Human-Computer interaction.

Artificial intelligence

Facial Expression Recognition for Multiple Faces in Less-Constraint Environment

Doctor of Philosophy (Computer Science)

By

Ms. Sadia Arshid

64-FBAS/PhdCS/F10

Supervisor

Dr. Ayyaz Hussain

Assistant Professor

DCS&SE, FBAS, IIUI

Department of Computer Science and Software Engineering
Faculty of Basic and Applied Sciences
International Islamic University, Islamabad

2017

INTERNATIONAL ISLAMIC UNIVERSITY ISLAMABAD
FACULTY OF BASIC & APPLIED SCIENCES
DEPARTMENT OF COMPUTER SCIENCE & SOFTWARE ENGINEERING

Date: ___________

Final Approval

It is certified that we have read this thesis, entitled "**Facial Expression Recognition for Multiple Faces in Less-Constraint Environment**" submitted by Sadia Arshid Registration No.64-FBAS/PHDCS/F10. It is our judgment that this thesis is of sufficient standard to warrant its acceptance by the International Islamic University Islamabad for the award of the degree of Doctor of Philosophy in Computer Science.

Committee

External Examiner:
Dr. Waseem Shahzad
Professor,
Department of Computer Science,
National University of Computer Emerging Sciences
(FAST), H-11/4, Islamabad
____

External Examiner:
Dr. Arif Jamal,
Assistant Professor,
Foundation University Institute of Engineering & Management Sciences (FUIEMS), Rawalpindi
____

Internal Examiner
Dr. Jamal Abdul Nasir,
Assistant Professor,
Department of Computer Science & Software Engineering
FBAS, IIUI
____

Supervisor
Dr. Ayyaz Hussain
Chairman,
Department of Computer Science & Software Engineering
FBAS, IIUI
____

Chairman:
Dr. Ayyaz Hussain
Chairman,
Department of Computer Science & Software Engineering
FBAS, IIUI
____

Dean:
Prof. Dr. Muhammad Arshad Zia
Dean, Faculty of Basic & Applied Sciences,
International Islamic University Islamabad
M. Arshad
____

Dissertation

This dissertation is submitted to

Department of Computer Science and Software Engineering,

Faculty of Basic and Applied Sciences,

International Islamic University, Islamabad

As a Partial Fulfillment of the Requirement for the Degree of

Doctor of Philosophy (*Computer Science)*

Publication Status:

1. Published paper

Arshid, Sadia, et al. "Multi-stage binary patterns for facial expression recognition in real world." *Cluster Computing* (2017): 1-9. (IF 1.5)

2. Publications in progress

Under Review

Illumination Invariant Robust Gradients Patterns for Expression Recognition—submitted in (AEU-INT J ELECTRON C) IF: 0.601

Declaration

I hereby declare that this thesis "**Facial Expression Recognition of Multiple Faces in Less Constraint Environment**" neither as a whole nor as a part has been copied out from any source. It is further declared that I have done this research with the accompanied report entirely based on my personal efforts, under the proficient guidance of my teachers especially my supervisor Dr. Ayyaz Hussain. I also declare that the work presented in this report has not been submitted in support of any other application or degree or qualification in any other University or Institute.

Sadia Arshid
(64-FBAS/PhdCS/F10)

Dedication

I dedicate this research to

Anum for continuous motivation and adoration

Sadia Arshid

(64-FBAS/PhdCS/F10)

Acknowledgement

Words are bound and knowledge is limited to praise ALMIGHTY ALLAH, the most Merciful who is the entire source of all knowledge and wisdom endowed to the mankind.

I would like to express my sincere appreciation to my learned supervisor, Dr. Ayyaz Hussain, for the time and supervision that he has given during my research.

I owe thanks to my loving parents, my husband Ahmad, Anum and my family for their ultimate care, support and prayers.

I am also thankful to all the faculty members of Computer Science department, especially to Mr. Asim Munir for his kind attention and encouragement.

I would like to express my thanks to all my friends and my students especially Ms. Sanneya Aziz for their reinforcement and cooperation.

May Allah bless them all (Ameen).

Sadia Arshid
64-FBAS/PhdCS/F10

Abstract

Affective computing is an active in the research from past several decades. It accentuates on the understanding and uncovering of human feelings and also exhibiting them up to some degree. This research focuses on understanding human emotions through facial expressions. In this study identification and perception of a human facial expressions are concentrated for environment with fewer requirements. Primary attention is given to real world images which are natural and vary from definition of expressions, with complex backgrounds and invariant lighting conditions and varied occlusions. To deal with real world expression having challenges of showing different expression for same emotion, and recognising expression of people belonging to different cultures, age group, ethnicity and gender a novel algorithm Multi-stage Binary Pattern (MSBP) is proposed. This is texture based technique in which gradient variance along with sign difference show promising result. As accuracy of expression recognition is highly effected due to both local and global light invariance so, both of these illumination variations are also tackled in this dissertation. Robust gradient patterns (RGP) are introduced to handle illumination issues. These patterns extract feature extraction of the basis of gradient difference thus balances light variations. Mathematical model shows RGP generates consistent patterns even image is distorted by light imbalance. Occlusion is another central point that abates execution of expression acknowledgment framework. Novelty in this work is that nine different types, size and shapes of occlusion simulations are adapted. This is initial effort to use Compound Local Binary Patterns (CLBP) for occluded data. Experiments showed that it performs best in maximum scenarios. Finally influence of facial expression in crowd behaviour recognition is also observed. Results show collective behaviour of more than one person can be determined successfully by using facial expressions using local binary patterns. Another associated factor is also investigated in this dissertation that is holistic approach works better for facial expression recognition in all conditions considered as less constraint environment i.e. expression recognition in real world, illumination variation handling, occlusion adaptation and crowd behaviour identification than zone based techniques.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AAM	Active Appearance Model
ACM	Active Contour Model
AFEW	Acted facial expression in wild
BPSO	Binary Particle swarm optimization
CK	Cohn Kanade Dataset
CLBP	Compound Local Binary Patterns
CNN	Convolutional Neural Network
DCT	Discrete Cosine Transforms
DWT	Discrete Wavelet Transforms
EBGM	Elastic Bunch Graph Matching
EMD	Empirical Mode Decomposition
EmtioW	Emotions in wild
FERET	Facial Recognition Technology
GLCM	Gabor filter and grey level occurrences
HDR	High Dynamic Range
HMM	Hidden Markov Model
HOG	Histogram of oriented Gradients
HWIN	Homomorphic Wavelet-based Illumination Normalization
ICA	Independent Component Analysis
JAFFE	Japanese Female Facial Expression
KNN	K Nearest Neighbors
KPCA	Kernel principle component Analysis
KSOM	Kohonen Self-Organizing Map
LBP	Local Binary patterns
LDA	Linear Discriminant Analysis
LGBPs	Local Gabor Binary Patterns
LGC HD	Local Gradient Coding Horizontal and Diagonal directions
LGCHVD	Local Gradient Coding Horizontal, Vertical and Diagonal directions
MSBP	Multi stage Binary Patterns
MSF	Masked Correlation filter

NDCT	Normalized Discrete Cosine Transform
PCA	Principle Component Analysis
PIF	Pose Invariant Flipping
PSO	Particle Swam Optimization
QMI	Quadratic Mutual Information
RCPR	Robust Cascaded pose Regression
RGP	Robust Gradient patterns
SFEW	Static facial expression in wild
SURF	Speed up robust features
SVM	Support Vector Machines

Chapter #1

Introduction

1. Introduction:

Affective Computing was introduced by Picard to refer to all "processing that identifies with, emerges from, or intentionally impacts feelings". According to Rosalin Picard [1], if we want computers to be truly intelligent and to communicate normally with us, we should give them the capacity to perceive, comprehend, and even to have and express feelings. Human-focused interfaces must be able to identify behavior and changes in the client's conduct, particularly his/her emotion, and to start collaborations in view of this data instead of basically reacting to the client's commands. Feelings themselves are extremely human matter, of which there is no evident hypothesis or comprehension. The fundamental foundation for understanding human emotion is the learning on feelings and their impact on human conduct and intellectual procedures.

This area of research is brimming with various difficulties and one among them is to impersonate human vision. Human vision is one out of numerous territories that need to comprehend the procedure of human intellect and duplicate that procedure with goal to supplement human existence with smart machines. For better human-PC communication it is vital for the machine to see individuals. This supported new explores with objectives to empower a PC to see individuals, remembers them and deciphers their feelings like motions, expressions and conduct. Expression recognition is a emerging field of research managing the issues with respect to feelings and machines. One approach to accomplish this objective is making computers sufficiently wise so they associate with clients in an indistinguishable path from people collaborate with each other. Hence human registering worldview recommends that UIs without bounds should be expectant and human focused, worked for people, and in view of normally happening human correspondence.

Affective computing consists of four related areas[2]. For correspondence, computers can both perceive and express feeling. Feelings can be communicated without truly having them, similar to an on-screen character assuming a part. Having feelings is a different, however exceptionally significant question. Last, computers might have passionate insight. Recently exceptionally in vogue in human brain research, it manages thinking and comprehension of feelings.

1.2 Role of Facial Expression in Emotion Recognition:

Humans naturally communicate with each other through vocal and non-vocal gestures[3]. Humans have God gifted talent to identify these gestures without any struggle or deferral but duplicating this skill in computers is still a stimulating issue because expressions are spawned by non-rigid muscle deformations which fluctuate from person to person.

Charles Darwin was one of the primary researchers to perceive that facial expression is a standout amongst the most intense and prompt means for people to impart their feelings, goals, and conclusions to each other[4]. With the progression of time the human personality has extended to such a degree that there is need to pass on this capacity to perceive and understand facial expressions to the computers.

Facial expressions assume a huge part in social and emotive lives. They are outwardly recognizable, conversational, and intuitive signs that elucidate current focus of consideration and regulate connections with the surroundings and other individuals in locality. They are straight and certainly distinguished means for imparting feelings[3]. Therefore, computerized analyzers of facial expressions seem to have characteristic place in different vision frameworks, including computerized apparatuses for behavioral research, lip perusing, bimodal discourse handling, videoconferencing, confront/visual discourse blend, affective computing, and perceptual man-machine interfaces. It is this extensive variety of guideline driving applications that has loaned a unique impulse to the examination issue of programmed of automatic facial expression analysis and d delivered a surge of enthusiasm for this exploration subject.

In the light of above discussion it is proved that facial expression assumes a critical part in smooth correspondence among people. The extraction and acknowledgment of facial expression has been the theme of different investigates subject to empower smooth communication amongst computer and their clients. Along these lines, computers later on will have the capacity to offer guidance in light of the mood of the clients. PC frameworks with this capacity have an extensive variety of utilizations in fundamental and connected research regions, including man-machine correspondence, security, law requirement, psychiatry, education, and broadcast communications. Prototype of most deliberated expressions is given in Figure 1.1.

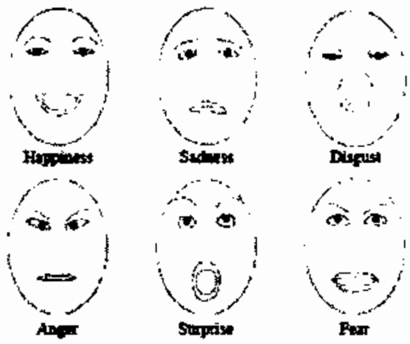


Figure 1. 1: Basic Template of Expressions [5]

1.3 Perception of Face Expressions:

In the elaboration of demeanors, the face with no appearance is said to be a neutral face. It is a casual face without constriction of facial muscles and without facial developments. It is the condition of a man's face more often than not, i.e., it is the facial appearance with no emotional demeanor. Interestingly, for a face with an expression, the facial muscles are by one means or another contracted or extended. Consequently, outward appearances are deformations of the unbiased face because of a man's psychological state. The comprehensive examples of expressions are given below:

Figure 1. 2: Facial Expression Examples from The Radboud Faces Database (Rafd) [6]

1.4 Facial Expression Analysis:

The common methodology to Automatic Facial Expression Recognition (FER) systems consists of following steps.

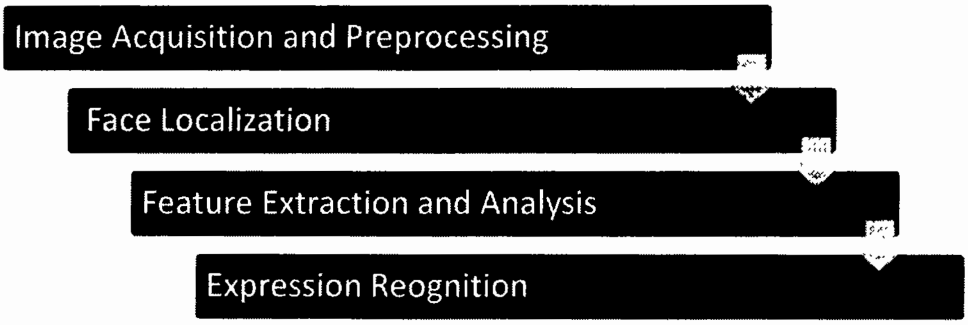


Figure 1. 3: Architecture of FER

1.4.1 Image Acquisition and Preprocessing:

In this method images are taken and preprocess for the improvement. Preprocessing phases usually contains standardization, contrast enhancement, noise removal, edge enhancement, or any measures that are needed to enhance the image. Better augmentation of the image gives better localization of face and facial features.

1.4.2 Face Localization:

As said above face detection and localization is the primary mechanism in the facial expression detection there are several techniques present for this task which are

- **Knowledge-based methods**[7]. These rule-based methods code human knowledge of what establishes a representative face. Usually, the rules apprehend the connections between facial structures.
- **Feature invariant methods**[8]. Objective of these algorithms is to discover structural features that exist even when the posture, perspective, or illumination variations, and then practice these to trace faces. The methodologies comprise facial features, skin color, textures and multiple features that is integration of color, size, and shape.
- **Template matching approaches**[9]. A number of regular patterns of a face are stowed to label the face as a whole or the facial features unconnectedly. The relationships between an input image and the stored patterns are computed for detection. It includes predefined face patterns and deformable templates.
- **Appearance-based approaches**[10]. In difference to template matching, the models are learned from a training images which must apprehend the illustrative changeability of facial form. These learned prototypes are then used for face localization. It encompasses Eigen face, Markov model, support vector machines, neural networks, distribution based, Naive based classifiers.

1.4.3 Feature Extraction and Analysis:

Expressions are mainly amalgamation of the facial features. So this deformation evidence can be taken from the features that are to be mined after face localization. The features that are significant and transformed due to any expression are location of eyebrows, size of eyes, wrinkles around the nose and locus and contours of the lips. All these features are to be extracted out either on the basis of landmark formation, Integral projections, fisher discriminant etc. Then these features are examined for the conclusion of suiTable classifier.

1.4.4 Expression Recognition:

When the features are mined they can be given either to the any appropriate classifier that either may learn and then classify expressions consequently or may match some statistical patterns and recognize expression accordingly. Broad techniques for expression recognition are Principle Component Analysis (PCA) [11] [12], neural nets [13] [14], Support Vector Machines(SVM) [15] [16], Bayesians Belief nets [17] etc.

1.5 Problem Domain and Research Motivation:

Research motivation is a step forward to make machines intelligent enough that they can understand human emotional state so that they can customize them self-according to the user. Less constraint environment means images are taken in real environment. In which we there are issues of imbalance illumination, occlusion, complex backgrounds. Further issues of gender, age and ethnicity are also handled in this thesis. Images are taken from SFEW which are closer to real world expressions. In this data set images of every age, gender and from different races are considered Main Challenges in this tasks are illumination imbalance and occlusion accommodation and recognition of real time expressions. Literature shows most of the work that has been done in illumination imbalance handling discusses global illumination issue only. Most software that are embedded in cameras or mobile phones use different filters to balance illumination which changes actual image value and highlight or suppress key points which might help in recognition. Whereas local illumination is generally either neglected or global and local illuminations are handled separately. This gives us inspiration to balance both local and global illumination simultaneously.

Another main issue is occlusion. Humans are capable to recognize emotions even some of facial points are occluded. We impart the same facility in learning algorithm by using local texture based methods. In this thesis we have done extensive experimentation to show high expression accuracy even if face is occluded from single part i.e. is single eye, lip, up to half face occlusion.

Expression recognition of multiple faces is valuable finding for asylums, customer satisfaction, prisoners, security risk in crowd. Very little work has been done in expression recognition of multiple faces. This motivated us to analyze expression of crowd.

1.6 Investigated Research Topics and Contribution:

In this research we analyze various factors that still make expression recognition challenging tasks for machine learning algorithms. Certain stimulating scenarios are developed and evaluated for efficient expression recognition. Our experiments are done on both simulated and real data. To best of our knowledge we are first one to do such an extensive work covering all issues with local texture based techniques.

1.6.1 Research Questions:

Research questions which are principally answered in this research are

1) Is it possible to recognize expressions with acceptable accuracy in real challenging dataset?
2) Can we enhance accuracy of facial expression recognition system by compensating impact of illumination variation (both global and local)?
3) Can we accommodate occlusion or pose variation in an automated system of facial expression recognition?
4) How correctly behavior of multiple faces can be identified by using facial expression recognition?

1.6.2 Handling Imbalance Illumination:

Illumination variation is still a hard problem in facial expression recognition research area. In the same person expression variation is kind of small change in facial points. When the lighting condition is changed, the same face appears variously. More specifically, the changes caused the by variation of lighting could be larger than the difference between of the appearance individuals. This causes inaccuracy of even same expressions of same individual. Existing research enforced that significant illumination changes can make dramatic changes in the projection coefficient vectors, and hence can seriously reduce the performance of the system. Fluctuation in appearance caused by illumination significantly influences the expression recognition performance and accuracy. Further very little work is earlier done in global and local illumination change where it has magnified impact in systems accuracy. In this research we developed an algorithm that generates consistent patterns for both local and global illumination changes thus improving learning capability of machines giving better recognition rate.

1.6.3 Accommodating Occlusion in FER:

Accurate localization of detailed facial features provides an important building block for identification and analysis of facial expressions. However, feature point localization tends to break down when applied to faces in real scenes where other objects in the scene (hair, sunglasses, and other people) are likely to occlude parts of the face [7]. As most of the work is done on the fontal face a little work is done on both frontal and profile view but in real time it is not necessary that end user is facing camera. So head pose variation and occluded components must be adapted for accurate feature extraction. In this research we focus on simulating different types of occlusion. These types include single eye occlusion, both eye occlusion, full eyes with eyebrows occlusion, mouth occlusion, and full mouth occlusion i.e. Mouth and side wise wrinkles occlusion, single eye and mouth occlusion, both eye and mouth occlusion, half face occlusion and finally miscellaneous in which all previously discussed occlusions are mixed. Satisfactory results are given by algorithm for all types of occlusion. Such extensive work is not done earlier.

1.6.4 Identifying Emotion of Crowd using FER:

Researchers should shift attention from the monitoring of a single person in a camera-monitored environment to that of groups and their behavior. This novel level of abstraction provides event descriptions which are semantically more meaningful, highlighting barely visible relational connections among people. Automatic Crowd Analysis is a research area which can be used for anomaly detection: panic scenarios, dangerous situations, illegal behaviors, etc. Little attention is given on the expression of the crowd based on facial expression. For doing so individual face must be extracted and then facial expression analysis is performed. Crowd expression recognition can be used to detect some malicious activity, or to identify if certain crowd is aggressive or peaceful. This research can be used to identify activity behavior of the crowd in a new dimension.

1.7 Thesis Organization:

This thesis is organized in following six following chapters.

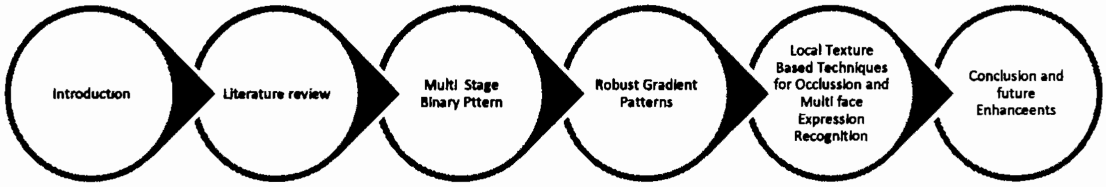


Figure 1. 4: Thesis organization

Chapter 1 gives brief introduction of research. It gives brief introduction of the domain knowledge. It give brief introduction of human emotions and how they can be perceived from facial expressions. Challenges in area of facial expression recognition are also discussed in this chapter. Chapter2 includes summary of existing work that exists literature review is divided into four sections. First section includes existing work in facial expression recognition. In second section existing work in imbalance illumination handling is discussed. Third section is about work that is conducted in occlusion handling for facial expression recognition. In fourth section it is discussed that little work is done in crowd behavior identification. Section five explains little about existing local texture based techniques. Chapter 3 addresses first module of the research that is about expression recognition in real world scenario. In real world issues of expression recognition are extensive than in lab based environment. Multistage Binary Patterns are proposed for FER. Chapter 4 discourses special issues like balancing illumination. There are two kinds of illumination changes that distort images. Global and local illuminations. Local illuminations are ignored in research that also have impact on result. So Robust Gradient Patterns (RGP) are proposed to handle both illuminations. Chapter 5 explains another major issue in FER that is occlusion. Facial features play important pat in recognizing expressions. If any part of face is missing it degrades accuracy of FER. In this chapter we introduced maximum sort of occlusion to verify their impact on FER. Chapter 5 presents novel idea of recognizing emotions of crowd by using expressions. Crowd behavior determination is important in many aspects i.e. customer satisfaction, students feedback, prisoner feedback and for security reasons. This chapter is initial effort for using expression recognition for crowd behavior identification. Chapter 6 concludes dissertation with contributions that are made in expression recognition in real world with MSBP, handling local and global illuminations with RGP, focusses on occlusions and crowd behavior identification along with results. It further explains future enhancements in this domain.

Chapter # 2

Literature Review

2. Literature Review

In this chapter detail study of existing techniques is provided. Further it is an effort to summarize published work that is done in FER. And gaps are identified which justify scope of the work presented in this thesis. Literature review of FER is divided into categories. Section 2.1 presents literature on Expression recognition. Section 2.2 claims issue in illumination imbalance and its solutions discussed to date. Section 2.3 is about work done so far on occlusion handling. Section 2.4 is literature about determining behavior of multiple people. Section 2.5 states problem statement that is evident from literature. Section 2.6 is about currently used datasets. Section 2.7 is about existing appearance based methods with which our techniques are compared.

2.1 Expression Recognition:

Many authors have contributed their efforts in automatic facial expression recognition, which is accompanied by several problems.

J.ou et.al [18] proposed Automatic detection of facial expressions is presented they have used 28 facial key points. These important feature points are further used by Gabor wavelet to locate other facial features such as forehead, eyes, eyebrows, nose etc. Gabor filter is used with 5 frequencies and 8 orientations of facial features. Features dimension is reduced using PCA and classified using K-Nearest neighbors (KNN). Cohn-Kanade (CK) database is used in this work. A new appearance based feature extraction technique local directional pattern is proposed by T. jbid et.al in [19]. This texture descriptor performs well as compare to Local Binary Patterns (LBP) and Gabor wavelet. Feature space is reduced using PCA and adaboost due to dimensionality reduction computation cost decreases and classification accuracy increases. Japanese Female Facial Expression (JAFFE) and CK database is used to perform experiments and images are classified using template matching and SVM.

L. Zhang et.al proposed a system in [20] in which face region is detected using adaboost framework. This work has two contributions first it used Active Appearance Model (AAM) for localization of the facial features in presence of large variations correctly. Secondly both appearance parameters and shape parameters are considered. They have introduced hybrid features base on AAM shape features, Appearance features, and geometric features. Further Quadratic Mutual Information (QMI) is used to select the optimum features. Cas-PAL facial expression database is used and expressions are classified using SVM.

Whereas D. Ghimire et.al in [21]worked on CK database in order to check the recognition accuracy of their system. An image or frame is taken into account for the landmark initialization and tracking and then the geometric features are extracted by applying elastic bunch graph (EBGM). Then these control points are reduced by using Multi class Adaboost, as a single feature corresponds to a weak classifier, Multiclass Adaboost method creates a subset of strong classifiers by boosting the weak ones which are useful for better classification. The classification process is done via SVMs and Dynamic time wrapping. According to this paper the best accuracy rate is achieved when the features are selected using Adaboost are classified by SVMs.

Later V.J. Mistry et.al performed a literature survey [22] on the Facial expression analysis, in which they have gone through some techniques e.g. PCA, LBP ,Independent Component Analysis(ICA) etc. and come to the conclusion that none of single technique is efficient in determining expressions. Each of them is having limitations in either recognition rate or timings. Better technique could be developed by combining at least two or more techniques for better results. Moreover, success rate can be improved by selecting good features via pre-processing.

M. Madhu et.al used multiple techniques to classify facial expression in [23]. In order to get a better accuracy Grey level co Occurrence is combined with kernel principal component analysis. Firstly, Grey level co Occurrence is used to extract features from the database images. Then sort these features and on basis of the similarity and calculation of weights of the images and then separate first 10 of the sorted features. Then on the basis of Kernel principle component Analysis (KPCA) the sorting of the first ten images will take place. After the completion of KPCA operation the features through local ternary pattern technique are extracted. Then SVM is used for the classification. For performing the experiments JAFFE dataset is used.

A.V. Subhi et.al used neural network technique in [24] i.e. the back propagation for the categorization of features. An active shape model and optical flow method are also used. The features are extracted using PCA as used by[22] [23]. Working under constrained environment can improve the accuracy.

N. Shukla proposed back propagation method for classification in [25] of features for the expression recognition, in former research paper features are extracted by using PCA while Shukla uses canny edge detector for identifying check lines, forehead mid forehead and mouth. It worked on the still images and the extraction of the desired features and emotion recognition in real time are also discussed. An automatic face detection is used that detects frontal face from the images. It also worked on features i.e. lips, eyes, then identify Bezier curves on the features, found longest binary patterns and then the classification is done on the basis of nearest neighbor. The system works well with single faces but it is not that promising for the multiple faces.

A. Saha et.al suggested contemporary approach of defining membership function of each input to some emotion [26]. And this membership function determines the emotional state of the face. They have obtained 38 facial features (Facial Action parameters). There fuzzy values with respect to the left, right and middle features are assigned to these parameters, which set by using sigmoid function. Centre of gravity method is used for defuzzification.

S.T. Roohi et.al proposed a novel fuzzy method for facial expression recognition on still images of the face in [27]. The new technique involves in extracting mathematical data from some special regions of the face including (eyes, region between eye and eye brows, lips, region between nose and lips) and fed them to a fuzzy rule-based system. Fuzzification operation uses triangular membership functions for both input and output. The Distinct feature of a system is its simplicity and high accuracy. Experimental results on JAFFEE database indicate good performance of the developed technique.

P.S. Subramanyan discussed in [28] that three issues in facial expression recognition. The first issue is related to feature selection, according to this research for better results the selection of the features should be good, the second issue is related to classification, recognition rate would increase when the data would be classified efficiently and for that an efficient machine learning algorithm should be used and experimentations should be done on diverse database. The paper further employed two more problems about scale and orientation of the images. More over the recognition rate also get affected by the uncontrolled lightening conditions. Complex background is another problem discussed, which can affect the required results. Eliminating these issues will increase the success rate.

