An Efficient Face Recognition Using Multi-Kernel Based Scale Invariant Feature Transform

CH. Hima Bindu *himabindu.chelluri@gmail.com*
Dept. of ECE,
GITAM University,
Hyderabad, INDIA

K. Manjunathachari *manjunath4005@gmail.com*
Dept. of ECE
GITAM University
Hyderabad, INDIA

Abstract

Face recognition has gained significant attention in research community due to its wide range of commercial and law enforcement applications. Due to the developments in the past few decades, in the current scenario, face recognition is employing advanced feature identification techniques and matching methods. In spite of vast research done, face recognition still remains an open problem due to the challenges posed by illumination, occlusions, pose variation, scaling, etc. This paper is aimed at proposing a face recognition technique with high accuracy. It focuses on face recognition based on improved SIFT algorithm. In the proposed approach, the face features are extracted using a novel multi-kernel function (MKF) based SIFT technique. The classification is done using SVM classifier. Experimental results shows the superiority of the proposed algorithm over the SIFT technique. Evaluation of the proposed approach is done on CVL face database and experimental results shows that the proposed approach has a recognition rate of 99%.

Keywords: SIFT, SVM, Multi-kernel SIFT, Face Recognition.

1. INTRODUCTION

In machine learning process, face recognition is explained as a process of extracting high dimensional features of a face image and then use these features in finding out the identity of subject. It is one of the implicit tasks of the daily activities performed by humans. However, for a machine, automatic recognition of face is a difficult and complex task. Hence, face recognition is one of the important computer vision problems. Since the past few decades, wide research is being conducted in this area by many researchers [1]. However, designing a reliable and a robust face recognition system remains an open problem to the scientific community. The various application areas in which face recognition can be used are: in security system, in video surveillance system, general identity authentication and smart card applications, gender classification[2], expression recognition [3].

Basically, face recognition schemes can be divided into two approaches [4], namely, face-based and constituent-based. Face-based approach uses global information of the face to perform recognition. This global information is directly obtained from the pixel information of face images and are represented by a small number of features. These features capture the variance among different faces. A few of the examples are Principle component analysis (PCA), Fisher linear discriminate (FLD) etc.

Constituent-based method, on the other hand, is proposed relatively more recently and currently an area of active research in the face recognition field. This method extracts local features from specific feature points of the face image. The success of this method depends on the accuracy of the facial feature extraction schemes. Few examples are Local Binary Patterns, Histogram of Oriented Gradients, and Gabor Wavelets etc. Usually, face based approaches attain better results on images captured in controlled conditions, whereas Constituent based approaches exhibit robustness to variations caused by expression or pose changes.

SIFT (scale invariant feature transform), proposed by Lowe [5], [6] is one of the constituent-based methods, which gained wide attention in the research community because of its robustness and affine invariance properties. This technique works by first detecting key points or interest points in the given face image and then computing feature descriptors at these keypoints. In performing recognition (or classification) process, each keypoint descriptor of the given image is matched independently with all descriptors that are extracted from the database images, and based on the result of the matching procedure; the image is allocated to a class present in the database.

2. RELATED WORK ON SIFT

Over the past few years research is being made in the direction to increase the performance and to decrease the complexity of the original SIFT algorithm. In 2004, PCA-SIFT was proposed by Y. Ke and Sukthankar [7] which not only addressed the high dimensionality problem of SIFT vector but also increased the speed and accuracy of classification. Each keypoint is represented by a 128 dimensional vector in SIFT which is of high dimensionality. To address this problem, Y. Ke used Principle component analysis that replaced histogram method in SIFT.

In 2006, Bay [8] proposed SURF algorithm to improve the detection and description efficiency of keypoints. The basic idea of surf is almost similar to SIFT. However, in SURF, a quick Hessian matrix is adopted for detection, which has competitive advantages on speed and accuracy. This technique describes each keypoint with a 64-dimensional feature descriptor.

Alhwarin et al. [9] proposed an improvised SIFT algorithm by splitting the features extracted from both the test and the database images into several subgroups before they are matched. This paper proposed two modifications to accelerate feature matching. In the first modification, the features are split into maximas and minimas and in the matching stage only the like features are compared. Secondly, a new element which is the angle between two independent orientations is added to the SIFT feature. This technique reduced the matching time by almost 40%.

