RESEARCH ARTICLE OPEN ACCESS

Enhancing Logo Matching and Recognition Using Local Features

S.Karthikeyan *, C.Kalaiyarasi**, P.Saranya***, V.Aishwarya Lakshmi****.
*(Assistant professor, Department of Computer Science and Engineering, Arunai Engineering College, Tiruvannamalai, Anna University-chennai)
**(PG Student, Department of Computer Science and Engineering, Arunai Engineering College, Tiruvannamalai, Anna University-chennai)
***(PG Student, Department of Computer Science and Engineering, Arunai Engineering College, Tiruvannamalai, Anna University-chennai)
****(PG Student, Department of Computer Science and Engineering, Arunai Engineering College, Tiruvannamalai, Anna University-chennai)

ABSTRACT
The design of visual feature extraction with scale invariant feature transform (SIFT) is widely used for recognition of object in logos. However, the real-time implementation suffers from heavy computation, and high memory storage, long latency because of its frame level computation, so we propose PCSIFT(principal component analysis SIFT) to over come all the drawback by recognition and matching multiple feature with multiple reference logos in real world application and in image archives by the help of designing a novel contribution framework. In Reference logos the test images are done on local features like, interest point, region etc. some terms of measuring feature like 1.feature matching quality is measured by the fidelity term, 2. feature co-occurrence/geometry is captured by neighbourhood criterian, 3.smoothness matching solution is done by the regularization term. We can use various logo set and real world logos to match and recognize the logos and the experimental result shows greater validity and accuracy.
***Keywords* -** Context-dependent kernel, logo detection, logo recognition, SIFT, PCSIFT.

I. INTRODUCTION

Logos are often appear in images/videos of real world by an indoor or outdoor scenes superimposed on objects of any geometry, ie by shirts of persons,boards of shops or posters in sports playfields, jerseys of players and billboards. In many cases perspective transformations and deformations are subjected and are some are corrupted by the help of noise present in the object or lighting effects, or some get partially occluded. Such images and logos are often having relatively low quality and low resolution. Regions with the logos may be very small and also contain only few information. Now-a-days logo detection and recognition scenarios has become very much important for a number of applications[1]. In logo detection and recognition the images taken in the real world environments with the help of generic system that must comply with contrasting requirements.

Logos are graphic productions that either recall some real world objects, emphasize a name, or simply display some abstract signs that have strong perceptual appeal by the popular logos in Fig1.1.(a).The pair of logos with malicious small changes and the different logos may have similar layout with slightly different spatial disposition of the graphic elements, orientation by localized differences of images in their size and shape Fig.1.1(b).by the help of perspective transformations and deformations are often corrupted by noise or lighting effects,or partially occluded[1]. Such images and logos thereafter have often relatively low resolution and quality. Regions that include logos might be small and contain few information inFig.1.1(c)]. Logo detection and recognition in these scenarios has become more important for a number of real world applications. . But the distinctiveness of logos is more often given by a few details carefully studied by graphic designers, emiologists and experts of social communication. Special applications of social utility have also been reported such as the recognition of groceries in stores for assisting the blind. A generic system for the logo detection and recognition in the images are taken from real world environments that must comply with contrasting requirements. Different logos have similar layout with a slightly different of spatial disposition and graphic elements, with localized differences in their orientation, size and their shape,of malicious tampering , differs by the presence and absence of one or few trait.

(a) (b) (c)

Fig:1.1(a) Examples of popular logos, (b) Pairs of logos with malicious small changes, (c) Examples of logos in bad light conditions.

On the one hand, the geometric and photometric transformations are taken in a large range and required to comply with all the possible conditions of image/video recording[2]. Since in real world images logos are not captured by the means of isolation, logo detection and recognition also be robust to partial occlusions[3]. At the same time, especially if we want to retrieve logos with some local features or discover malicious tampering we must obtain the small differences in their local structures and then are captured in the local descriptor and are sufficiently distinguishing for recognition. Then various techniques has been performed some of them are (1) scale-space extrema detection,(2)accuratekeypointlocalization,(3)orientati on assignment, and (4) the local image key point descriptor.

