CHAPTER 2

LITERATURE SURVEY AND PROBLEM DEFINATION

Basic purpose of literature survey is developing understanding that what has been done and what needs to be done. It demonstrates the knowledge of the research field and understanding of different methodologies used in early researches. Basic methodology used for reviewing knowledge in this thesis is described as following (figure 2.1).

Fig. 2.1: Basic methodology used for literature survey
There exist several techniques for the recognition of human emotions. Emotion Recognition methods vary according to used features for feature extraction, feature optimization process and classification techniques. For instance, Discrete Cosine Transform (DCT), Discrete Wavelet Transform, Fast Fourier Transform (FFT), Gabor Filter, Eigen Face and Principle Component Analysis are some different methods of emotion recognition which are applied to detect the functional features of face for the emotion recognition of facial expressions. Researchers are experimenting in above mentioned techniques, and their work can be divided into two parts, first the Basic Feature Extraction Techniques for FER, and second, the Optimized Feature Extraction Techniques for facial emotion recognition.

In the literature, review of various research papers in the area of facial gesture recognition is described. There are many research papers published on facial gesture recognition. Detailed literature review of some papers is explained in following sections. Different approaches for feature extraction methods used for facial gesture recognition are:

1. Transform based feature extraction
2. Gabor Filter
3. Principle Component Analysis
4. Local binary pattern.

Researchers are experimenting on various techniques which are above mentioned, the review of their work is as follows:

2.1 Literature Review of Different Feature Extraction Techniques

- Literature based on Transform feature extraction techniques

DCT feature extraction technique also extracts characteristics of expression from low frequency region. Sameer S. Kulkarni and John Moriarty [18] propose the impact of image block size on discrete cosine transform based feature extraction for facial expression recognition. The objective of the paper is to determine the principle which gives best block size for increasing the recognition accuracy with DCT. The results show that accuracy is increased with block DCT compared to without DCT.
Xiaolong Fan and Brijnesh Verma [34] suggested in their research that extracting the features about the structures of face and classification procedures is find out using neural network. The distinguished and connected features vector set is used to analyze facial images which are experimented on FERET data set. This theory has been proved better than many of the already existing theories and computes better detection rate.

The techniques of DCT and Principle component analysis are used for feature extraction process, but the question however remains that which technique gives the better result. To get an answer to this question, the comparative analysis of DCT feature extraction technique and 2DPCA feature extraction technique was done by Bin Jiang et al. [35]. The results show that DCT based feature extraction techniques have more accuracy compared to the 2D Principle component analysis feature extraction technique for facial expression recognition.

N. U. Khan [36] has proposed analysis of different feature extraction techniques for facial expression recognition and analyzed comparatively with different approaches at feature extraction and classification level. This document studies the research effort and attempts to estimate each on composite parameter basis that extends from performance to identifying facial gestures. There is an attempt to estimate the best approach that could be used that satisfies all the indicated parameters.

Fengjun Chen, Zhiliang Wang [16] used Wavelet Packet Decomposition based technique of feature extraction for facial expression analysis. On the extracted features of wavelet packet decomposition, K-L transform feature extraction is applied to optimize the features and classified with back propagation Neural Network Ensemble technique. For acquiring more shape information of expression, they have proposed facial gesture analysis using wavelet packet decomposition and neural network. This is geometric feature extraction process using wavelet packet decomposition.

Wang, Jiang and Li [37] have suggested a combined two phase DWT-DCT technique of feature extraction for face recognition. First DWT is processed on image. DWT shows the characteristics of feature official expression and other pattern into Low frequency coefficients,
so low frequency coefficients are extracted then DCT transform is applied on extracted Low frequency coefficients of DWT. The final features are passed to support vector machine classifier and the results show that the proposed technique achieved better accuracy than PCA.