In 2013, Saleem and Sablatnig [10] proposed modifications in SIFT algorithm in a way to improve the contribution of edges during the construction process of SIFT descriptors in multispectral imaging. In 2015, A.T. Tra et al. [11], proposed another modified SIFT algorithm where Dominant gradients were identified in the sub histogram of the SIFT descriptor. With this modification the dimension length of the descriptor drastically reduced from 128-bit in the original SIFT to 48-bit.

2.1 Related Work On Face Recognition Based On SIFT

Mohamed et al. [12] proposed a SIFT approach for face recognition and compared it with other techniques like eigen faces and fisher faces. The results show the superiority of SIFT over these two methods, especially for smaller training face images. In 2008, Yanbin et al. [13] compared real extracted features with training sets to recognize the face. ORL face database is used for experimentation and achieved recognition rates of 96.3% for SIFT, 92.5% for PCA, 91.6% for ICA, and 92.8% for FLD. C Geng et al. [14] proposed two new approaches: Volume-SIFT (VSIFT) and Partial-Descriptor-SIFT (PDSIFT) for face recognition based on the original SIFT algorithm. Experimentation on the ORL and AR databases depicted that the performance of PDSIFT is significantly better than the SIFT approach.

Majumdar and Ward [15] proposed an approach for face recognition which reduce the number of SIFT descriptors by discriminative ranking of SIFT feature. This method not only reduced computational complexity but also showed that it is robust to changes in head pose, illumination, facial expression and partial occlusion. In 2010, Križaj et al. [16] proposed fixed key point SIFT (FSIFT). According to this technique, SIFT descriptors are computed at predefined locations learned during the training stage. Experimental results on extended Yale database concludes this method performs better than the techniques such as LDA, PCA, and other techniques based on SIFT. In 2014, Dagher et al. [17] proposed face recognition approach using SIFT in combination with K-means algorithm. The proposed algorithm provides better recognition rate than LDP and other SIFT based techniques. Lenc and Kral [18] proposed a corpus creation algorithm to extract faces from the database of photos of different individuals taken in an uncontrolled environment and to create a facial corpus. The faces are extracted using SIFT based Kepenekci face recognition approach. The experimental results show that their approach outperforms the original Kepenekci method.

3. SIFT TECHNIQUE

The SIFT algorithm identifies key/interest points in an image which are invariant to scaling, translation and rotation. These keypoints are represented by orientation invariant feature descriptor. These features are highly distinctive. SIFT features are commonly used for the object recognition and have hardly been used for face recognition. SIFT features are invariant to scale, rotations, affine transformations and partially invariant to illumination changes. Basically, this algorithm has four computational stages: i) Keypoint detection, ii) removal of unwanted key points, iii) orientation assignment, and iv) descriptor calculation.

3.1 Keypoint Detection

In the first stage, the location of the interest points called keypoints, are identified in the scalespace. Generally, Pixels with maximum or minimum intensities form keypoints in the scalespace. The scale space denoted by L(x,y,σ) is formed by convolution of variable-scale Gaussian function G(x,y,σ) with the given image I(x,y).

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

Where

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

And * represents convolution in the equation.

The difference of Gaussian (DOG) images are found by the difference of two nearby scaled images whose scales are separated by a multiplicative factor of k.

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) \qquad (3)$$

$$= L(x,y,k\sigma) - L(x,y,\sigma) \qquad (4)$$

Then for keypoint identification, in DoG pyramid, each pixel is compared with its neighboring pixels, that is, with 8 pixels in the same level and nine pixels in each of the lower and higher levels, that is, a total of 26 neighboring pixels. If the pixel intensity is maximum or minimum of all the neighboring pixels, then it forms a keypoint.

3.2 Removal of Unwanted Keypoints

Not all the keypoints found in the previous stage are reliable. In this stage, the best keypoints are selected based on their stability, by discarding the points with low contrast and that are poorly localized to edge. Keypoints with low contrast are eliminated if the intensity of the pixel is below some threshold. To eliminate the keypoints with poor edge response, the ratio of principal curvatures of each candidate keypoint is checked. For keypoint with good edge response, the principal curvature across the edge will be much larger than the principal curvature along it. The principle of curvature is computed by 2 X 2 Hessian matrices at the location, and scale of key points. If the ratio is below a prescribed threshold, the keypoint is retained, otherwise it is removed.