II. CONTEXT-DEPENDENT SIMILARITY

The list of interest points for matching reference logo that are taken to test image. We get the definition of context dependent 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.

2.1 CONTEXT FOR MATCHING: Context is mainly used to find interest point similarity between two images in order to tackle logo detection while in, context was also used for kernel design,by help of support vector machine it classifies and to handle object[1].

2.2 DESIGN MODEL UPDATION: spatial and geometric relationships (context) are done by adjacency that finds the difference between interest points that belonging to two images (a reference logo and a test image)[1]. The interactions between interest points at different orientations and locations resulting into an anisotropic context,by the adjacency matrices model while in, context was isotropic.

2.3 THE DIFFUSION PROCESS SIMILARITY: Resulting from the definition of context, similarity between interest points is anisotropically and recursively diffused.

2.4 MODEL OF INTERPRETATION: The designed similarity may be interpreted as a joint distribution. 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 (and not over all the objects in a training database as in).

Fig2.1: Results when using a context-free strategy

While matching with context-free strategy it matches all the position available in the objects. There is no particular position or feature points for that strategy. The fig2.1 represents context free strategy it matches all the positions and points in
the object by these the logos are not been detected properly,and it took long time but the result is not accurate.
Where as in context dependent similarity it matches only the particular points which has similar funcitonalities ie.feature points the fig2.2 represent context dependent similarity it matches only the particular points in the given objects,it matches large number of features and the result is accurate and it works faster compared to context-free strategy.

Fig2.2: Results when using a context-dependent strategy

III. OVERVIEW OF PCA-SIFT

PCA-SIFT which consists of scale space extrema detection, accurate keypoint localization, orientation assignment and the local image descriptor.

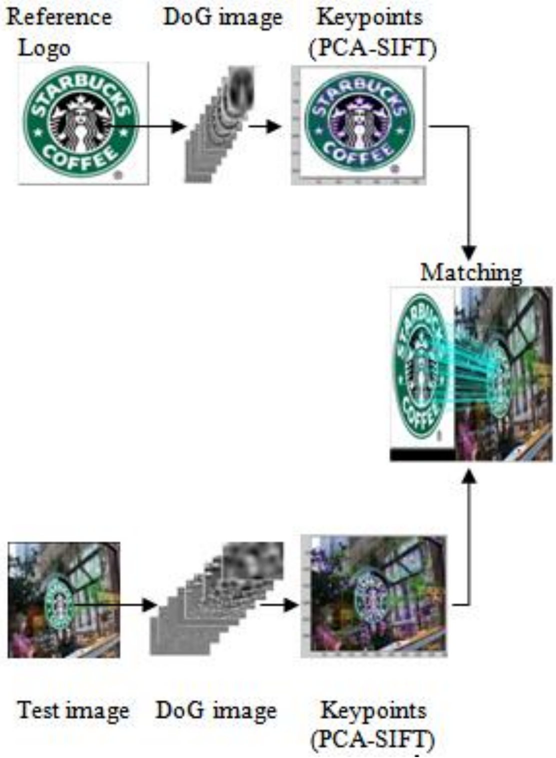

Fig3.1: Working of PCA-SIFT

IV. PRINCIPLE COMPONENT ANALYSIS

The algorithm for local descriptors, PCA-SIFT accepts the same input as the standard SIFT descriptor, the sub-pixel location, scale and dominant orientations of the keypoint.

A 41×41 patch is extracted at the given scale, centered over the keypoint and rotated to align its dominant orientation to a canonical direction. PCA-SIFT work in the following steps.

(1) An eigen space is pre-computed to express the gradient images of local patches.(2) Given a patch, its local image gradient is computed.(3) The gradient image vector is projected using the eigen space to derive a compact feature vector. The feature vector is significantly smaller than the standard SIFT feature vector.

4.1 SCALE SPACE EXTREMA DETECTION

The original image is taken and the blurred out images are generated progressively. The original image is resized to its half size. The blurred out images are generated and again it is repeated. Blurring is often referred to as the convolution of the gaussian operator and their image.

