# Expression Invariant Face Recognition using DWT SIFT Features

Amrita Biswas
Associate Professor
EC Department,SMIT
Majitar,Rangpo-737132
East Sikkim

M.K.Ghose
Dean Academics& Professor
CSE Department,SMIT
Majitar,Rangpo-737132
East Sikkim

## ABSTRACT
Expression variations pose a serious problem to automatic face recognition. In this paper a robust expression invariant face recognition algorithm is proposed using a single image for training. Discrete Wavelet Transform (DWT) using Coiflets as the basis function is applied to the face images. Then SIFT descriptors of the combination of Approximation and Detail subbands are computed for each of the face images. Matching is done using Angle Distance Similarity measure. The algorithm has been tested on the Essex Grimace database and only a single image has been used for training. The results are promising and have been compared with standard SIFT based method, standard DWT based method, Eigenfaces & Fisherfaces method with single training images.

## General Terms
PatternRecognition, ComputerVision, ArtificialIntelligence.

## Keywords
Face Recognition, DWT, SIFT, Expression Invariant.

## 1. INTRODUCTION
Automatic Face Recognition has huge no. of applications in the real world and hence a lot of research has been devoted to the development of automatic face recognition systems. However face recognition has lot of issues which are still unsolved as human face is not a rigid object and it can be varied easily under different situations. It has been seen that facial expression of a person has a major impact on the face recognition systems. To cope with it, some algorithms rely on extracting stable face features, e.g., extracted line segment [1] and geometric invariants [2]. [3] is based on dynamic link matching which is invariant under face rotation and deformation. One of the setbacks with most of the work done is that there is less reliable invariants when faces carry major expression variations. Few currently developed face recognition systems are based on video sequences so they can handle a variety of facial expressions [4, 5] because video sequences will consist of more number of successive images and hence contain more information compared to the case when only a single image is available. However, a traditional face recognition system based on single training image is still important because in many practical situations, video sequences continuously recording a face image may not be available. Moreover, video sequences will need additional storage space which may be a constraint in certain applications.

Researchers have proposed different methods for face recognition in general.Scale Invariant Feature Transform(SIFT**)** [6] proposed by Lowe has attracted a lot of attention among researchers for pattern recognition because of its excellent performance on object recognition. SIFT features have been investigated for face recognition problems [7] and shown to provide promising results. Variants of SIFT like VSIFT & PDSIFT have also been investigated [8].FSIFT which involves computation of SIFT descriptors at fixed locations was proposed in [9].The ASIFT descriptor proposed in [10] has also been used by researchers for face recognition [11].The rest of the paper is organized as follows. The feature extraction methods used namely SIFT and DWT have been described briefly in Sec.2 and Sec.3contains the details of the implemented algorithm and the results obtained. We conclude in Sec.4

## 2. FEATURE EXTRACTION METHODS
Face Recognition algorithms in general are divided into three main parts, namely: Preprocessing, Feature Extraction &Classification. Preprocessing generally involves noise removal, face detection and resizing operations. Since we are using a standard face recognition database where the face images are of uniform size we have omitted the preprocessing stage. For feature extraction DWT of the images are first computed and linear combination of all the obtained subbands are taken. SIFT descriptors of the fused DWT coefficients are computed to form the final feature vector set. SIFT Descriptors & DWT have been explained briefly as follows:

### 2.1 SIFT
The scale invariant feature transform, called SIFT descriptor, has proved to be invariant to image rotation, scaling, translation, partial illumination changes, and projective transform. The main idea of the SIFT descriptor is detecting feature points efficiently through a staged filtering approach that identifies stable points in the scale-space. Local feature points are extracted from the following steps:

- Select candidates for feature points by searching peaks in the scale-space from a difference of Gaussian (DoG) function.
- Localize the feature points by using the measurement of their stability.
- Assign orientations based on local image properties.
- Calculate the feature descriptors which represent local shape distortions and illumination changes.

After candidate locations have been found, a detailed fitting is performed to the nearby data for the location, edge response, and peak magnitude. To achieve invariance to image rotation, a consistent orientation is assigned to each feature point based on local image properties and describes it relative to this orientation. The histogram of orientations is formed from the gradient orientation at all sample points within a circular window of a feature point. Peaks in this histogram correspond to the dominant directions of each feature point. For illumination invariance 8 orientation planes are defined. Towards this end, the gradient magnitude and the orientation are smoothed by applying a Gaussian filter and then sampled over a 4 x 4 grid with 8 orientation planes.