3.3 Orientation Assignment

An orientation is assigned to each keypoint by constructing histogram of gradient orientation θ(x,y) weighted by gradient magnitude m(x, y) from the keypoint's neighborhood:
Where

$$\mathrm{m(x,y)} = \sqrt{\left(L(x+1,y) - L(x-1,y)\right)^2 + \left(L(x,y+1) - L(x,y-1)\right)^2} \quad \text{----------------(5)}$$

$$\theta(\mathrm{x,y}) = \tanh(L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y)) \quad \text{--------------------(6)}$$

The idea is to accumulate all the gradient directions and magnitudes of the pixels in a patch around each keypoint. Then the most prominent orientation in that patch is found by constructing histogram of gradient orientations and assign this orientation to the keypoint. This stage of assigning orientation provides rotation invariance.

3.4 Keypoint Descriptor Calculation

A keypoint descriptor is the last stage and is generated to identify a keypoint. It is a kind of fingerprint which must be unique to each keypoint and it enables to distinguish one keypoint from the other. It is created by computing the gradient magnitude and orientation at each pixel in a 16X16 Gaussian weighted window with keypoint at the center of the window. This 16X16 window is further sub-divided into sixteen 4X4 windows. The gradient magnitudes of each the sixteen windows are accumulated to an 8-bin orientation histogram. Thus, each keypoint descriptor consists of sixteen histograms with eight bins in each histogram, i.e., 16X8=128 elements

4. PROPOSED APPROACH

In this paper, an adaptive SIFT-based approach for face recognition is proposed to enhance the accuracy of face recognition system. The feature descriptors are calculated at predefined image locations that are learned during the training stage. In the proposed approach we use a novel weightage function in the keypoint description stage of the SIFT algorithm instead of the Gaussian function. A multi-kernel function (MKF) consisting of a linearly added logarithmic kernel and an exponential kernel is used in the proposed approach. Once the descriptors are found, the features undergoes classification. Classification is done using an SVM classifier. SVM is a supervised learning model used for data classification. Experimental results shows that the proposed method outperforms the SIFT method.

The algorithm for the proposed approach of face recognition is as follows:
Algorithm:
Step 1. Load the test image.
Step 2. Convert the image to gray scale.
Step 3. Preprocess the given image using Gaussian filter.
Step 4. Calculate feature descriptors of the image using proposed approach.
Step 5. Repeat 1–4 steps for all the database images.
Step 6. The feature descriptors of the test image and database images are given as inputs to the SVM classifier.
Step 7. Go to Step 1 to repeat the same procedure for next test image.

Block diagram for the proposed approach is shown in fig (1):

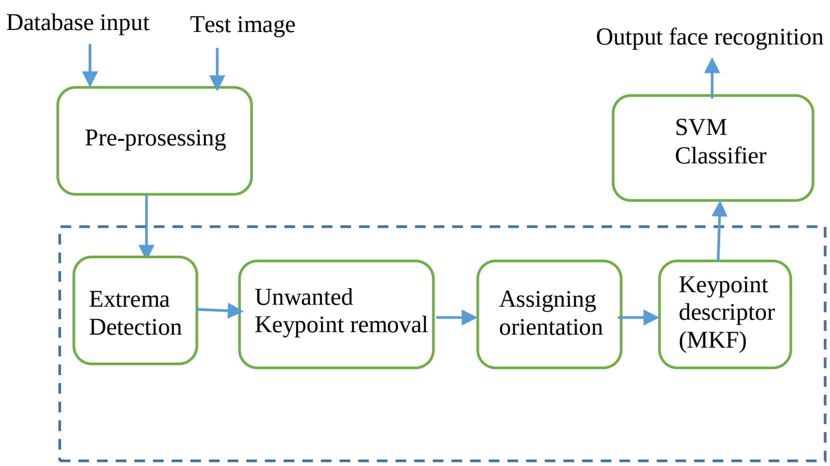

FIGURE 1: Block diagram of the proposed approach.