If the face recognition system is robust on different facial pose then the requirement of a capture device to capture a person’s face for identification process will no longer be mandatory, as stated by Omid, Shahdi. The method of face recognition proposed in this paper will remain authentic with a very large head pose variation. The features will be extracted from local regions of the face after applying both discrete wavelength transform and discrete cosine transform. To deduce the connection between face in a given pose and its frontal view, learning strategy is applied. According to the results they obtained that their proposed method only relies on a single gallery image. When the face images are of low resolution quality then also the method gave its high performance with the recognition method. They proposed a novel approach in this paper to tackle the very complex problem in face recognition like identifying non frontal face based on frontal face gallery. First several patches were obtained from the given face and then DWT and DCT were applied on those patches. After, it converted the resulted matrix into row vectors and then it combined them into one long vector which was named face feature vector. In the learning face the relationship between non-frontal feature vector for every pose and its relevant frontal feature vector stabilized as the coefficient map. The frontal feature was estimated and its relevant match was found by Euclidean distance when the coefficient map along with the non-frontal feature vector was used in the testing period. In results it was concluded that the suggested method applied very well in comparison with other methods with poor quality images too. In future the recognition rates can further enhance by adding geometrical features. For enhancing the rates of identification further the geometrical features will be added in future [38].
• Literature based on Gabor Filter feature extraction technique

Behnam Kabirian Dehkordi [39] et al. presented a Gabor Filter Technique for geometric part of facial expression, rather than the whole space of face. For this purpose a suitable mask is designed using Gabor Filters, and then it is convolved with the original image. Main components of facial expression, such as nose, ear, mouth and eyebrows (with eye) are characterized and then defined distances and points are selected automatically on these components. This method has a higher accuracy in comparison to other methods.

Xiaoli Li, Qiuqi Ruan, Chengxiong Ruan [40] have identified the high dimension and high redundancy of Gabor Filter Feature Extraction Technique. They added that local Gabor Filter is more beneficial for less running time of feature extracting process. They have given the reasons of above logic that local Gabor filters with lower scales achieved higher power than higher scales Gabor filter. Filtering process is used to optimize the feature of Gabor and size of feature vector is reduced. Results also show that the suggested scheme provides better and fast result in terms of running time of application.

LinLin Shen and Li Bai [31] have proposed a review analysis on different Gabor optimization technique as Gabor-Kernel PCA, Gabor ICA etc. They proposed that Gabor-Kernel PCA methods are combined in order to achieve highly discriminative features for recognition. The proposed method is robust to variations in illumination changes also.

R. Prema et al. [32] had proposed a system for face recognition in which they used approaches of facial expression recognition with Gabor Filter and Principle Component Analysis. The features of Gabor are reduced by principle component analysis. They have proposed that their proposed system is better compared to individual Gabor filter technique only.

Quan-You Zhao et al. [33] had presented new Gabor optimized feature extraction technique for gesture recognition algorithm based on fusion of Gabor and LBP (local binary pattern) Features and show that Gabor + LBP optimization is better than Gabor+ PCA optimization.
According to Md. Tajmilur Rahman face recognition algorithm work using neural networks trend by Gabor features. With Gabor filter co-efficient at different orientation and scale the system will commence on convolving some different sized images of particular face. Scaling of RMS contrast and contribution of morphing are two novel contribution of this paper as an additional tool of image recognition perfection. The neural system utilized for face acknowledgment depends on the Multi Layer Perception (MLP) architecture with back-propagation calculation and fuses the convolution channel reaction of Gabor. They have justified the effectiveness of the algorithm on a transformed facial image database with image taken in different illumination conditions. A neural network based face identification system is presented using Gabor filter coefficient that can face with illumination changes. By implication of contrast equalization by using the RMS value of the image pixels the recognition performance has improved significantly and by adding the intermediate revolutionary snapshots between two different poses by utilizing image morphing in the processing step. A face from new image that neglects the variation between images of the same face should be identified by a face recognition system. A common approach to overcome the impediments in image variations according to changes in illumination conditions is to use image demonstration that is relatively insensitive with those variations. Images convolved with Gabor-like filters, Edge maps, and image intensity derivatives are some examples of such representations. A practical study that evaluates the effectiveness of the suggested method for facial images at different view angles and different illuminations is focused upon, in this paper [41].