J. Kalita in [29] cropped features as left eye, right eye, lip, nose and lip together and nose. Then these cropped images are resized in order to calculate the Eigen vector of these features and the eigenvectors and Eigen values are then stored. The input image also gets processed the same as training image then classification is done on the basis of nearest neighbors. Former work has been done by using Eigen face method and creating the Eigen space for the whole face but in this the Eigen face calculation in done by creating different Eigen space for different expressions. It performs a literature review on the performance of various techniques used in emotion recognition system, certain suggestions were also given for acquiring betterment. There must be meta-analysis of data so that different scheme should be compared. Facial expression recognition color component should be analyzed. Then algorithms should be robust to handle alignment and illumination,

The Emotion Recognition in the Wild Challenge and Workshop (EmotiW) 2013[30] expressed grand Challenge consists of an audio-video based emotion classification challenges, which mimics real world conditions. Traditionally, emotion recognition has been performed on laboratory controlled data. While undoubtedly worthwhile at the time, such lab controlled data poorly represents the environment and conditions faced in real-world situations. The goal of this Grand Challenge is to define a common platform for evaluation of emotion recognition methods in real-world conditions. The database in the 2013 challenge was the Acted Facial Expression in Wild (AFEW), which has been collected from movies showing close-to-real-world conditions. EmotiW challenge was platform for researchers to compete with their emotion recognition methods on in the wild data. The audio-visual challenge data is based on the AFEW database. The labelled Train and Val sets were shared along with unlabeled Test set. Meta-data containing information about the actor in the clip were shared with the participants. The performance of the different methods will be analyzed for insight on performance of the state-of-art emotion recognition methods on `in the wild' data.

W.L.Chao et.al in [31] facial expression recognition based on modified version of conventional LBP operator called ES-LBP is proposed. Expression specific LBP extract only the facial expression related essential features and ignores unrelated features. Further a denoising mechanism is proposed in this step only expression specific features dimensions are considered so dimensionality of feature space is reduced. JAFFE images are classified using SVM and CNN.

L. Zhang et.al in [32] gave a detailed comparison of facial expression techniques in realistic data and lab controlled data. Realistic data used in this paper is collected from World Wide Web and TV broadcast. Classification of emotions in three categories positive negative and neutral is contrary to be claimed in lab based data. Nature of data has more impact on fear and sad expressions. From the literature it is evident that most of the work has been done on the datasets that are taken under controlled environment in one way or other. Important feature set varies from minimum 4 features lips, left eye, right eye, mouth to 38 or facial points. In some systems data is manually cropped that reduces effectiveness of the system.

A. Majumder et.al in [33] discussed appearance based features are classified using Kohonen Self-Organizing Map (KSOM). LBP uniform patterns are considered instead of all non-uniform patterns. Feature space is reduced to 59 dimensions instead of 944 dimensions. LBP features vector dimensionality is further reduced using PCA so computation cost reduced. KSOM based results are compared with widely used classification techniques.

Compound local binary patterns (CLBP) are used for facial expression recognition by F. Ahmed et.al in [34] . In this technique, two-bit code is generated instead of one bit. In code, first bit represents sign difference between the neighbor and center pixel and second bit represent magnitude difference. This sixteen-bit code is then divided in to two halves to reduce dimensionality. Histograms are then generated from extracted codes. SVM is used as classifier to classify six expressions. JAFFE and CK are used dataset. And result accuracy is 90.4%.

Local Gradient Coding (LGC) is a different version of LBP, used for facial expression recognition investigated by Y. Tong et.al in [35] Instead of using sign difference between central and neighboring pixel, it explores grey level relationship among pixels in horizontal and vertical directions. Then it is further improved by using horizontal and diagonal relationship. On the basis of Euclidean distance histograms are matched and it classifies expressions. It gives 90% results on JAFFE data set.

V. Ramachandran et.al in [36] principle component analysis is used for feature extraction. Particle swarm optimization techniques is then applied on extracted features for advance feature selection. The training is done by emotional back propagation neural network. Used dataset for experiment is CK. Generated results with feature selection gained more recognition rate.

M.H. Siddiqi et.al in [37] proposed a normalized technique based on mutual information for selection of features. The components of this mutual information are max redundancy and max relevance. In order to eliminate any dominance factor of the two components, the mutual information is normalized. Curvelet transform is the feature extraction method used. Linear Discriminant Analysis (LDA) is chosen for the reduction of feature vector space. For classification of the expressions Hidden Markov Model (HMM) is employed JAFFE, Yale B and CK are the datasets utilized to experiment.

H. Qayyum et.al discussed frequency domain technique in [38] for expression recognition. Features are extracted by stationary wavelet transform and combination of both horizontal and vertical sub-bands is pass on to the classifier. To reduce feature dimensionality Discrete Cosine Transforms (DCT) is used. Neural network is the selected classifier and training is done by back propagation. Experiments are performed on JAFEE and CK datasets and a local dataset is also recorded named MS-Kinect. A maximum of 98.83% accuracy is achieved with JAFFEE dataset.

A.T. Lopes et.al in [39][40] proposed a Convolutional Neural Network to recognize facial expression. As neural networks work best, when trained with big data so to overcome this problem certain pre-processing methods are adopted. Through this pre-processing step only expression specific feature image is extracted. It also involved rotation correction. Sampling is done between layers of CNN training. CK+, JAFFE and BU-3DFE are the chosen datasets. Result showed that with normalization combinations better accuracy is attained.

Table 2. 1: Summary of Expression Recognition Literature

Sr#	Technique	Paper & Methodology	Results	Comments
1	Gabor wavelet PCA KNN	Automatic Facial Expression Recognition Using Gabor Filter And Expression Analysis [18] 28 geometric facial key-points PCA for dimension reduction. KNN for classification	Ck dataset is used for experiments. System gives 80% accuracy	Efficiency of Feature extraction depends on preprocessed images. Effective measures are needed to increase recognition rate
2	LDP AdaBoost	Robust facial expression recognition based on local directional pattern [19] Feature extraction is done by LDP. Adaboost for feature reduction. SVM classifier.	CK and JAFFE dataset is used 82.5% accuracy for expression recognition	Insufficient directional information
3	AdaBoost AAM QMI	Automatic Facial Expression Recognition Based on Hybrid Feature [20] Face detection by adaboost. AAM for hybrid features. QMI is then used for optimum feature selection. SVM classifier is used.	CAS-PEAL dataset for five expressions. 87.33% accuracy	Use only frontal view images. Number of expressions used are five

4	EBGM, Multiclass Adaboost	Geometric Feature based Facial expression Recognition in Image Sequences Using Multiclass Adaboost and support vector machines [21] EBGM for landmark localization. Multiclass Adaboost for Feature Enhancement. Using SVM boosted Feature gives better result as compared to DTW (Dynamic Time Wrapping)	Ck dataset is used 95% accuracy is claimed	Facial land marks localization is challenging task
5	KPCA (Kernel Principal Component Analysis), LTP (Local Ternary Pattern)	A combinatorial approach to human face recognition and expression identification [23] Combination of grey level co-occurrence matrix and KPCA in preprocessing. LTP for feature extraction. SVM for classification.	JAFFE dataset is used. 99% accuracy	
6	PCA & Neural network	The Facial expression detection from human facial image by using neural network [24] Features are extracted from full face image by using PCA. Expressions are then recognized by neural network.	JAFFE dataset is used.	Its test and training database is same. It works for controlled environment

7	Neural networks	Using Back Propagation Recognition of Facial Expression [25] Forehead & Mid forehead wrinkles. Cheek Wrinkle, Mouth length, are used facial features for expression recognition.	Manual data 90% accuracy	Neural networks require large training data.
8	SIFT, FAP	Facial expression recognition experiments with data from television broadcasts and the World Wide Web [32] SIFT and FAP for global feature extraction and SVM for classification	Data set from TV drama.. TV news and web and NVIE and FEEDTUM datasets Accuracy for NVIE is 83%, for FeedTUM is 63% And form realistic data 26%	Low accuracy for realistic data. This is because light variations and pose variations are not handled.
9	Uniform LBP PCA	Local Binary Pattern based Facial Expression Recognition using self-organized Map [33] Only uniform LBP are used to reduce feature vector length. KSOM is used as classifier.	MMI dataset is used. 69.81% accuracy	Uniform LBP reduced performance as it skips some important information.
10	Compound LBP and SVM	Person-Independent Facial Expression Recognition Based on Compound Local Binary Pattern (CLBP) [34] LBP is modified from one bit code to two bit code. One bit is to encode sign difference and second bit for magnitude difference of neighboring pixel.	CK and JAFEE dataset 90.4% accuracy	Feature vector length is increased.

11	LGC and Euclidean distance	Facial expression recognition algorithm using LGC based on horizontal and diagonal prior principle [35] This approach exploits relationship neighboring pixels in horizontal and diagonal directions.	JAFFE dataset is used 88.89% accuracy	
12	Particle swarm optimization, PCA Back propagation neural network	Facial expression recognition with enhanced feature extraction using PSO & EBPNN [36] Firstly features are extracted by PCA, then selection is done by Particle swarm optimization and neural network is used for classification.	Ck dataset is used 88% accuracy	Training and testing sample size scan be increased to improve results
13	Mutual information (max redundancy & max relevance), Curvelet transform, LDA. HMM	Human facial expression recognition using curvelet feature extraction and normalized mutual information feature selection [37] Curvelet transform is used for feature extraction. Proposed technique then selects features. LDA for feature reduction. HMM for classification.	JAFFE, CK, Yale B 99% accuracy	Not test for real time data.
14	Stationary wavelet transform, DCT	Facial Expression Recognition Using	JAFFE, CK+, MS-Kinetic(local dataset) 98%, 96% and 94% respectively	Dimensions of DCT wavelet transforms are increased.

	& Neural network	Stationary Wavelet Transform Features [38] Worked in frequency domain. Combined both horizontal and vertical sub bands. Generated a local dataset MS Kinect. Neural network as classifier		Neural networks need more data for training.
15	Convolution Neural Network	Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order [39] Pre-processing in done by applying rotation correction and normalization. Only expression specific feature image is fed to CNN	CK, JAFEE, 3BUDEF 86%	Eye detection necessary to perform pre-processing step. All images are attained in controlled environment. Only frontal face images are used for experiments.

From the above literature it is concluded that existing techniques of FER are tested in lab controlled environment. In controlled environment there is balanced illumination and expressions are exhibited according to definition of Ekman [77]. But if expressions have to be seen in real environment, situation is quite different. Even expression of happiness varies among individuals. Similarly all expressions have variation among persons and more complex scenario same person can express expressions differently in different situations. In literature [34-39] neural nets are used for training which can perform well only if there is large training data. For smaller training set it shows degraded results. Geometric techniques [24, 25, 32] are discussed which need exact localization of features thus more complex and more prone to light variations and occlusion. Appearance based techniques[18,19,33,34,35] seems to be efficient and less sensitive to noise but still these techniques are not tested against real world dataset. Thus there is need to propose a technique that can recognize real time expressions. Major issues that affect performance of FER in real environment are illumination and occlusions thus these should also be gripped.

2.3 Literature on handling illumination Imbalance:

A lot of work has been done in balancing illumination for various applications. But most the work has been done on global illumination issue. Some researchers have been discussed in this section.

M. emadi et.al in [41] used Local binary pattern to extract static facial features. LBP has stronger discriminant power to illumination changes it is used to identify facial expressions from low resolution images. Further boosted LBP is proposed to extract the most discriminant features with adaboost. Boosted LBP provide best recognition accuracy using SVM classifier this technique gives 89% accuracy for sadness and some of the expressions are classified incorrectly. 2D wavelet is used by V.Vidya et.al and Y.luo et.al to resolve illumination problem [42] [43]. For this, they used Yale database B and XM2VTS human face databases. This proposed system is very effective for normalizing shadows and can easily be implemented in real time face recognition systems but this process is not very efficient and is time consuming.

V. Vidya et. al in [43] Proposed a selective illumination enhancement technique to resolve illumination issue. The technique uses a correction factor to selectively illuminate the dark regions. For improving the performance of the system, Threshold based Discrete Wavelet Transform feature extraction is proposed by V. Vidya and X. Yuan in [42] [43]. For finding the feature vector space for the optimal feature subset, a Binary Particle Swarm Optimization (BPSO)-based feature selection algorithm is used. Proposed system gives good results on illumination variant databases like Extended Yale B and Color FERET.

A Homomorphic Wavelet-based Illumination Normalization (HWIN) by X. Yuan et.al [43] method which focuses on one of the problem of expression recognition i.e. the uneven illumination which is caused by varying lightening conditions. For this purpose, image analysis is done using wavelet transform, after normalization detail components are enhanced for further processing. For reducing the noise caused by the variance in the lightening effects of different images, a Difference of Gaussian filter is used. In this paper they worked on Yale B and Extended Yale B datasets for finding the recognition rate. With this proposed preprocessing, and 2DPCA and Euclidean distance recognition more accurate results are achieved. As compared to other techniques it gives 10% more accurate results.

S.T. Roohi et.al in [45] Proposed an illumination normalization method for face recognition. An illumination component is replicated by mean estimation method. For the elimination of illumination component, mean estimation is subtracted from the original image. Ratio of quotient image and modulus mean value is used to enhance the face texture features. For evaluation Yale B, Extended Yale B, CMU-PIE, CAS-PEARL-R1 etc. databases are used and for illumination testing PCA and LDA are used as they are highly sensitive to light changes.

Y. Liu et.al in [46] Proposed better face recognition method that uses a combination of two techniques: Uniform Morphological Correction (which uses morphological opening which involves erosion followed by dilation) and Pose Invariant Flipping (PIF). Discrete Wavelet Transforms (DWT) is applied in which detailed component is enhanced and approximation component is suppressed. BPSO (Binary Particle Swarm Optimization) algorithm for feature selection it gives good results for Color FERET, Extended Yale B, Pointing Head Pose and CMU PIE databases.

C. Jaya Mohan et.al proposed in [47] methods deals with illumination, pose variations and occlusion issues. It localizes facial landmarks, extracts facial components, preprocessing, estimate face pose, extracts features using LBP Histograms, fusion of pose adaptive classification. Preprocessing involves illumination normalization. For removal of darker regions and shadows gamma correction and Gaussian filtering is used respectively. Equalization of normalization is done to rescale the image intensity. This method gives more accurate results as compared to other holistic methods.

LBP is another technique that is proposed a 2-D expression invariant face recognition system which uses two types of optical flow techniques and image synthesis is discovered by S.Lee et.al in [48]. For each candidate only one neutral image is used. This system is computationally costly as compared to the previous systems. It also presented other methods such as robust preprocessing and an extension of the LBP local texture descriptor etc. for illumination variation problem.

J. Y.Zuhu et.al [49] Proposed a High Dynamic Range (HDR) schemed images, representing cruel lightening conditions. In preprocessing different HDR tone mapping methods i.e. exponential (exp-tmo), logarithmic (log-tmo), and Mantiuk's method (Manttmo) are compared against Normalized Discrete Cosine Transform (NDCT). Results showed that HDR tone mapping methods given enhanced results. SVM is used for expression classification. Features extraction is done by LBP as well as with SURF (speed up robust feature) and their results are compared. For LBP face image is divided into 12 parts whereas, for SURF 8×8 grid of face image with varying block width is taken. Results show that SURF with SVM generated better recognition accuracy. JAFFEE and Static facial Expression in wild (SFEW) datasets are also used in the experimentation. Texture features as well as small-scale features such as eyes, eyebrows, mouth etc. are of great importance with recognizing face or facial expressions. J. Y. zuho et.al in [50] Proposed a method for illumination compensation in face recognition, incorporating both types of features. To solve the problems of previous morphological techniques for handling illumination effect, a generalized dynamic morphological quotient image (GDMQI) method based on retinex theory is projected. Texture enhancement is achieved by histogram equalization. Results showed that combination of GDMQI and histogram equalization generated better accuracy. CMU PIE, Yale-B, Extended Yale-B and AR databases are hired foe experiments.

V. Vidya et.al in [51] Proposed a new technique to handle single image based illumination in face recognition, Logarithm Gradient Histogram (LGH) which incorporates spectral wavelength along with magnitude and direction. To eliminate lightening effect LMSN-LoG filter is used to extract components named, logarithm gradient orientation (LGO) and logarithm gradient magnitude (LGM). Post processing stage then generate the histograms of LGO & LGM, and integrate them. Datasets used for experimentation are Extended YaleB, CMU-PIE, HFB and FRGC.

Table 2. 2: Summary of Illumination Imbalance literature review

Sr#	Technique	Paper & Methodology	Results	Comments
1	LBP, Adaboost and SVM	Facial expression recognition based on Local Binary Patterns: A comprehensive Study [41] LBP features are extracted then boosted by adaboost. SVM is used for classification.	CK dataset 88.9% accuracy	Handled limited lightening variation.
2	2D Wavelets & Low frequency part->0	Illumination Normalization using 2D Wavelet [42]	Yale database 90% accuracy	Effective to compensate shadows

		Multi resolution method is used for illumination invariant. Only single image of individual is required for training.		
3	Correction factor to selectively illuminate the dark regions & Binary Particle Swarm Optimization (BPSO)-based feature selection algorithm	Face recognition using threshold based DWT feature extraction and selective illumination enhancement technique[44] Dark side and light side images separated and removal of noise. Illuminate dark regions by using energy function	Yale database B 90%	promising technique under arbitrary variations in illumination and expressions
4	HWIN, Wavelet transform & Difference of Gaussian filter.	Illumination Normalization Based on Homomorphic Wavelet Filtering for Face Recognition[43] Wavelet transform for image analysis, then normalization. Difference of Gaussian filter for noise reduction. 2DPCA for extraction. Euclidean distance for recognition.	Yale database B 78%	

5	Mean estimation method. PCA & LBP. SVM	Facial expression recognition based on fusion feature of PCA and LBP with SVM [45] Mean estimate is subtracted from original image to elimination illumination. PCA, LBP for extracting features. SVM for classification.	Live images. 85%accuracy
6	Unique combination of UMC and PIF DWT & BPSO	Feature Accentuation using Uniform Morphological Correction as Pre-processing Technique for DWT based Face Recognition [46] Morphological opening, then background subtraction. DWT is applied in which detailed component is enhanced and approximation component is suppressed. BPSO (Binary Particle Swarm Optimization)	FERET, CMU PIE, Yale B 67%, 23%, 97%

		algorithm for feature selection		
7	Retinography theory. V component of HSV & Sobel operator to find gradient image.	Illumination Normalization Using Weighted Gradient Integral Images[47] Gradient information has been used for normalizing appearance of image under varying lightning condition. Integral image enhanced	Manual images, Quantitative results	Visually good results for cannot be used for further processing. No standard dataset
9	Optical flow computation. Image synthesizing. Variation of LBP i.e. Local ternary pattern	Face recognition under expressions and lighting variations using artificial intelligence and image synthesizing [48] Optical flow & image synthesizing for handling illumination. LTP texture descriptor is used.	Masked image is used.face recognition system	Facial points are labeled manually. High computation cost.
10	LBP & SURF. SVM. NDCT, exp-tmo,log-tmo & Manttmo	Towards HDR Based Facial Expression Recognition under Complex Lighting [49] HDR dataset is created. Preprocessing by NDCT, exp-tmo, log-tmo & Manttmo and	JAFFE,SFEW, HDR lab generated dataset 60% accuracy	Comparatively small dataset.

		there results are compared. Feature extraction by LBP and SURF. SVM for classification.		
11	GDMQI & histogram equalization	Multiscale morphology based illumination normalization with enhance d local textures for face recognition [50] GDMQI is used for illumination compensation. Histogram equalization is then applied to enhance texture information.	CMU-PIE, AR and Yale database B 99%, 98% 99% Accuracy	
12	LGH & LMSN-LoG filter	Illumination Invariant Single Face Image Recognition under Heterogeneous Lighting Condition. Pattern Recognition [51] Two components extracted are logarithm gradient orientation (LGO) and logarithm gradient magnitude (LGM).	CMU-Pie, Yale database B 98% 97%	Worked on synthetic data.

The histograms of LGO
& LGM are then
integrated.

It is clear from the literature, illumination is generally handled in image recognition or face recognition application [42, 43, 44, 50, 51]. Facial expression recognition itself is challenging task from face recognition. Light variation makes this task more challenging. [41, 45 and 49] are paper dealing with facial expression recognition. Generally global illumination is handled and little attention is given to local illuminations. Local illuminations generally occur due to shadow, hair, accessories, side line wrinkles, and edges. These local illuminations affect performance of FER as well. There is need to develop an algorithm that can successfully handle local and global illuminations. In this thesis, this study is also under consideration.

2.3 Literature on Occlusion Handling:

Occlusion is always an issue for computer vision applications. Work done on occlusion handling in face recognition and facial expression recognition is conversed in this section.

R. Azmi et.al in [52]Proposed a method that resolves pose and illumination variation issues. Pose estimation method is for 2D images that classify face pose. Then from classification rules that image is classified to pose class in a low dimensional subspace by using feature extraction method. Proposed system also uses shadow compensation method for illumination variation issue. This system gives better face recognition results.

Local Gabor binary patterns (LGBPS) are used for facial expression recognition in presence of occlusion by J. Shermina et.al in [53] . Features extraction is done by combination of Gabor and LBP which enhances spatial histogram. KNN is used as classifier. It gives 89% results if lower face is occluded and 90% results if upper face is occluded.

X. P. Burgos et.al in [54] proposed system for partial occlusion, illumination and pose variations. It identifies similar blocks in the face image and then these blocks are used in removing occlusion through block matching technique. By extracting Empirical Mode Decomposition (EMD) features expressions are detected, ANN is combined to form a better recognition technique and PCA is used for face recognition.

J. Yan proposed in [55] a Robust Cascaded Pose Regression (RCPR) method which detects occlusions explicitly and uses robust shape-indexed features. System gives improved results on LFPW, LFW and HELEN datasets. It reduces failure cases by half and it detects face occlusions with an 80/40% precision. For detection of face and localization of key points this paper.

P. Vandana et.al explored in [56] a method on cascade regression framework to overcome the problems in face detection due to pose variations. Large pose validation is handled by using HOG feature. To handle large face pose variations they proposed to generate multiple hypotheses and to rank them to get final results. These methods can be learned in a structural SVM framework. This method gives good performance on LFPW dataset while gives poor results on 300-W dataset.

G. Ghiasi et.al in [57] Proposed a system for resolving occlusion based problems. They use Iterative Closest Point (ICP) technique for surface registration of face image and Principal Component Analysis (PCA) to deal with missing or incomplete data due to occlusion.

R. Min at. al explored in [58] a deformable part model that explicitly models the occlusion of parts. The model structure allowed to increase the training set of data having large number of instances which are synthetically occluded. The model is tested on different sets of data having occlusion.

Local Gabor binary patterns (LGBPHS) are used to handle occlusion in face recognition. Face is divided into parts depending on nature of occlusion by L. Zhang et.al in [59]. In this paper, only two types of occlusions are handled sunglasses and scarf so face is divided into two zones. Gabor filters are used to extract features. SVM is used to detect occlusion and selective LBP is used for face recognition.AR dataset is used for experimentation and it gives 92% correct results for face recognition. It does not handle expressions.

Gabor filters are used to extract features on multi resolution and multi scale by A. Vyas et.al in [60]. Monte Corla algorithm is used to extract set of features from templates being generated via Gabor filters. These features are extracted from the partially occluded faces. All these feature vectors are compared to find distance template matrix which includes most similar features from all the templates. Support vector machine is then used to determine most related distance template matrix. Similarity of his template is measure against test image template by using Support Vector Machine to classify expression in one of seven classes. This approach can achieve 75% accuracy result on JAFFE and 90% accuracy for CK. Issue in Gabor filter is its multi scale and multi resolution filter, generating high dimensional feature vector.

LBP can be used to extract most discriminative features as it works on local characteristics of the image is proposed by V. P. Vishwanath et.al in [61]. Feature vector is the histogram that is concatenated after whole execution. These histograms can be compared with test image template by using chi square distance matrix. JAFFE data set was used. From the confusion matrix it is evident that this technique cannot perform well in case of happy, neutral, fear and sad expressions. They have also suggested that mouth plays most important role, as by occluding mouth recognition rates are highly affected.

Active Contour Model (ACM) is used for pedestrian detection by using Histograms of oriented Gradient (HOG) and LBP by R. Li et.al in [62]. Feature vector is generated by combining HOG and LBP features. SVM is used for training. Issue in HOG is once again high dimensionality of feature vector length.

Gabor filter and grey level occurrences (GLCM) have been used for facial expression recognition under partial occlusion by E. J. He et.al in [63]. In this approach, Gabor filter bank is applied on the image for feature extraction, GLCM is used to identify association gap among the pixels. Gaussian normalization is applied at the end to generate feature vector.

Masked correlation filters(MSFs) have been used for partially occluded face recognition by P. Ramalingam et.al in [64]In this technique masked correlation filters are designed having prior knowledge about type of occlusions. Test image are compared with most relevant occluded training image. Drawback of this approach is there must be prior knowledge of occlusion type, which is not possible in real scenarios.

Fuzzy based rule learning and HMM have been used for expression variant face recognition P. Moallem et.al in [65]. In this paper PCA is used for face detection. Haar classifier are used for feature extraction. HMM is used for learning and maximum likelihood is used by fuzzy logic to determine either person is authorized or unauthorized. But major disadvantage of fuzzy rule based system is these are not robust or required large dataset for training and expert opinion to train the system.

Pyramid histograms of oriented gradients (PHOG) and KNN is used for facial expression recognition against occlusion P. ji et.al in [66]. In this paper face is detected by using voila jones algorithm. Four levels of pyramids of HOG are calculated resulting in four bins of histograms. KNN with Euclidean distance is used for classification.

Skin color, HOG and LBP are used for face occlusion detection by W.J. Li in [67]. In this paper four step strategy is used to detect face occlusions. First of all foreground is extracted. Head is then estimated by HOG. Features are extracted by using LBP. SVM is trained on these features. Skin color features and SVM results are merged to detect occluded face.

Table 2. 3: Summary of Literature Review of Occlusion

Sr#	Technique	Paper & Methodology	Results	Comments
1	Gabor LBP and KNN	Facial Expression Recognition in the Presence of Occlusion Using Local Gabor Binary Patterns [53] Images are convolved with Gabor filter. LBP is then calculated from GMPs and LGBPS histograms are concatenated.	JAFFE Dataset, 89% accuracy	Use limited range of occlusion. High dimensions of feature vector.
2	Block matching technique, EMD (Empirical Mode Decomposition), PCA & ANN	Face Recognition System with various Expression and Occlusion based on a Novel Block Matching Algorithm and PCA [54] Block matching technique for occlusion. EMD features are then extracted for detecting expression. PCA for face recognition.	Yale Database, JAFFEE 96%, 95%	Used for face Recognition, controlled occlusion is handled & lower value of FAR and FRR error rate in both the databases
3	RCPR (Robust Cascaded Pose Regression) which is enhanced version of (CPR)	Robust face landmark estimation under occlusion [55] Reduces exposure to outliers by detecting	LFPW, HELEN and LFW	40% precision

	Cascaded Pose Regression	occlusions explicitly and using robust shape-indexed features. introduction of a challenging face landmark dataset i.e. Caltech Occluded Faces in the Wild (COFW)		
4	Iterative Closest Point (ICP) & Principal Component Analysis (PCA)	Face Recognition in Presence of Occlusion Using Machine Learning Classifier [57] ICP technique for surface registration of face image and PCA to deal with missing or incomplete data due to occlusion	Live Face	
5	hierarchical deformable part model & Star topology	Occlusion coherence: Localizing occluded faces with a hierarchical deformable part model [58] a hierarchical deformable part model for face detection then key point localization that explicitly models occlusions of parts and Star topology	LPFW, COFW	Model implicitly represents the pattern of Part occlusions, it does not integrate local image evidence For the occlude itself.

		to identify missing components		
6	Local Gabor binary patterns	Efficient detection of occlusion prior to robust face recognition[59] Gabor filter bank is used for feature extraction and SVM for classification.	AR dataset with eye and mouth occlusion only 92% accuracy	High dimensional feature vector.
7	Gabor filter and Monte Corla Algorithm	Random Gabor based templates for facial expression recognition in images with facial occlusion [60] Monte Corla Algorithm extract features on Gabor templates. SVM is used for classification	CK, JAFFE 80%	Feature vector length is more.
8	Uniform LBP and Template Matching	Effect of Different Occlusion on Facial Expressions Recognition[61] Face is divided in to subparts such as forehead, eyes, nose & mouth. Uniform LBP is calculated for subparts and their histogram is concatenated. Template matching is used	CK dataset is used, with mouth and eye occlusion 97% accuracy	Use limited range of occlusion type.