5. IMPLEMENTATION AND RESULTS

The implementation of the proposed technique is done in matlab. CVL face database (CVL,1999) [19] is used to perform the experimentation of the proposed method. This database consists of face images of 114 persons with seven images per person. Six samples are collected from the database for further evaluation. Fig.2 and fig.3 shows the sample database images and their corresponding multi kernel based SIFT features. To perform the analysis, the evaluation metrics used are false acceptance rate (FAR), false recognition rate (FRR) and accuracy. The evaluation of these parameters are done with respect to percentage of training data. In addition, the performance of the proposed method is compared with the SIFT technique.

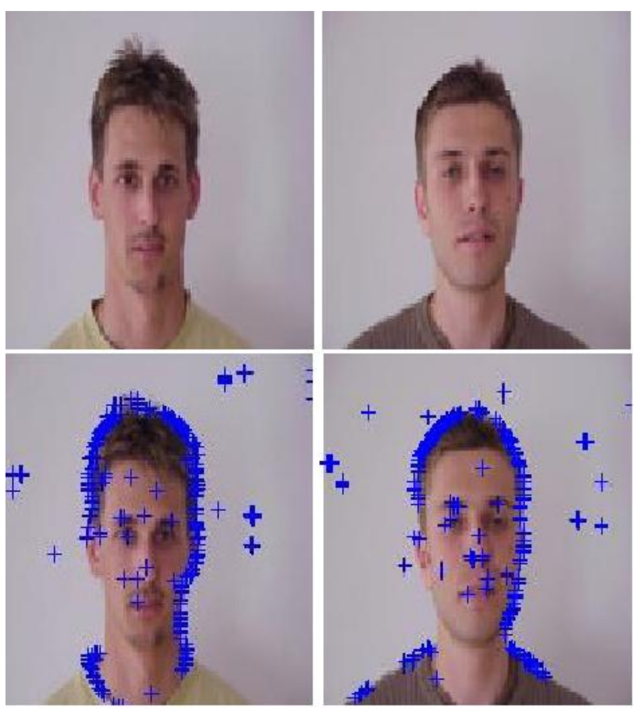

FIGURE 2: Multi-Kernel Function Based SIFT Features.

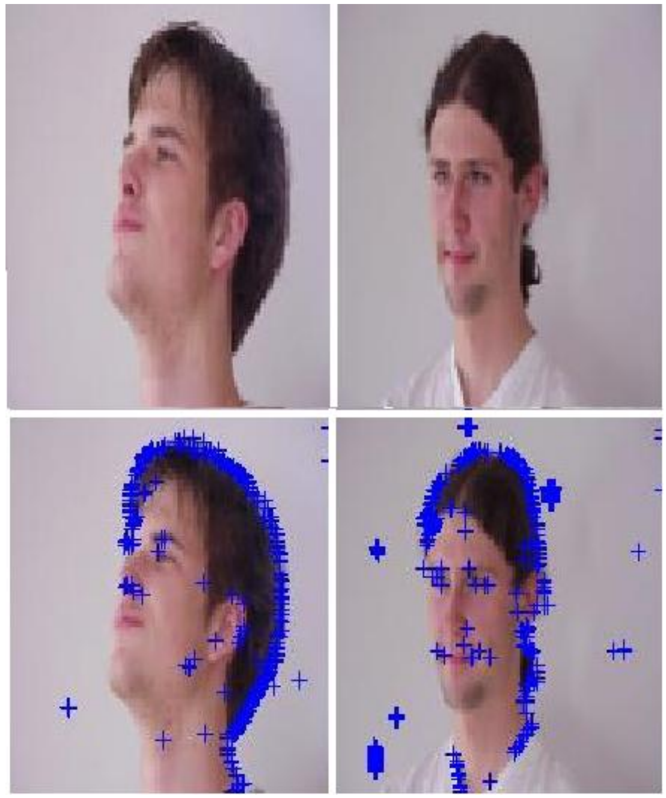

FIGURE 3: Multi-Kernel Function Based SIFT Features.