It was suggested that a Gabor Based face representation achieved enormous success in face identification, by Md. Tajmilur Rahman and Md. Alamin Bhuiyan. A practical study focused in this paper evaluates the effectuality of the approached method for facial images at different view angles and different illuminations. The dispensation time of Gabor Transformation on a 1.2 GHz system is decreased to below 1 second with the size 100*100 pixels of image. The effectiveness of this approach would justify by them by testing this approach with more people face pictures and some common databases. No of Gabor feature for each sample is 100x100x15=150,000 in the magnitude’s pixel corresponding to a Gabor Feature. That’s why, there next step will be to upgrade the algorithm that could be able to
employ some more complicated classifier and distance majors to show Gabor faces with spatial and frequency features [42].

It was stated by Ngoc-Son Vu and Alice Caplier that the performance could be dramatically reduced of face recognition system if poses of a face are different from the face in gallery. They present a pose robust model of face recognition, also focus on how face patches change as the view point varies in appearance in this paper. They also explain a new model which based on two robust local appearance descriptor, local binary patterns (LBP), and Gabor Wavelet. For face recognition these two descriptors have been widely exploited and different strategies investigated for combining them. According to their knowledge all previous combination methods developed for frontal face identification. A local statistical framework is introduced for face recognition by them in a cross pose variation and worked only with one frontal reference image. They evaluate this method on the FERET (Face Recognition Technology) pose dataset and the result of experiment show that very high recognition rate is achieved over a large scale of pose variations available in that challenging data set. They present a new pose stable face recognition model focused on modelling and change of face patches in visual according the variations of view point. They proposed a new descriptor model instead of using pixel based appearances as they combined two local feature descriptor Local Binary Patterns (LBP) and Gabor wavelets. They also proposed the Retina filter to be used as a pre-processing technique. And the results show that there method was efficient over a large scale of pose variations obtained on the FERET pose dataset. In future work when the face area is cropped using automatic detection of the mouth and eye region then they will evaluate the effectiveness of their model. They focused the effectiveness for only pose variations but it is expected that the proposed model will be applicable for other variations like aging and expression [43].

Jun Ou, Xiao-Bo Bai, Yun Pei, Liang Ma and Wei Liu presents a system that uses 28 facial feature key-points in images detection and Gabor wavelet filter provided with 5 frequencies, 8 orientations. They have presented an automatic facial expression recognition system utilizing visual C++, which adopts Gabor wavelets to extract facial feature and a KNN classify the facial expression emotion. The proposed method is also limited: the effectiveness
of extraction expression feature is completely dependent on the effectiveness of pre-
processing of the raw image. Further work involves taking many effective measures to
improve the recognition accuracy. Such as creating a custom-built image database to conduct
extensive experimental studies, and designing effective primitive features without any pre-
processing on the facial expression images. [44]

Praseeda Lekshmi.V and Dr. M. Sasikumar suggested Gabor wavelet filter based
feature extraction for analyzing facial expression using faces and classify class labels using
Support Vector Machine. They reduce the 40 Gabor bank into less feature matrix for better
result. This method exhibits a better performance ratio and has very low computational
complexity [45].

Thomas S. Huang proposed a technique to recognize the facial expression in the video
input. Author said that basically all the facial expression techniques can be classified into two
categories namely static technique and dynamic technique. The static method recognizes the
frame based on the tracking result of that frame while dynamic technique create a Hidden
Markov Model (HMM) which then matches the frame in overall video. But Author analysed
that the static technique are more reliable and easy to train and implement. The author used
Bayesian network classifier for class identification He integrated the classifier and facial
tracking system together to build a real-time facial expression recognition system [46].

According to Jen-Tzung there are 3 basic elements in the face recognition system which
is Feature extraction, Discriminant Analysis and Classification rules. For Feature extraction
we apply the multiresolution wavelet transform technique to extract wavelet-faces. Linear
discriminant analysis is applied on wavelet-faces to reinforce discriminant power. In the
classification the nearest future plane and nearest feature space classifiers are for good
decision in wide variety of faces. And combining the discriminant wavelet-faces and nearest
future space classifier achieves best face recognition method. This resulted in the output
which says that feature extraction received higher recognition rates and low recognition costs
than eigen-face using Principal Component Analysis. And adding Linear Discriminant
Analysis enhances the performance and reduction in dimensions as well as effective
discriminability enlargement. The experiments show the effectiveness of a combined approach using discriminant wavelet-face for feature representation and Nearest Future Plane and Nearest Future Space for robust classification [47].