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \quad (1)$$

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

L is an blurred image, G is an gaussian blur operator, I is the image, x and y are their location coordinates, σ is the scale parameter (greater the value and greater the blur). The * is an convolution operation in x and y. It applies gaussian blur G onto the image I. The scale spaces are used to generate DoG images. Two consecutive images in an octave are picked and one is subtracted from the other.

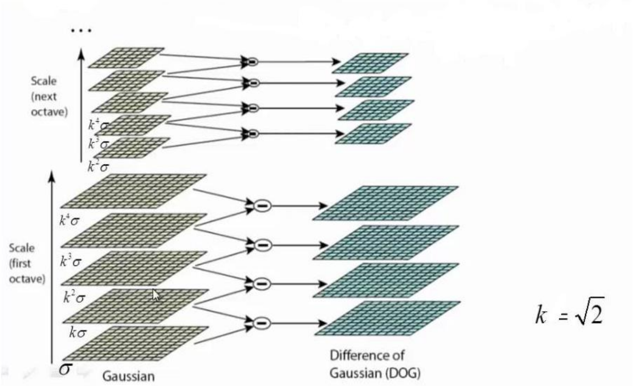

Fig4.1: Scale space extrema detection

Above formula used to get an difference of Gaussian image.The convolved images are grouped and represented as octave (an octave means corresponds to doubling the value of σ), and the value of k is represented as an selected fixed number values of convolved images per octave.

The frame at the highest scale per octave is downsampled by four as the input frame for next octave, as depicted in Fig.4.later the Gaussian image is converted into difference of Gaussian by combining two Gaussian frame an single difference of Gaussian is obtained.

4.2 KEY POINT LOCALIZATION

The keypoint candidates are detected by comparing each point to its 8 neighbors on the same scale and each of its 9 neighbors one scale up and down. Every point that is bigger or smaller than each of its neighbors is a keypoint candidate. The candidates that are located on an edge or have poor contrast are eliminated. The hessian matrix is used to eliminate edge responses.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (3)$$

The ratio of the principle curvatures can be found by calculating the trace $Tr(H)$ and determinant $Det(H)$ of the hessian matrix.

$$Tr(H) = D_{xx} + D_{yy} \quad (4)$$

$$Det(H) = D_{xx}D_{yy} - D_{xy}^2 \quad (5)$$

The principle of curvatures is calculated by,

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad (6)$$

Keypoints that have a curvature threshold less than r are rejected, as well as points that have a negative value for $Det(H)$.

4.3 ORIENTATION ASSIGNMENT

To assign an orientation we use a histogram and a small region around it. Using the histogram, each key points are represented as an eight pairs, each pairs has an separate meaning. The most prominent gradient orientation(s) are identified.The orientation assignment step, all the keypoint is assigned there must be more than one orientations based on local image gradient directions that is used to achieve invariance to rotation of images. In that, the SIFT selects the most largest vector named as *S*-vector, this to represented as the orientation assignment. These steps obtains invariance to image location on the object, and also in scale and rotation. Finally, while obtaining the local image detector step, a keypoint has computed a descriptor vector for each such that

it is highly distinctive, fully invariant and partially invariant to other variations on the local images.

4.4 KEY POINT DESCRIPTOR

Keypoint descriptor is mainly used to construct the feature vector for the keypoint which has already localized .To do this, a 16×16 window around the keypoint is obtained, in each window the eight pair of histogram has been represented tat forms an 128 dimensional vector. This 16×16 window is broken into sixteen 4×4 windows.Within each 4×4 window, gradient magnitudes and orientations are calculated. These orientations are put into an 8 bin histogram.Any gradient orientation in the range 0-44 degrees add to the first bin, 45-89 add to the next bin. These 128 numbers form the "feature vector".