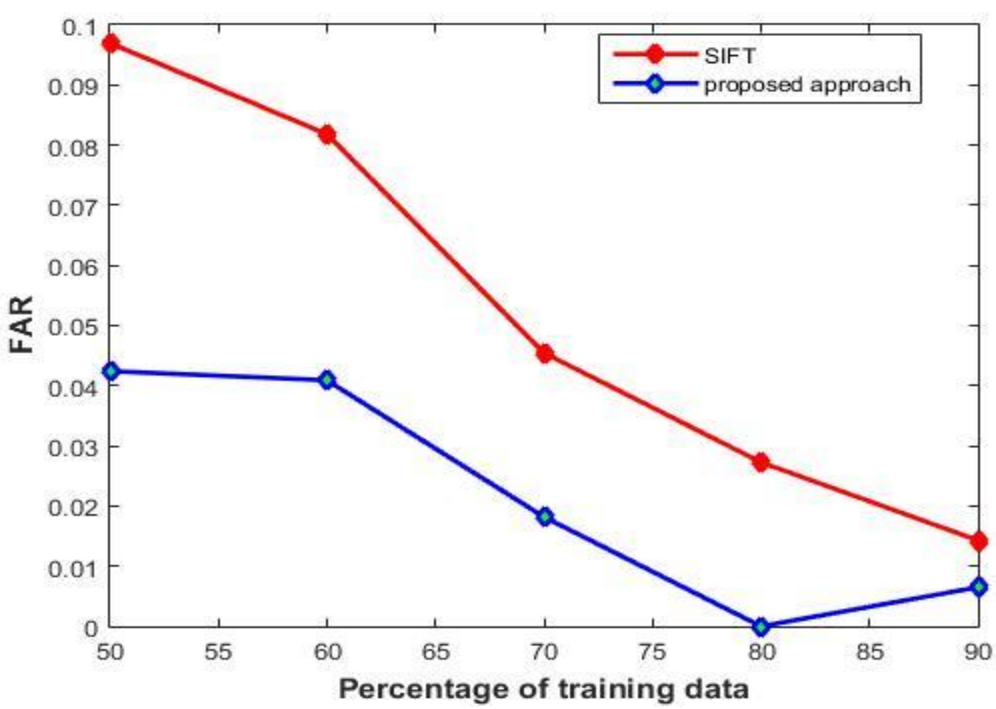

FIGURE 4: FAR.

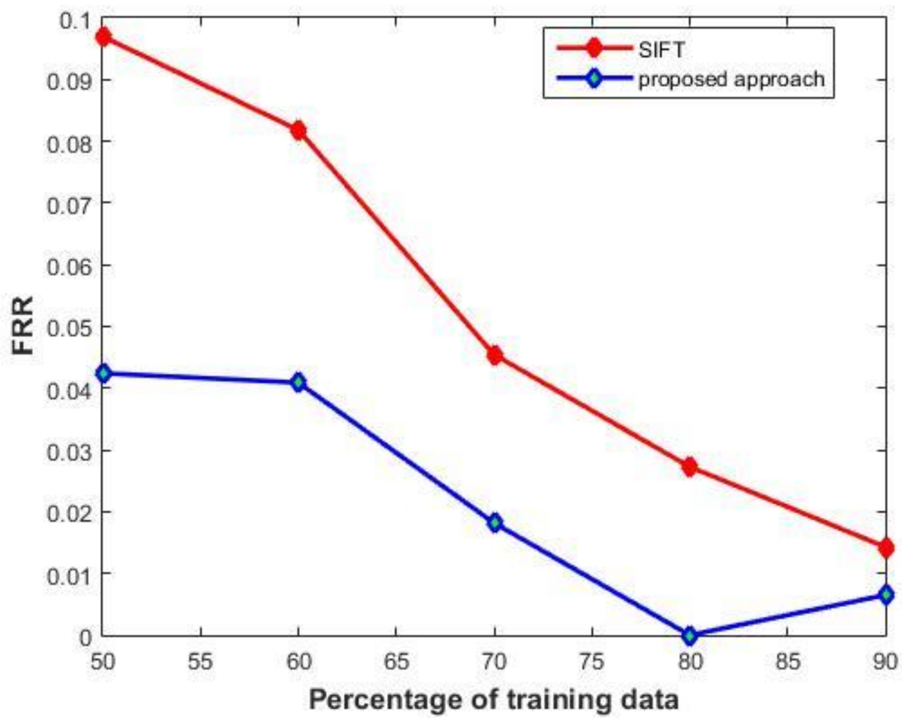

FIGURE 5: FRR.