9	ACM	ACM Based ROI Extraction for Pedestrian Detection with Partial Occlusion Handling, [62] HOG and LBP are used to generate ACM	INRIA dataset 96%	Pedestrian detection High dimensional feature vector.
10	Gabor filter and GLCM	Facial Expression Recognition under Partial Occlusion Based on Gabor Filter and Gray-Level Co-occurrence Matrix, [63] Gabor filters are used for feature extraction and GLCM to identify occlusion.	JAFFE dataset with eye and mouth occlusion. 83% accuracy with eye occlusion, 87% accuracy with mouth occlusion	High dimensional feature vector.
11	MCfs	Masked correlation filters for partially occluded face recognition [64] Masked correlation features of trained occlusion images are calculated against test image.	CMU-PIE, AR database	Prior knowledge is important.
12	Fuzzy rule based knowledge and HMM	Recognizing Expression Variant and Occluded Face Images Based on Nested HMM and Fuzzy Rule Based Approach [65]	JAFFE, AR dataset 97% accuracy	Fuzzy is not robust. Only glasses are used as occlusion.

		PCA is used for face detection and Haar classifiers are used for feature extraction. HMM is used for learning and fuzzy rules for classification.		
13	PHOG and KNN	Optimized PHOG and KNN for Robust Frontal Facial Expression Recognition against Occlusion [66] PHOG is used for feature extraction and KNN for classification.	JAFEE,CK 97.73% accuracy	HOG results in larger feature vector
14	Skin color, HOG and LBP	Face occlusion detection using skin color ratio and LBP features for intelligent video surveillance systems[67], HOG is used for head estimation. LBP features are used to train SVM which is merge with skin color features to detect occlusion	95%	Larger feature vector length

Literature Review of occlusion handling reveals that there is very little work done in the occlusion handling in FER. Major work in occlusion handling in done in Face Recognition [53, 54, 55, 58, 59, 60, 64]. In case of facial expression recognition mostly eyes and mouth is occluded. There is need to determine result of occlusion handling with different type and size of occlusion this issue is also tackled in this thesis.

2.4 Literature in Crowd Behavior Identification through FER:

Researchers have shifted their attention from the monitoring of a single person in a camera-monitored environment to that of groups and their behavior. for determining crowd behavior more work is done anomaly detection by S. Mohammad in [68] and movements in crowd by J. M. Grant et.al in [69], A. Das et.al in [70] Little attention is on the expression of the crowd on the basis of facial expression. In this literature we are going to discuss importance of identifying crowd behavior by A. Patwardhan et.al and A. Das in [71] [72] respectively , detection of faces in crowd and then facial expression recognition for single person that need to be extended to crowd.

A. Das et.al in [72]edge based detection method is used for crowd behavior identification. Edge detection is used with grid lines. Canny edge detector is used to obtain edges. Movement from reference point was tracked across the frame sequences. Accuracy of group emotion detection is 70.9%.

D. Y. Chen et.al in [73] visual saliency is discussed i.e. how one face get attention of rest of people in crowd. Some faces are dominant on the basis of some features like color, texture etc. Major parameters determined in this paper are intensity difference, spatial difference, area of each face and camera distance. Voila jones algorithm is used for face detection. Then each face is cropped and copied to avoid background impact. Images are converted to gray scale and its distances are measured. Viewer's point of view is added to observe saliency of face.

Background subtraction by B. Solemaz et.al [74] is used to find the position of an isolated region that comprises an individual person or a set of occluded persons is detected. Each isolated region is considered a vertex and a human crowd is thus modeled by a graph. To construct a graph, Delaunay triangulation is used to systematically connect vertices and therefore the problem of event detection of human crowds is formulated as measuring the topology variation of consecutive graphs in temporal order. To effectively model the topology variations, local characteristics, such as triangle deformations and eigenvalue-based sub graph analysis, and global features, such as moments, are used and are finally combined as an indicator to detect if any anomalies of human crowd is present in the scene. Experimental

results obtained by using extensive dataset show that our system is effective in detecting anomalous events for uncontrolled environment of surveillance videos.
Linear dynamical systems and the Jacobin matrix defined by the optical flow, are used to study crowd flow in videos by J. Foytik et.al in [75]. Time integration of the dynamical system provides particle trajectories that represent the motion in the scene; these trajectories are used to locate regions of interest in the scene. The eigenvalues are only considered in the regions of interest, consistent with the linear approximation and the implicated behaviors. The algorithm is repeated over sequential clips of a video in order to record changes in eigenvalues, which may imply changes in behavior. The method was tested on over 60 crowd and traffic videos.

Kalman filter is used for creating a system for multiple face recognition and tracking by P. Ekman in [76] . Low Level system is formed using a face database of twenty distinct people trained using Modular Principal Component Analysis (MPCA), which is used to discriminate between multiple faces. Classification is done by using a feature correlation metric. After tracking the faces, they are then evaluated by a high-level face recognition subspace which is formed using a huge database of people and processed using Adaptive MPCAs.

Table 2. 4: Summary of literature review in Crowd Behavior

Sr #	Technique	Paper & Methodology	Results	Comment
1	Canny edge Detector	Edge Based Grid Super-Imposition for Crowd Emotion Recognition [72]. Optimized grid is superimposed on each frame and intersection points are matched with edges detected	Manual data set prepared.	Movement of crowd is identified in matches.

2	Difference based Approach	A Novel Approach to Attend Faces in the Crowd through Relative Visual Saliency[73] Intensity difference, spatial difference, distance from camera, and area of each face	Subset of FEI dataset.	This approach worked for face salience in crowd
3	Graph Modeling	Visual based human crowd behavior analysis based on graph modeling and matching[74] Delaunay triangulation is used to systematically connect vertices	Generic videos	Facial expressions are not used for crowd behavior identification
4	Kalman Filter and Modular PCA for tracking and recognition of multiple faces	Tracking and Recognizing Multiple Faces Using Kalman Filter and Modular PCA[76] Tracker continuously locate faces in the input video	Generic dataset	Facial expressions are not used for crowd behavior identification

Crowd behavior identification is important are of research in which interaction sof individual sin crowd are observed. Little attention is given to identify behavior of crowd on the basis of expression. In this thesis, this is also validated that either crowd behavior can be identified by determining expression of crowd or not.

2.5 Research Problems:

The problems which are evident from the literature survey are characterized in three categories given below.

2.5.1 Facial Expression in Less Constraint Environment:

From the literature discussed it is evident that most of the work done in expression recognition used synthetic datasets which is free of illumination and occlusion issues. Moreover expressions portrayed by the subjects are mimicked according to the template of expressions as by Ekman in [77]. While in real data expressions vary among individuals and according to situations. Then in real data we have issues of illumination, shadows, occlusion etc.

2.5.2 Balancing Illumination:

Illumination variation is still a hard problem in facial expression recognition research area. When the lighting condition is changed, the same face appears variously. More specifically, the changes caused the by variation of lighting could be larger than the difference between of the appearance individuals. Research enforced that significant illumination changes can make dramatic changes in the projection coefficient vectors, and hence can seriously reduce the performance of the system. Fluctuation in appearance caused by illumination significantly influences the face detection performance and accuracy.

This Problem becomes more severe in real scenarios. As in that case brightness is not the only factor disturbing illumination, rather one person's shadow can effect illumination factor of others. And the surroundings may also cause illumination variations.

Figure 2. 1:Example of illumination imbalance[78]

2.5.3 Accommodating Occlusion and Adapting Pose Variation in FER:

Accurate localization of detailed facial features provides an important building block for many applications including identification and analysis of facial expressions. However, feature point localization tends to break down when applied to faces in real scenes where other objects in the scene (hair, sunglasses, and other people) are likely to occlude parts of the face [7]. More over Face alignment is one important component in face based applications, such as face attribute and expression analysis [2]. As most of the work is done on the fontal face a little work is done on both frontal and profile view but in real time it is not necessary that end user is facing camera. So head pose variation and occluded components must be anticipated for accurate feature extraction.

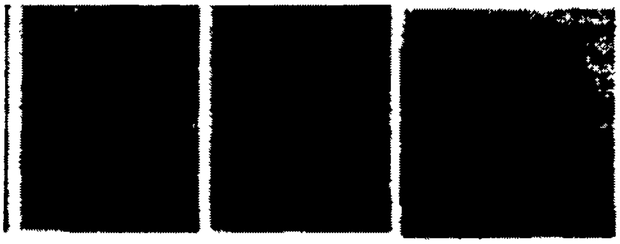


Figure 2. 2Examples of occlusion

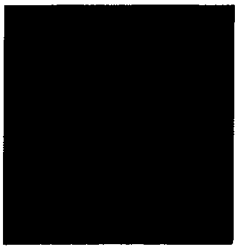


Figure 2. 3: Required feature for Expression Recognition

2.5.4 Identifying Emotion of crowd using FER:

Researchers in video-surveillance shifted the attention from the monitoring of a single person in a camera-monitored environment to that of groups and their behavior. This novel level of abstraction provides event descriptions which are semantically more meaningful, highlighting barely visible relational connections among people. Automatic Crowd Analysis is a research area which can be used for anomaly detection: panic scenarios, dangerous situations, illegal behaviors, etc. Little attention is on the expression of the crowd on the basis of facial expression. For doing so individual face has to be extracted and then facial expression analysis would be performed. Crowd expression recognition can be used to detect some malicious activity, or to identify if certain crowd is aggressive or peaceful. This research can be used to identify activity behavior of the crowd as well.

2.6 Existing Techniques:

In this section a discussion of state of art appearance techniques is given. Previous works in feature extraction can be divided into two categories: geometric feature-based methods and appearance-based methods as described by C. Shan et.al [79]. In geometric based technique geometric features such as lines, curves, blobs or any other geometric shape features are detected. This involves shape and localization of facial features and their tracking. Geometric features based technique shows sensitivity to noise and illumination. While appearance based techniques exhibits the texture, shape, color features of image. This technique mainly focuses on the apparent changes in image. Appearance based technique shows less sensitivity to noise.

In Figure 2.4 basic feature extraction algorithms that are most widely used in both approaches are shown.

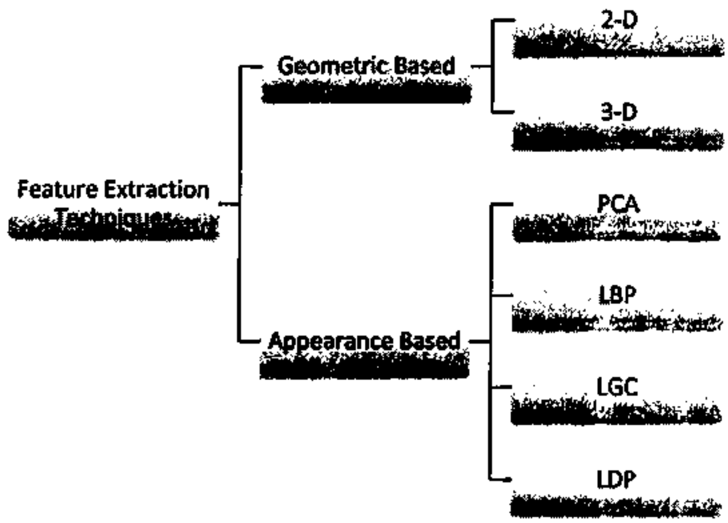


Figure 2. 4: Techniques for Feature Extraction

Though Appearance based methods are less subtle to noise and quite simpler and efficient than geometric based methods we adapted certain variation of these methods in this thesis which result in higher accuracy. In this section we are discussing three basic state of art techniques LBP, LGC and CLBP which are then used as comparison as well.

2.6.1 LBP:

LBP is state of art technique for facial expression recognition used by B. Jun et. al in [80]. It gives considerably good feature extraction of wrinkles bulges and other changes in facial muscles and classification results of LBP histograms give around 90% accuracy. In Conventional LBP image is divided in 3x3 window of grey values. Comparison of central pixel is made with its 8-neighborhood. This comparison is generally in circle, we followed clockwise rotation. Another comparison window is created in which value 1 is stored if neighborhood pixel is greater than center pixel else it receives value 0. Once this comparison is complete 8-bit binary code is generated. This 8-bit binary code is converted to decimal and central pixel is replaced by this decimal value in LBP image. Then histogram of LBP image is generated as feature vector. Mathematically

$$LBP(X_c, Y_c) = \sum_{p=0}^{7} s(i_p - i_c)2^p \qquad (2.1)$$

$$S(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \qquad (2.2)$$

Here i_c is value of central pixel and i_p is value of neighboring pixel. Illustration of LBP encoding is given

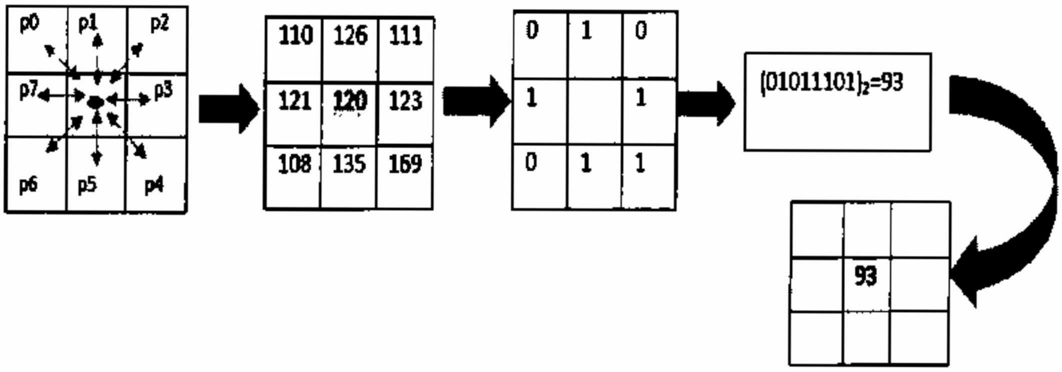


Figure 2. 5: LBP Feature descriptor

2.6.2 Local Gradient Coding in Horizontal vertical and diagonal direction (LGCHVD): LGCHVD explores different relationship in the neighborhood [81]. Instead of having difference from the central pixel in LGC coding horizontal difference and vertical difference is calculated around the central pixel. Basic idea is to discover texture properties in gradient direction as well. Mathematically

$$LGC(x_c, y_c) = s(g_0 - g_2)2^7 + s(g_7 - g_3)2^6 + s(g_6 - g_4)2^5 + s(g_0 - g_6)2^4 + s(g_1 - g_5)2^3 + s(g_2 - g_4)2^2 + s(g_0 - g_4)2^1 + s(g_2 - g_6)2^0 \quad (2.3)$$

Where these pixels are considered as

g_0	g_1	g_2
g_7	x_c, y_c	g_3
g_6	g_5	g_4

Figure 2. 6: LGCHVD window

Illustration of LGCHVD is

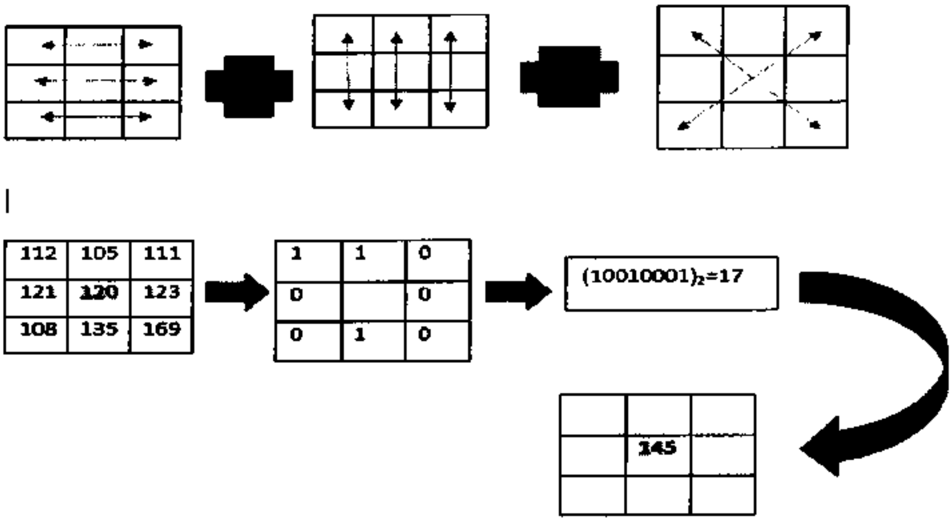


Figure 2. 7: LGCHVD descriptor

There is another version of LGC which is local gradient coding in horizontal and diagonal direction (LGC HD). In LGCHD feature vector length is reduced to 32. Mathematically,

$$LGCHD(x,y) = s(g_0 - g_2)2^4 + s(g_7 - g_3)2^3 + s(g_6 - g_4)2^2 + s(g_0 - g_4)2^1 + s(g_2 - g_6)2^0 \quad (2.4)$$

Illustration of LGCHD is

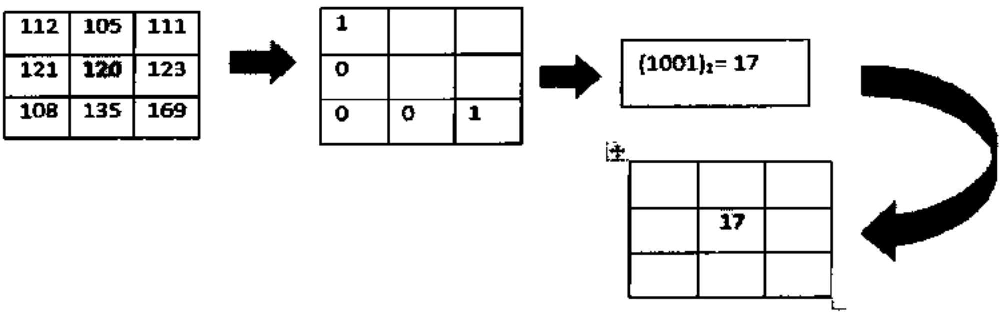


Figure 2. 8: LGCHD descriptor

2.6.3 CLBP:

In CLBP along with sign difference as LBP among the neighborhood of a pixel CLBP keeps magnitude difference as well. Thus, in CLBP instead of one bit code for each neighborhood, 2 bit code is generated. Least bit is for sign change as in LBP, i.e. if neighboring pixel value is greater than central pixel it attains value 1 in LBP space else value 0. Second bit is for magnitude difference. Magnitude difference is calculated by taking average of neighborhood say P_{avg}. Then neighboring pixel values are compared with P_{avg} in the same way as LBP coding. If pixel value is greater than P_{avg} it attains value 1 else 0. By combining both these sign and magnitude difference a 16- bit code is generated which results increase in dimension. To reduce dimension two sub codes CLBP1 and CLBP2 are generated.

Mathematically,

Least bit is generated as mentioned in equation (1).

For second bit

$$\bar{P} = \left(\frac{1}{9}\right)\sum_{p=0}^{7} i_p \tag{2.5}$$

$$bit2(x_c, y_c) = \sum_{n=0}^{7} s\left(i_p - \bar{P}\right)2^n \tag{2.6}$$

Where,

$\bar{P}$ *Is* average of pixel values in neighbourhood

$\left(i_p - \bar{P}\right)$ Comparing neighborhood pixel with average.

Illustration of CLBP is given below in example.

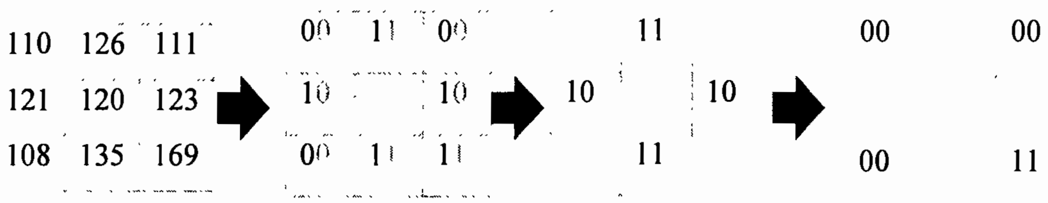


Figure 2. 9: CLBP descriptor

2.7 Chapter Summary:

In this chapter literature of Expression Recognition, handling imbalance illumination and accommodating occlusion. In Expression Recognition, extensive work is done on synthetic datasets like JAFFE and CK, but very little attention is given to real data. Likewise, illumination is handled as while but global and local illuminations are overlooked so far. There is very little work done on accommodating occlusion handling in facial expression recognition. Further we extend literature from one person to crowd and to best of our knowledge no such work has been focused by any researcher in this area. Then over view of state of art techniques is given as these would be used as comparison with our techniques.

Chapter # 3

Multi Stage Binary Patterns

3. Multi-Stage Binary Patterns:

In this chapter, we proposed a novel technique for expression recognition in real world. This is bi-level pattern generation technique that helps in learning real world expressions. This chapter covers details and mathematical model of MSBP with examples. Section 3.1 discuss motivation of developing MSBP. Section 3.2 is about detail modeling of MSBP and its pseudo code is provided. Section 3.3 is about Experimental setup and in section 3.4 is about performance measures. In section 3.5 results of MSBP are discussed in detail with accuracy charts and confusion matrix and section 3.6 concludes the chapter.

3.1 Motivation:

In this chapter, we proposed a robust method that generates binary patterns in spatial domain to explore gradient difference and sign difference among neighborhood. These binary patterns give promising results with SFEW and other datasets than state of art techniques. It maintains local texture differences along with gradient variance. This resolves the issues of intricate background and neighborhood enlightenment that happen because of specific embellishments as scenes or hair, fluctuating and uncalled light reflections that are not skillfully handled in regular LBP and the greater part of its varieties. MSBP is 16-bit encoding plan, 2-bit per neighborhood. Least bit of MSBP reserves local texture alteration. In local texture alteration, central pixel works as focal pixel. Window distinction framework is ascertained by taking contrast of neighboring pixels with focal pixel. While second bit administers magnitude change. For magnitude change, difference grid is designed by calculating absolute difference of neighboring pixels with central pixel. Then average of this difference grid is used as threshold to determine second paired bit of MSBP scheme. This procedure helps in compensation of global and local illumination.

3.2 Multi-Stage Binary Patterns (MSBP):

MSBP is 16-bit encoding scheme: two bits for every pixel in window. Least bit is the sign variance as it is in LBP. Modification is in the second bit. For magnitude, primarily difference-grid is premeditated. Difference grid sustains inconsistency of neighboring pixels with central pixel. Average of this difference grid say D_avg is considered as threshold. At that point, neighboring pixels are contrasted with this threshold to produce twofold encoding. If pixel esteem is more noteworthy than the limit, it accomplishes esteem 1 else 0. This aide in limiting the effect of neighborhood poor enlightenment. By consolidating both these – sign and magnitude contrast – a 16-bit code is produced. The 16-bit encoding brings about high dimensional element vector having length 65536. Thus, with a specific end goal to diminish measurement, two sub codes: MSBP1 and MSBP2 each of length 8-bit, are produced which are joined to accomplish feature vector having length 256.

Mathematically,

Let δ_1 is the least significant bit and δ_2 is the most significant bit.

$$\delta_1(x_c, y_c) = \sum_{p=0}^{7} s(i_p - i_c)2^p \tag{3.1}$$

For δ_2

Firstly, Gradient value is calculated between focal pixel and its neighborhood as

$$g_n = |i_p - i_c| \tag{3.2}$$

$$\bar{g} = (1/p)\sum_{n=0}^{p-1} g_n \tag{3.3}$$

$$\delta_2(x_c, y_c) = \sum_{p=0}^{7} s(g_n - \bar{g})2^p \tag{3.4}$$

Where,

i_n *is neighborhood pixel and* i_c *is central pixel*

$\bar{g}$ *Finding average of gradient*

$(g_n - \bar{g})$ *Comparing neighborhood pixel with average gradient*

S(x) is same as in equation 2.

$\delta_1\delta_2$ are paired to spawn binary encoding which is then fragmented into MSBP1 and MSBP2 code to condense dimensions. MSBP1 is acquired by linking 4-associated neighbors of focal pixel. On the off chance that (x, y) is focal pixel then (x ±1, y) and (x, y±1) pixels will produce MSBP1 code. MSBP2 is acquired by linking the 8-associating neighbors barring 4 interfacing neighbors. For (x, y) as focal pixel (x±1, y±1) will produce MSBP2 code.

Illustration of MSBP is given below. Sixteen-bit code is generated and sub codes are given as MSBP1 and MSBP2 are given below in Figure 3.1:

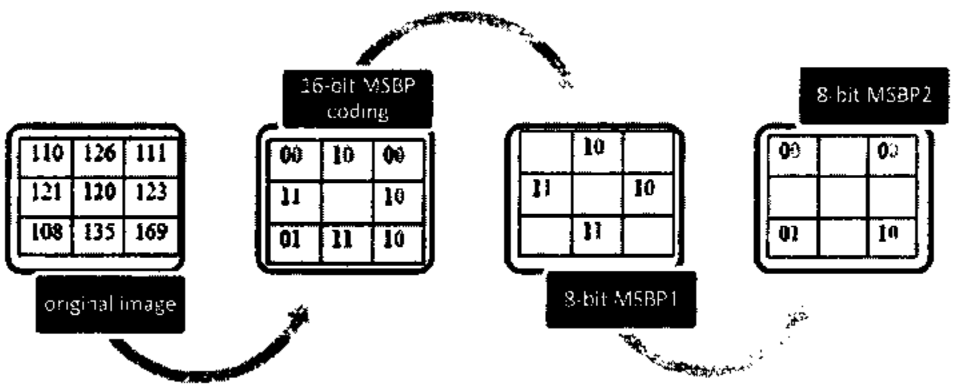


Figure 3.1: MSBP descriptor

3.2.1: Pseudo code for MSBP:

In this section step by step procedure of MSBP is explained.

Algorithm MSBP:

Input: set of Training images (TI = 1...213), set of Class of training Images (C = 1...7)
Output: CDLBP histogram

```
For i= 1 to size of Training Images
Image=TI(i)
Calculate size of training Image in rows(i), cols(i) where i ε TI .. range of I varies with size of training set
For row=1 to rows(i)
  For col=1 to cols(i)
        [p0, p1, p2, p3, p4, p5, p6, p7] = Compare central pixel with its neighborhood to generate adjacency matrix
        [d0,d1,d2,d3,d4,d5,d6,d7] = Find out difference matrix by taking difference of central pixel with its neighborhood
        [diff avg] = Find out average of difference matrix.
        [da0,da1, da2, da3, da4, da5, da6, da7]=Compare diff avg pixel with its neighboring pixels to generate adjacency matrix
        Generate sub codes of MSBP
        MSBP1(row, col) = convert this binary string [p6da6p4da4p2da2p0da0] to decimal
        MSBP2(row, col) = convert this binary string [p5da5p3da3p1da1p7da7] to decimal
  End
End
Generate histogram of MSBP1
Generate histogram of MSBP2.
Concatenate both histograms.
End
```

3.3 Experimental Set-up:

An arrangement of tests has been performed utilizing Static Facial Expressions in the Wild (SFEW) database [82]. It consolidates seven expressions (anger, disgust, fear, happy, neutral, and sad and surprise). Static Facial Expressions in the Wild (SFEW) database has been made by choosing video outlines from AFEW informational index. Informational index contains 700 pictures with unconstrained facial expressions like that of real world environment i.e. distinctive head postures, huge age-go, impediments, changed concentration, distinctive determination of face and assorted enlightenment. SFEW contains six basic expressions (angry, disgust, fear, happy, sad, surprise and the neutral) given in Figure 3.2.

