International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622
International Conference on Humming Bird (01st March 2014)

RESEARCH ARTICLE OPEN ACCESS

Logo Matching and Recognition with Interest Points Using Context–Dependent Similarity

Nishanth.T.R ,Jemimah Simon
P.G. Scholar,Dept of Computer Science and Enginering,
Vins Christian College of Engineering,Nagercoil. Email:-t.r.nishant@gmail.com
Asst.Proffessor,Dept of Computer Science and Engineering,VCCE.

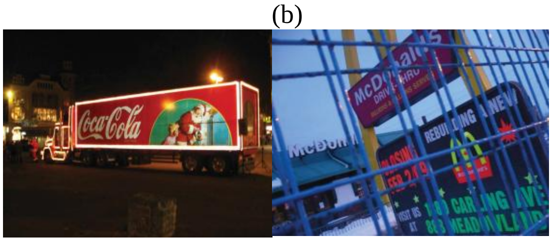
Abstract:
Logo matching and recognition isimportant for brand advertising and surveillance applications and it discovers either improper or non-authorized use of logos.An effective logo matching and recognition method for detect logos in a high-motion sports videos.The central issues of this technology are fast localization and accurate matching and unveil the malicious use of logos that have small variation with respect to the original. 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 image archives. Query logo and target logo images consist of spatial context of local features like interest points, regions.The similarity matching is based on three terms:1) Minimization terms refers to the distance between two points. 2) Context criterion refers to coherence of alignments in the target logo and query logo images. High alignment scores are considered. 3) Regularization criterion refers to the probability distribution of interest points. Near duplicate logos as well as logos with some variability in their appearance can be recognized. This method has high probability of success of matching and recognition.

I. INTRODUCTION

The expanding and massive production of visual data from companies, institutions and individuals, and the increasing popularity of social systems like Flickr, YouTube and Facebook for diffusion and sharing of images and video,have more and more urged research in effective solutions for object detection and recognition to support automatic
annotation of images and video and content-based retrieval of visual data[1],[3]. Graphic logos are a special class of visual objects extremely important to assess the identity of something or someone. In industry and commerce, they have the essential role to recall in the customer the expectations associated with a particular product or service. This economical relevance has motivated the active involvement of companies in soliciting smart image analysis solutions to scan logo archives to find evidence of similar already existing logos, discover either improper or non-authorized use of their logo, unveil the malicious use of logos that have small variations with respect to the originals so to deceive customers, analyze videos to get statistics about how long time their logo has been displayed.Logos are graphic productions that either recall somereal world objects, or emphasize a name, or simply display some abstract signs that have strong perceptual appeal. Color may have some relevance to assess thelogo identity. But the distinctiveness of logos is more often given by a few details carefully studied by graphic designers,semiologists and experts of social communication

(a)

(b)



(c)

Fig 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.

Different logos may have similar layout with slightly different spatial disposition of the graphic elements, localized differences in the orientation, size and shapeor in the case of malicious tampering differ by the presence or absence of one or few traits.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 containfew information. Logo detection and recognition in these scenarios has become important for a number of applications. Among them, several examples have been reported in the literature, such as the automatic identification of products on the web to improve commercial search-engines [4], the verification of the visibility of advertising logos in sports events [5],[7], the detection of near-duplicate logos and unauthorized uses[5],[7]. Special applications of social utility have also been reported such as the recognition of groceries in stores for assisting the blind[10].

A generic system for logo detection and recognition in images taken in real world environments must comply with contrasting requirements. On the one hand, invariance to a large range of geometric and photometric transformations is required to comply with all the possible conditions of image or video recording. Since in real world images logos are not captured in isolation, logo detection and recognition should also be robust to partial occlusions. At the same time, especially if we want to discover malicious tampering or retrieve logos with some local peculiarities, we must also require that the small differences in the local structures are captured in the local descriptor and are sufficiently distinguishing for recognition.

