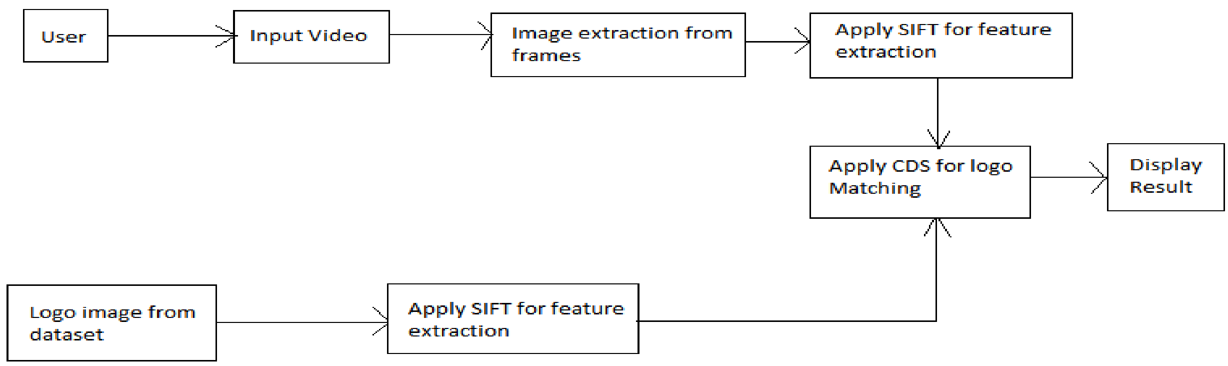


Fig 3.2.1 Architecture Design

3.3 Process Flow

Returns the user input the reference image and test image detection, we first find out that you want to process both images, all the features their key points out of both finding images, and we will remove the image using SIFT features and using the features we then explore the images we object descriptors that match or logo image to detect the CDs algorithm applied.

3.4 Context-Dependent Similarities

The use of context for matching: Context is used to find interest point correspondences between two images in order to tackle logo detection; context was used for kernel design in order to handle object classification using support vector machines.The update of the design model: Adjacency matrices are defined in order to model spatial and geometric relationships between interest points belonging to two images (a reference logo and a test image). These adjacency matrices model interactions between interest points at different orientations and locations resulting into an anisotropic context, context was isotropic.The similarity diffusion process: Resulting from the definition of context, similarity between interest points is recursively and anisotropic ally diffused.

The interpretation of the model: Our designed similarity may be interpreted as a joint distribution (pdf) which models the probability that two interest points taken from $S_x \times S_y$ match. In order to guarantee that this similarity is actually a pdf, a partition function is used as a normalization factor taken through all the interest points in $S_x \times S_y$.

3.4.1 Context

The context is defined by the local spatial configuration of interest points in both S_x and S_y. Formally, in order to take into account spatial information, an interest point S_x *and* S_y is Defined as $x_i = (\varphi_g(x_i))$, $(\varphi_f(x_i))$, $(\varphi_o(x_i))$, $(\varphi_s(x_i))$, $(\varphi(x_i))$ where the symbol $(\varphi_g(x_i)) \epsilon R^2$ stands for the 2*D* coordinates of *xi* while $(\varphi_f(x_i)) \epsilon R^c$ corresponds to the feature of *xi* (in practice *c* is equal to 128, i.e. the coefficients of the SIFT descriptor and extra information about the orientation of *xi* (denoted $\psi o(xi) \in [-\pi,+\pi]$) which is provided by the SIFT gradient and about the scale of the SIFT descriptor (denoted $\psi s(xi)$). Finally, use $\omega(xi)$ to identify the image from which the interest point comes from, so that two interest points with the same location, feature and orientation are considered different when they are not in the same image; this is motivated by the fact that we want to take into account the context of the interest point in the image it belongs to. Let $d(xi , y j) = \|\psi f(xi) - \psi f(y j)\|2$ measure the dissimilarity between two interest point features, where $\|\cdot\|2$ is the "entrywise" *L*2-norm (i.e. the sum of the square values of vector coefficients).The context of *xi* is defined as in the following:

