ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

An Efficient Face Recognition and Retrieval Using LBP and SIFT

Yogesh R. Tayade[1,] Prof. S.M. Bansode[2]

ME Scholar, Comp. Sci. & Engg. Government College of Engineering, Aurangabad, Maharashtra, India[1]

Asst. Prof. Dept. of Comp. Sci. & Engg. Government College of Engineering, Aurangabad, Maharashtra, India[2]

Abstract: In this paper, we have proposed a method for face recognition and retrieval. In most of the cases various methods are unable to increase retrieval rate of face images especially LFW images, with the help of proposed system the retrieval rate drastically increased. In face recognition, inter class objects should have larger distance than intra class objects ideally. By extracting LBP & SIFT features of training images and arranging them in sparse representation; shape context and inner distance shape contexts methods are applied on test image for deriving relevant images with better performance.

Keywords: LFW, Inner distance and shape context,SIFT, LBP.

I. Introduction

Content based image retrieval (CBIR) is the basic idea behind image retrieval techniques in image processing. Providing a test image or query image, extracting feature from it on the basis of shape and color, comparing with images from database using different techniques is challenging work. While comparing the test image feature with the dataset, interclass and intraclass distance should be taken into account for retrieval of images.

In Co-transduction for shape retrieval framework[1], this combines two different distance metrics. With the same spirit as co-transduction, tri-transduction combines three different distance metrics. The significant performance improvement on four large data sets has demonstrated the effectiveness of co-transduction/tri-transduction for shape/object retrieval. But it did not performed significantly on LFW dataset2], which contains face images with different lightning pose, illuminations. We used Scale Invariant Feature Transform (SIFT) algorithm for corner detection on the face image, Local Binary Points (LBP) and Inner Distance and Shape Context (IDSC) to compare feature of test image with images in dataset.

II. Previous Work

To extend the idea of the traditional content based image retrieval systems to face images automate the face retrieval system is the one solution. The users can ask for relevant images through query to the system by providing face images. Systems like QBIC (Query by image content)[3] and MIT Photobook [4] have been used for such applications. The QBIC is a classic CBIR system which allows querying by simple query images and image properties like average color, color distribution and texture to retrieve the face images the user is looking for.

Photobook [5] is a content-based retrieval system that provides methods for searching several types of related image databases including faces. The key advantage of Photobook is that it uses methods like PCA (eigenfaces) as one of the key image feature for performing the retrieval. PCA based compression of the images is statistical in nature this restricts it to some extent.

The authors of LFW[2] proposed two different methods. Face recognition method of Eigen-faces is first approach and second, they used the recognition system implemented by Nowak and Jurie[6]. These methods sample randomly corresponding patch pairs from the image and uses randomized tree to assign the patch pair in cluster encoded with binary vectors. Support Vector Machine (SVM) is used to classify these binary vectors.

Sanderson and Lovell[7] used symmetric division for patches. Which divides each face image in adjacent region (2*2) and again these regions are divided into blocks of (8*8) size. To extract descriptive feature from these blocks 2D Discrete Cosine Transform (DCT) applied and from these features probabilistic histogram is generated for every region. Image comparison is done based on the difference of histogram[8].

Pinto uses in [9] a pixel representation, a V1-like representation and a V1-like+ representation. Enhancing their work in [10], they added six new instances with little variations in the form of size and resolution of the images. Classification is done by multiple kernels learning (MKL) associated with SVM.

Wolf[11] tried with descriptor based methods. Using Local Binary Pattern (LBP) operator and two novel patch

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

based descriptor, three patch LBP and four patch LBP. They extend their work One Shot Similarity Score[12] by providing unrestricted dataset.

III. Proposed Work

In proposed system we worked on the limitations mentioned in paper Co-Transduction in Shape Retrieval[1] and we come up with relevant output. Modular flow of proposed system is shown in fig. 1 below.

The proposed system takes an image as input filters it and represent in sparse matrix deriving SIFT and LBP features, the detail process is described below.

A. *Preprocessing*

In this module we removed the noise if any. Noise is some unwanted things that contaminate an image. This is achieved with the use of median filter.