## 2.2 DWT

Discrete wavelet transform is a popular method in signal processing and has been used in various research fields. Wavelet transform has advantages of multi-resolution & multi-scale decomposition. In frequency domain, when the facial image is decomposed using two dimensional wavelet transform, we get four sub regions. These regions are: one low-frequency region LL (approximate component), and three high-frequency regions, namely LH (horizontal component), HL (vertical component),and HH (diagonal component), respectively. In wavelet packet decomposition [12], we divide each of these four regions in a similar way.With each level of decomposition, the number of pixels in each subband greatly reduces into packets while the important features of the underlying image are retained. The low frequency region in decompositions at different levels is the smoothed version of the input image, while the high frequency regions contain the finer detail or edge information contained in the input image. To ensure almost nearly rotational invariance, the linear combination of the four subbands (LL,LH,HH & HL) can be taken. This combination provides the sum of approximation & detail energy bands.[14] Coiflets are a particular class of wavelets which are near symmetric. Also, note that coiflet has zero wavelet and zero scaling function moments in addition to one zero moment for orthogonality. As a result, a combination of zero wavelet and zero scaling function moments used with the samples of the signal may give superior results compared to wavelets with only zero wavelet moments. [13].

**Fig 1: Sample Images of Essex Grimace Database**

# 3. WORK DONE

In our work the 2D Discrete Wavelet Transform of the input image is computed. Coiflets are used as the basis of wavelet transform. Four sets of subbands are obtained. The approximation subband and three detail subbands namely horizontal detail, vertical detail and diagonal detail subbands. First the four subbands obtained are fused. Then SIFT descriptors of the fused wavelet coefficients are computed to obtain the feature vectors. The Angle Distance similarity matching was used for comparison of the test images with the training images. Let $x=\{x_1,x_2,..x_N\}$ be the training image vector and $y=\{y_1,y_2,..y_N\}$ be the testing image vector. Then the Angle Distance similarity measure is given as :

$$d(x,y)=cos^{-1}(x.y)$$

**(a) (b) (c)**

**Fig 2: (a) Original Image (b)DWT Image(c) SIFT Features of DWT Image**

The algorithm was tested on the Essex-Grimace database which contains a sequence of 20 images (180x200), each of 18 individuals consisting of male and female subjects, taken with a fixed camera. During the sequence, the subject moves his/her head and makes grimaces which get more extreme towards the end of the sequence. Sample images of this database are shown in Fig. 1.After application of DWT the size of the fused wavelet transform coefficient obtained is 102x92.Wavelet transformed image sample have been shown in fig.2(b)

SIFT Descriptors of the wavelet transformed image is computed and sample images have been shown in fig.2(c).Each Keypoint is represented by a feature vector of size 1x128.The average number of keypoints obtained from each image is approximately 50.So average size of the feature descriptors is 50x128.

The first image of each class of the Grimace database has been used for training and the remaining images used for testing. The angle distance similarity measure has been applied on the SIFT Descriptors for matching, since for efficiency in

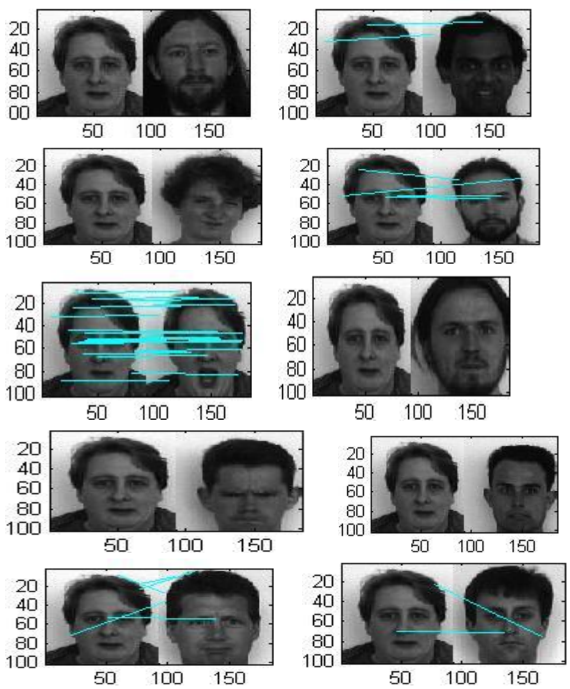

**Fig 3: SIFT Feature Matching of test image with training images**

Matlab, it is cheaper to compute dot products between unit vectors rather than Euclidean distances. Also the Inverse Cosine of dot products of unit vectors is a close approximation to the ratio of Euclidean distances for small angles. The SIFT feature matching of some sample images have been shown in Fig.3 .In the figure a test image has been compared with 10 different training images and maximum match is obtained with the correct class image with variation of expression.The results obtained have been compared with standard SIFT(feature extraction by SIFT alone) as well as standard DWT(feature extraction by DWT alone) for face recognition. The results have also been compared with standard Eigenfaces & Fisherfaces method for face recognition.