$$N^{\theta,\rho}(x_i) \models \{x_j : \omega(x_j) = \omega(x_i), x_j \neq x_i \text{ s.t. (i), (ii)}$$

hold with

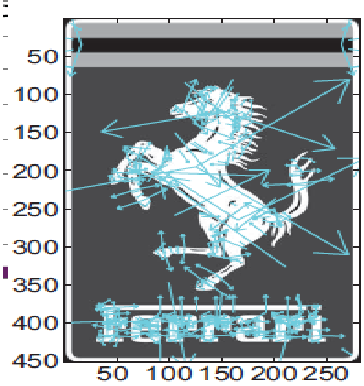
$$\frac{\rho-1}{N_r}\epsilon_p \leq \|\psi_g(x_i) - \psi_g(x_j)\|_2 \leq \frac{\rho}{N_r}\epsilon_p \qquad \text{(i)}$$

and

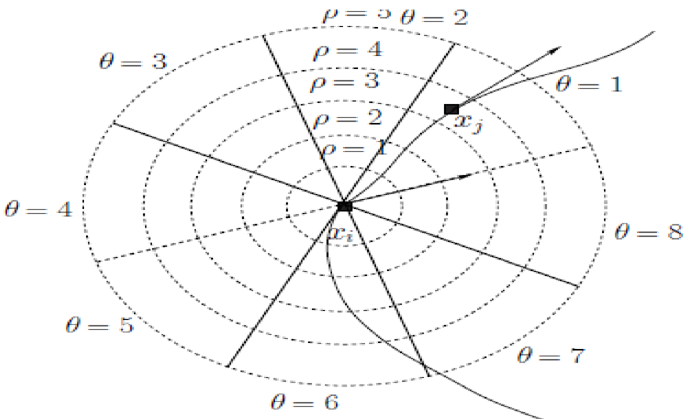
$$\frac{\theta-1}{N_a}\pi \leq \angle(\psi_o(x_i), \psi_g(x_j) - \psi_g(x_i)) \leq \frac{\theta}{N_a}\pi \qquad \text{(ii)}$$

Where, = $(\varphi_g(x_j)) - (\varphi_g(x_i))$ *is the vector between the two points coordinate* $\varphi_g(x_j)$ *and* $\varphi_g(x_i)$.

In the above definition, $\theta=1,\ldots,N_a$,$p=1,\ldots N_r$ correspond to indices of different parts of that disk.



(a)



(b)

Fig 3.4.1(a) Collection of SIFT points with their locations, orientations, and scales. (b) Definition and partitioning of the context of an interest point *xi* into different sectors (for orientations) and bands (for locations).

3.4.2 Similarity Design

Here *k* as a function which, given two interest points (x,y) ϵ S_x , S_y. Provides a similarity measure between them. For a finite collection of interest points, the sets S_x , S_y are finite. We can view function *k* as a matrix K, i.e. K*x,y* = *k(x, y)*, in which the "*(x, y)*–element" is the similarity between *x* and *y*. Also represent with $P_{\theta,p}$, $Q_{\theta,p}$ the intrinsic adjacency matrices that respectively collect the adjacency relationships between the sets of interest points *SX* and *SY* , for each context segment; these matrices are defined as $P_{\theta,p,x,x'}$ $=g_{\theta,p}$(x,x'), $Q_{\theta,p,y,y'}$ = $g_{\theta,p,}$(y.y') where g is the decreasing function.

The similarity K between the two objects *SX* , *SY* is obtained by solving the following minimization problem

$$\begin{aligned} \min_{\mathbf{K}} \quad & \mathrm{Tr}(\mathbf{K}\ \mathbf{D}') + \beta\ \mathrm{Tr}(\mathbf{K}\ \log\ \mathbf{K}') \\ & -\alpha \ \textstyle\sum_{\theta,\rho} \mathrm{Tr}(\mathbf{K}\ \mathbf{Q}_{\theta,\rho}\ \mathbf{K}'\ \mathbf{P}'_{\theta,\rho}) \qquad (1) \\ & \text{s.t.} \begin{cases} \mathbf{K} \geq 0 \\ \|\mathbf{K}\|_1 = 1. \end{cases} \end{aligned}$$

Here α and $\beta \geq 0$ and operations log is $\geq$ applied individually to every entry in the matrix for instance K.