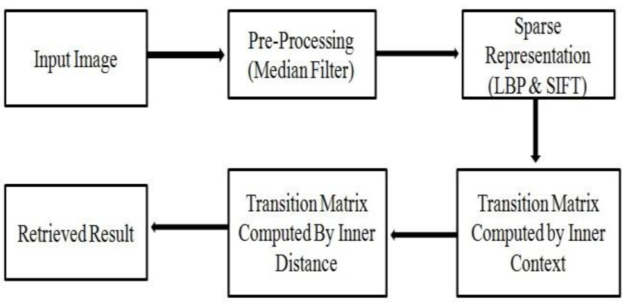

Fig. 1 Modular Flow of Proposed Work

B. *Scale Invarient Feature Transformation (SIFT)*

Collection of local features vectors transformed from the image is main motto of the SIFT. Feature vectors acquired from the images have no effect of scaling, rotation or transformation on the image. Any object represents many features, SIFT image feature vector provide a set of features that are not changed under any conditions[13]. Feature vectors calculated using SIFT can identify various objects in different images. Transformations are used to match faces images in the dataset.

While allowing an object to be recognized in a larger image, SIFT image features also allow objects in multiple images of the same location taken from different positions within the environment, to be recognized.

To obtain those features, SIFT algorithm works in steps we use feature extraction equations from [13][14]. At first step, "scale function" is used to identify featured locations and scales that are identifiable from different views of the same object.

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

Where, * is the convolution operator, G(x, y, σ) is a variable-scale Gaussian and I(x, y) is the input image.

To obtain stable key-points locations in scale-space difference of Gaussians, locating scale-space extrema, D(x, y, σ) by computing the difference between two images, one with scale k times the other. D(x, y, σ) is then given by:

$$D(x,y,\sigma) = L(x,y,k\sigma) - L(x,y,\sigma) \qquad (2)$$

To detect the local maxima and minima of D(x, y, σ) each point is compared with its 8 neighbors at the same scale, and its 9 neighbors up and down one scale.

To eliminate more points from the list of key-points by finding those that have low contrast or are poorly localized on an edge are calculated by the Laplacian,. The location of extremum, **z**, is given by:

$$z = -\frac{\partial^2 D^{-1}}{\partial x^2}\frac{\partial D}{\partial x} \qquad (3)$$

If the function value at **z** is below a threshold value then this point is eliminated.

C. *Local Binary Pattern (LBP)*

The Local Binary Pattern (LBP) operator basically designed for extracting features of textures[15]. Later it was also shown that this operator performs well on other tasks[16]. The basic LBP operator assigns to every pixel a pattern. For each pixel location its intensity value is compared to every pixel in its neighborhood. The result of this comparison is a set of binary values. These binary values are converted to a decimal number. Originally a 3×3 neighborhood is considered. Later, the LBP was extended to use a bigger neighborhood. In this extension a circular neighborhood is considered which is determined by the radius of the ring and the number of samples lying on this ring. Mostly it is sufficient to use the latter. The pattern P on location (x, y) in the image I given the neighborhood $\{(x_0, y_0), \ldots (x_{S-1}, y_{S-1})\}$ can be calculated[8] using the equation(4) :

$$P(x,y) = \sum_{i=0}^{i=S-1} 2^i \, sign\big(I(x,y) - I(x_i,y_i)\big) \qquad (4)$$

D. *Inner Distance and Shape Context (IDSC)*

In our proposed work to build shape descriptor we added the inner-distance with SIFT and LBP, defined as the length of the shortest path within the shape boundary. Inner-distance is rigid to shape articulations. For example, in Fig. 2, although the points on shape (a) and (b) have similar spatial distributions, they are quite different in their part structures. On the other hand, shapes (c) and (b) appear to be from the same category with different articulations. The inner-distance between the two marked points is quite different in (a) and (c), while almost the same in (c) and (b). Intuitively, this example shows that the inner-distance is insensitive to articulation and sensitive to part structures, a desirable

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

property for complex shape comparison [17]. In this example, it is clear that the inner-distance reflects part structure and articulation without explicitly decomposing shapes into parts.

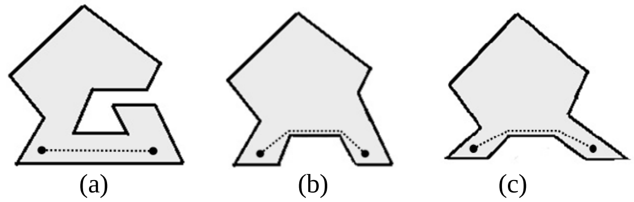

Fig.2 Three objects with the dashed lines denote shortest paths within the shape boundary that connect landmark points.

E. *The Inner-Distance*

First, we define a shape O as a connected and closed subset. Given a shape O and two points x; y ϵ O, the inner-distance between x; y, denoted as d(x,y; O), is defined as the length of the shortest path connecting x and y within O. One example is shown in Fig. 3.