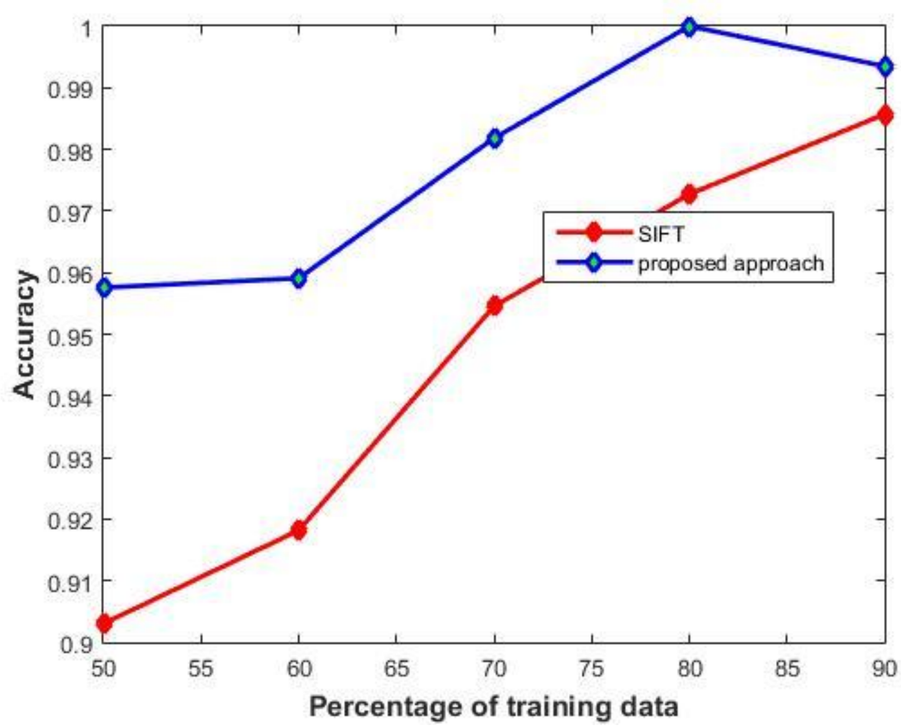

FIGURE 6: Accuracy.

For a face recognition system, FAR indicates the rate of acceptance of an unauthorized face as an authorized face. Low the FAR, better is the system. Fig.4 shows the performance analysis for FAR parameter. For lower percentage of training data, the figure shows an FAR of about 0.05. There is an improvement seen in FAR as the percentage of training data is increased. For 90% of training data, FAR is as good as 0.01.

FRR indicates the rate of rejection of an authorized face as an unauthorized face. Similar to FAR, FRR must be as low as possible for a good face recognition system. Fig.5 shows the performance analysis for FRR parameter. For lower percentage of training data, the figure shows an FRR of about 0.05. As the percentage of training data is increased, FRR also shows an improvement. For 90% of training data, FRR is 0.01.

The parameter accuracy in fig. 6 also shows a significant increase in the recognition rate as the percentage training data is increased. The recognition rate of the proposed method is more than 97% for 70% training data and above as compared to SIFT which is only 95%.

For both techniques, the evaluation parameters show good results at higher percentage training data. It is also clearly understood by the figures that the proposed method outperformed the SIFT technique in all the three parameters.

% Training Data	Recognition rate in %	
	SIFT	Proposed approach
50	90	96
70	96	98
80	97	100
100	98	99

TABLE 1: Table Illustrating The Recognition Rates of SIFT and Proposed Approach.

6. CONCLUSION

Now-a-days face recognition technique based on local feature keypoints have significant theoretical and practical importance. These methods are also used for image registration, image retrieval, and so on. In this paper a new technique is proposed to recognize faces by improving SIFT features. It is an image local feature description algorithm based on scale-space. This paper analyzed the SIFT and the proposed approach using the CVL face database. We have shown that the performance of the proposed method is significantly better than the performance of popular SIFT technique. The descriptors of the identified features of an image are written using the proposed approach. The classification of the features is performed with an SVM classifier. Results shows that the proposed technique has an accuracy of more than 99% as compared to SIFT which exhibits 97%.