According to Irene Kotsia and Ioannis Pitas there are two main methods for the facial expression recognition in Image Sequences Using the Geometric Deformation Features and Support Vector. Now the main elements for facial expression are (anger, disgust, fear, happiness, sadness, and surprise) which are set of muscle movements known as Feature Action Units (FAU), which thus result in Facial Action Coding System (FACS). Now the user has to set the candied nodes to the depicted at the first frame of Image Sequence and now the grid tracking and deformation system will detect the frames in the consecutive video over time as the expression evolves. This approach is achieved higher accuracy of facial expression recognition using the proposed multiclass on Cohn–Kanade database [48].

- Literature based on Local Binary Pattern feature extraction technique

Varsha et al. [49] (Local Binary Pattern) had performed Local Binary Pattern feature extraction technique for facial expression recognition application and suggested that experimental results of local binary pattern is better than principle component analysis for the accuracy of system.

Boulbaba Ben Amor [50] has suggested that Geometric Motion Feature Extraction is better than tracking model distance based feature extraction process. Moreover, the feature extraction process, Shuyang Wang et al. [22] worked on preprocessing and image capturing from video. They have proposed a scheme of Localization with Motion capture data to generate the frame from video for application of facial expression recognition.

Muzammil Abdul rahman [51] has performed Gabor + LBP feature extraction technique for facial expression recognition and concludes that Gabor + LBP are better than Gabor + PCA in terms of accuracy. They apply pre-processing and then extracted Gabor feature and LBP features separately for generating combined texture and edge feature vector
Principal Component Analysis (PCA) is used to reduce dimensionality of the feature matrices. Results shows; GW+LBP have outperformed than individual Gabor and LBP.

Jun Wang et al. [52] has worked on Markov model feature extraction process which is based on the concept of Markov model. Using the Markov model, the accuracy of recognition system increases but the accuracy label is very less than other model of feature extraction as Gabor+ LBP.

Seyed Mehdi Lajevardi [53] has presented a new two tier feature extraction for the facial expression identification. The method proposed a new scheme of feature reduction generated through Gabor Filter for facial expression recognition. The paper is based on the concept that an average of vector set represents all the values of the set. The overall features of Gabor filter is reduced by 1/7 using the evolution of the average of corresponding Gabor coefficients. This is two phase optimization of Gabor Filter which is first optimized using average filtering then PCA. The result shows that two phase filtering (average-PCA) is better than one phase filtering (PCA only).

It was stated by Lemley [54] et al. that there are several machine learning methods for the gender classification by using face images but the most of them do not suggest any exclusive or standard solution for the problem. Some variety of benchmarks is used to access these methods. Here they used the benchmark to select and concise but no winner was found because the accuracy of the classification depends on the benchmark used. They have compared nine different machine learning methods to classify gender on two benchmarks, by using indistinguishable research methodology to allow direct comparison among efficiencies of different classifiers and feature extraction methods. They have also provided comparable results for effectiveness of the algorithm. Their main goal was to explore the gender classification with the help of learning algorithm. Several experiments on state of art gender classification were performed and then accuracy of these experiments was later compared on two datasets. They used DTCWT on a very huge scale i.e. 15000 database for ‘in the wild’. The accuracy achieved was 98% (FERET) and 96% (Audience) which was better than the
previous ones. They have successfully proved that DTCWT work better for the images which are more similar to FERET.

For face and facial gender classification based on face images a technique was proposed by Leiw [55] et al. with the use of convolution neural network (CNN). When it was applied in pattern recognition it produced less design complexity in comparison with other CNNs. By fusion of convolution and sub sampling layers, the processing layers were reduced only four. In contrast to conventional CNN, convolution operation was replaced with cross correlation thereby resulting in reduced computational load. Network was skilled with the help of second order back propagation algorithm. Two database SUMS and AT&T is used to evaluate the performance of proposed CNN, achieving 98.75% accuracy on SUM and 99.38% on AT&T.