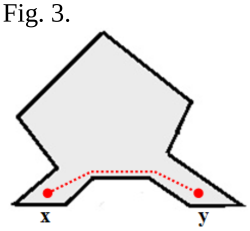

Fig. 3. Definition of the inner-distance. The dashed line shows the shortest path between point *x* and *y*.

F. *Sparse Representation:*

Extracting the LBP and SIFT features of images and represented by sparse matrix. A basic problem in object recognition is to use labeled training samples from distinct object classes to correctly determine the class to which a new test sample belongs. We arrange the training samples in a matrix such that every column represents features of distinct images.

G. *Transition matrix computed by Shape Contexts and IDSC:*

STEPS:

- Finding a list of points on shape edges
- Computing the shape context
- Computing the cost matrix
- Finding the matching that minimizes total cost
- Modeling transformation
- Computing the shape distance

IV. Experimental Result

We load complete training dataset into a database in classified format. Then test image or query image or input image is passed through proposed modular flow such as preprocessing, LBP and SIFT feature extraction operations and we successfully retrieve relevant images to test image from the database.

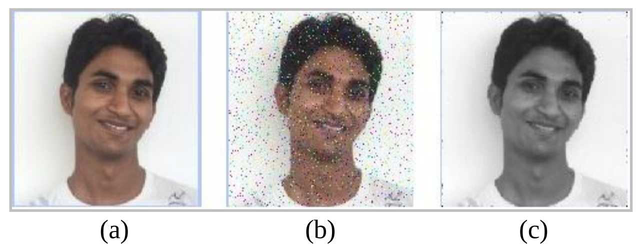

Fig. 4 Image after preprocessing operations image (a) Test Image, (b) Noise Detected on image and (c) Image after median filter.

Here, figure 5 shows some sample test images with their LBP featured images, SIFT image with corner points on the images.

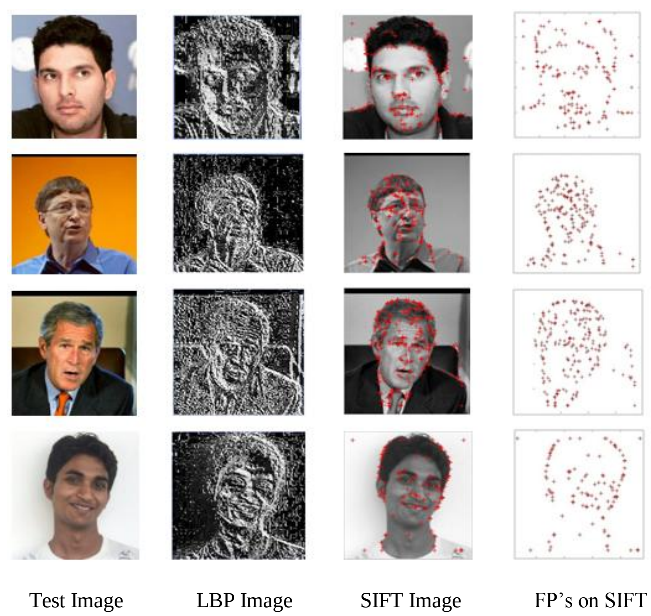

Fig. 5 Sample Images with there LBP and SIFT Features.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