Figure 3. 2: Facial Expression in SFEW

Viola Jones' face location calculation is utilized to concentrate confront from the chose pictures. Resultant Face pictures are then scaled and normalized. An example of the localized picture after applying face localization is given in the Figure 3.3.

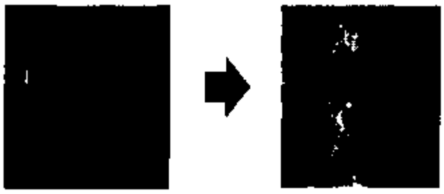


Figure 3.3: Face Localization and Normalization

After face confinement, components are extricated by utilizing all above-discussed methods. All local texture based methods are implemented are tested for the said dataset. Feature vectors are created in type of histograms which are sent to the classifiers for expression characterization. There are two situations in which examinations are led shown in Figure 3.4.

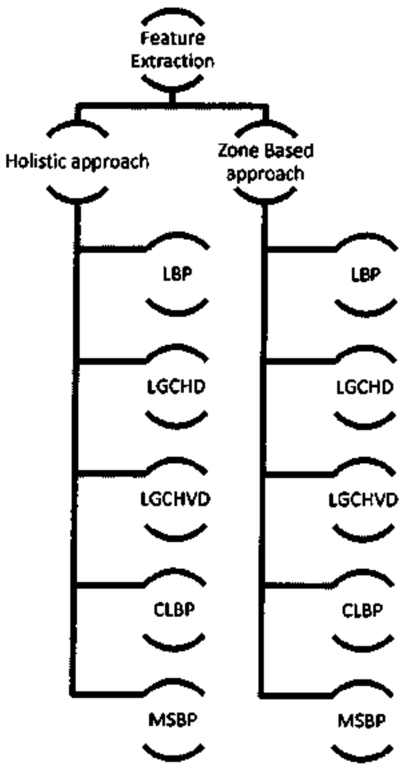


Figure 3.4: Experimental Setup

3.3.1: Experiments in Holistic Approach:

In holistic approach impact of any change in facial features is treated as global change. Fundamental thought encompass holistic approach is it that normally whole face contributes in expression exemplification. In this approach MSBP is smeared on full face to obtain MSBP1 and MSBP2 image. Then these MSBP1 and MSBP2 images are transformed into particular histogram as histograms comprehend occurrences of micro information. These two histograms are linked to acquire spatial feature vector. Length of feature vector is 512. Figure 3.5 contains portrayal of MSBP.

Figure 3.5: MSBP in holistic Approach

3.3.2: Experiments in Zone based Approach:

In zone base approach change in facial point is taken as local change. Additionally with a specific end goal to have location data of facial elements in histograms zone based approach is successful one. Localized image is separated into equivalent size. MSBP and MSBP2 is computed for every zone. These processed pictures are then deliberate to their particular histograms. Every one of these histograms are linked to acquire single element vector. Criticalness of this element vector is to union location and spatial data. And to observe that either feature deformation in expression local impact is more effective than global impact.

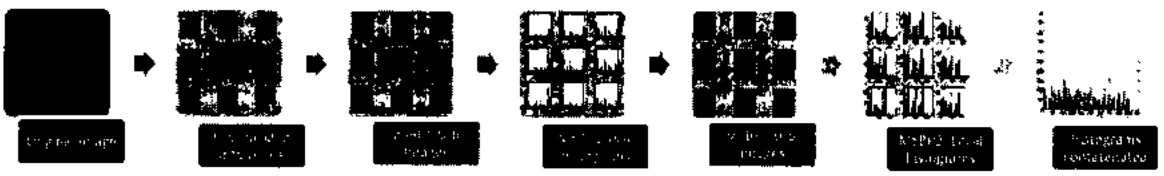


Figure 3.6: MSBP in Zone Based Approach

3.4 Performance Measures:

Quantitative research is the approval of framework utilizing numerical, factual or computational methods. These techniques can be utilized to check that whether the outcomes are valid as for the prepared informational indexes. The quantitative measures incorporate Accuracy, Precision and Recall. These measures are utilized as a part of order segment to discover the precision of proposed framework.

3.4.1 Accuracy:

Accuracy is the global correctness of the system and is calculated by adding all of the correct classifications divided by total

Classifications

$$\text{Acuracy}=\frac{\text{no of true positives+no of true negatives}}{\text{number of true positives+number of true negatives+no of false positives+no of false negatives}}$$

3.4.2 Precision:

Precision is the degree to which repetitive measurements under unchanged conditions show the identical results. It is the measure of accuracy given that a particular class is predicted.

$$\text{Precision} = \frac{\text{No of true positives}}{\text{no of true positives+no of false positives}}$$

3.4.3 Recall:

The determination of the ability of system to select the occurrences of a certain class from a dataset is called Recall. It corresponds to the true positive rate and called sensitivity.

$$\text{Recall}=\frac{\text{no of true positives}}{\text{no of true positives+no of false negatives}}$$

3.5 Results:

To manifest the impact of smaller scale data, trials are performed in all-encompassing methodology and zone based approach. In holistic manner MSBP is applied on full face while in zone based approach, confront picture is separated into nine equivalent amounts of, extraction strategy is connected to each part, and the resultant histograms of all parts are linked to shape a solitary histogram to be utilized as contribution by the classifier.

Ten times cross approval plan is utilized to decide the characterization rate. In ten times, cross-approval, the genuine example is unpredictably divided into 10 measure up to estimated sub-tests. Among these 10 sub-tests, a solitary sub-test is held as approval information for testing the model, and the rest of the 9 sub-tests are utilized as preparing information. This procedure is then duplicated 10 times with each of the 10 sub-tests utilized precisely once as the approval information.

SFEW database is utilized to test the working of the component extraction strategies with changing head-posture development and enlightenment conditions.

3.5.1: Results in Holistic Approach:

Firstly, the extraction procedures are connected on full face pictures. MSBP highlights set with straightforward strategic classifier produced the finest outcome i.e. 96.47% which is the best among every single other system. As discussed earlier, all these methods are implemented. Than SFEW is tested on these methods for the results.

Table 3. 1: Results of LBP, CLBP, LGC HD, LGC HVD and MSBP for Full Face Image with SFEW database

Techniques →	LBP			CLBP			LGC-HD			LGC-HVD			MSBP		
Classifiers ↓	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)
BF Tree	0.72	0.71	71.17	0.79	0.79	79.41	0.89	0.88	88.82	0.88	0.88	88.23	0.88	0.88	88.23
Simple Logistics	0.78	0.78	78.23	0.85	0.85	85.29	0.92	0.92	92.35	0.93	0.92	92.94	0.96	0.96	96.47
KNN	0.71	0.70	70	0.84	0.84	84.11	0.9	0.92	92.35	0.94	0.93	93.52	0.95	0.95	95.29
Bagging	0.68	0.69	69.41	0.78	0.78	78.82	0.88	0.88	88.23	0.87	0.87	87.05	0.91	0.91	91.76
Naïve Bayes	0.68	0.65	65.29	0.72	0.70	70.58	0.83	0.81	81.76	0.84	0.80	80.58	0.90	0.88	88.82

Accuracy chart of these expression is given in following Figure 3.7.

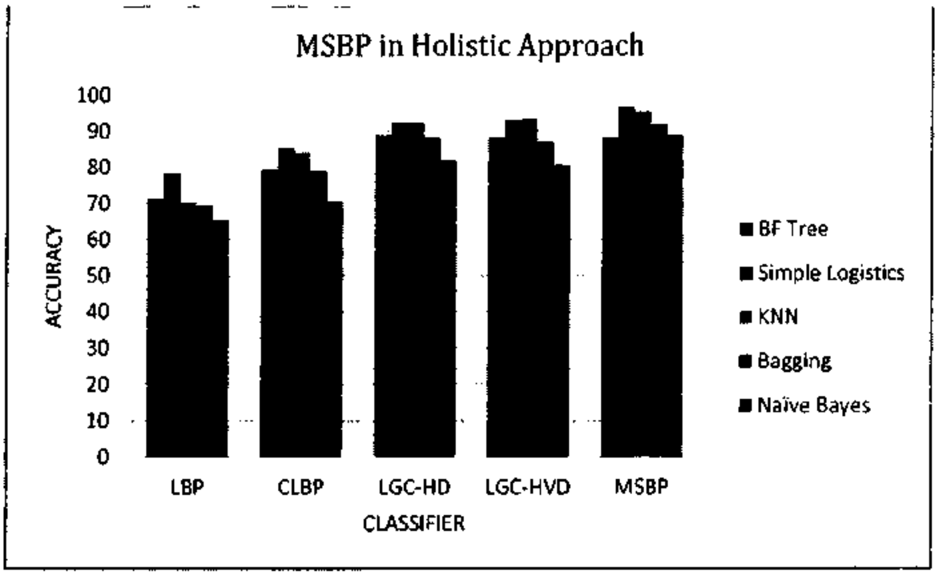


Figure 3. 7: Accuracy Chart of MSBP in holistic Approach

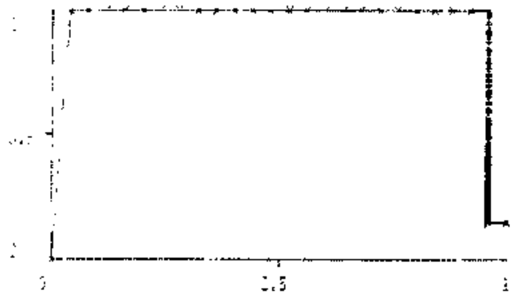
Findings clearly explicate that single LBP is not enough measure for FER in challenging dataset as it missed some important relationship of gradient difference. Similarly, LGC and LGCHVD with gradient difference only provide results better than LBP but still improvement is required. In CLBP along with sign difference magnitude difference is also observed but results could not improve much. Thus, in MSBP due to gradient difference with sign difference it outperforms than rest of the feature extraction techniques.

Results can be minutely observed in the confusion matrix given in Table 3.2. Expression precision is given in the following confusion matrix. Disgust, Fear and Sad are completely perceived. But Anger, Happy and Surprise demonstrate little misclassifications.

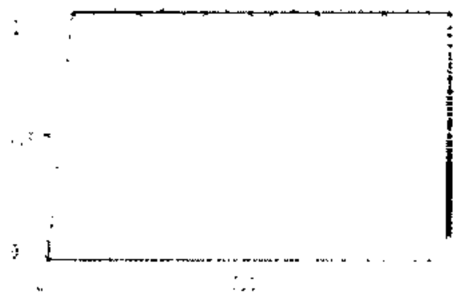
Table 3.2: Confusion Matrix of MSBP in Holistic Approach

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	96.00	0	0	0	0	0	4 00
Disgust	0	100	0	0	0	0	0
Fear	0	0	100	0	0	0	0
Happy	0	0	0	93.10	6.90	0	0
Neutral	0	0	0	3.35	93.30	3.35	0
Sad	0	0	0	0	0	100	0
Surprise	0	0	0	0	0	4.17	95.83

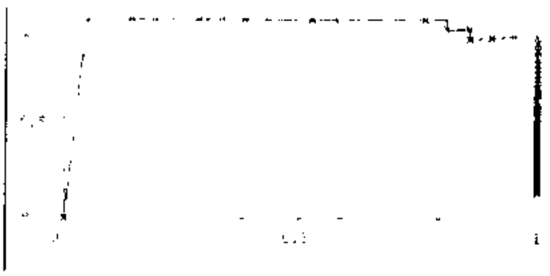
For further details precision recall curves are given. Figure 3.9 shows precision recall curve for anger class and from area under the curve (AUC) it is very clear that results are satisfactory.



(a)

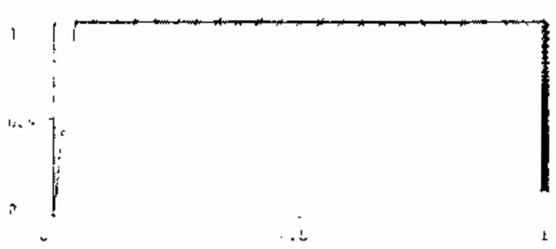


(b)



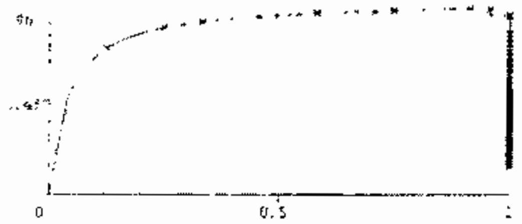
(c)

(d)



(e)

(f)



(g)

Figure 3. 8: Precision recall Curve of MSBP for Anger (a), Disgust (b), Fear(c), happy (d), Neutral (e), sad (f), Surprise (g)

For generating Precision Recall Curve Weka is used. For evaluating precision recall curve we have used AUC. For anger, disgust, fear, happy and neutral there is nearly optimal precision recall curve. AUC shows expressions recognition rate is satisfactory.

3.5.2 Results in Zone Based Approach:

Results are significantly reduced in zone based approach. Reason for decline in rates is in FER all features play important role. While this information is missing in zone based strategy. In this situation, nearby effect of elements distortion is considered however it dropped comes about because of 96.7 % to 60%, as essential relationship among components is missed when face is separated in equivalent size zones. However, in the event that exclusive zone based situation of all methodologies is measured it is unmistakably demonstrated that MSBP works superior to LBP, LGCHD, LGCHVD and CLBP. The accompanying chart demonstrates that execution of MSBP in zone based approach is likewise better to the condition of texture based techniques methods. Table 3.2 shows result of MSBP in zone based approach.

Table 3. 3: Results of LBP. LGCHD. MSBP, CLBP and LGCHVD in zone based Approach

Techniques → / Classifiers ↓	LBP			LGC-HD			MSBP			CLBP			LGC-HVD		
	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)
BF Tree	0.318	0.324	32.35	0.437	0.424	42.35	0.602	0.6	60	0.364	0.376	37.64	0.403	0.406	40.58
Simple Logistics	0.361	0.4	40	0.436	0.418	41.76	0.485	0.471	47.05	0.522	0.506	50.58	0.445	0.424	42.35
KNN	0.454	0.441	44.11	0.555	0.512	51.17	0.592	0.559	55.88	0.466	0.465	46.47	0.567	0.512	51.17
Bagging	0.366	0.4	40	0.426	0.459	45.88	0.574	0.576	57.64	0.452	0.471	47.05	0.495	0.518	51.76
Naive Bayes	0.264	0.306	30.58	0.293	0.359	35.88	0.355	0.376	37.64	0.266	0.306	30.58	[illegible]	0.406	40.58

From Table 3.3 it is clear that over all performce of FER is decreased in zone based approach, when compared to hloistic approach given in Table 3.1. This is because for expression all features play important part. zone based approach work good in face recognition techniques but in expression recognition more details are required. among all these implemented techniques MSBP still perform better. it gives 60% results while rest of techniques lied on 32%, 42%, 37% and 40%. Thus gradient difference along with sign difference make MSBP riscriminant than state of art techniques.

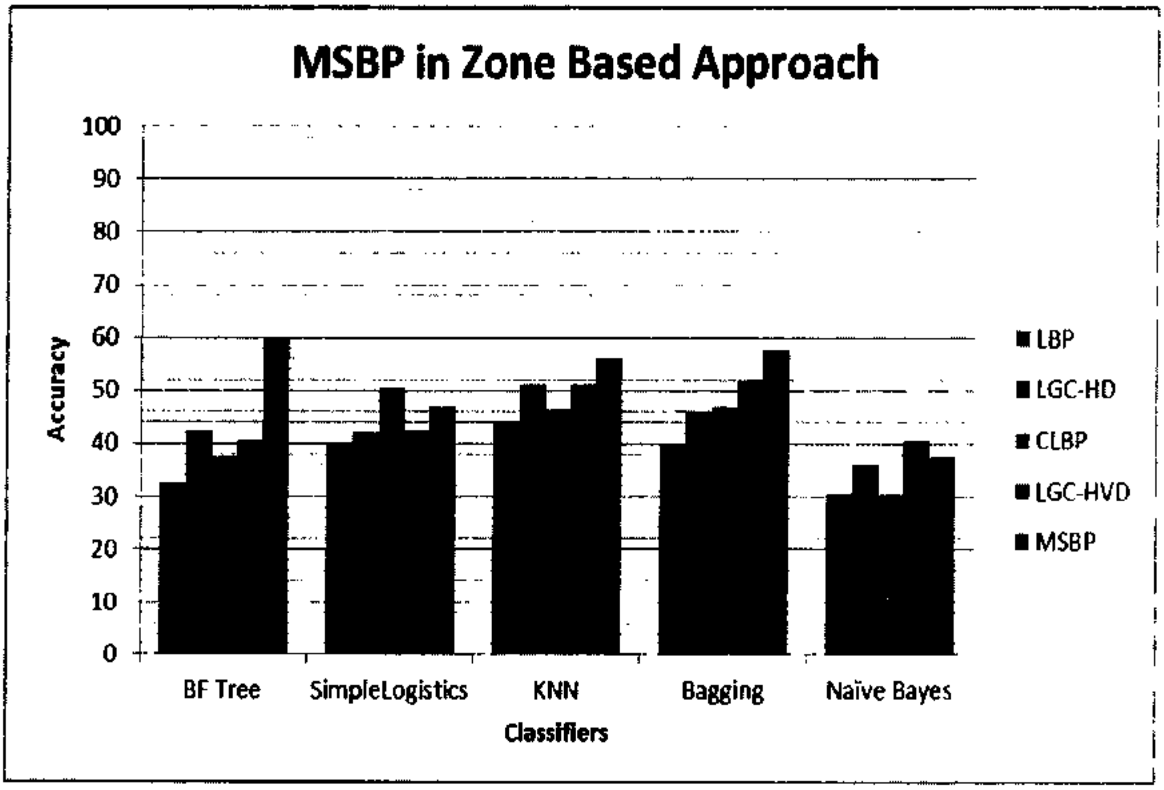


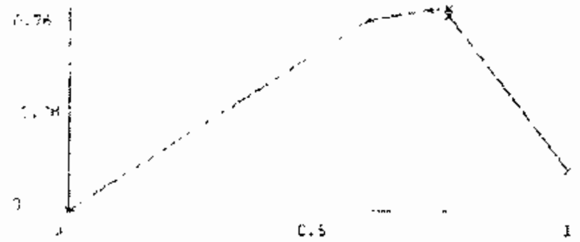
Figure 3. 9: Results of LBP, CLBP, LGC HD, LGC HVD and MSBP for zone based approach

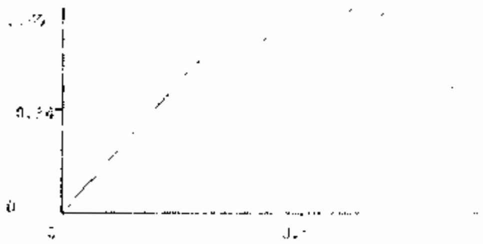
Accuracy chart in Figure 3.10 highlighted results of MSBP in zone based approach. It can be seen clearly in Figure 3.10 that MBP perform better with all classifier sin such challenging dataset. For further investigation of results on expression level confusion matrix is given in Table 3.4.

Table 3. 4: Confusion matrix of MSBP in Zone based Approach

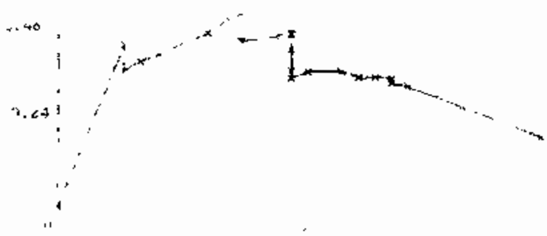
	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	76.00	16.00	0	4.00	0	4 00	0
Disgust	16.67	66.67	5.55	11.11	0	0	0
Fear	0	0	54.17	20.85	16.67	8.33	0
Happy	6.89	3.46	20.70	37.93	24.13	6.89	0
Neutral	0	0	6.89	20.70	65.51	3.45	3 45
Sad	4.76	4.76	0	4.76	14.30	52 38	19.04
Surprise	0	0	0	0	12.50	16.67	70 83

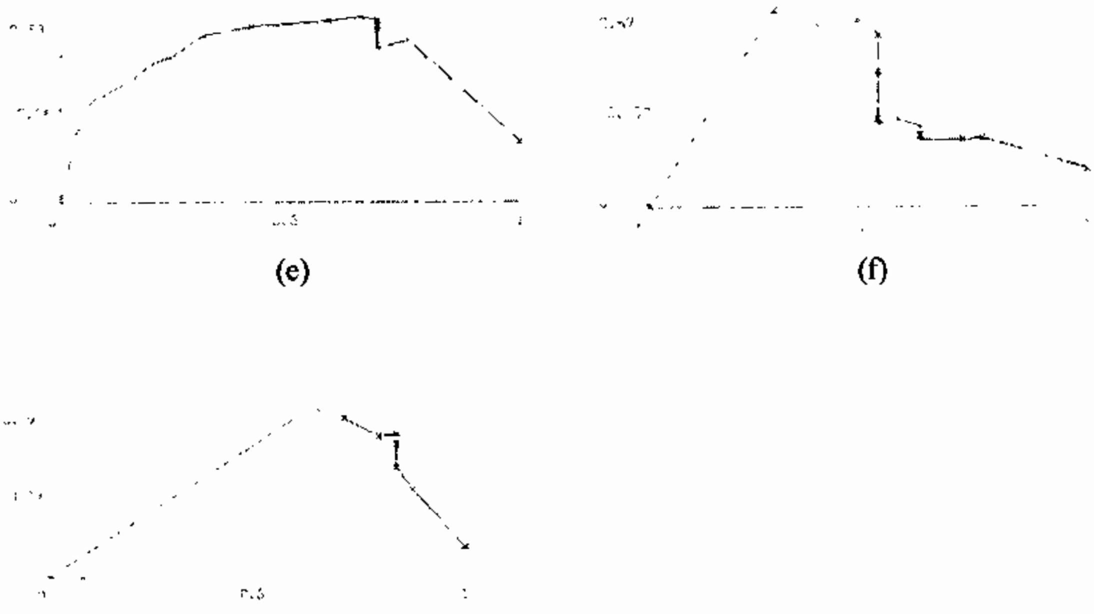
Confusion matrix explicate recognition rate of expressions. It is plainly delineated in confusion matrix that all-encompassing methodology is superior to anything zone based approach in expression acknowledgment.


(a)


(b)


(c)


(d)



(e) (f)

(g)

Figure 3. 10: Precision recall Curve of Anger (a), Disgust (b), Fear(c), Happy (d), Neutral (e), Sad (f), Surprise (g)

Figure 3.10 shows precision recall curves for expressions evaluated in zone based approach. AUC shows same results as given by confusion matrix in Table 3.4. There are some inaccuracies in determining expression recognition. When compared with holistic it clearly shows that holistic approach is better for expression recognition than zone based approach.

3.5.1 Significance Testing (T-Test)

This is statistical significance test. In T-Test conducted for feature extraction techniques LBP and MSBP, Null hypothesis is LBP=MSBP that is there is no relationship between LBP and MSBP. If $p<0.05$ it means there is significant difference and MSBP is better than LBP.

Table 3.5: Significance Test of MSBP and LBP

t	T Test is used
(298)	degree of freedom
3.414	t statistics
p	0.0003; as $p < 0.05$ it shows
	Difference is significant; null is rejected.

Among feature extraction techniques (N=300), there was statistically significant difference among LBP (M=0.354, SD=0.43) and MSBP (M=0.442, SD=0.12), $t(298)=3.414$, $p<0.05$,

$CI_{0.95}$ -0.13,0.03. Thus we reject null hypothesis. There is significant difference between MSBP1 and LBP.

3.6 Chapter Summary:

In this part a novel strategy MSBP is recommended that is powerful in true situation for facial expression recognition. It is invigorated to deal with unevenness and neighborhood enlightenment by utilizing gradient difference and sign distinction. The combination of gradient difference and sign magnitude makes MSBP robust in real time environment with challenging light conditions, real expression depiction and complex backgrounds. Promote, examination of MSBP is made with LBP, LGC-HD, LGC-HVD and CLBP. These all are local texture based techniques and implemented to simulate same conditions for evaluating results. Examinations are directed utilizing two unique situations: holistic and zone based. This examination fundamentally utilized holistic approach for LBP, LGC-HD, LGC-HVD and CLBP. It is also main contribution of the work that even existing techniques are tested on real world data-set for the first time. Comes about presume that proposed method outperforms all the surface based systems looked at in the paper for both holistic and division based approaches. In holistic approach, it gives 96.7% results which are better than existing state of art techniques. While in zone base approach over all accuracy is declined to 60% but among other techniques MSBP still performs best. Results of FER declined in zone based approach as it loses full description of face that is necessary for FER.

Chapter # 4

Handling Imbalance Illumination in FER

4. Handling Imbalance Illumination in FER

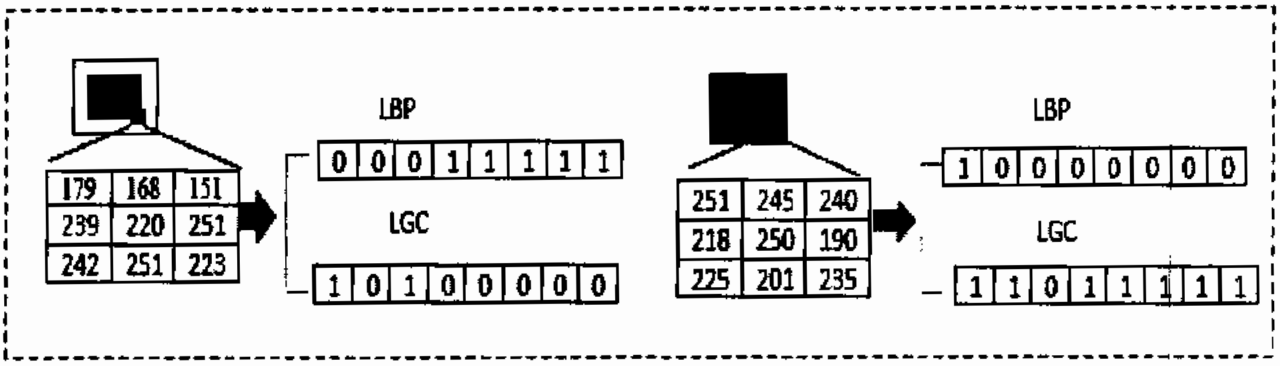
In this chapter novel RGPs are proposed for handling imbalance light variations in FER. In section 4.1 detailed mathematical analysis of influence of global and local illuminations on images is given. It clearly shows non-consistent patterns of existing local texture based techniques. This discrepancy makes learning FER difficult process. In section 4.2 novel RGPs are proposed. These RGPs generate gradient difference patterns which are consistent in case of global and local light variations. In section 4.3 experimental setup is discussed in detail. In section 4.4 it is proven that RGPs provide good results in FER in challenging light conditions. For this it is initial effort to simulate data with global and local illuminations. Further real noisy data is also tested by using RGP. Section 4.5 give brief conclusion of the chapter.