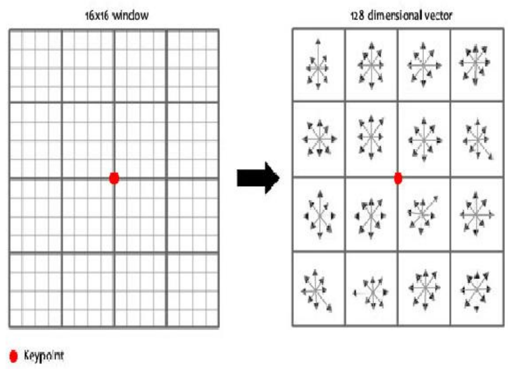

Fig4.2: key point descriptor

In the orientation assignment, keypoint localization, and the local frame descriptor this are used for the calculation to use a lot of dividers mainly used to compute Taylor expansion(used for eliminating edge response) and to perform normalization where transformation and detransformation also performed. In this formation it causes hardware implementation these implementation is difficult. Thus, problems solving is an important issue for SIFT hardware design that is implemented.Another problem for using SIFT hardware is its highly computational data dependent structure. In the flow of data in SIFT, a new scale image obtained has to wait for the completion of the previous scale image that is obtained by the Gaussian pyramid.

V. COMPUTING A PROJECTION MATRIX

For each keypoint, image patches are extracted around it with size 41 x 41pixels. The horizontal and vertical gradients are calculated resulting in a vector of size 39 x 39 x 2 = 3042. All these vectors are made into a k x 3042 matrix A where k is the number of keypoints detected. The covariance matrix of A is calculated. The eigen vectors and eigen values of covA are calculated. The first n eigenvectors are selected and the projection matrix is an n x 3042 matrix composed of these eigenvectors. The projection matrix is only computed once and saved. It results in a PCA-SIFT descriptor of the size n.

VI. LOGO DETECTION AND RECOGNITION

Application of CDS to logo detection and recognition requires to establish a matching criterion and verify its probability of success. False Rejection rate and False acceptance Rates is denoted as (FRR and FAR, respectively)

$$\text{FAR} = \frac{\text{No of incorrect logo detection}}{\text{No of logo detections}} \quad (7)$$

$$\text{FRR} = \frac{\text{No of unrecognized logo appearance}}{\text{No of logo appearances}} \quad (8)$$

Below dataset represent these false acceptance rate and false rejection rate results; For these result we can use any type of logo sets. It may be an dataset, we can use any company logos,college logos or else we may get from MICC logo set(which is an pre defined logo set that contains more than 1000 logos) by the help of FRR and FAR guarantees a high detection rate at the detriment of a small increase of false alarms. Diagrams in Fig.6.1 show FAR and FRR for the different classes using MICC-

Logos dataset. Each dataset has different classes.According to the classes the result is obtained.

Fig6.1: MICC-Logos dataset. Logo classes

By these SIFT based logo detection and recognition logos can be easily detected where in reference logo by the help of testing the images, so the overall number reaches the above fixed threshold of SIFT. SIFT matches are obtained by help of the each interest point in *SA* its Euclidean distance to all interest points in *SB* ,and keep matching only the nearest-neighbors. RANSAC is also one of an algorithm used mainly on logo recognition and detection and it follows the same idea of SIFT but it introduces a model called (transformation) based criterion but not much more important consistent in practice[4]. This algorithm selects only the certain criterion that matches and satisfy an affine transformation obtain between the reference logos and test images.

The (iterative) RANSAC matching process, is applied as a "refinement" of SIFT matching a similar approach is used in both cases of algorithm while matching the Lowe's second nearest neighbor test is satisfied. Secondly,CDS logo detection algorithm that use context in their matching procedure. The Video Google [5]approach is closely related to SIFT method as it introduces a spatial consistency criterion, according to that only it matches and share similar spatial layouts and then they are selected.

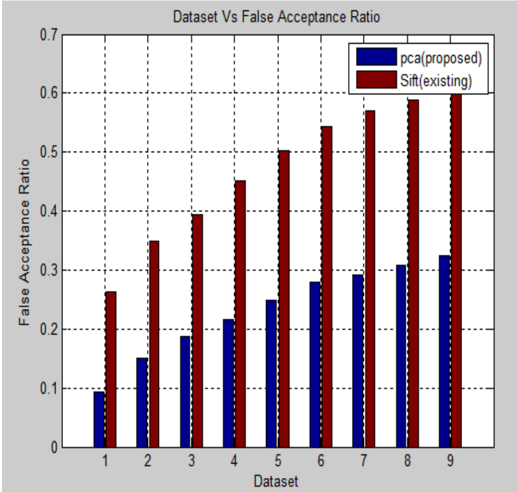

Fig6.2: False acceptance rate of (SIFT AND PCA-SIFT)