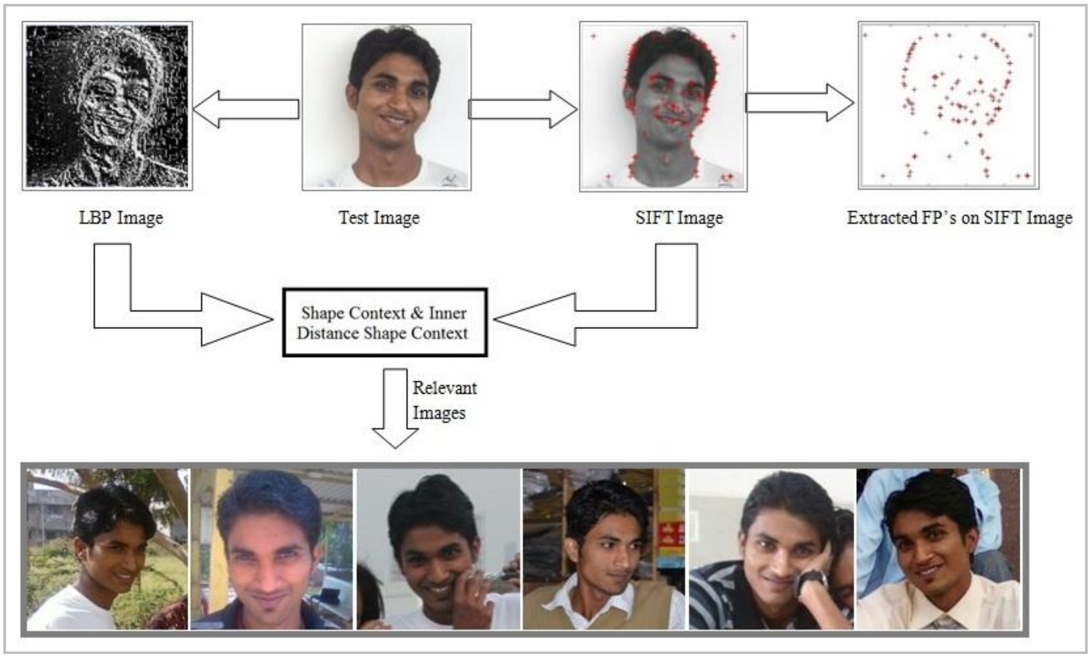

Fig. 6 Relevant Images

We generate a performance graph on the basis of time required for retrieval of relevant images from the dataset corresponding to the total number of images in to our dataset. In fig. 7 the graph shows the time required for retrieval of images from the dataset which is directly proportional to the total number of images in the dataset. As the dataset grows time gradually goes up. For this result, we tested retrieval time for 1000 images in regular addition of 100 images per iteration.

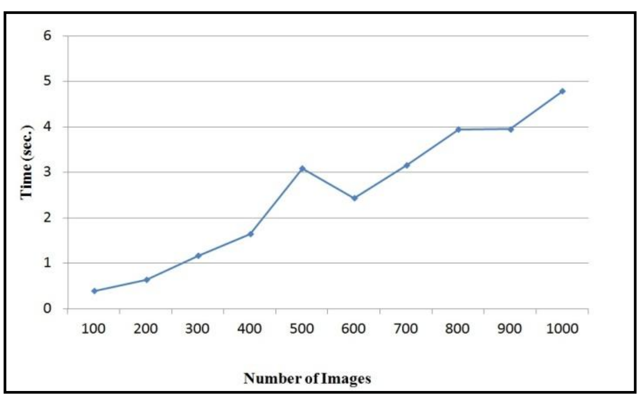

Fig. 7 Performance Graph

In performance analysis, our work efficiently reduced recognition and retrieval time as compare to some of previous methods like DWT(Discrete Wavelet Transformation), PCA(Principle Component Analysis), LDA(Linear Discriminant Analysis) for face recognition and retrieval[18]. Table 1 shows the comparison of different methods with respect to time vs number of images in dataset. Graphical representation of table one is also shown in the underlying fig. 8

Table 1 Comparison table

Number of Images →	20	40	60	80	100
Methods	Time (sec.)				
DWT	1	1.52	1.79	2.4	2.58
PCA	0.7	0.753	0.823	0.854	0.97
LDA	0.5	0.64	0.72	0.81	0.87
SPARSE	0.105	0.163	0.32	0.28	0.39

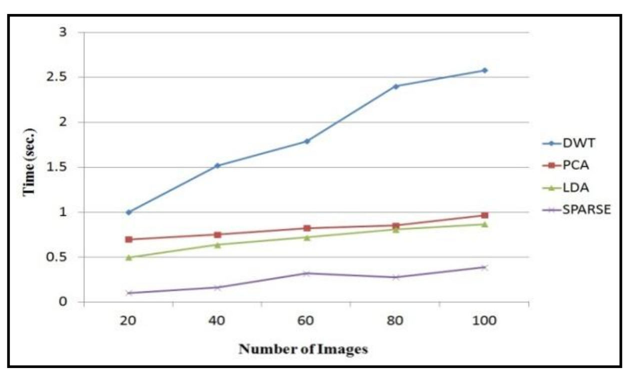

Fig. 8 Comparison Graph

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

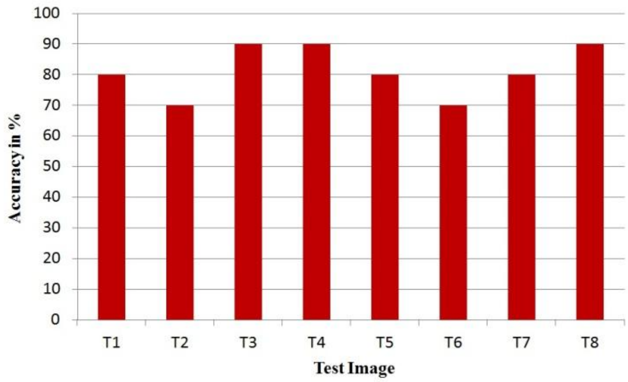

Fig. 9 Accuracy chart

Proposed system has achieved accuracy about 81.25 %. Fig. 9 shows the accuracy in percentage for different 8 sample test images.