4.1 Motivation:

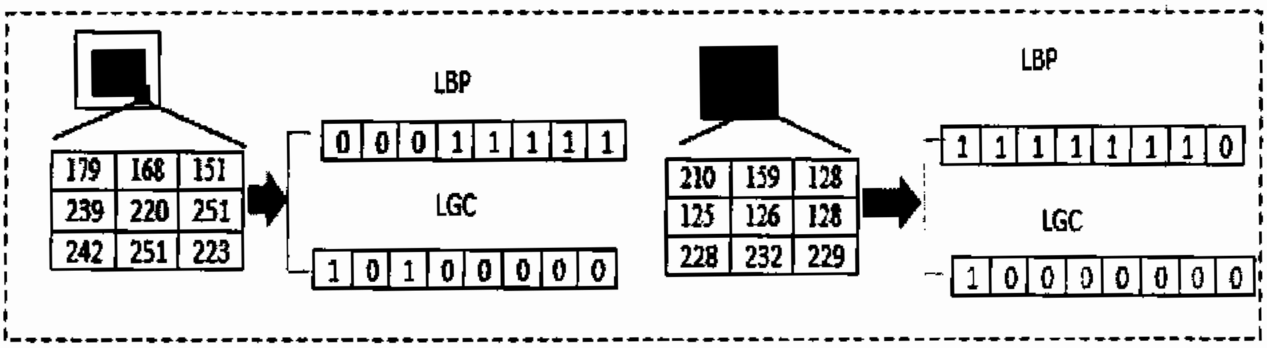
Illumination variation is a challenging issue in facial expression recognition research area. With slightest change in light that might be even because of shadow or some accessories pixel values changed. This makes task of learning expression recognition extremely difficult for machines. More specifically, the changes caused the by variation of light could be as large as recognizing between different individuals. Illumination changes can make histrionic changes in the projection coefficient vectors, and hence can seriously reduce the performance of the system. Reflections and shadows are two main categories of appearance variation due to illumination variation.

For improvising accuracy of the system, change in illumination must be tackled carefully. In this study feature extraction is focused for undertaking light variations.

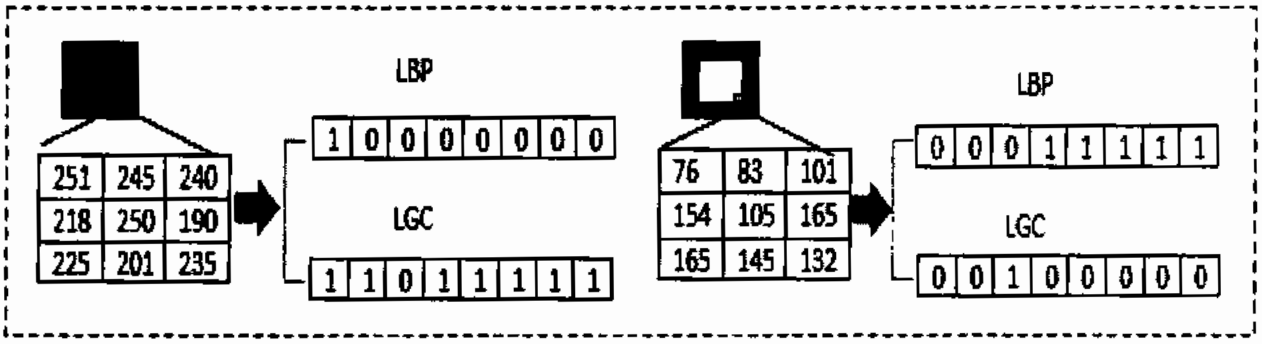
From extensive literature review, it is apparent that LBP and LGC are invariant to the global intensity variations. On the other hand these techniques are sensitive to local intensity variations that occur commonly along edge components such as eyes, eyebrows, nose, mouths, whiskers, beards, or chins due to internal factors (eye glasses, contact lenses or makeup) and external factors (different background). This sensitivity of the existing FER method generates different patterns of local intensity variations and makes learning of the face expression difficult. This can be shown in Figure 4.1. In this Figure 4.1, Image is shown in boxes (gray and white box, black, gray box) while Table below the box show pixel values. Along with image and pixel window its respective LBP and LGC patterns are generated by using equation #1 and equation #2. Four different scenarios of light variation are discussed in Figure 4.1 parts (a, b, c &d). In 4.1(a) global light changes are reflected, changes in light has same effect on background and foreground. Global light variations are shown in image and in pixel window as well with changes their respective LBP and LGC codes. In Figure 4.1 (b) and (c) local light variations are shown. In 4.1(b) only background change is shown in same pattern as discussed above. In 4.1(c) local change in term of foreground change is shown only. In 4.1(d) most complicated case of light variation is discussed. In this case foreground has four different sort of local light variations. It is seen very clearly that in each case LBP, and LGC generated patterns are different. Though in all scenarios we have same image with some light variations only. But due to these light variations each time LBP, LGC generated patterns vary.



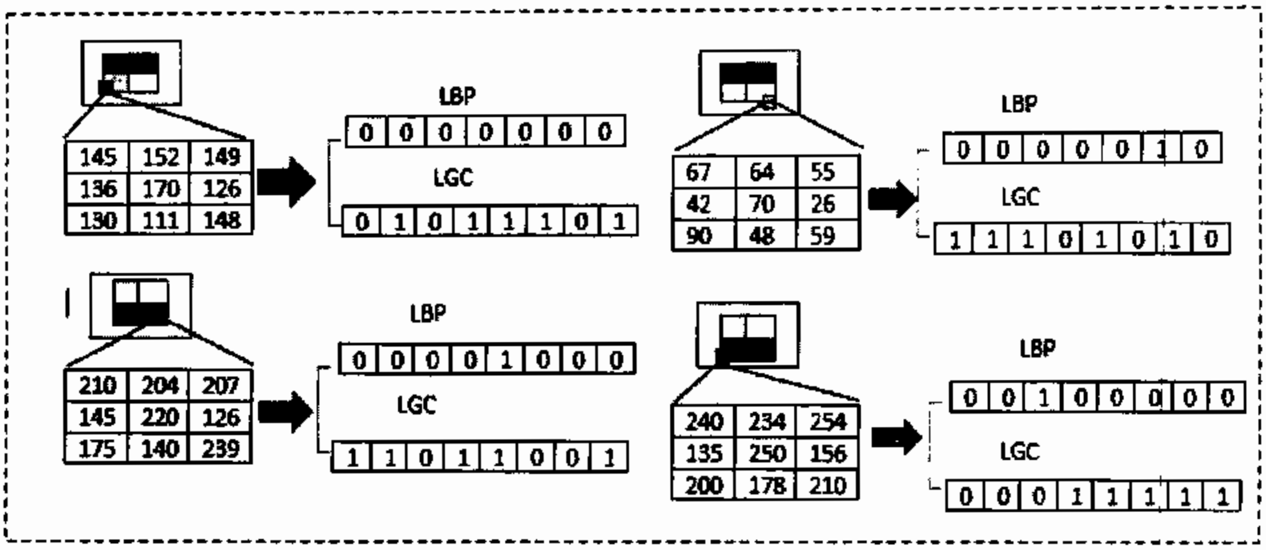
(a) LBP & LGC in case of global illumination



(b) LBP & LGC when background changes locally



(c) LBP and LGC patterns when the gray level of foreground changes locally



(d) LBP and LGC patterns when the gray level of image changes locally

Figure 4. 1:(a) (b)(c)and (d) are examples of inconsistent Patterns of LBP and LGC in case of local illumination variation

4.2: Robust Gradient Patterns (RGP):

RGP is proposed in this dissertation for handling uncontrolled illumination variations in images. It has stronger distinctive power than other texture based techniques. This is local spatial window based technique. For each window RGP obtain absolute gradient values by taking absolute value of intensity differences between central pixel and its neighboring pixels. Then average of these gradient values is used as a threshold value. Gradient values are compared with threshold to produce binary encoding and calculate decimal value from this binary encoding to acquire RPG patterns.

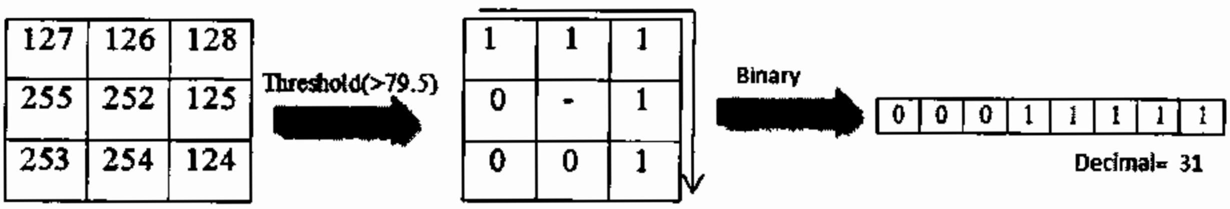


Figure 4. 2: RGP Descriptor

RGP operator works on $2\times r+1$ by $2\times r+1$ template to represent the image local structure. RGP is calculated as:

$$Threshold_{i,j} = \frac{1}{m.n}\sum_{i=1}^{m}\sum_{j=1}^{n}\left|p_{i,j} - p_{\left\lceil\frac{i}{2},\frac{j}{2}\right\rceil}\right| \tag{4.1}$$

$$p_{\left\lceil\frac{i}{2},\frac{j}{2}\right\rceil} = s\left(\left|p_{i,j} - p_{\left\lceil\frac{i}{2},\frac{j}{2}\right\rceil}\right| - Threshold_{i,j}\right).2^{p} \tag{4.2}$$

$$s(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases} \tag{4.3}$$

Where, $p_{i,j}$ is the total no of sampling points and r is the radius of the neighboring $p_{\left\lceil\frac{i}{2},\frac{j}{2}\right\rceil}$ is the center pixel point, $s\left(\left|p_{i,j} - p_{\left\lceil\frac{i}{2},\frac{j}{2}\right\rceil}\right| - Threshold_{i,j}\right).2^{p}$ is the absolute difference gradient values, $Threshold_{i,j}$ is average of difference value.

Once RGP image is obtained, its histogram is calculated which is used as feature vector. As histogram skips location information, it only records occurrence of values. To incorporate location information along with frequency of values extended histograms are calculated. Initially image is divided in equal size blocks. RGP image of each block is obtained by applying RGP encoding. Histograms for each RGP image are generated. Then all these histograms are concatenated to obtain extended RGP histogram.

Performance of RGP is tested in three different scenarios, first is RGP is tested on original images. Images are then subjected to global and local illumination variance, and RGP is applied in each state. Results show that RGP performs well in all setups. RGP Generate consistent pattern when intensity level of image is changed globally as given in Figure 4.2. Figure 4.2 (a, b, c and d) show RGP for image subjected to global changes, effected with local change with background change only, local change with foreground change only and finally when foreground is effected by different local light variations. in each case it is clear that for same image RGP remain consistent.

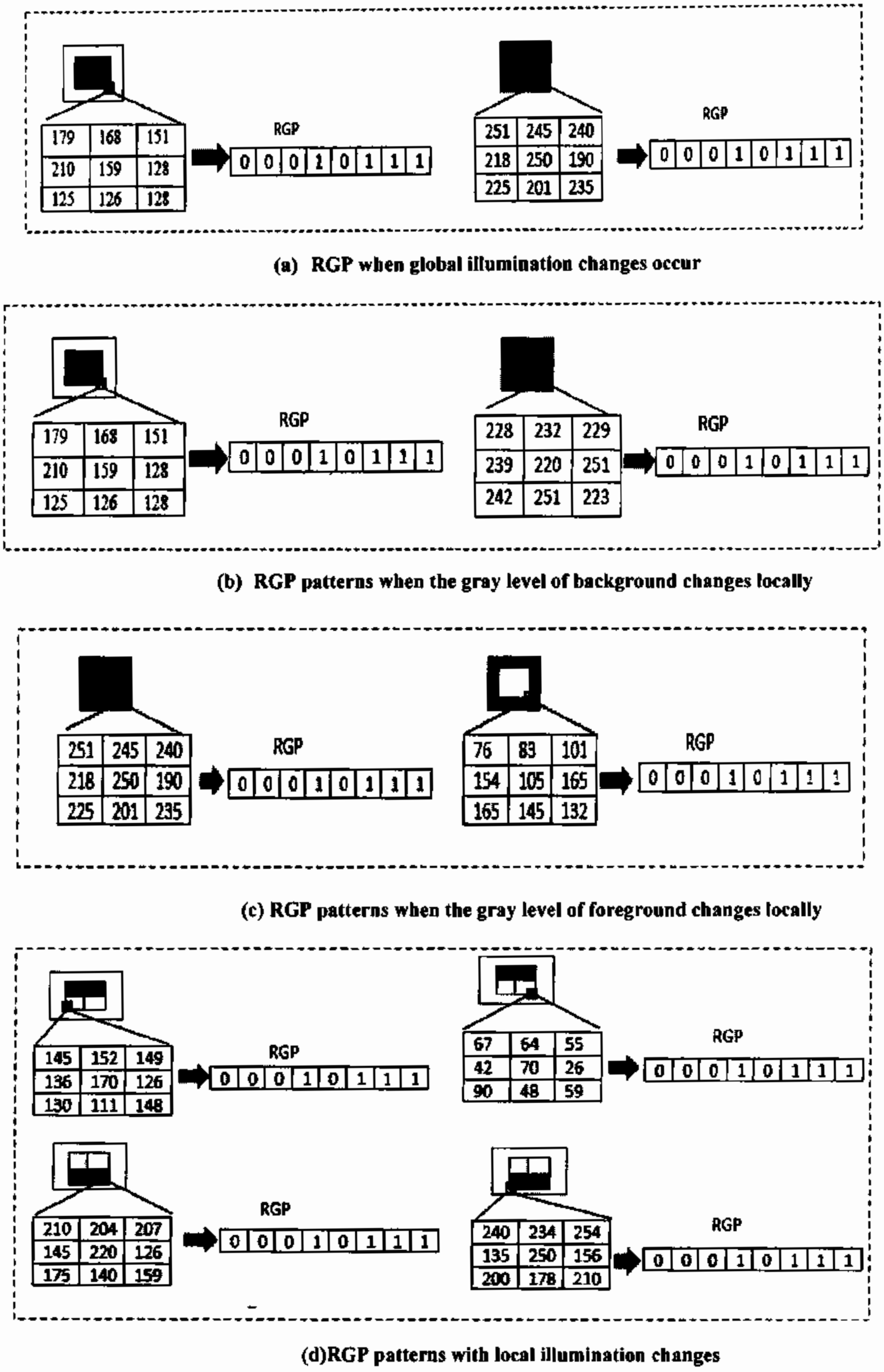


Figure 4. 3: Consistent patterns generated by RGP with global and local illumination changes

4.2.1: Pseudo code for RGP:

Algorithm RGP:

Input: set of Training images (TI=1..213), set of Class of training Images (C=1..7)
Output: RGP histogram

```
For i= 1 to size of Training Images
Image=TI(i)
Calculate size of training Image in rows(i), cols(i) where i ε TI .. range of I varies with size of training set
For row=1 to rows(i)
   For col=1 to cols(i)

            [r0,r1, r2, r3, r4, r5, r6, r7]=calculate difference matrix by taking absolute difference of central pixel in
neighborhood
            [avg]=find mean of difference window
            [rd0, rd1, rd2, rd3, rd4, rd5, rd6, rd7]= generate Boolean matrix by comparing pixels with avg
            RGP= convert [rd0, rd1, rd2, rd3, rd4, rd5, rd6, rd7] to decimal
    End
End
Generate RGP histogram
End
```

4.3 Experimental Setup:

Series of experiments are conducted in two datasets JAFFE [83] and SFEW[82]. JAFFE is synthetic dataset of 10 Japanese females exhibiting seven expressions including neutral. This is lab-controlled dataset in which images are taken in with ideal light variations. To validate impact of light variations images are degraded by applying global and local illumination changes.

RGP is further compared with local texture based techniques in two scenarios, holistic and zone based approach. In Holistic method feature extraction technique is applied on whole image. In zone based technique image is divided in nine equal parts and then each part is separately processed to extract features and finally histograms of all parts are concatenated.

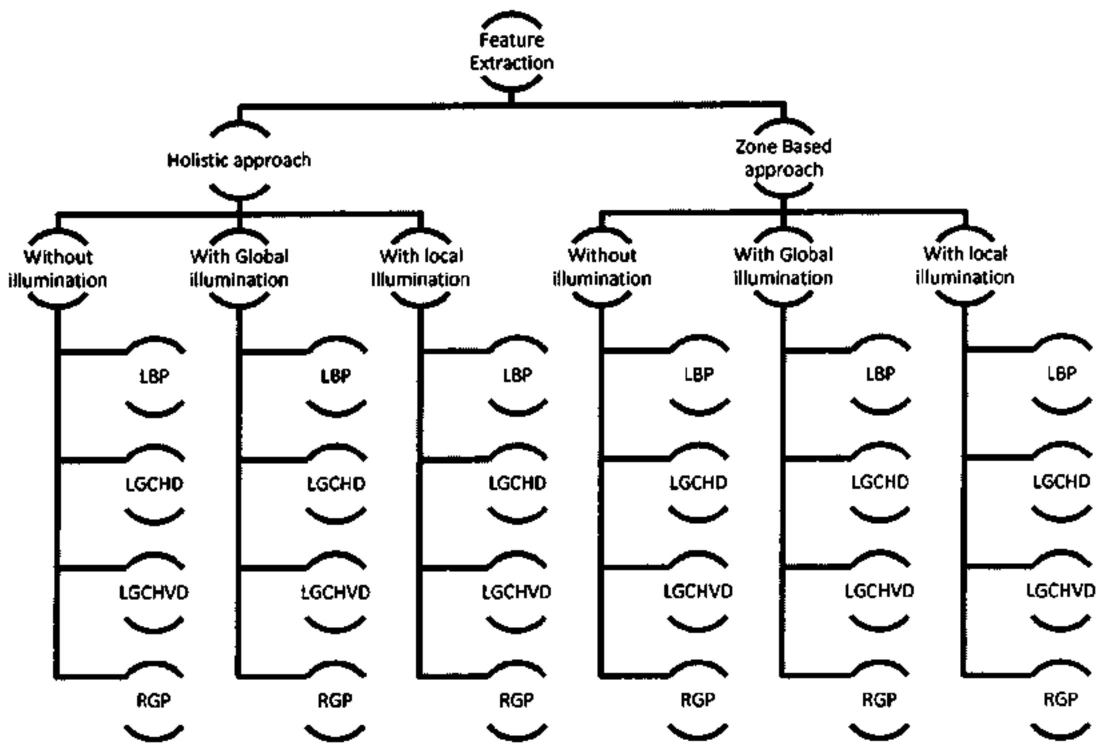


Figure 4. 4: Flow of RGP

4.3.1 Facial expression recognition on images without illumination changes

In the first experimental setup, JAFFE database is used. Initially face is detected from the background by using voila Jones algorithm. Then normalization is performed on images as preprocessing. Features of these normalized images are extracted using LBP, LGC, LGC-HD and RGP texture based techniques. Histograms are these features are constructed to generate feature vector of length 256. These feature vectors are then used in different classifiers to classify them in different expressions.

Figure 4. 5: Images without any illumination changes

4.3.2 Facial Expression Recognition Based On Images with Global Change:

In the second experimental scenario, global illumination change is applied on all normalized images with equal intensity. Features are extracted and feature vector is created in same way as explained in section 4.3.1.

Figure 4. 6: Images with global illumination Changes

4.3.3 Facial Expression Recognition Based On Images with Local Change:

In the third experimental setup, first of all image is divided in to four sub regions and then local illumination changes are applied on each of these sub parts with different intensities. These subparts are concatenated to form single image again. Now we have an image which is distorted by four different intensities. Feature vector is generated in the same way as given in section 4.3.1.

Figure 4. 7: Local Illumination changes in zone based Approach

4.4 Results:

RGP results are discussed by using two datasets, JAFFE and SFEW. JAFFE is used to verify results of RGP is all three scenarios i.e. with original image, with global illumination changes and with local illumination changes. As JAFFE is lab based data so there is uniform light in original image. Than JAFEE is distorted with global illumination and RGP is tested for its results. Finally, JAFFE is simulated with local variations in which light intensity is varied on each part of face as discussed in Figure 4.1 and 4.2 and then its results are generated. Further these set of experiments are tested in holistic approach and in zone based approach.
Another set of experiments are conducted on real noisy dataset that is SFEW.

4.4. 1: Results in Holistic Approach:

Experiments performed in holistic approach can be demonstrated in Figure 4.9. In Figure 4.9 explaining methodology in detail, three scenarios are discussed, RGP with original image, RGP with global illumination change and RGP with local illumination change. In case of global illumination change intensity of whole image is changed. While in local illumination changes image is divided in 2x2 sub parts and different proportions of light variations are introduced on these parts and these subparts are again combined to make a full image with illumination change in local parts called local change in full face image

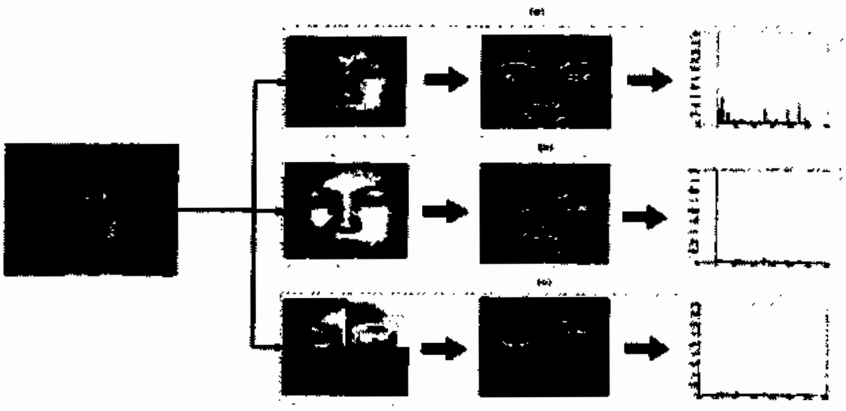


Figure 4. 8: Experiments in Holistic Approach

Results of LBP, LGCHVD, LGCHD and proposed RGP on full face image in above mentioned scenarios, It can be seen that RGP outperforms in all scenarios; without any illumination, it gives same performance as state of art technique, in case of local illumination changes and global illumination changes it give performance better from state of art LBP, LGC-HVD and LGC -HD. When there is not any illumination change in gray level of images LBP gives 96.71% accuracy LGCHVD gives 94.84% accuracy LGCHVD gives 96.71% accuracy and RGP 96.24% accuracy as shown in Table 4.1.

Table 4. 1: Results of LBP, LGCHVD, LGCHD and RGP in Holistic Approach (without illumination changes)

Technique	Accuracy (%) without illumination Changes	Accuracy (%) with global illumination changes	Accuracy (%) with local illumination changes
LBP	96.7136	89.2019	92.9577
LGCHVD	94.8357	91.5493	95.7746
LGCHD	96.7136	92.0188	92.4883
Proposed RGP	96.2441	95.7746	96.2441

When global changes occur in images LBP gives 89.20% accuracy, LGCHVD gives 91.55% accuracy, LGCHD gives 92.02% accuracy and RGP gives 95.77% accuracy as given in Table 4.1, with local change occur accuracy of LBP reduces to 92.96% accuracy of LGCHVD reduces to 95.77% and accuracy of LGCHD is 92.49% while accuracy of RGP is 96.24% accuracy as demonstrated in Table 4.1.

Thus it can be seen that RGP has stronger discriminate power than other texture based technique. It increases the ability to detect facial expression in illumination variance, because of its gradient preservation. Gradient difference make it robust to noise and light variations. This result can be summarized in chart in Figure 4.10. Further it is clear from the Figure 4.10 that for same images RGP generate consistent patterns that is why performance rate of RGP is same in all three scenarios (original image, global illumination and local illumination), while in case of LBP and LGC inconsistent patterns are generated, thus performance of these vary with different light variations. Further it can be observed that LBP and LGC give good results with original image only. With light variations, their accuracy degrades with light variation either it is global or local.

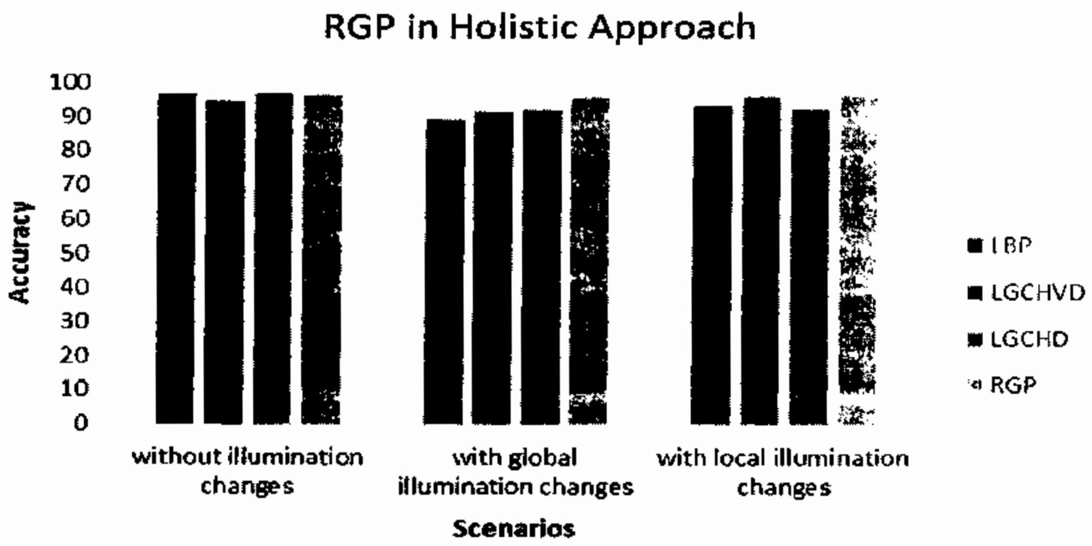


Figure 4. 9: Accuracy chart of LBP, LGCHVD, LGCHD and RGP in Holistic Approach

Confusion matrix of RGP in all these tree scenarios is given in Table 4.4, 4.5 and 4.6.

Table 4. 2: Confusion Matrix of RGP (without illumination changes)

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	29	1	0	0	0	1	0
Disgust	0	28	1	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	0	29	2	0	0
Neutral	0	0	0	1	29	0	0
Sad	0	0	0	0	1	29	1
Surprise	0	0	0	0	0	1	29

This confusion matrix shows that true positive rate of RGP is good in almost all expressions of the database. It is very clear that fear can be classified with 100% accuracy. While there are few misclassifications in sad surprise, neutral, happy, disgust and anger.

Table 4. 3: Confusion Matrix of RGP (with global illumination changes)

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	28	1	0	0	1	0	0
Disgust	1	26	1	1	0	0	0
Fear	0	2	28	2	0	0	0
Happy	0	0	1	28	2	0	0
Neutral	0	0	1	3	23	3	0
Sad	0	0	0	0	4	24	3
Surprise	0	0	0	0	0	7	23

Table 4.6 shows that results are slightly varied in case of light variations but still RGPs True positive rate us satisfactory.

Table 4. 4: Confusion Matrix of RGP (with local illumination changes)

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	1	28	0	0	0	0	0
Fear	0	1	31	0	0	0	0
Happy	0	0	2	29	0	0	0
Neutral	0	0	0	0	30	0	0
Sad	0	0	0	0	2	29	0
Surprise	0	0	0	0	0	2	28

Main contribution of RGP is in case of local illumination changes. It is very clear from the confusion matrix that even in case of local changes gradient difference can still measure expressions with high accuracy. Where results of LBP, LGC and LGC-HVD reduce considerably. Though LGC and LGCHVD also uses direction difference but gradient difference and using same as threshold have much better impact in FER

Confusion matrices show that accuracy of RGP in classification of individual expression is best in case of local illumination changes where state of art techniques cannot perform well.

4.4.2: Results in Zone Based Approach:

Experiments performed in zone based approach are demonstrated in following Figure 4.14. In this approach, we divided the face image in to 3x3 local parts and use these sub-regions as input then illumination change is applied on the whole image and divide this image in sub regions for further processing. For local illumination change we have changed the intensity level of 3x3 sub regions with different proportions.