3.4.3 Solution

Let's consider the adjacency matrices $\{P_{\theta,p}\}$ θ, p ,$\{Q_{\theta,p}\}$ θ, p related to a reference logo S_x and a test image S_y respectively, each of which collects the adjacency relationships between the image interest points for a specific context segment θ, ρ.

Proposition 1: Let u denote the matrix of ones and introduce

$$\zeta = \frac{\alpha}{\beta} \sum_{\theta,\rho} \|\mathbf{P}_{\theta,\rho}\mathbf{u}\mathbf{Q}'_{\theta,\rho} + \mathbf{P}'_{\theta,\rho}\mathbf{u}\mathbf{Q}_{\theta,\rho}\|_{\infty}$$

Where $\|\ .\|_\infty$ is the "entry wise "L_∞-norm. Provided that the following two inequalities hold

$$\zeta \exp(\zeta) < 1 \qquad (2)$$

$$\|\exp(-\mathbf{D}/\beta)\|_1 \geq 2 \qquad (3)$$

the optimization problem (1) admits a unique solution ˜K, which is the limit of the *recursive form*

$$\mathbf{K}^{(t)} = \frac{G(\mathbf{K}^{(t-1)})}{\|G(\mathbf{K}^{(t-1)})\|_1} \qquad (4)$$

with

$$G(\mathbf{K}) = \exp\left\{ -\frac{\mathbf{D}}{\beta} + \frac{\alpha}{\beta} \sum_{\theta,\rho} (\mathbf{P}_{\theta,\rho}\mathbf{K}\mathbf{Q}'_{\theta,\rho} + \mathbf{P}'_{\theta,\rho}\mathbf{K}\mathbf{Q}_{\theta,\rho}) \right\} \qquad (5)$$

And

$$\mathbf{K}^{(0)} = \frac{\exp(-\mathbf{D}/\beta)}{\|\exp(-\mathbf{D}/\beta)\|_1}.$$

Beside $K^{(t)}$ satisfy the convergence property

$$\|\mathbf{K}^{(t)} - \tilde{\mathbf{K}}\|_1 \leq L^t \|\mathbf{K}^{(0)} - \tilde{\mathbf{K}}\|_1 \qquad (6)$$

With $L = \zeta \exp(\zeta)$.

3.5 Algorithm

CDS Logo Detection and Recognition
Input : {reference logo image : lx Test image:ly , CDS
parameter:Є,α,τ,β,Na,Nr }

Processes:
Extract SIFT from lx, lyand let Sx= {x1, xn},
Sy= {y1, yn} be respectively the list of interest points
taken from both images;

Step 1:
For i=1 to n
find context matching for xi where it is key point of
referral image.
End for

For i=1 to n
find context matching for yi where it is key point of
test image.
End For

Step 2: Set t=1 to max=30
Step 3:
For i=1 to n
For j=1 to m
Compute CDS matrix
Increment t i.e. does t++;
End for
End for
Repeat step 3 until t > max or convergence.

Step 4:
For i=1 to n do
For j=1 to m do
ComputeKyj|xi
Match between xi andxjis declared only if
Kyj|xi ≥ $\sum_{s=j}^{m} K_{ys\,|\,xi}$

Step 5:

If number of matches in $Sy>\tau|Sx|$
Then logo matched i.e. detected
Otherwise
Logo not detected.
Output: A Boolean value determining whether the reference logo in lx is detected inly.