Juan E. Tapia and Claudio A. Perez [56] have proposed fusion based feature extraction including three groups of features, three spatial scales and four different mutual information measures to select features. They have shown enhanced solution by the fusion of LBP with various spatial scales & radii, and select the features with the use of mutual information. They have used conditional mutual information maximization (CMIM) as measures of mutual information. The gender classification accuracy was boost up by selecting the features along with fusion of LBP features in compare with the previous ones. In addition to this, processing time was reduced to some extent by selection of feature which results in making the application feasible.

Caifeng Shan, Shaogang Gong and Peter W. Mc Owan labels the pixel of an image by holding a 3x3 neighbourhood of each pixel with centre value and consider the result as a binary pattern and 256-bit histogram of the LBP labels computed over a region is used as a texture descriptor. Since the performance of the boosted strong classifier originates in the characteristics of its weak hypothesis space, then evaluate other kind of weak classifier as alternative to template matching in order to achieve better classification performance [57].
• Literature based on PCA based feature extraction

More over feature extraction process, Hongbo Deng et al. [58] works on classification stage of facial emotion recognition. According to him, the Adaboost algorithm is better than back propagation classifier used for classification which is used boosting approach. It built a series of weak classifier and strong classifier suing HAAR features which minimize the generalized training errors.

Rammohan Mallipeddiet et al. [59] have suggested Principal Component Analysis (PCA) based feature extraction technique which is a subspace projection method for face identification. In this technique, feature extraction is a method for good accuracy of face recognition. Heuristic approach is used for feature selection and experimental results shows that it outperformed in terms of accuracy.

An innovative method was proposed by Hunchuan Lu and Hui Lin [60] from Dalian University of Technology China to predict gender of a person with the combination of ellipse face images, Gabor filters, Adaboost learning and SVM classifier. As face description based on HAAR like feature, Gabor feature or ICA is proved as an effective way to fetch expression of face, they compared three types of features which are selected by Adaboost method using FERET database. Firstly, they have performed comparison of various pre-processing methods which are Face detector, wrap face images and ellipse face images. In the meantime different methods (Gabor wavelets, HAAR-like wavelets, PCA, ICA) to extract features were also compared. The outcomes obtained from the test proved that this technique (combination of ellipse face images, Gabor wavelets and Ada+SVM classifier) receives enhanced performance. Secondly, experiment was performed on ICA and PCA feature extraction method. The result of this experiment proves that ICA works steadily than PCA method.
Literature based on Facial colour feature extraction technique

Yamada [61] proposed that colour information of emotions is the information for the unique identification of expression. Authors have studied how the facial colour of human beings affects emotion changes. They studied the effectiveness of facial colour image while laughing. Such emotion is based on synthesizing a dynamic facial colour and expression. Dynamic facial colour has been analyzed on the basis of their synthesized average facial colour image thus; it is representative of its effectiveness.

Various researchers have optimized the basic features of above discussed techniques by suggesting the new optimization techniques for feature extraction process; out of those techniques, I have reviewed some techniques which are given in table 2.1.
2.2 Summary

The brief summarized review of some Gabor Filter based feature extraction techniques is given below in table 2.1.

TABLE 2.1 Review Summary of Gabor Filter based feature extraction techniques

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Author</th>
<th>Solution Approaches/Methods</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sagar, Ganapathi V., Savita Y. Barker, K. B. Raja, K. Suresh Babu, and K. R. Venugopal [62]</td>
<td>DWT based feature extraction</td>
<td>Simple to eliminate noisy and unnecessary pattern in hh frequency domain.</td>
<td>Accuracy level is low.</td>
</tr>
<tr>
<td>2</td>
<td>Thamizharasi, A., and Jayasudha JS [63]</td>
<td>DCT based feature extraction</td>
<td>Better than DWT.</td>
<td>Accuracy level is low.</td>
</tr>
<tr>
<td>3</td>
<td>S. Nazari, M.S Moin [64]</td>
<td>Gabor Filter</td>
<td>Accuracy level better than DWT and DCT</td>
<td>High redundancy and huge dimension.</td>
</tr>
<tr>
<td>4</td>
<td>Jiying Wu, Gaoyun An, And Qiuqi Ruan [65].</td>
<td>Gabor filter + PCA</td>
<td>Better than simple Gabor filter</td>
<td>High redundancy and huge dimension.</td>
</tr>
</tbody>
</table>