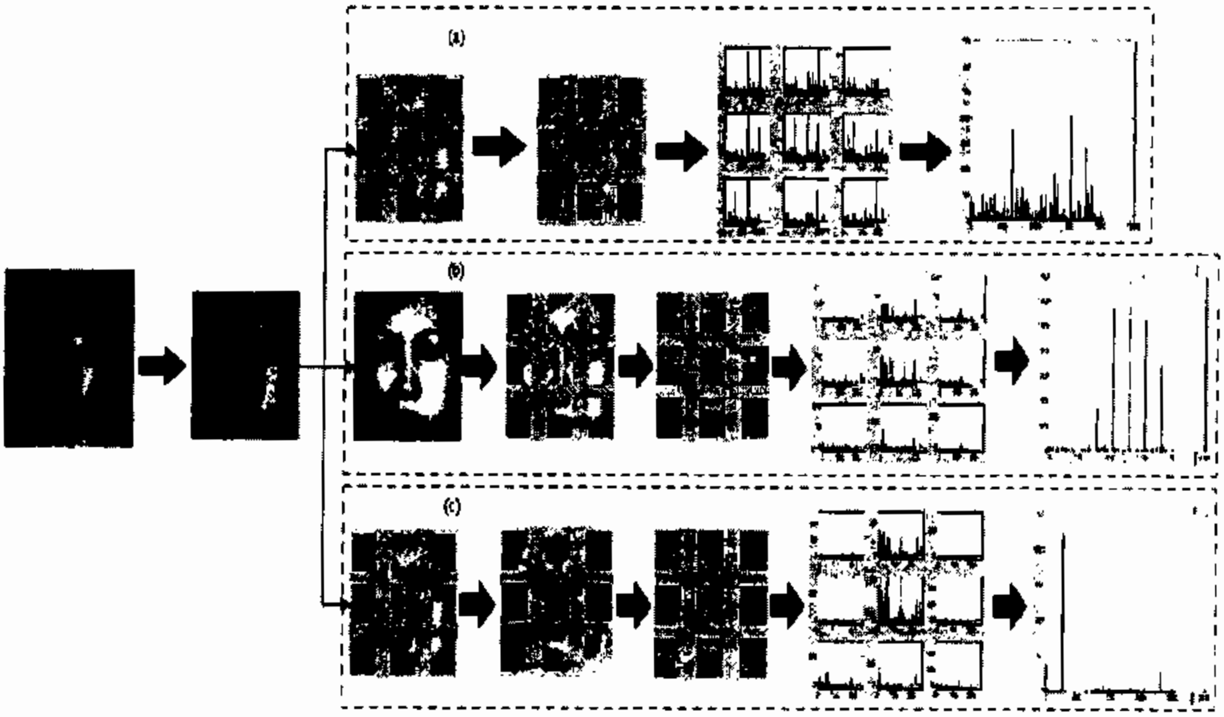


Figure 4. 10: Experiments in Zone Based Approach

Result of LBP, LGCHVD, LGCHD and proposed RGP in zone based approach in all three scenarios as in holistic approach. RGP outperforms in all scenarios; without any illumination, it gives same performance as LBP and LGC, in case of local illumination changes and global illumination changes RGP perform best. It gives performance better from state of art LBP, LGC-HVD and LGC -HD. When there is not any illumination change in gray level of images

LBP gives 96.71% accuracy LGCHVD gives 94.84% accuracy LGCHVD gives 96.71% accuracy and RGP 96.24% accuracy as given in Table 4.7

Table 4. 5: Results of LBP, LGCHVD, LGCHD and RGP in Zone Based Approach

Technique	Accuracy(%) without illumination changes	Accuracy (%) with Global illumination changes	Accuracy (%) with Local illumination changes
LBP	93.4272	89.2019	89.2019
LGCHVD	93.4272	87.7934	88.2629
LGCHD	93.4272	88.7324	90.1408
RGP	97.1831	88.7324	96.2441

When global changes occur in images LBP gives 89.20% accuracy, LGCHVD gives 91.55% accuracy, LGCHD gives 92.02% accuracy and RGP gives 95.77% accuracy as given in Table 4.7.

when local change occur accuracy of LBP reduces to 92.96% accuracy of LGCHVD reduces to 95.77% and accuracy of LGCHD is 92.49% while accuracy of RGP is 96.24% accuracy so our purposed technique has stronger discriminant power than other texture based technique and RGP can increase the ability to detect facial expression better as it has given the greater accuracy in all scenarios.

Summary of above results is given in the form of Figure 4.14. It is shown in the Figure 4.14 that in zone based approach RGP perform better than LBP and LGC , this due to fact that RGP maintain Gradient difference that minimize impact of light variations.

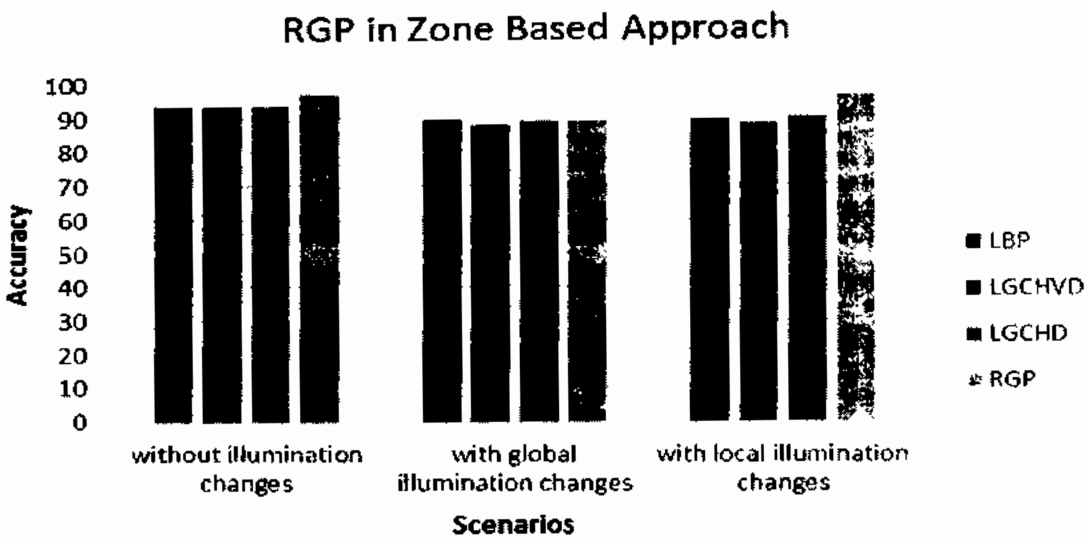


Figure 4. 11: Accuracy Chart of LBP, LGCHVD, LGCHD and RGP in Zone Based Approach

Expression level accuracy of RGP is depicted in confusion matrix given in all scenarios.

Table 4. 6: Confusion Matrix of RGP (without Illumination Changes)

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	28	1	0	0	0	1	0
Disgust	0	27	2	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	0	31	0	0	0
Neutral	0	0	0	1	28	1	0
Sad	0	0	0	0	0	31	0
Surprise	0	0	0	0	0	0	31

Table 4.7 shows that RGP can accurately identify expression in zone based approach as well.

Table 4. 7: Confusion Matrix of RGP (with Global Illumination Changes)

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	28	1	0	0	0	0	1
Disgust	0	29	0	0	0	0	0
Fear	0	6	22	4	0	0	0
Happy	0	0	2	29	0	0	0
Neutral	0	0	0	5	22	3	0
Sad	0	0	0	0	0	30	1
Surprise	0	0	0	0	0	1	29

Table 4.8 reflect same results as discussed in Table 4.11, there are misclassifications in case of neutral and fear that has impact on overall accuracy as well. While anger, disgust, sad and surprise are classified correctly.

Table 4. 8: Confusion Matrix of RGP (with Local Illumination Changes)

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	29	0	0	0	0	0	1
Disgust	0	26	3	0	0	0	0
Fear	0	0	31	1	0	0	0
Happy	0	0	2	29	0	0	0
Neutral	0	0	0	1	29	0	0
Sad	0	0	0	0	0	31	0
Surprise	0	0	0	0	0	1	30

RGP in case of local illumination perform better than that of global illumination. It can be strongly agreed from Figure 4.8 which shows that true positive rates of almost all expression are better than global illumination case. In case of global illumination in Figure 4.17 there were maximum misclassification in case if fear and neutral expression. While in case of local illumination there is only one misclassification in both cases which shows that results are improved remarkably. From expression level distinction, it is obvious that RGP perform good in all scenarios.

4.4.3: Results in SFEW

In section 4.4.1 and 4.4.2 results are discussed with simulated data. JAFFE dataset is lab based dataset. In which global and local illumination changes are introduced to verify result of proposed technique in all scenarios of light variations. Results are further investigated with real noisy data that is SFEW. This dataset is already explained in chapter 3. As these images are taken from the movies so it has almost same light challenges as in real world.

Table 4. 9: Results of RGP with SFEW

Techniques →	LBP	LGC-HD	LGC-HVD	RGP
↓ Classifiers	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
Simple Logistics	78.23	92.35	92.94	94.15
KNN	70	92.35	93.52	95.9
Bagging	69.41	88.23	87.05	91.81

Gradient of pixel window carry very strong information about the features of face. It is -evident by testing RGP in real noisy data as well. SFEW is challenging dataset in terms of expression description and light variations. Results in RGP shows that if we preserve gradient information and extract features gradient then it is more useful than straight sign difference or simple directional difference. Findings of RGP are also shown in accuracy chart given in Figure 4.12.

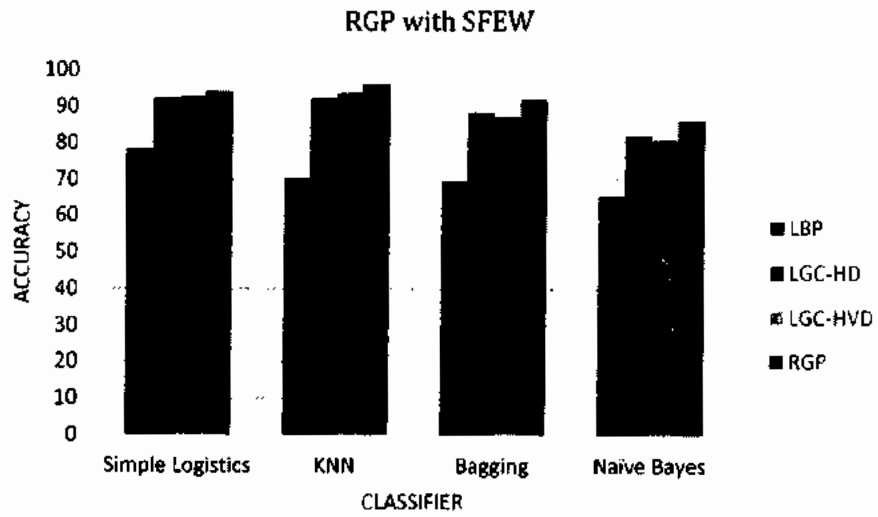


Figure 4. 12: Accuracy Chart of RGP with SFEW

Figure 4.12 explicitly show that performance of RGP is better in real noisy data than LBP, LGC and its variations. This emphasis the fact that gradient difference play very important role in expression recognition. Confusion matrix of RGP when tested on SFEW is given in Table 4.10

Table 4. 10: Confusion Matrix of RGP with SFEW

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	25	0	0	0	0	0	1
Disgust	1	15	2	0	0	0	0
Fear	0	0	23	1	0	0	0
Happy	0	0	0	28	1	0	0
Neutral	0	0	0	0	28	1	0
Sad	0	0	0	0	0	21	0
Surprise	0	0	0	0	0	0	24

Confusion matrix show that RGP can identify anger, sad and surprise with 100% accuracy in challenging light conditions with real noisy data. There are a few misclassifications in case of neutral happy and fear.

Precision recall curves of RGP with SFEW is also given to validate results.

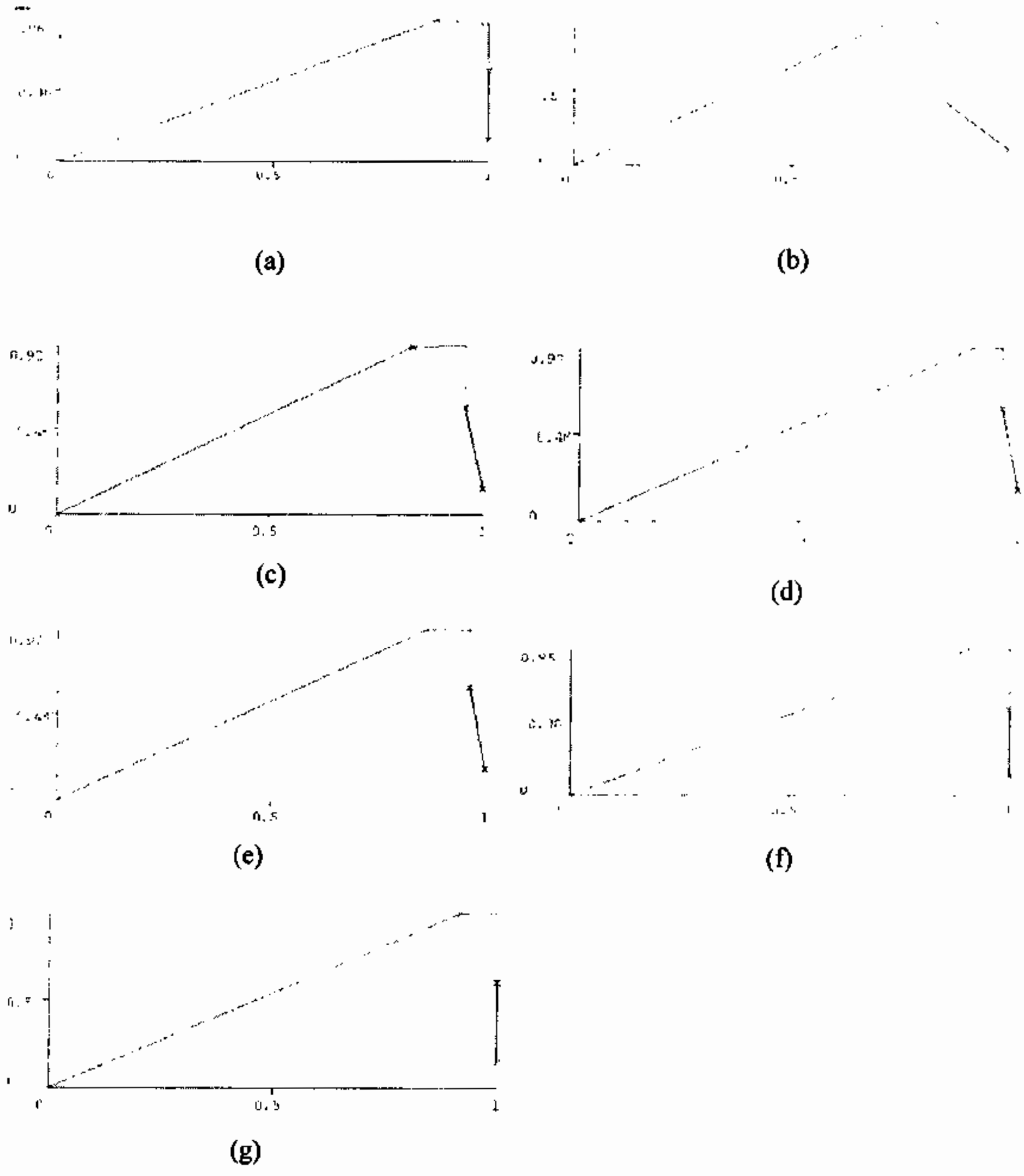


Figure 4. 13: Precision Recall Curves for Anger (a), Disgust (b), Fear(c), Happy (d), Neutral (e), Sad (f) and Surprise (g)

AUC is used to determine precision recall curve. It is shown in Figure 4.13 that precision recall curves fall the expressions give satisfactory results.

4.4.4 T-Test RGP

This is statistical significance test. In T-Test conducted for feature extraction techniques LBP and RGP, Null hypothesis is LBP=RGP that is there is no relationship between LBP and RGP. If $p<0.05$ it means there is significant difference and RGP is better than LBP.

Table 4. 11: Significance Test of LBP and RGP

t	T Test is used
(298)	degree of freedom
2.18	t statistics
p	0.014; as $p < 0.05$ it shows
	Difference is significant; null is rejected.

Among feature extraction techniques (N=300), there was statistically significant difference among LBP (M=0.354, SD=0.43) and MSBP (M=0.099, SD=0.27), $t(298)=2.18$, $p<0.05$, $CI_{0.95}$ -0.12,0.0365. Thus we reject null hypothesis. There is significant difference between RGP and LBP.

4.5 Chapter Summary:

Light variation has serious impact on the accuracy of expression recognition. These light variations can be further categorized as global and local light variations. Global light variations are due to some issues in source of light whereas local variations are due to shadows, around the edges as eye brows, lines or wrinkles. RGP is novel approached proposed, that is invariant to the both global and local intensity variations because of its gradient preservance. Though LBP, LGCHVD and LGCHD are also insensitive to global illumination but in case of local illumination these techniques failed to perform. By taking gradient difference minute details like eye brows, side wrinkles lines such information is also retained. It is because RGP maintain gradient difference, which generates consistent pattern regardless of changing situations of light intensity variations in facial expressions. The experimental results confirm the effectiveness of RGP. The results show that RGP detect expression with more accuracy when tested on JAFFE database with three scenarios, on original images with accuracy of 96%, images with global illumination changes 95% and images with local illumination changes 96%. Further results are investigated with SFEW which shows that RGP outperforms with noisy data as well with accuracy of 91%.

Chapter # 5

Local Texture Based Techniques for Accommodating occlusion and Multi Face Facial Expression Recognition

5. Local texture Based Techniques for Occlusion Accommodation and Multi Face Expression Recognition:

This chapter covers two challenging issues in FER in innovative dimensions. These two issues are Accommodating occlusion and recognizing crowd behavior by using FER. Occlusion accommodation is given in Section 5.1 and Crowd behavior recognition is given in section 5.2.

5.1: Accommodating Occlusion in FER:

Most important component in determining FER are facial features. If these features are hidden because of any reason, accuracy of FER is decreased. In this chapter these features are emphasized by hiding them and then classifying expressions. Novelty in this work is introducing nine different types of occlusions simulation. Main contribution in this chapter is on occlusion simulation. It includes different type, size and shape of occlusions are introduced. Section 5.1.1 is about motivations of investigating occlusion. Section 5.1.2 explain details of occlusion simulation. Majorly eye, mouth and miscellaneous occlusions are simulated. The toughest of occlusion simulation that will cover real data is miscellaneous one. Section 5.1.3 is about image preprocessing. In section 5.1.4 feature extraction is discussed. Three local texture based techniques are used for feature extraction LBP, LGC and CLBP. This is also novel in this dissertation that CLBP and LGC are tested first time for occlusions in FER. Section 5.1.5 contains detailed results of all these feature extraction techniques with occlusions.

5.1.1 Motivation:

Occlusion in an image refers to difficulty in the view of an object. Occlusion generally observed in face images can normally be either natural or synthetic. Natural occlusion indicates to obstruction of views between the two image objects without any intension while synthetic occlusion refers to artificial barrier of intentionally covering the image's view with a white/black solid rectangular block [84]. Partial occlusion has been observed in many areas of image processing. Even in real time application face image intentionally becomes occluded via accessories such as:

- Sunglasses, scarf, beards, hat
- Hand on face
- Face dirt
- Face behind Fence
- Texture on Face images [85]

The research for using an efficient feature extraction technique, which can extract a good feature subset. So, that recognition rate can be increased even in case of occlusion.

5.1.3 Occlusion Simulation:

Occlusion simulation is first step in which different types of occlusions are introduced in existing dataset, JAFFE [11]. In this thesis nine types of occlusion are inserted in the images. These occlusions are broadly categorized as eye occlusions, mouth occlusion and face occlusion. Finally all type of occlusions are merged to make single miscellaneous occlusion dataset on which algorithm is trained and satisfactory results are found. Moreover most research in occlusion done so far is rectangular shape but in this research another challenge of different shapes of occlusions are also introduced. This is because of different shapes of masks, hairstyles, glasses and other type of accessories that might be used. In Figure 5.21 simulation of occlusion types are discussed. Mainly occlusion is divided into four main categories Eye occlusion, Mouth occlusion, face occlusion and miscellaneous occlusion. In eye occlusion we further have three categories, single eye occlusion both eye occlusion and full eye occlusion. In case of single and both eye occlusion only eyes are occluded, it mainly cover the scenarios of using masks of various filters having glasses on eyes. Further in different types all styles of glasses can be accommodated. In full eye occlusion eyes along with eye brows are occluded. This case includes covering head with scarf and glasses and also there are some sort of hair styles covering this area. Next category of occlusion is Mouth occlusion. In mouth occlusion further two types of occlusion is handled. It is only lip area or lip area with side wrinkles are also hidden. Lips are very important in determining expression of happiness, sadness and surprise. Then along with lips either side wrinkles are also important or not for this such investigation is also conducted. Next type of occlusion is Face occlusion in which left or right profile of the face is occluded. Then miscellaneous occlusion covers three types of occlusions. Single eye and lip, both eyes and lips these are also some styles or masks that can be used. Most challenging simulation of occlusion is all above mention occlusions are randomly added in a dataset. All these types of occlusions are given with examples in subsections 5.1.3.1 to 5.1.3.4

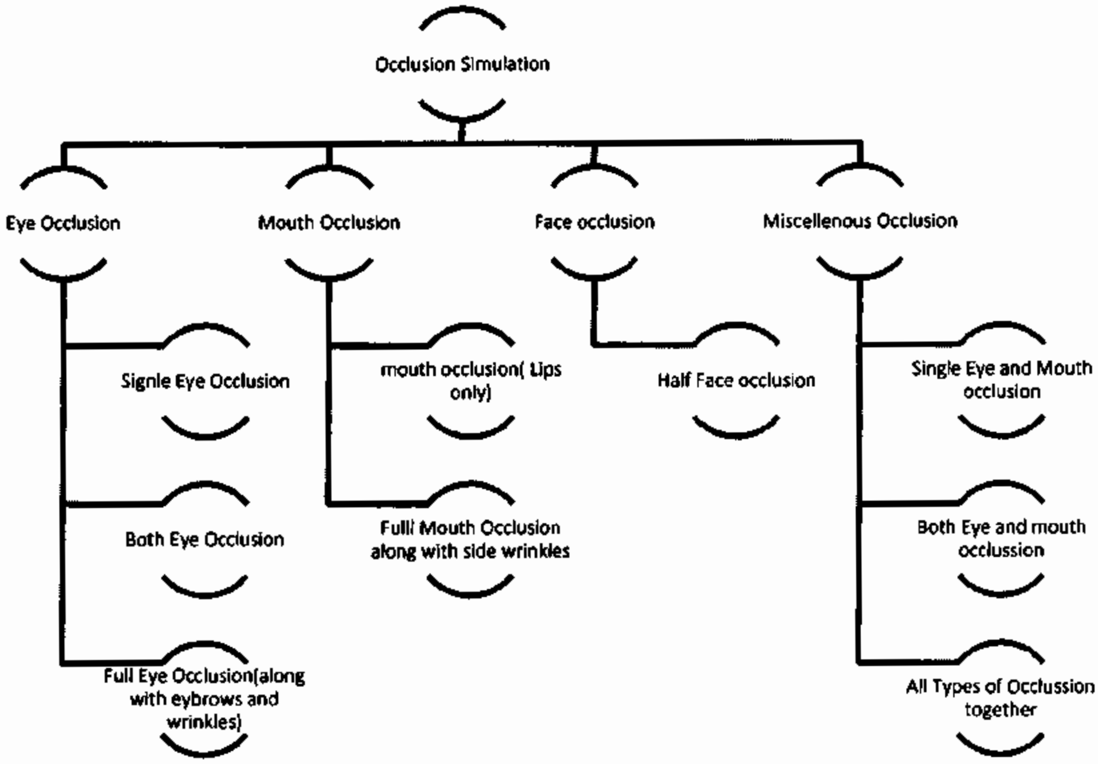


Figure 5. 1: Types of occlusion

5.1.3.1 Eye Occlusion:

In eye occlusion single eye, both eyes and full eye occlusion are included. In single eye one of the either right or left is occluded with different styles of obstacles. In both eye occlusion both eyes are covered .in full eye both eye along with eye brows and wrinkles between the eyebrows are covered with different shapes. All these three kinds of occlusions shown in following figures.

Figure 5. 2: Examples of Single Eye Occlusion

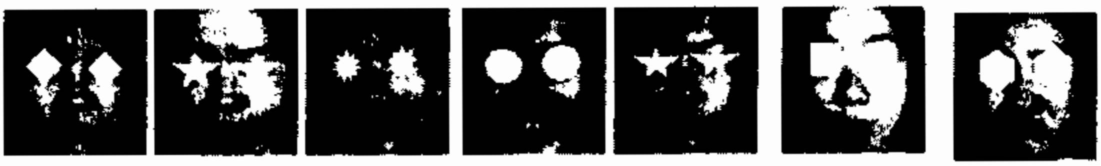


Figure 5. 3: Example of both eye occlusion

Figure 5. 4: Example of full eye occlusion

5.1.3.2 Mouth Occlusion:

In mouth occlusion only lip region and lip area along with side wise wrinkle both are occluded to see impact of lips and wrinkles in expression recognition. These simulations are shown below

Figure 5. 5: Example of Mouth Occlusion

Figure 5. 6: Example of Full Mouth Occlusion

5.1.3.3 Face Occlusion:

In face occlusion half face is occluded in order to compensate profile view.

Figure 5. 7: Example of half face occlusion

5.1.3.4 Miscellaneous Occlusion:

In miscellaneous occlusion three types of occlusion are discussed. Single eye and lip occlusion, both eye and lip occlusion and in third type all previously discussed types of occlusions are mixed in one data set. Simulations of these occlusions are given below.

Figure 5. 8: Example of both Eye and Mouth Occlusion

Figure 5. 9: Example of Single Eye and Lip Occlusion

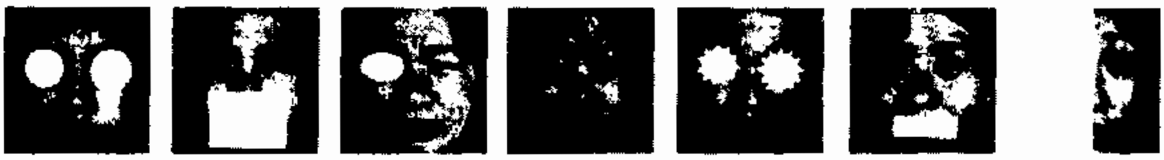


Figure 5. 10: Example of Miscellaneous Occlusion

5.1.4 Image Preprocessing:

In Image preprocessing techniques are applied to equalize these parameters at general on whole frame. In image preprocessing image is normalized.

5.1.5 Feature Extraction:

Successful feature extraction is possible only when images are taken in controlled conditions. In this thesis we are comparing different variation of LBP to extract features. Novelty of work is using LGC and CLBP first time for handling occlusion. And robustness and accuracy of these techniques will also be compared to state of art techniques i.e. LBP.

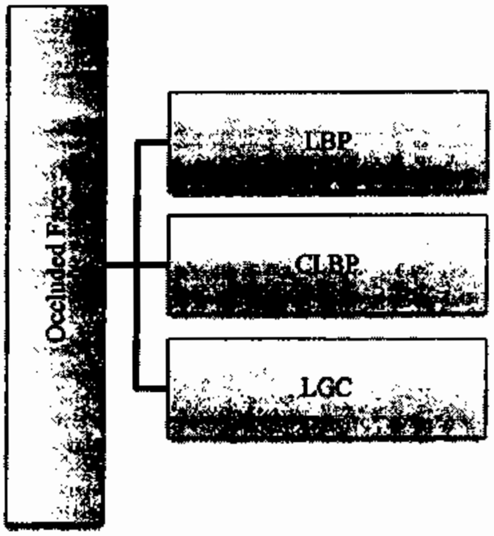


Figure 5. 11: Block Diagram of feature extraction methods in occluded face images

All these feature extraction techniques are already discussed in section 2.6. Methodology and pseudo code of CLBP is given.