3.6 Logo Detection And Recognition

Application of CDS to logo detection and recognition requires establishing a matching criterion and verifying its probability of success. Detecting and Recognition free-form graphical patterns such as logos are challenging. Large variations in logo style and low quality videos can make detection difficult. Complicating matters the foreground content of frames generally includes a mixture of machine printed text, diagrams, tables and other elements. From the application perspective, accurate localization is needed for logo recognition. Logo detector must consistently detect and extract complete logos while attempting to minimize the false alarm rate. Treat a logo as a non-rigid shape, and represent it by a discrete set of 2-D feature points extracted from the object. 2-D point features offer several advantages compared to other compact geometrical entities used in shape representation, because it relaxes the strong assumption that the topology and the temporal order of features are well preserved under image transformations and degradations. For instance the same portion of contours in one logo sample may overlap, while appearing separated in other cases. Represented by a 2-D point distribution, a shape is more robust under image degradations and noise while carrying discriminative shape information.

Let $R \subset$ R2 × R128 × $[-\pi,+\pi]$ × R+ denote the set of interest points extracted from all the possible reference logo images (see Section II-A) and X a random variable standing for interest points in R. Similarly, we define $T \subset$ R2 × R128× $[-\pi,+\pi]$ × R+ as the set of interest points extracted from all the possible test images (either including logos or not) and Y a random variable standing for interest points in T . X and Y are assumed drawn from existing (but unknown) probability distributions. Let's consider $SX = \{X1, . . . , Xn\}$, $SY = \{Y1, . . . , Ym\}$ as n and m realizations with the same distribution as X and Y respectively. To avoid false matches we have assumed that matching between YJ and X is assessed iff

$$K_{Y j l X} = \geq \sum_{j \neq J}^{m} K_{Y j l X} \qquad (7)$$

The intuition behind the strong criterion above comes from the fact that when KYJ $|X$ $\gg \sum_{j \neq J}^{m} K_{Y j l X}$, the entropy of the conditional probability distribution K.$|X$ will be close to 0, so the uncertainty about the possible matches of X will be reduced. The reference logo SX is declared as present into the test image if, after that the match in SY has been found for each interest point of SX , the number of matches is sufficiently large (at least τ $|SX|$ for a fixed $\tau \in [0, 1]$, being $1 - \tau$ the occlusion factor tolerated). We summarize the full procedure for logo detection and recognition in Algorithm.

IV. CONCLUSION

We have a new logo detection and localization approach referred to as reference likeliness reliant of a new classroom was founded on the strength of the suggested process in many aspects. (i) equality in addition to information about the spatial configuration that design as well as Visual characteristics, (ii) the ability to control the influence of the context and the regularization of the solution via our energy function, (iii) the tolerance to different aspects including partial occlusion, makes it suitable to detect both near-duplicate logos as well as logos with some variability in their appearance, and (iv) the theoretical roundedness of the matching framework which shows existence of a reference logo into a test image from video, the probability of success of matching and detection is high.

REFERENCE

[1] HichemSahbi, LambertoBallan, Member, IEEE,Giuseppe Serra, and Alberto Del Bimbo, Member, IEEE, “Context-Dependent Logo Matching and Recognition”, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 22, NO. 3 MARCH 2013.

[2] A. Bagdanov, L. Ballan, M. Bertini, and A. Del Bimbo.Trademark matching and retrieval in sports video databases. In Proceedings of the International Workshop on Multimedia Information Retrieval, ACM, 2007.

[3] L. Ballan, M. Bertini, A. Del Bimbo, L. Seidenari, and G. Serra, “Event detection and recognition for semantic annotation of video,” Multimedia Tools Appl., vol. 51, no. 1, pp. 279–302, 2011.

[4] A. Joly and O. Buisson. Logo retrieval with a contrario visual query expansion. In Proceedings of the seventeen ACM international conference on Multimedia, ACM, 2009.

[5] D. Lowe. Distinctive image features from scaleinvariant keypoints. International journal of computer vision, 60(2):91{110, 2004.

[6] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Scha_alitzky, T. Kadir, and L.V. Gool. A Comparison of Affine Region Detectors. International Journal of Computer Vision, 65(1-2), 2005

[7] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, “Contentbased image retrieval at the end of the early years,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 12, pp. 1349–1380, Dec. 2000.

[8] R. Datta, D. Joshi, J. Li, and J. Z. Wang, “Image retrieval: Ideas, influences, and trends of the new age,” ACM Comput. Surv., vol. 40,no. 2, pp. 1–60, 2008.