## 4. CONCLUSION

In this paper we have implemented a novel method for robust expression invariant face recognition using DWT SIFT features and the results obtained are promising. There is remarkable improvement in face recognition results when SIFT Descriptors are applied to the fused wavelet transform coefficients. The proposed algorithm has sufficient scope of improvement. Someappropriate normalization step could be included in the preprocessing stage to make the algorithm suitable for illumination invariant face recognition, since the SIFT features are partially invariant to illumination changes & the keypoint detector is not. All the images in the Essex Grimace database are frontal with minimal pose variations. Moderate pose variations of face images could be taken care of by further refining the algorithm since it is a well-known face that DWT as well as SIFT features have properties of rotation invariance. In fact investigating normalization methods and using rotation invariance properties of DWT and SIFT for modifying the algorithm are the future plans of our research.

**Table 1.Results Table depicting the Method & Recognition % Obtained**

| Method | % Recognition |
|---|---|
| Eigenfaces | 78 |
| Fisherfaces | 85 |
| DWT | 88 |
| SIFT | 90 |
| DWT+SIFT | 100 |

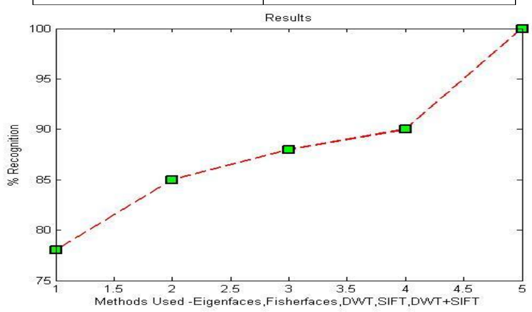


**Fig 4: Results Obtained**

## 5. REFERENCES

[1] B.Park,K.Lee and S.Lee,, 2005,Face recognition using face-arg matching. Pattern Analysis and Machine Intelligence, IEEE Transactions, 27(12):1982–1988

[2] A.M.Bronstein,M.MBoronstien and R.Kimmel,2003,Expression-invariant 3d face recognition. Proc. Audio and Video-based Biometric Person Authentication (AVBPA),Lecture Notes in Comp., (2688):62–69.

[3] L.Wiskott, and C.von der Malsburg,1995,Face Recognition by dynamic link matching,Lateral Interactions in the Cortex:Structureubs/htmlbook96,1995

[4] D.O.Gorodnichy,2005,Video based framework for face recognition in video,The 2nd Canadian Conference on Computer and Robot Vision(CRV'05),pages 330-338.

[5] U.Park,H.Chen and A.Jain,2005,3d model assisted face recognition in video,The 2nd Canadian Conference on Computer and Robot Vision(CRV'05),pages 322-329.

[6] David G.Lowe,Distinctive Image Features from Scale Invariant Keypoints,2004,International Journal of Computer Vision,60(2)

[7] H.RChennamma,LalithaRangarajan,Veerabhadrappa, 2010,Face Identification from Manipulated Facial Images using SIFT,Emerging Trends in Engineering & Technology(ICETET),2010,3rdInternational Conference,pp 192-195

[8] Cong Geng,Xudong Jiang,2009,Face Recognition using SIFT Features,Image Processing(ICIP),16th IEEE International Conference,pp 3313-3316

[9] JanezK,V.Struc,N Pave si,Adaptaion of SIFT features for Robust Face Recognition,2010,ICIAR 10,Proceedings of the 7th International Conference on Image Analysis and Recognition,PartI,Springer-Verlag Berlin Heidelberg © 2010,pp 394-404

[10] GuoshenYu,Jean M Morel,ASIFT:An Algorithm for fully Affine Invariant Comparison,Image Processing OnLine,Vol.2011,http://dx.doi.org/10.5201/ipol.2011.my-asift.

[11] LifangWu,etal,A Face Authentication Scheme based on Affine-SIFT(ASIFT) and Structural Similarity(SSIM),2012,Biometric Recognition Lecture Notes in Computer Science ,Vol 7701,pg-25-32

[12] R.C.Gonzalez and R.E.Woods ,2006,Digital image Processing Dorling Kindersley,India,Pearson Prentice Hall.

[13] C.SydneyBurrus and A. Gopinath and HaitaoGuo, 1998,Introduction to Wavelets and Wavelet Transforms, Prentice Hall, N.J 07458, USA,.

[14] S.Biswas and A.Biswas,2012,Face Recognition Algorithms Based on Transformed Shape Features,IJCSI,International Journal of Computer Science Issues,IJCSI,Vol.9,Issue.3,No.3,pg-445-451