II. RELATED WORKS

Kato's system[8],was among the earliest ones. It mapped a normalized logo image to a 64 pixel grid, and calculated a global feature vector from the frequency distributions of edge pixels.Zernike- logos were described by global Zernike moments, local curvature and distance to centroid .Sivic and Zisserman-bag[8], of visual words approach to represent affine covariant local regions from a codebook of SIFT descriptors; visual words were weighted with tf-idffor large-scale retrieval.Joly,Buisson and Costantinopoulos[7].To discard the outliers they performed geometric consistency checking, assuming the presence of affine geometric transformation between query and target images. Particularly, in the authors applied the standard RANSAC algorithm[5], to refine the initial set of feature matches.Fergus.The joint distribution of the geometry of object parts was considered.Carneiro and Jepson - suggested to group local image features in flexible spatial models to improve matching accuracy between image.Chum-accounted for spatial proximity between visual words by performing spatial geometric hashing.Chum and Matasconsidered a special case where feature appearance is ignored and only spatial relations between pairs of features are used.Pantofaru introduced a method for object detection and localization which combines regions generated by image segmentation with localpatches.Mortensen combined SIFT descriptors with the shape descriptor of the point neighborhood very similar to shape context.Bronstein-defined spatially sensitive bags of pairs of features the distribution of near pairs of features. In this approach, the representation is only affine covariant and has a very high dimensionality.Gao proposed a two-stage algorithm that accounts for local contexts of keypoints. They considered spatial-spectral saliency to avoid the impact of cluttered background and speed up the logo detection and localization.Kleban-employed a more complex approach that considers association rules between frequent spatial configuration of quantized SIFT features at multiple resolutions

III.CONTEXT-DEPENDENT SIMILARITY

Let$S_X = \{x_1 \dots x_n\}$,$S_Y = \{x_1 \dots x_n\}$ be respectively the list of interest points taken from a reference logo and a test image.To borrow the definition of context and similarity design from in order to introduce a new matching procedure applied to logo detection. The main differences with respect to reside in the following.

1) *The use of context for matching:* Context is used to find interest point correspondences between two images in order to tackle logo detection while in context was used for kernel design in order to handle object classification using support vector machines.
2) *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, while in context wasisotropic.
3) *The similarity diffusion process:* Resulting from the definition of context, similarity between interest points is recursively and anisotropically diffused.
4) *The interpretation of the model:* Our designed similarity may be interpreted as a joint distribution whichmodels 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 $S_X \times S_Y$.

A.Context

The context is defined by the local spatial configuration of interest points in both S_X and S_Y.

$$N^{\theta,\rho}(x_i) = \{x_j : \omega(x_j) = \omega(x_i), x_j \neq x_i \text{ s.t. } (1), (2) \text{ hold}\}$$

with

$$\frac{\rho - 1}{N_r} \epsilon_p \leq \| \psi_g(x_i) - \psi_g(x_j) \| 2 \leq \frac{\rho}{N_r} \epsilon_p \quad (1)$$

and

$$\frac{Q-1}{N_a}\pi \leq \angle\left(\psi_0(x_i), \psi_g(x_j) - \psi_g(x_i)\right) \leq \frac{\theta}{N_a}\pi \quad (2)$$

where $\psi_g(x_j) - \psi_g(x_i)$ is the vector between the two point coordinates $\psi_g(x_j)$ and $\psi_g(x_i)$.

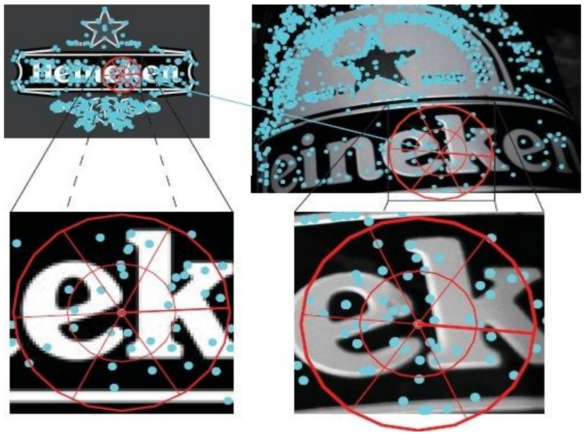


Fig 2.Example of real context definition. The two columns show thepartitioning of the context of two corresponding interest points, which belong to two instances of "Heineken."

B. Similarity Design

Define k as a function which, given two interest points $(x, y) \in S_X \times S_Y$ provides a similarity measure between them. For a finite collection of interest points, the sets S_X, S_Y are finite.Function k as a matrix **K** that is $\boldsymbol{K}_{x,y} = k(x, y)$ in which the "$(x.y)$element" is the similarity between *x* and *y*.

$$\min_K T_r(KD') + \beta T_r(K \log K') - \alpha \sum_{\theta,\rho} T_r\left(K Q_{\theta,\rho} K' P'_{\theta,\rho}\right)$$

$$s.t. \begin{cases} K \geq 0 \\ \| K \|_1 = 1 \end{cases} \quad (3)$$

This criterionalso makes it possible to consider the spatial configuration of the neighborhood of each interest point in the matching process. This minimization problem is formulated by adding an equality constraint and bounds which ensure a normalization of the similarity values and allow to seeK as a probability distribution.