7. REFERENCES

[1] He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang. “Face recognition using Laplacian faces”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(3):328–340, 2005..

[2] E. Makinen, R. Raisamo, Evaluation of gender classification methods with automatically detected and aligned faces. IEEE Trans. Pattern Anal. Mach. Intell. **30**(3), 541–547, 2008

[3] Y. Shinohara and N. Otsu, "Facial Expression Recognition Using Fisher Weight Maps," in Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Vol.100, pp.499-504, 2004.

[4] R. Chellappa, C. L. Wilson and S. Sirohey, ”Human and machine recognition of faces: a survey”, Proceedings of the IEEE, Vol. 83, No. 5,705-740, 1995

[5] Lowe, D.G. “Object recognition from local scale invariant features”. In Proceedings of the 7th IEEE International Conference on Computer Vision, IEEE, Vol. 2, 1150-1157, 20-27 September, 1999

[6] Lowe, D.G. “Distinctive image features from scale-invariant keypoints”. International Journal of Computer Vision, 60 (2), 91-110, 2004

[7] Ke, Y., Sukthankar, R.. “PCA-SIFT: A more distinctive representation for local image descriptors”. In Computer Vision and Pattern Recognition (CVPR 2004), 27 June – 2 July 2004. IEEE, Vol. 2, 506-513.

[8] Bay, H., Tuytelaars, T., Gool, L.V. “SURF: Speeded up robust features”. In Computer Vision – ECCV 2006 : 9th European Conference on Computer Vision, 7-13 May 2006. Springer, Part II, 404-417.

[9] F. Alhwarin, C. Wang, D. Ristić-Durrant, and A. Gräser, “Improved SIFT-features matching for object recognition,” in Proceedings of the International Conference on Visions of Computer Science: BCS International Academic Conference (VoCS '08), pp. 178–190, London, UK.

[10] S. Saleem and R. Sablatnig, “A modified SIFT descriptor for image matching under spectral variations,” in Image Analysis and Processing—ICIAP 2013: 17th International Conference, Naples, Italy, September 9–13, 2013. Proceedings, Part I, vol. 8156 of Lecture Notes in Computer Science, pp. 652–661, Springer, Berlin, Germany, 2013.

[11] A. T. Tra, W. Lin, and A. Kot, “Dominant SIFT : A Novel Compact Descriptor,” in Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on, May 2015.

[12] A. Mohamed, “Face recognition using SIFT features,” Tech. Rep., Caltech, Pasadena, Calif, USA,2006.

[13] H. Yanbin, Y. Jianqin, and L. Jinping, “Human face feature extraction and recognition base on SIFT,” in Proceedings of the International Symposium on Computer Science and Computational Technology (ISCSCT '08), vol. 1, pp. 719–722, Shanghai, China, December 2008.

[14] Geng, Cong, and Xudong Jiang. "Face recognition using SIFT features." In Image Processing (ICIP), 2009 16th IEEE International Conference on, pp. 3313-3316. IEEE, 2009.

[15] A. Majumdar and R. K. Ward, “Discriminative SIFT features for face recognition,” in Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering (CCECE '09), pp. 27–30, St. John's, Canada, May 2009.

[16] J. Križaj, V. Štruc, and N. Pavešić, “Adaptation of SIFT features for robust face recognition,” in Image Analysis and Recognition: 7th International Conference, ICIAR 2010, Póvoa de Varzim, Portugal, June 21–23, 2010. Proceedings, Part I, vol. 6111 of Lecture Notes in Computer Science, pp. 394–404, Springer, Berlin, Germany, 2010.

[17] I. Dagher, N. El Sallak, and H. Hazim, “Face recognition using the most representative SIFT images,” International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 7, no. 1, pp. 225–236, 2014.

[18] L. Lenc and P. Kral, “Automatic face recognition system based on the SIFT features,” P

[19] Peter Peer, CVL Face Database, http://www.lrv.fri.uni-lj.si/facedb.html.