2.3 Research Gaps

From the above literature review it is observed that:-

1. Accuracy and running time are the research issues in Facial Expression Recognition System.

2. Dependence of accuracy and running time of system is based on preprocessing, feature selection and classification.

3. Feature extraction as well as classification is responsible for accuracy of system.

4. The key dependence of running time of application is based on feature vector such that if feature vector size is reduced, then classification process will take less time.

2.4 Problem Identification

In area of pattern recognition, the problem of feature extraction using Gabor filter can be viewed as huge dimensionality and high redundancy. So for better results dimensionality and redundancy should be reduced. It refers to converting the information into a compressed set of features which represent to a unique class for the purpose of identification. Gabor filter has following problem issues [41]:

(1) High dimension

(2) High redundancy

(3) The size of Gabor feature vector is high so it takes much time in computation and classification process of expression identification.

Advantage of Feature Extraction Technique using Gabor Filter: Accuracy of Gabor filter has strong robustness against variations or changes made in illumination. Gabor filter has good characteristics of selecting information from spatial position and different orientation. Orientation in Gabor filter has a greater significance than the amplitude for acquiring shape information of expressions. So it is widely used in pattern recognition [33].
1. In the frequency domain and spatial domain, the Gabor filter technique is a good for localization.

2. Gabor Filters is better compare to DCT in the case of light variation images.

3. Angle and frequency have valuable information regarding prominent pattern and specific features of expression in digital image.

- Gabor Filters generate edge information but with redundant features in huge dimension which is reduced by feature reduction techniques. Over the years these Gabor features extraction as improved many proposals which is classified as one level feature reduction or two level feature reduction in order to improve Gabor features.

- Gabor–PCA, Gabor-Sampling, Gabor-LDA are one level optimization of Gabor filter using PCA, Sampling and LDA respectively.

- Gabor–Mean-PCA, Gabor-Mean-DCT is two level optimization of Gabor filter using Mean-PCA, Mean-DCT respectively.

- The one phase optimization provides better unique features and characteristics of expression compared to basic feature extraction process and the two tier optimization provides better unique features and characteristics of expression compared to one tier feature extraction process.

- Excessive optimization reduces the feature vector size and may increase dominant features. So we motivate to apply proposed three level optimization of Gabor filter.

- In current research the redundancy of Gabor features is reduced using 3 tier filtering techniques.
2.5 Research Objectives

After going through relevant available literature, extracting gaps in present study and problem identification it is concluded that accuracy of Gabor filter has strong robustness against variations or changes made in illumination and it has good characteristics of selecting information from spatial position and different orientations. So for handling illumination changes and different orientation issues Gabor filter is used but high dimension and huge redundancy are the problems of features generated with Gabor filter. These features need to be optimized. In the present research the redundancy of Gabor features is reduced using proposed three tier filtering techniques.

The objectives of this research are as following:

Objective 1: Analysis of different existing feature extraction approaches for facial emotion analysis system and conclude the best feature extraction technique from these.

Objective 2: Design and analysis of different three tier optimization approached which will be able to optimize recognition rate and reduce the confusion rate of system. Features are optimized by applying all the three techniques optimally in order to generate compact and concise feature.

Objective 3: Validate the proposed approaches with different datasets.

Objective 4: Validate and conclude the best three tier approach for facial gesture recognition system.
To meet the above objective, first step was to explore all relevant research papers from various available resources. Quality metrics proposed by researchers was analysed on different components and following pattern recognition methodology has been followed with optimization of feature vector (fig 2.2):

Fig. 2.2: Pattern Recognition with optimization of feature vector