IV.LOGO DETECTION AND RECOGNITION

Application of CDS to logo detection and recognition requires to establish a matching criterion and verify its probabilityof success.Detecting and Recognition free-form graphical patternssuch as logos is challenging. Large variations in logo style and low quality images can make detection difficult. Complicating matters the foreground content of documents generally includes a mixture of machine printedtext, 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 itby 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.

A. Matching and Retrieval

Given a query logo instance and a database of detectedlogos, our goal of logo matching is to compute an effective ranked list for logos in the database. By constructing the list of best matching logos, we effectively retrieve the set of documents from the same organizational entities.

1) *Corner regions:* Each frame is subdivided in 3:5:3 proportions horizontally and vertically into nine regions and the four corner regions are selected.
2) *Edge Detection:* We opted for Canny edge detectionmethod especially due to its twotierthresholding. Canny's method uses two thresholds to detect strong and weak edges, and includes the weak edges in the output only if they are connected to strong edges.
3) *Logo Persistence:* The presence of a logo is corroborated if the edge persist from frame to frame. To this effect, a given percentage of the edge pixels comprised in the mask region at time t-1 should survive at time t.
4) *Thresholding:*The time-averaged edge field is binarizedvia hysteresis thresholding method. First strong edges are obtained with a high threshold value, then weak edges are included provided they are connected to strong edges.
5) *Morphological Operations:* We apply closing to merge neighboring pixel groups, hole filling to prevent deformation of logo mask after opening, and finally opening to remove noise in the background.
6) *Shape Constraints:* TV logos possess typical shape characteristics the basic ones being the limited ranges of their area and aspect ratio. These constraints are used to eliminate improbable shapes. Furthermore logos should be sufficiently distanced from frame boundaries.
7) *Logo Mask Stability:* The final check consists in the stability of the logo which means that the candidate mask should not change beyond a tolerance in area, in its coordinates and in the size of the bounding box throughout the logo search sequence.

B.System Design

The system modules are,

- Preprocessing
- Feature extraction
- Interest point recognition
- Logo matching

1)Pre-processing

Consists of processes aimed at the geometric and radiometric correction, enhancement or standardization of imagery to improve our ability to interpret qualitatively and quantitatively image components.

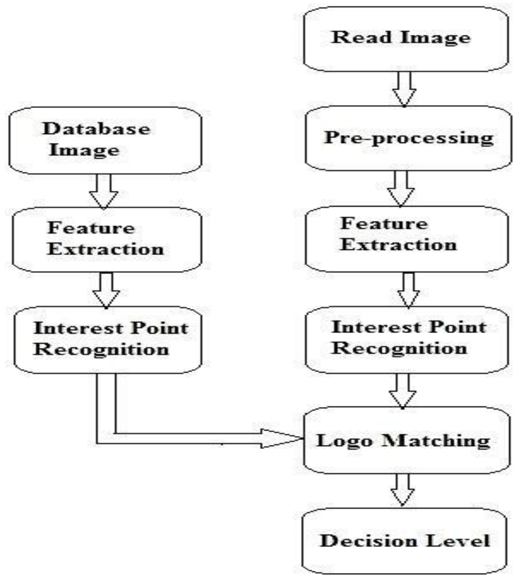


Fig3. Architectural Design

- *Radiometric Enhancement:*The main purpose for applying radiometric corrections is to reduce the influence of errors or inconsistencies in image brightness values.
- *Spatial Enhancement:*Used to improve the visual quality analytical properties and extract biophysical/landscape parameters.
- *ContrastEnhancement:*Contrastenhancemen t used to brighten the image that appear dark or hazy. Used to deliver an image with optimal quality and clarity.

2)Feature extraction

- *Color:*Calculate percentage of color present in image.

- *Text:*Find an unique underlying characteristics of textures.
- *Edge:*Edges correspond to large discontinuities in the image.

3)Interest point recognition

Intersection point between two or more edge segments.The context and orientation of the interest points are considered.Context refers to the 2D spatial coordinates and Orientation refers to the angle of theinterest points.Interest point recognition is based on edges and curvature of the logo images.

3)Logo matching

Detect the same feature points independently in both logo images.Reliable matching of a corresponding point.Localization is used to find where exactly is a point.

VI.EXPERIMENTS AND RESULTS

The experiments consisted of using the implemented algorithmfor a hockey game video. The source video was downloaded from YouTube at 720p High Definition resolution at 30 frames per second. The VLFEAT MATLAB-based implementation of SIFT was employed and the whole system was implemented and tested under MATLAB.