CONCLUSION

In this proposed work, we provide an efficient method for face retrieval by combining three different algorithms SIFT, LBP and IDSC. Through this work we successfully retrieve the face images from trained dataset of Labeled Faces in Wild (LFW) images efficiently achieving better retrieval rate .

As we worked on static dataset, further we can enhance our work by providing dynamicity to the database and test image by adding modules for image acquisition through camera and cropping techniques for resizing them.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers for their comments that helped to improve this paper. The authors would also thank Prof. S. M. Bansode for their valuable guidance and suggestions for paper revision.

REFERENCES

[1] "Co-Transduction for Shape Retrieval" Xiang Bai, Bo Wang, Cong Yao, Wenyu Liu, and Zhuowen Tu, IEEE Transactions on Image Processing, vol. 21, no. 5, May 2012

[2] [2]G. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, Labeled faces in the wild: A database for studying face recognition in unconstrained environments Univ. Massachusetts, Amherst, MA, Tech. Rep. 07- 49, 2007.

[3] M. Flickner, H. Sawhney,W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. Query by image and video content: The qbic system. *Computer*, 28(9):23–32.

[4] A. Pentland, R. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases, 1994.

[5] A. Pentland, R. Picard, and S. Sclaroff. Photobook: tools for content based manipulation of image databases. *Proc. SPIE: Storage and Retrieval for Image and Video Databases II*, 2185.

[6] E. Nowak and F. Jurie, "Learning visual similarity measures for comparing never seen objects," in *Conference on Computer Vision & Pattern Recognition*, jun 2007.

[7] C. Sanderson and B. C. Lovell, "Multi-region probabilistic histograms for robust and scalable identity inference." in *ICB*, ser. Lecture Notes in Computer Science, M. Tistarelli and M. S. Nixon, Eds., vol. 5558. Springer, 2009, pp. 199–208.

[8] Computer Vision for Human-Computer Interaction Research Group Institute for Anthropometrics University Karlsruhe (TH) " Evaluation of Local Descriptors on the Labeled Faces in the Wild Dataset" , Masoud Roschani

[9] N. Pinto, J. DiCarlo, and D. Cox, "Establishing good benchmarks and baselines for face recognition," in *Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition*, 2008.

[10] J. DiCarlo, N. Pinto, and D. Cox, "How far can you get with a modern face recognition test set using only simple features?" in *Computer Vision and Pattern Recognition (CVPR)*, 2009.

[11] L. Wolf, T. Hassner, and Y. Taigman, "Descriptor based methods in the wild," *Faces in Real-Life Images Workshop in European Conference on Computer Vision (ECCV)*, 2008.

[12] Y. Taigman, L. Wolf, and T. Hassner, "Multiple one-shots for utilizing class label information," *British Machine Vision Conference (BMVC)*, 2009.

[13] "Implementing the Scale Invariant Feature Transform(SIFT) Method" YU MENG and Dr. Bernard Tiddeman

[14] D. G. Lowe, "Object recognition from local scale-invariant features," *Computer Vision, IEEE International Conference on*, vol. 2, pp. 1150–1157 vol.2, 1999.

[15] T. Ojala, M. Pietikäinen, and D. Harwood, A comparative study of texture measures with classification based on feature distributions., 1996, pp. 51–59.

[16] T. Ahonen, A. Hadid, and M. Pietikãinen, *Face Recognition with Local Binary Patterns*, 2004, pp. 469–481.

[17] "Shape Classification Using the Inner-Distance" Haibin Ling David W. JacobsUS-Israel Binational Science Foundation grant number 2002/254.

[18] "Performance Alalysis of Face Matching and Retrieval In Forensic Applications", N. Lavanyadevi**,** SP. Priya & K. Krishanthana , International Journal of Advanced Electrical and Electronics Engineering (IJAEEE) ISSN (Print) : 2278-8948, Volume-2, Issue-2, 2013.

BIOGRAPHY

Yogesh R. Tayade, he is currently pursuing M.E. in computer science and engineering from Government College of Engg, Aurangabad, Maharashtra, India.

Prof. S. M. Bansode, she is currently working as assistant professor in Department of Computer Science and Engineering, Govt. College of Engineering, Aurangabad, Maharashtra, India. She has 14 year of experience in the field of teaching.