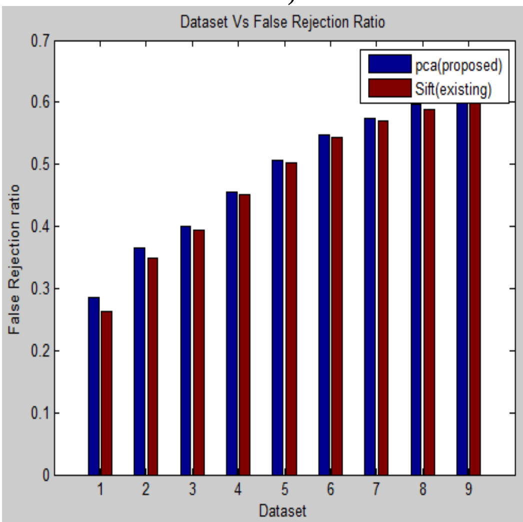

Fig6.3: False rejection rate of (SIFT AND PCA-SIFT)

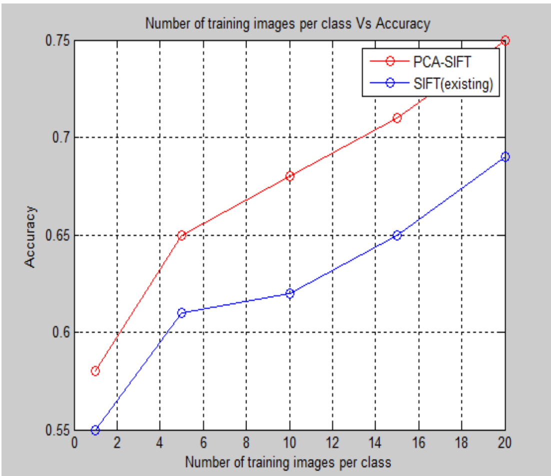

Fig6.4: Comparison of accuracy (SIFT AND PCA-SIFT)

VII. FUTURE ENHANCEMENT

The future work is to extend the PCA to local feature descriptor Speeded Up Robust Feature descriptor (SURF). The PCA-SIFT can be extended to other object recognition problems such as object tracking and object matching with large dataset.

VIII. CONCLUSION

The proposed PCA-SIFT algorithm is quite simple and compact. The PCA is applied to the normalized gradient patch. The PCA based local descriptors are more distinctive, more robust to image deformations and more compact than the standard SIFT representation. The result shows that using these descriptors in logo matching results in increased accuracy and faster matching. The PCA-SIFT feature vector is significantly smaller than the standard SIFT feature vector and can be used with the matching algorithms. The euclidean distance between two feature vectors is used to determine whether the two vectors correspond to the same keypoint in different images.

The proposed system solves the problem of high dimensions of the feature vectors extracted from the existing SIFT algorithm and addresses the computation latency. The performance evaluation is done based on the parameters like FAR and FRR. The simulation result shows that the proposed scheme performs well with number of training images in the dataset with increased accuracy.

REFERENCES

Proceedings Papers:

[1] H. Sahbi, L.Ballan, G.Serra, and A.Del Bimbo, “Context- Dependent Logo Matching and Recognition” *IEEE trans, vol.22, no. 3, 2013*.

[2] Liang-Chi Chiu, Tian-Sheuan Chang ”Fast SIFT Design for Real-Time Visual Feature Extraction” *IEEE VOL. 22, NO. 8, AUGUST 2013.*

[3] Tsung-Jung Liu,Weisi Lin, and C.-C. Jay Kuo “Image Quality Assessment Using Multi-Method Fusion” *IEEE VOL. 22, NO. 5, MAY 2013.*

[4] George E. Dahl, Dong Yu, Li Deng, and Alex Acero, “Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition” *IEEE VOL. 20, NO. 1, JANUARY 2012*

[5] 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.*

[6] 8.J. Sivic and A. Zisserman “Efficient visual search of videos cast as text retrieval “*IEEE trans, vol.20, no. 3, 2009*