Initially, the video was labeled with five categories representing the visibility of the logos, as shown in Figure 3. Categories 1 to 3 are the most important ones, because a human viewer can clearly identify a logo pertaining to each of those. Then, we consider that a miss occurs if one frame in this category is not found. Categories 4 and 5 represent instances of the logo that can hardly be recognized by a human viewer. Though they are not taken into account for the recall calculation, to do not consider that a false positive occurs if the method declares that a match is found for it. Figure 4 illustrates the detection of the logo in frame 929, with blue lines connecting the pairs of matching features.

The Basic Algorithm performs poorly, due to many of the reasons that were already mentioned in this report: blurring, frequent perspective changes and occlusion. The logos are not the focus of attention of the sports event and is always on the background. The performance for the first pass achieves less than 30% of recall for the HSBC logo and less than 10% for the SONY logo. In both cases, no false positives were detected.

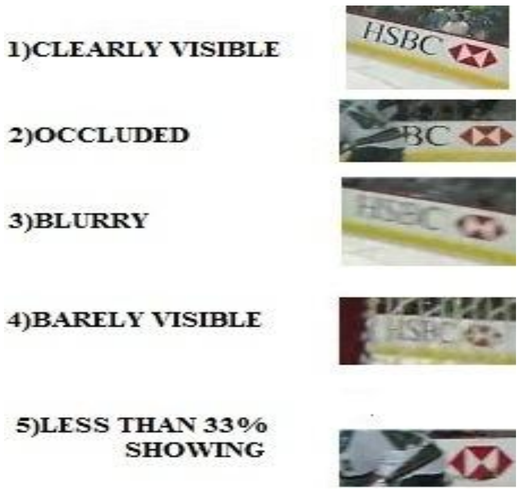


Fig. 4. Categories of the logos present in the video

Along with the bar graphs for the two versions of the RANSACAlgorithm, two vertical bars are placed on the top of the columns that represent the aggregate results for the algorithm. These two bars indicate the percentage of misses from the first pass that cannot be achieved by the strategy employed in the RANSAC Algorithm. This situation occurs when a sequence of frames in which a logo is present does not have a match for any of its frames. In this case, the RANSAC Algorithm is not able to propagate frames and improve the performance. Using these two bars, it is easy to visually understandthe maximum recall that can be achieved by the second pass. It is then clear that the second pass performs well and gets close to the maximum that it could possibly achieve.It is clear that the results for the implemented system are not very good: the best-case recall is below 50%. However, this is mostly due to the poor performance of the first pass. The RANSACgives a performance enhancement in the order of 20%, capitalizing efficiently on the temporal redundancy inherent to a video source.

Frame Category	1		2		3		TOTAL	
HSBC	Count	Precentage	Count	Precentage	Count	Precentage	Count	Precentage
Growth truth	314	61.57%	146	28.63%	50	9.80%	510	100%
HSBC	Count	Recall	Count	Recall	Count	Recall	Count	Recall
Basic Algorithm	112	35.67%	16	10.96%	0	0%	128	25.10%
RANSAC Algorithm	175	55.73%	42	28.77%	1	2.00%	218	42.75%

Table 1. Recall results for the algorithms.