Algorithm for CLBP:

Input: set of Training images (TI=1..213), set of Class of training Images (C=1..7)
Output: CLBP histogram

```
For i= 1 to size of Training Images
Image=TI(i)
Calculate size of training Image in rows(i), cols(i) where i ε TI .. range of I varies with size of training set
For row=1 to rows(i)
  For col=1 to cols(i)
        [p0, p1, p2, p3, p4, p5, p6, p7]=compare central pixel with its neighborhood
        Find out average of selected window
        [a0, a1, a2, a3, a4, a5, a6, a7]=Compare average value of window with boundary pixels in window
        Generate sub codes of CLBP on basis of 4 neighbor and 8 neighbor
        CLBP1(row,col)=convert this binary string [p6a6p4a4p2a2p0a0] to decimal
        CLBP2(row,col)=convert this binary string [p5a5p3a3p1a1p7a7] to decimal
   End
End
Make histogram of the CLBP1 obtained.
Make histogram of CLBP2 obtained.
Concatenate both histograms.
End
```

5.1.5 Experimental Setup and Results:

Experiments are conducted in both scenarios of holistic face and zone based approach. For experimentation we have to use simulated dataset as in all previous studies simulated dataset is used with limited occlusion. In this work we introduce nine different types of occlusions. Further shape and size of occlusion also vary to see performance of CLBP in these scenarios. Different shape and size of occlusion also give impact of random occlusion and it covers real time occlusions as well. In holistic approach feature extraction is applied on whole face while in zone based approach face image is divided in 3x3 zones and features of each zone are extracted and then concatenated to form feature vector. Results are discussed categorically as occlusion simulation discussed earlier.

5.1.5.1 Single Eye Occlusion:

A type of occlusion is introduced where single eye is hindered with different shapes. Covering only single eye is subject to find that how much an eye contributes to expression recognition. Figure 5.12 has a sample image to show single eye occlusion used in experimentation.

Figure 5. 12: Single Eye Occlusion

***a) Examining* Feature Extraction Techniques on Full Face Image:**

A full face image is taken, feature extraction techniques are applied on it and their histograms are taken to the classifiers as input.

Table 5. 1: Results of LBP, LGC-HD, LGC-HVD and CLBP with Single Eye Occlusion for Full Face Image

Technique	**LBP**	**LGCHD**	**LGCHVD**	**CLBP**
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	98.12%	91.07%	92.01%	99.53%
Simple Logistics	97.65%	93.42%	93.42%	98.12%
KNN	97.18%	95.77%	92.01%	98.59%
Bagging	99.06%	92.48%	96.24%	99.53%
Naïve Bayes	76.99%	84.97%	84.50%	77.46%

CLBP features attained highest classification rate of 99.53% against other extraction techniques. Accuracy of LBP features cannot be ignored i.e. 99.06%, closest to the best results. Hence, CLBP features for full face work best when only single eye is hindered. This is because gradient magnitude might be misled by occluded features but intensity magnitude works better along with sign difference.

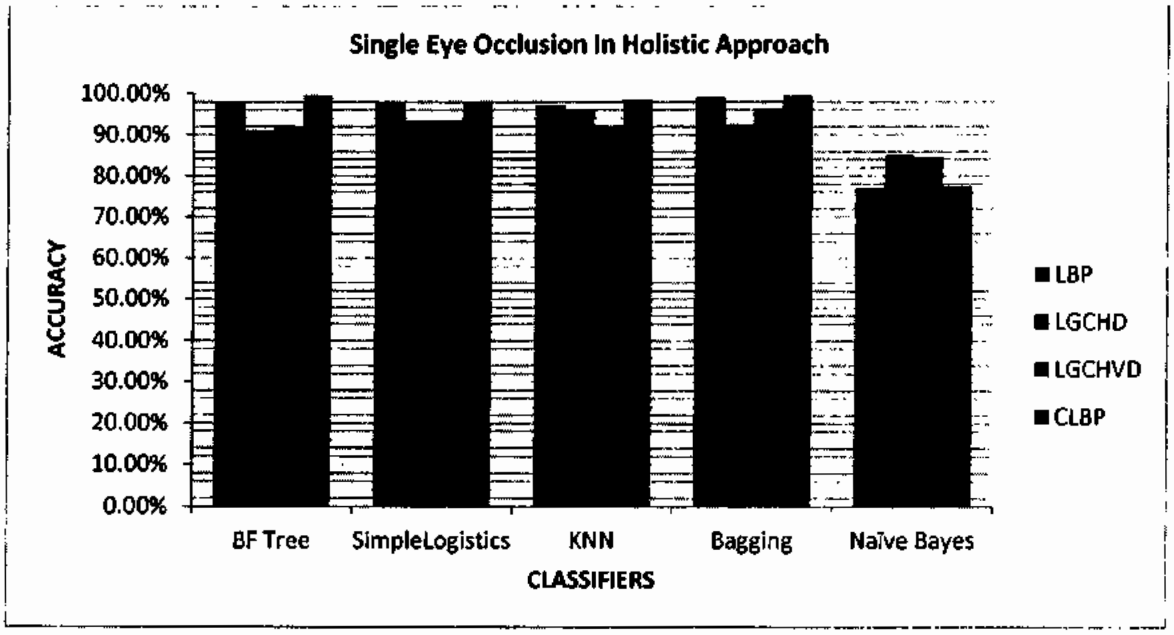


Figure 5. 13: Results of LBP, LGC-HD, LGC-HVD and CLBP with Single Eye Occlusion for Full Face Image

Confusion matrix of CLBP is given in Table 5.2.

Table 5. 2: Confusion Matrix of CLBP classified with Bagging

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	29	0	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	1	30	0	0	0
Neutral	0	0	0	0	30	0	0
Sad	0	0	0	0	0	31	0
Surprise	0	0	0	0	0	0	30

Above matrix gave 100% recognition for expressions like anger, disgust etc. whereas, in comparison to other expressions minor decline in accuracy occur when instance of happy is classified as fear.

b) Examining Feature Extraction Techniques in zone based Approach:

Face image is first divided in nine parts; extraction technique is than applied on every part and histograms of all parts are concatenated to form a single histogram for classification. The process is repeated for all four techniques.

Table 5. 3: Results of LBP, LGC-HD, LGC-HVD and CLBP with Single Eye Occlusion for Zone based approach

Technique□ Classifier□	LBP Accuracy (%)	LGCHD Accuracy (%)	LGCHVD Accuracy (%)	CLBP Accuracy (%)
BF Tree	83.56%	80.75%	78.40%	88.73%
Simple Logistics	91.54%	85.91%	86.85%	91.07%
KNN	93.42%	71.83%	60.56%	92.95%
Bagging	89.20%	86.38%	84.50%	91.07%
Naïve Bayes	86.38%	86.85%	79.81%	84.03%

From the Table 5.3 showing results of all four extraction methods, LBP and CLBP perform better with most classifiers.

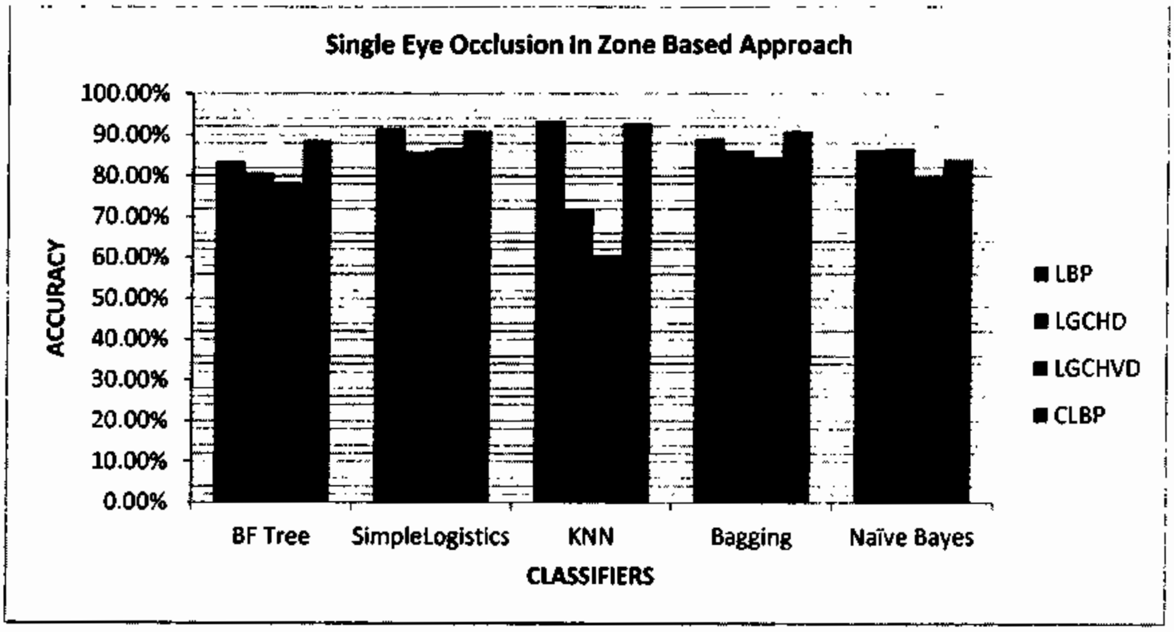


Figure 5. 14: Results of LBP, LGC-HD, LGC-HVD, MSBP and CLBP with Single Eye Occlusion for Zone based approach

In the Table 5.4 confusion matrix is discussed in zone based approach.

Table 5. 4: Confusion Matrix of LBP classified with KNN

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	28	1	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	0	29	2	0	0
Neutral	0	0	0	0	28	2	0
Sad	0	0	0	0	2	25	4
Surprise	0	0	0	0	0	3	27

Misclassifications degrading the overall accuracy for example, neutral and surprise as sad, happy as neutral, sad as neutral and surprise could be observed in the confusion matrix in Table 5.4.

With single eye occlusion both CLBP in holistic approach and LBP in division approach, gave the best accuracies. Comparing these portrays that holistic approach with CLBP extracted features generated the best recognition rate i.e.99.53%.

5.1.5.2 Both Eyes Occlusion:

Eyes are incredibly important when identifying a person's expression. As most of the emotions could be identified via eyes, that makes them a key feature in terms of expression recognition. Different size and shapes of occlusion is introduced around eyes to get the importance of features other than eyes.

Figure 5. 15: Both Eyes occlusion

a) Examining Feature Extraction Techniques on Full Face Image:

Features are extracted from the face image with both eyes occluded and a histogram representing features of the whole face image are given to the classifier to estimate expression.

Table 5. 5: Results of LBP, LGC-HD, LGC-HVD and CLBP with Both Eyes Occlusion for Full Face Image

Technique	LBP	LGCHD	LGCHVD	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	97.65%	91.54%	88.73%	97.18%
Simple Logistics	96.24%	92.01%	91.07%	97.65%
KNN	95.30%	92.01%	92.01%	95.77%
Bagging	98.12%	90.14%	92.01%	97.18%
Naïve Bayes	76.99%	84.50%	86.38%	78.40%

Features extracted by LBP, attained maximum accuracy i.e. 98.12% Thus, LBP features extracted from full face gained maximum success rate with both eyes occlusion.

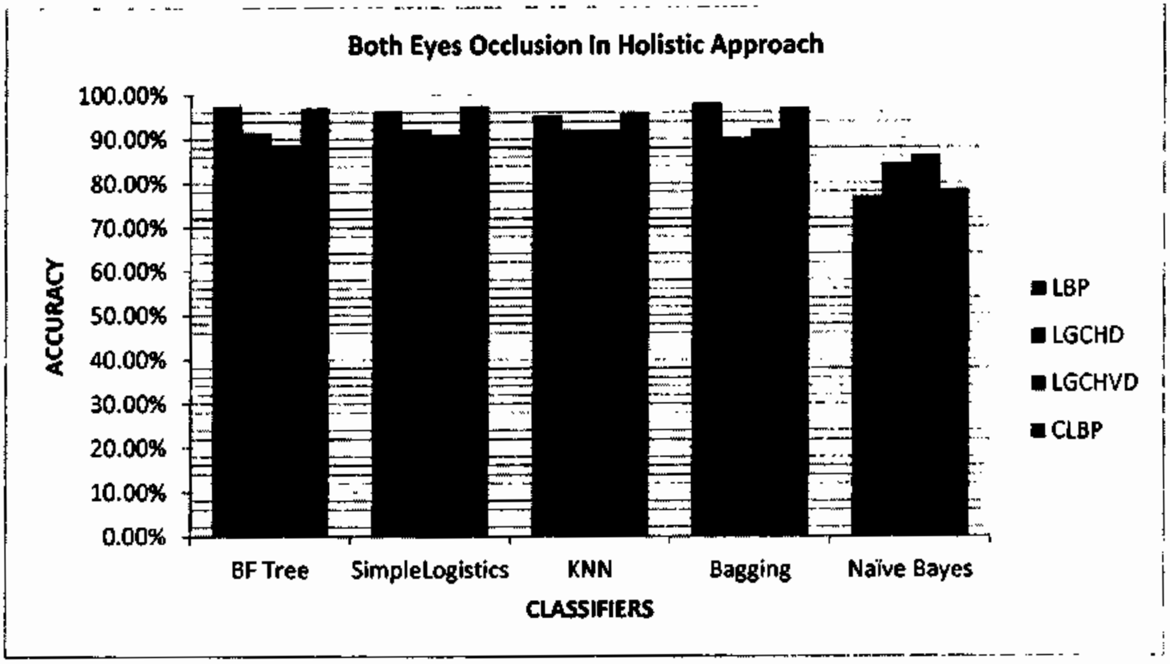


Figure 5. 16Results of LBP, LGC-HD, LGC-HVD, and CLBP with Both Eyes Occlusion for Full Face Image

Confusion matrix is given in Table 5.6.

Table 5. 6: Confusion Matrix of LBP classified with Bagging

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	29	1	0	0	0	0	0
Disgust	0	29	0	0	0	0	0
Fear	0	0	31	1	0	0	0
Happy	0	0	1	30	0	0	0
Neutral	0	0	0	0	30	0	0
Sad	0	0	0	0	0	30	1
Surprise	0	0	0	0	0	0	30

Eyes help a lot in expressing feelings, yet occluding them resulted 100% accuracy for surprise and disgust which lead to the fact that other texture feature also contain a lot of information. Other expressions contain minor misclassifications such as happy as fear, sad as surprise etc.

b) Examining Feature Extraction Techniques on Zone based approach:

For classification, a concatenated histogram containing information of all the nine parts is given to the classifier as input.

Table 5. 7: Results of LBP, LGC-HD, LGC-HVD and CLBP with Both Eyes Occlusion for Zone based approach

Technique	LBP	LGCHD	LGCHVD	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	85.44%	81.22%	81.69%	90.14%
Simple Logistics	92.48%	86.38%	88.26%	89.20%
KNN	95.30%	72.77%	61.97%	92.48%
Bagging	87.79%	89.67%	90.14%	92.01%
Naïve Bayes	85.91%	84.50%	81.69%	80.28%

From the results in Table 5.7, it is observed that among other extraction techniques, LBP assessed best result i.e. 95.30%. On the other hand CLBP reached a decline and achieved an accuracy of 92.48%. Hence LBP gives best result, when image is divided into zones.

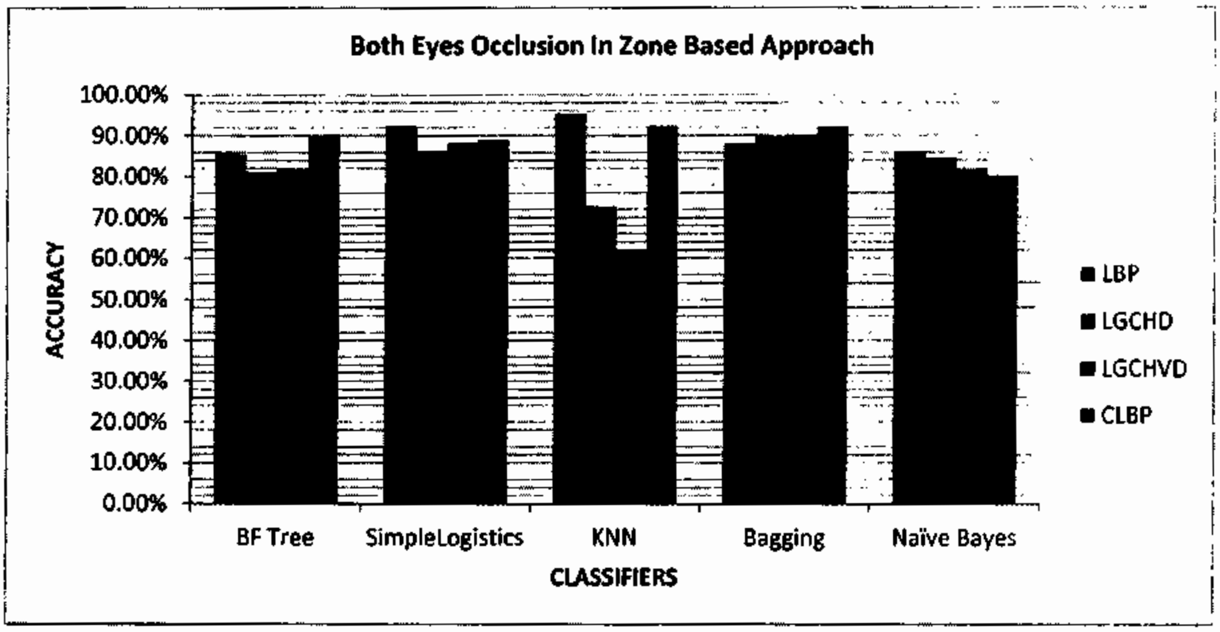


Figure 5. 17: Results of LBP, LGC-HD, LGC-HVD and CLBP with Both Eyes Occlusion for Zone based approach

Confusion matrix in Table 5.8 explains that feature other than eyes also contribute a lot in expression recognition. Dataset containing both eyes occlusion with LBP recognized anger, fear and happy to exactness. There are also some misclassifications e.g. disgust as fear and surprise as sad. But majorly sad as neutral and surprise leaded to a decrease in accuracy rate.

Table 5. 8: Confusion Matrix of Both Eye Occlusion in Zone based Approach

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	28	1	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	0	31	0	0	0
Neutral	0	0	0	0	29	1	0
Sad	0	0	0	0	2	25	4
Surprise	0	0	0	0	0	2	28

On the whole, comparing success rate of holistic and division based approach LBP gained highest recognition rate with both eyes occluded.

5.1.5.3 Full Eyes Occlusion:

To find the importance of eyes along with eyebrows and wrinkles, occlusion is introduced to hide complete eyes area. Wrinkles between eyes also help to differentiate between expressions. Thus, hiding this area provide significance of other facial parts.

Figure 5. 18: Full Eye Occlusion

a) Examining Feature Extraction Techniques on Full Face Image:

A histogram representing extracted features of full face is advanced towards the classifier for evaluation of results. Images contain full eyes occlusion are processed to extract feature and their histograms are generated and analyzed in Table 5.9.

Table 5. 9: Results of LBP, LGC-HD, LGC-HVD and CLBP with Full Eyes Occlusion for Full Face Image

Technique□	LBP	LGCHD	LGCHVD	CLBP
Classifier□	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	96.24%	90.61%	86.38%	98.59%
Simple Logistics	93.89%	94.36%	92.48%	96.71%
KNN	92.48%	96.24%	94.83%	96.24%
Bagging	96.71%	92.95%	92.48%	98.12%
Naïve Bayes	72.30%	84.50%	86.85%	78.40%

CLBP extracted features when employed as input to BF Tree classifier achieved maximum accuracy, leading 98.59% recognition rate. LBP features attained 96.24% accuracy.

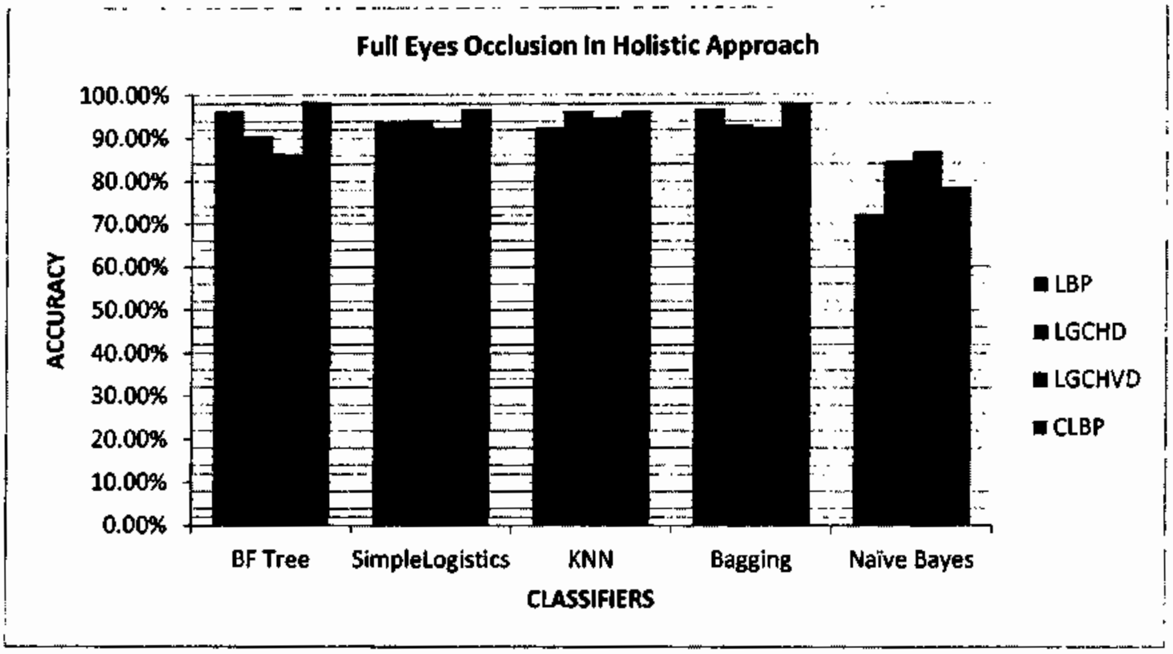


Figure 5. 19: Results of LBP, LGC-HD, LGC-HVD, and CLBP with Full Eyes Occlusion for Full Face Image

Expressions like neutral and sad acquired 100% recognition. Minor misclassifications e.g. fear as disgust, surprise as sad etc. has been observed in the confusion matrix given in Table 5.10.

Table 5. 10: Confusion Matrix of CLBP classified with BF Tree

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	29	0	0	0	0	0
Fear	0	1	31	0	0	0	0
Happy	0	0	0	31	0	0	0
Neutral	0	0	0	0	30	0	0
Sad	0	0	0	0	0	30	1
Surprise	0	0	0	0	0	1	29

b) Examining Feature Extraction Techniques on Zone based approach:

Image is divided in nine parts and features are extracted from each part. A histogram is generated representing features for each part and resultant histogram after the concatenation of all nine histograms is passed on to the classifier to get the results.

Table 5. 11: Results of LBP, LGC-HD, LGC-HVD and CLBP with Full Eyes Occlusion for Zone based approach

Technique	LBP	LGCHD	LGCHVD	MSBP	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	87.79%	83.56%	80.28%	77.46%	92.48%
Simple Logistics	92.48%	92.48%	87.32%	84.03%	94.36%
KNN	95.77%	74.17%	57.74%	79.34%	94.83%
Bagging	90.14%	87.79%	91.07%	86.85%	91.07%
Naïve Bayes	84.97%	85.91%	80.28%	81.22%	81.22%

It can be examined through the findings that 95.77% classification rate for full eyes occlusion is highest between others and this percentage of accuracy is obtained by state of the art techniques LBP. CLBP achieved 94.83% accuracy closed to the highest.

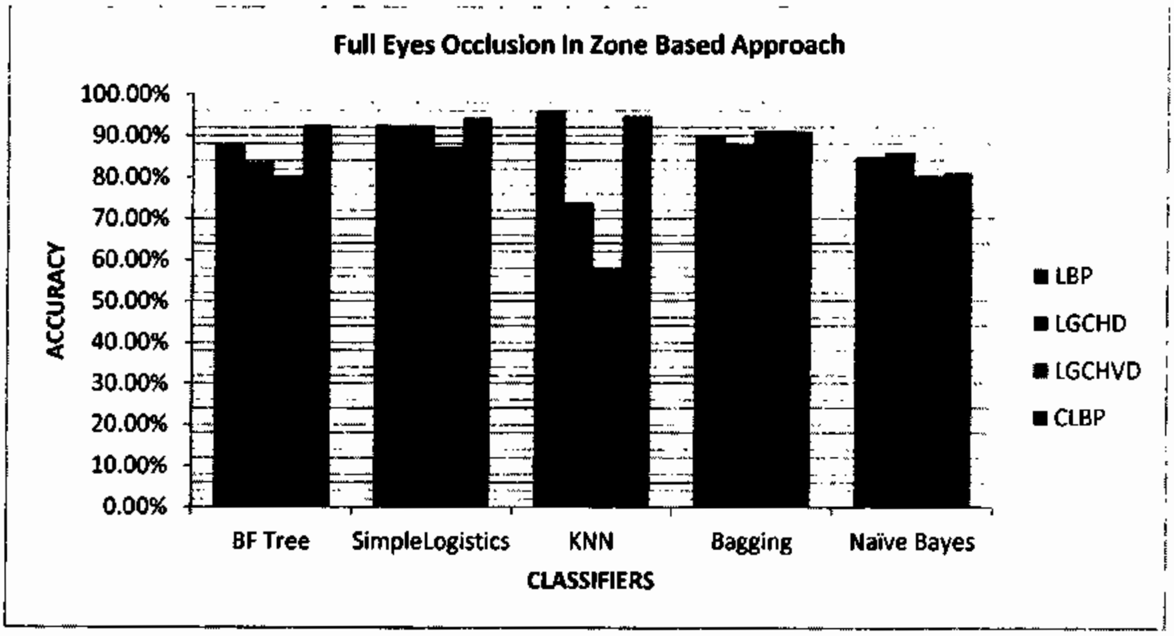


Figure 5. 20: Results of LBP, LGC-HD, LGC-HVD and CLBP with Full Eyes Occlusion for Zone based approach

Confusion matrix illustrates the misclassification of sad as neutral and surprise, surprise as sad majorly caused a decline in accuracy. Whereas, expressions like anger, disgust and fear achieved 100% accuracy.

Table 5. 12: Confusion Matrix of LBP classified with KNN

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	29	0	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	0	30	1	0	0
Neutral	0	0	0	0	29	1	0
Sad	0	0	0	0	1	27	3
Surprise	0	0	0	0	0	3	27

Among both holistic and zone based methods CLBP attained highest recognition.

5.1.5.4 Mouth Occlusion:

In expressions, mouth, position of lips has been considered as an important feature. To differentiate the importance of mouth with and without side lines in texture based techniques, two different types of occlusion are introduced. First is occlusion of different size and shapes covering only the mouth is introduced, see Figure 5.27.

Figure 5. 21: Mouth Occlusion

a) Examining Feature Extraction Techniques on Full Face Image:

Features are extracted from the face image with mouth occlusion and a histogram is generated, representing features of the whole face image. Classifier than acquired the histogram for classification of the expression.

Table 5. 13: Results of LBP, LGC-HD, LGC-HVD and CLBP with Mouth Occlusion for Full Face Image

Technique	LBP	LGCHD	LGCHVD	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	97.65%	91.54%	94.36%	99.06%
Simple Logistics	98.59%	96.71%	93.89%	97.18%
KNN	96.24%	95.77%	95.30%	99.06%
Bagging	98.59%	91.54%	94.36%	98.12%
Naïve Bayes	77.46%	84.03%	86.38%	78.87%

A recognition rate of 99.06% is observed in the Table 5.13 which is achieved through the classification of CLBP features by KNN, describing the importance of mouth in comparison to other features.