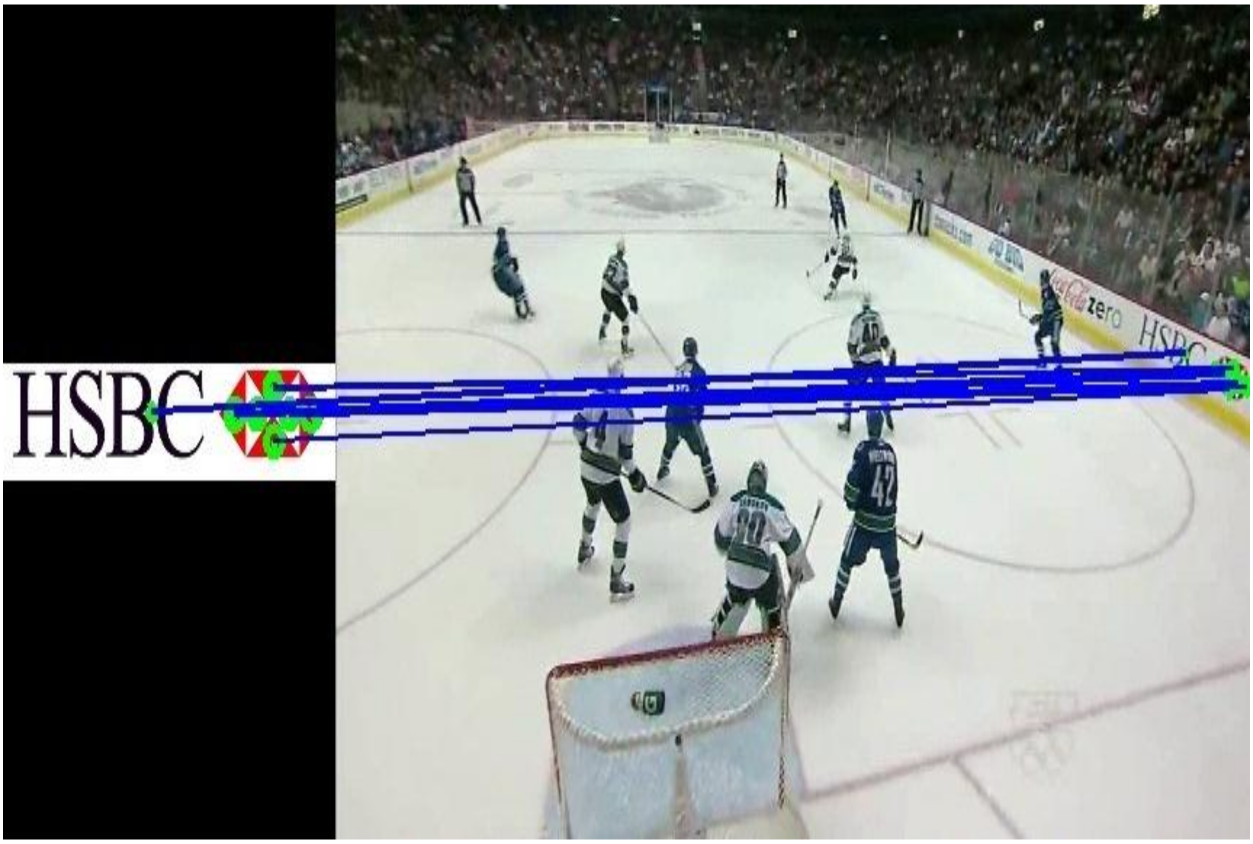


Fig. 5. Example of a correct match: The blue lines connect matched features (green) from the images

VII. CONCLUSION AND FUTURE WORK

A novel logo detection and localization approach is introduced on a new class of similarities referred to as context dependent. The strength of the proposed method resides in several aspects: (i)The inclusion of the information about the spatial configuration in similarity design as Toll as visual features, (ii) The ability to control the influence of the context and the regularization of the solution via our energy function.

Logo matching is important nowadays to detect non-authorized use of logos. Logo detection used to be done in high quality images only. But using the proposed method, logos with partial occlusion can also be detected and the accuracy of logo recognition is also high.

REFERENCES

[1] S. Belongie, J. Malik, and J. Puzicha, “Shape matching and object recognition using shape contexts,” IEEE Trans. Pattern Anal.Mach.Intell., vol. 24, no. 4, pp. 509–522, Apr. 2002.

[2] L. Ballan, M. Bertini, and A. Jain, “A system for automatic detection and recognition of advertising trademarks in sports videos,” in Proc.ACMMultimedia, Vancouver, BC, Canada, 2008, pp. 991–992.

[3] J. P. Eakins, J. M. Boardman, and M. E. Graham, “Similarity retrieval oftrademark images,” IEEE Multimedia, vol. 5, no. 2, pp. 53–63, Apr.–Jun.1998.

[4] J. Luo and D. Crandall, “Color object detection using spatial-color jointprobability functions,” IEEE Trans. Image Process., vol. 15, no. 6, pp.1443–1453, Jun. 2006.

[5] M. Merler, C. Galleguillos, and S. Belongie, “Recognizing groceries in situ using in vitro training data,” in Proc. IEEE Comput. Vis. PatternRecognit. SLAM Workshop, Minneapolis, MN, May 2007, pp. 1–8.

[6] J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide-baseline stereo from maximally stable extremal regions,” Image Vis. Comput.,vol. 22, no. 10, pp. 761–767, 2004.

[7] J. Schietse, J. P. Eakins, and R. C. Veltkamp, “Practice and challenges intrademark image retrieval,” in Proc. ACM Int. Conf. Image Video Retr.,Amsterdam, The Netherlands, 2007, pp. 518–524.

[8] J. Sivic and A. Zisserman, “Efficient visual search of videos cast as textretrieval,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 4, pp.591–606, Apr. 2009.

[9] H. Sahbi, J.-Y.Audibert, and R. Kerivan, “Context-dependent kernels for object classification,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 33,no. 4, pp. 699–708, Apr. 2011.

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