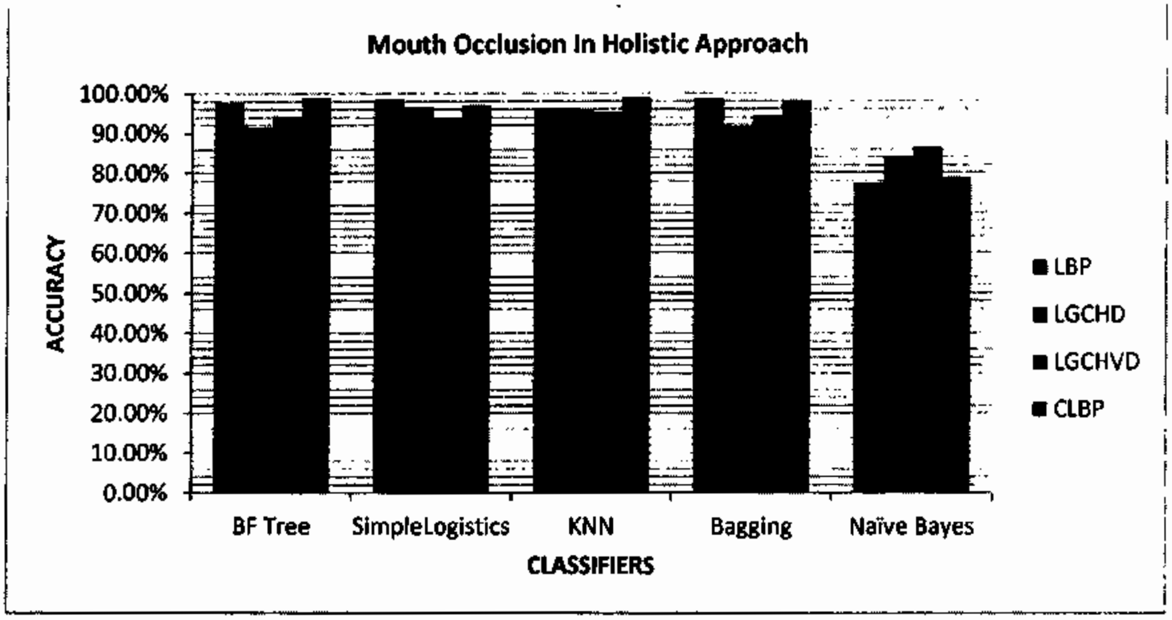


Figure 5. 22: Results of LBP, LGC-HD, LGC-HVD and CLBP with Mouth Occlusion for Full Face Image

Confusion matrix in Table 5.14 shows that almost all expressions gained 100% accuracy. Minor misclassifications do exist i.e. sad as surprise.

Table 5. 14: Confusion Matrix of CLBP classified with KNN

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	29	0	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	0	31	0	0	0
Neutral	0	0	0	0	30	0	0
Sad	0	0	0	0	0	29	2
Surprise	0	0	0	0	0	0	30

b) Examining Feature Extraction Techniques on Zone based Approach:

Image is first divided in nine parts and extraction technique is applied on each part. Histogram of each part is generated. For classification, a concatenated histogram containing information of all the nine parts is given to the classifier as input.

Table 5. 15: Results of LBP, LGC-HD, LGC-HVD and CLBP with Mouth Occlusion for Zone based approach

Technique	LBP	LGCHD	LGCHVD	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	92.48%	81.69%	79.81%	92.01%
Simple Logistics	94.83%	87.32%	83.09%	92.48%
KNN	98.12%	84.97%	64.31%	94.83%
Bagging	90.14%	88.73%	83.56%	89.67%
Naïve Bayes	81.69%	85.91%	80.28%	83.56%

LBP achieved 98.12% accuracy which is highest among others. CLBP achieved 94.83% and MSBP gained 89.67% success.

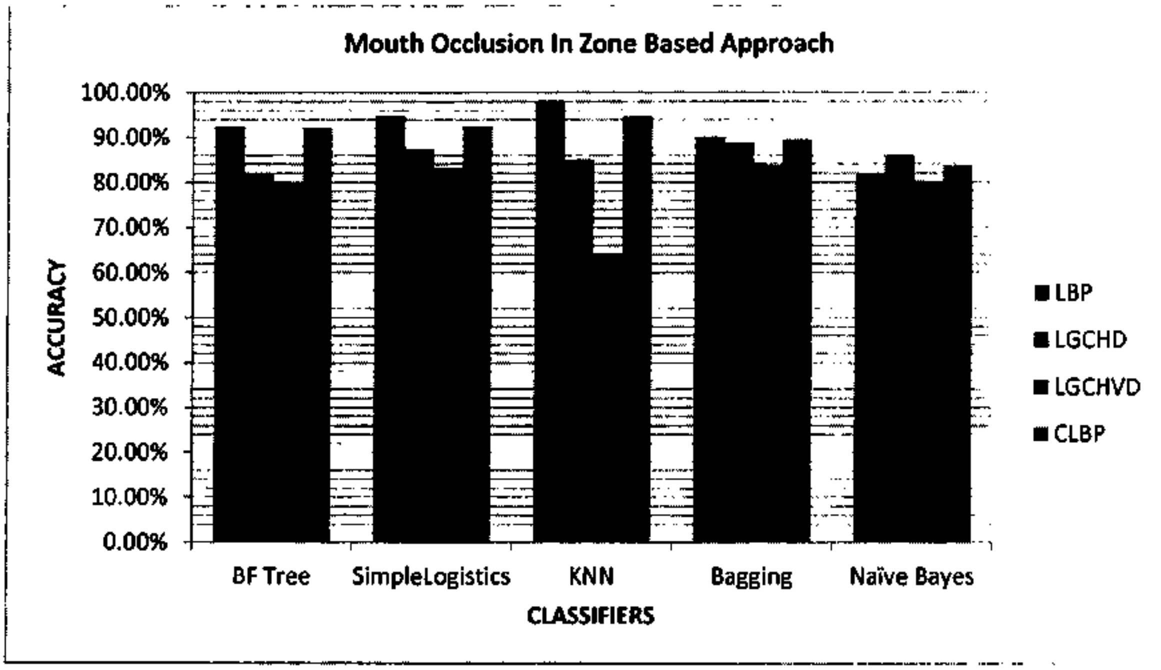


Figure 5. 23: Results of LBP, LGC-HD, LGC-HVD and CLBP with Mouth Occlusion for Zone based approach

According to the results of confusion matrix in Table 5.16 anger, fear, happy, neutral and surprise are fully classified but other expressions like sad as neutral and surprise and disgust as fear degraded the performance.

Table 5. 16: Confusion Matrix of Mouth Occlusion in Zone Based Approach

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	28	1	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	0	31	0	0	0
Neutral	0	0	0	0	30	0	0
Sad	0	0	0	0	1	28	2
Surprise	0	0	0	0	0	0	30

With hiding only the mouth area without side wrinkles CLBP features in holistic approach generated the best results.

5.1.5.5 Full Mouth Occlusion:

As a significant part of human's face mouth with its side wrinkles and bulges contribute much in judging a person's emotion. To identify the importance of texture based extraction techniques mouth and wrinkles are occluded. An image in Figure explains occlusion around whole mouth area. This is second type of occlusion around mouth.

Figure 5. 24: Full Mouth Occlusion

a) Examining Feature Extraction Techniques on Full Face Image:

Feature extraction techniques are applied on full face image and their histograms are forwarded to the classifiers to estimate the identification rates.

Table 5. 17: Results of LBP, LGC-HD, LGC-HVD and CLBP with Full Mouth Occlusion for Full Face Image

Technique	LBP	LGCHD	LGCHVD	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	95.77%	93.24%	88.73%	99.06%
Simple Logistics	95.30%	98.12%	94.36%	98.12%
KNN	96.24%	96.24%	92.01%	99.06%
Bagging	96.71%	94.83%	91.54%	99.06%
Naïve Bayes	76.05%	82.62%	82.62%	77.46%

Best classification rate i.e. 99.06% is achieved when features are extracted via CLBP. Both LBP and LGC HD gained 96.24%.

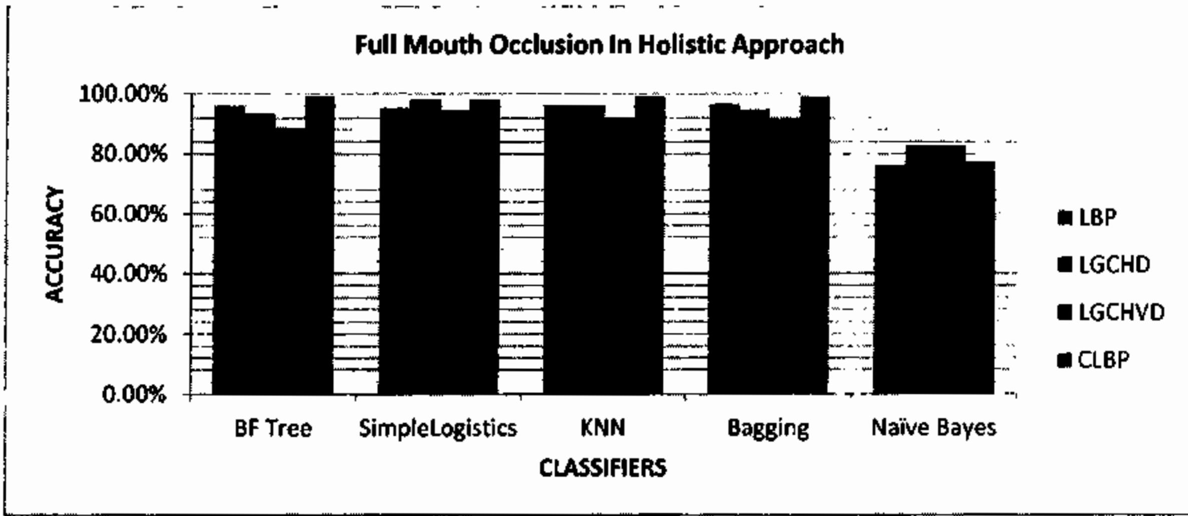


Figure 5 25: Results of LBP, LGC-HD, LGC-HVD and CLBP with Full Mouth Occlusion for Full Face Image

Confusion matrix in Table 5.18 achieved 100% accuracy for expressions like happy, disgust and neutral. Overall 99.06% accuracy has been achieved when CLBP is applied, indicating the importance of features other than mouth.

Table 5. 18: Confusion Matrix of CLBP classified with KNN

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	29	0	0	0	0	0
Fear	0	0	31	1	0	0	0
Happy	0	0	0	31	0	0	0
Neutral	0	0	0	0	30	0	0
Sad	0	0	0	0	0	31	0
Surprise	0	0	0	0	0	1	29

b) Examining Feature Extraction Techniques on Zone based approach:

Concatenated histogram of all the parts is given to the classifier for expression recognition.

Table 5. 19: Results of LBP, LGC-HD, LGC-HVD and CLBP with Full Mouth Occlusion for Zone based approach

Technique	LBP	LGCHD	LGCHVD	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	92.01%	80.75%	74.64%	90.61%
Simple Logistics	90.14%	87.32%	83.56%	91.54%
KNN	92.48%	74.64%	50.23%	94.83%
Bagging	89.20%	90.14%	79.81%	87.79%
Naïve Bayes	84.03%	84.03%	77.46%	82.62%

Classification of CLBP features with KNN achieved 94.83% accuracy. For full mouth occlusion, dividing image then extracting features via LBP gives fewer results in comparison to CLBP.

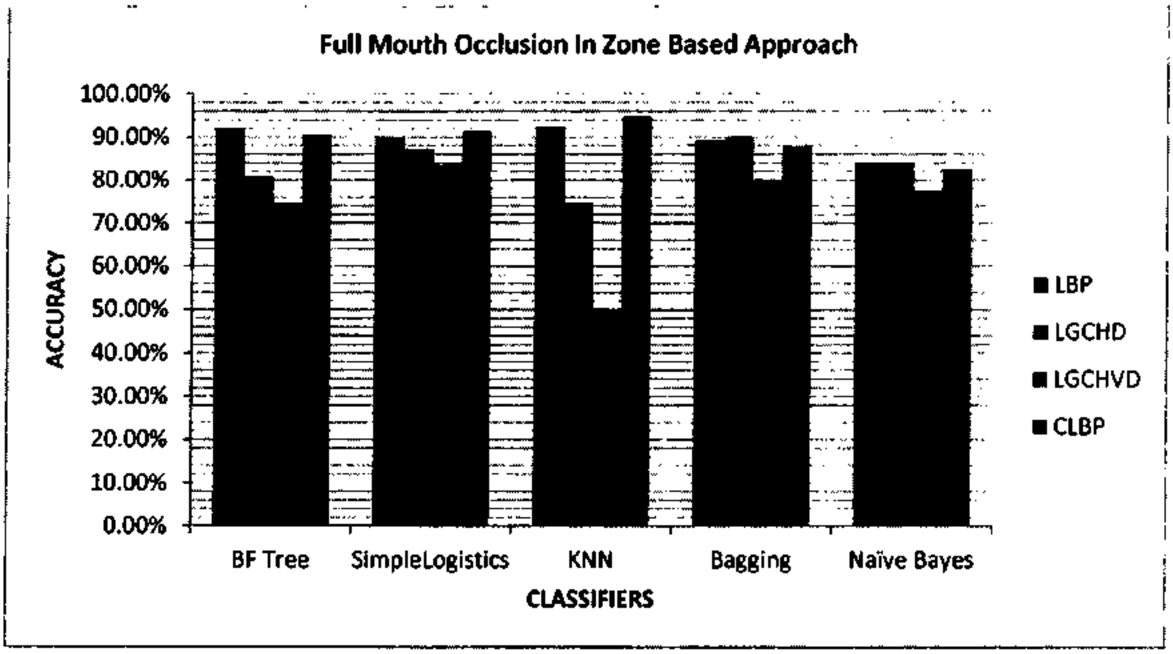


Figure 5. 26: Results of LBP, LGC-HD, LGC-HVD and CLBP with Full Mouth Occlusion for Zone based approach

Anger and fear achieved 100% accuracy but on the other hand some misclassification e.g. sad as neutral and surprise etc. can be observed in the confusion matrix in Table 5.20.

Table 5. 20: Confusion Matrix of CLBP classified with KNN

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	28	1	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	0	30	1	0	0
Neutral	0	0	0	0	29	1	0
Sad	0	0	0	0	3	25	3
Surprise	0	0	0	0	0	2	28

It can be observed while working with zones or taking face as a whole CLBP features were best with hindered mouth and side wrinkles.

5.1.5.6 Single Eye & Mouth Occlusion:

In order to see effect of occlusion further, single eye and mouth images are occluded. To observe the information of texture with single eye such type of occlusion is introduced as shown in Figure 5.27.

Figure 5. 27: Single eye and mouth occlusion

a) Examining Feature Extraction Techniques on Full Face Image:

A full face image with single eye and mouth occluded with different shapes is taken and feature extraction is applied. A histogram of extracted features is used as in input to the classifier.

Table 5. 21: Results of LBP, LGC-HD, LGC-HVD and CLBP with Single Eye & Mouth Occlusion for Full Face Image

Technique	LBP	LGCHD	LGCHVD	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	95.77%	88.26%	92.48%	99.59%
Simple Logistics	96.24%	94.83%	92.01%	95.30%
KNN	95.77%	94.36%	92.48%	97.65%
Bagging	96.71%	92.01%	91.54%	97.65%
Naïve Bayes	76.99%	84.97%	85.56%	76.99%

Analyzing results of the Table 5.21, CLBP features attained highest classification rate of 99.59% against other extraction techniques. On the other hand LBP and LGC-HD gained 95.77% and 94.36% respectively.

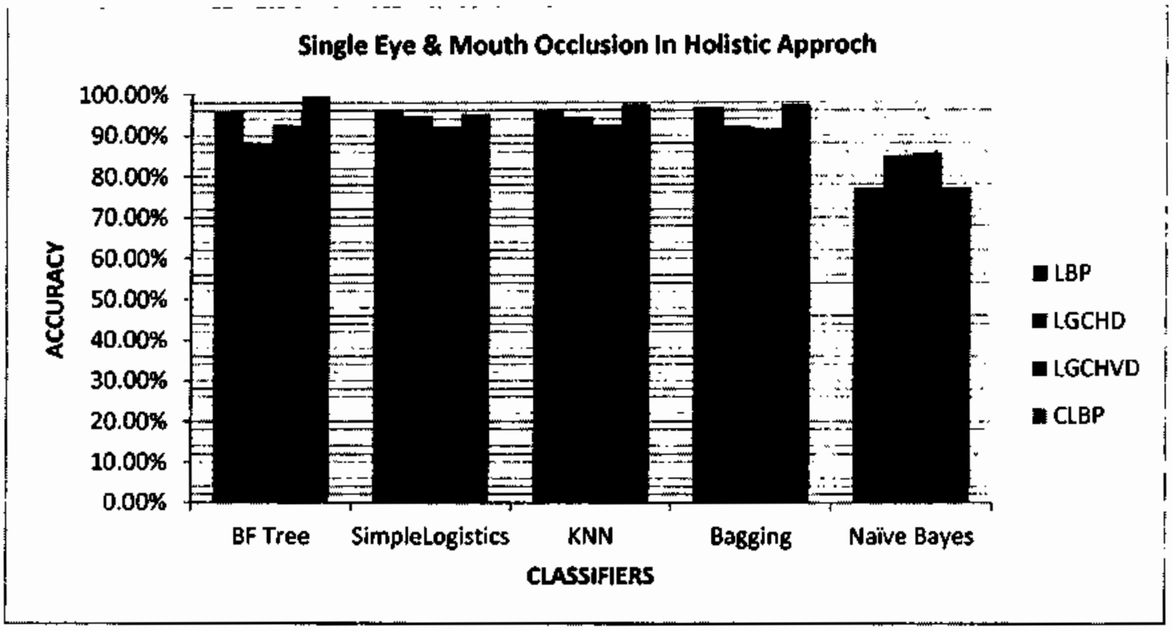


Figure 5. 28: Results of LBP, LGC-HD, LGC-HVD and CLBP with Single Eye & Mouth Occlusion for Full Face Image

Misclassifications exist in the above confusion matrix for example sad as neutral but other expressions like anger, happy, surprise etc. are 100% recognized.

Table 5. 22: Confusion Matrix of CLBP classified with BF Tree

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	1	28	0	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	0	31	0	0	0
Neutral	0	0	0	0	29	1	0
Sad	0	0	0	0	1	30	0
Surprise	0	0	0	0	0	0	30

b) Examining Feature Extraction Techniques on Zone based approach:

A concatenated histogram of all parts is given to the classifier as input.

Table 5. 23: Results of LBP, LGC-HD, LGC-HVD and CLBP with Single Eye & Mouth Occlusion for Zone based approach

Technique	LBP	LGCHD	LGCHVD	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	86.38%	81.22%	76.52%	90.14%
Simple Logistics	92.48%	85.91%	82.62%	89.67%
KNN	92.01%	72.77%	55.39%	90.14%
Bagging	88.26%	86.38%	83.09%	92.01%
Naïve Bayes	85.91%	86.38%	79.34%	85.91%

Features of both LBP and CLBP attained good results but LBP achieved maximum correctness of 92.48% in comparison to other techniques.

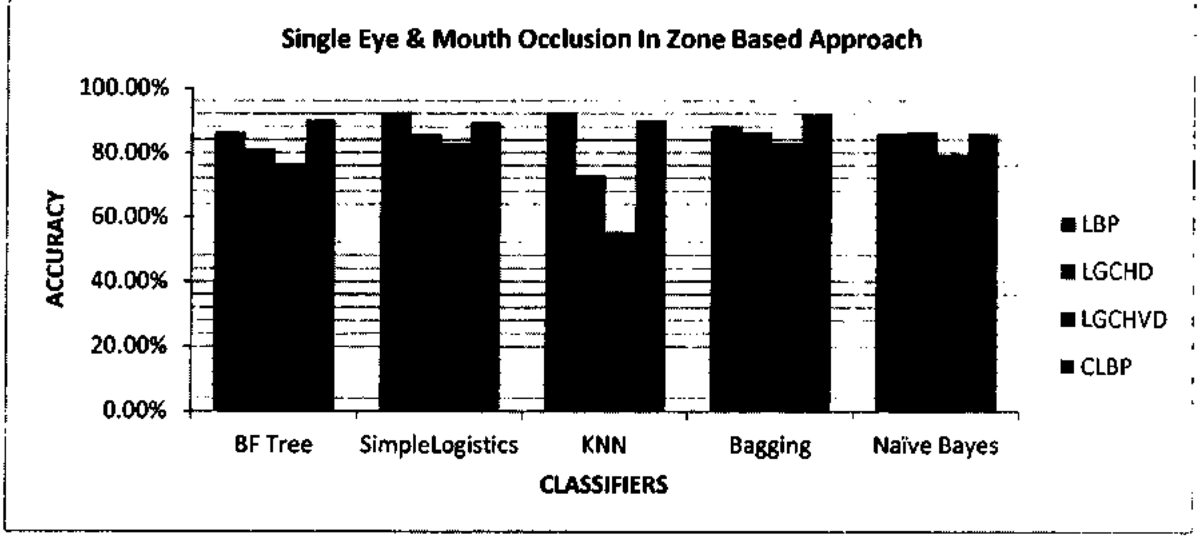


Figure 5. 29: Results of LBP, LGC-HD, LGC-HVD and CLBP with Single Eye & Mouth Occlusion for Zone based approach

Confusion matrix in Table 5.24 shows a lot of misclassification for example happy as neutral, sad as neutral and surprise etc. featuring the importance of both mouth and eyes with texture based techniques. Only anger and fear have achieved 100% accuracy.

Table 5. 24: Confusion Matrix of LBP classified with Simple Logistics

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	28	2	0	0	0	0	0
Disgust	0	27	2	0	0	0	0
Fear	0	0	31	1	0	0	0
Happy	0	0	0	31	0	0	0
Neutral	0	0	0	0	27	3	0
Sad	0	0	0	0	3	25	3
Surprise	0	0	0	0	0	2	28

Comparison of both holistic approach and zones shows that CLBP gives best results with single eye and mouth occlusion.

5.1.5.7 Both Eyes & Mouth Occlusion:

Eyes and mouth play a vital role in recognizing expressions. An occlusion of different shapes and size is applied on both eyes and mouth in order to find the significance of other features for example wrinkles and face bulges etc. in texture based techniques. Figure 5.30 shows an image, screening a sample of occlusion around eyes and mouth.

Figure 5. 30: Both eyes and mouth occlusion

a) Examining Feature Extraction Techniques on Full Face Image:

Feature extraction has been done on full face image via all four techniques and the histograms of resultant features have been used as input to different classifiers in order to get the optimum recognition rate with dataset containing hindrance of both eyes and mouth region.

Table 5. 25: Results of LBP, LGC-HD, LGC-HVD and CLBP with Both Eyes & Mouth Occlusion for Full Face Image

Technique□	LBP	LGCHD	LGCHVD	CLBP
Classifier□	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	95.30%	88.26%	89.67%	97.18%
Simple Logistics	96.24%	94.83%	89.67%	96.24%
KNN	95.30%	92.48%	91.07%	93.89%
Bagging	96.71%	88.73%	93.42%	97.18%
Naïve Bayes	76.99%	84.50%	86.85%	76.99%

From the results above it has been observed that CLBP provides best result i.e. 97.18% as compared to other extraction techniques. LBP slightly lowers the average accuracy to 96.71%.

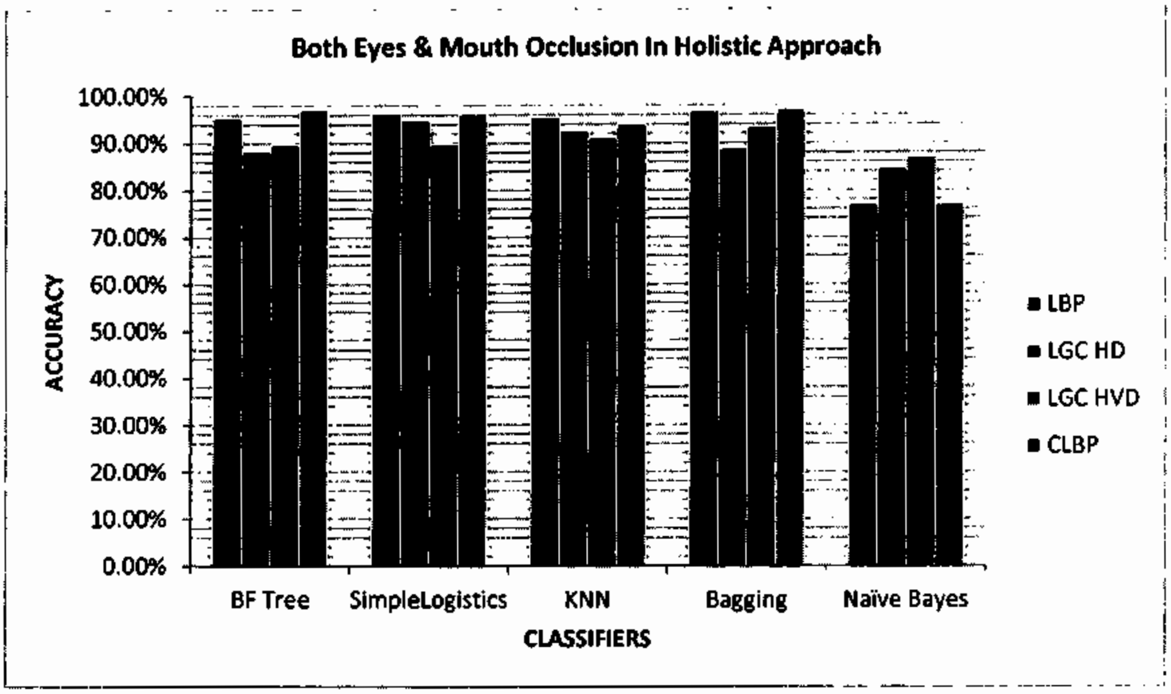


Figure 5. 31: Results of LBP, LGC-HD, LGC-HVD and CLBP with Both Eyes & Mouth Occlusion for Full Face Image

The confusion matrix shows that face parts like wrinkles and bulges also contain a lot of information as expression i.e. surprise has achieved 100% recognition accuracy while other expressions are slightly misclassified for example sad as surprise, neutral as happy, etc.

Table 5. 26: Confusion Matrix of LBP classified with Bagging

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	29	1	0	0	0	0	0
Disgust	1	28	0	0	0	0	0
Fear	0	1	31	0	0	0	0
Happy	0	0	1	30	0	0	0
Neutral	0	0	0	1	29	0	0
Sad	0	0	0	0	0	30	1
Surprise	0	0	0	0	0	0	30

b) Examining Feature Extraction Techniques on Zone based approach:

A concatenated histogram of all the parts of the image is passed on to the classifier for expression evaluation.

Table 5. 27: Results of LBP, LGC-HD, LGC-HVD and CLBP with Both Eyes & Mouth Occlusion for Zone based approach

Technique	LBP	LGCHD	LGCHVD	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	89.67%	82.15%	88.03%	91.54%
Simple Logistics	94.83%	88.26%	83.09%	91.07%
KNN	95.30%	74.17%	60.09%	91.07%
Bagging	89.20%	88.73%	88.26%	92.95%
Naïve Bayes	85.91%	85.44%	83.56%	81.69%

An average recognition rate of 95.30% has been achieved when features are extracted by LBP which gives highest accuracy.

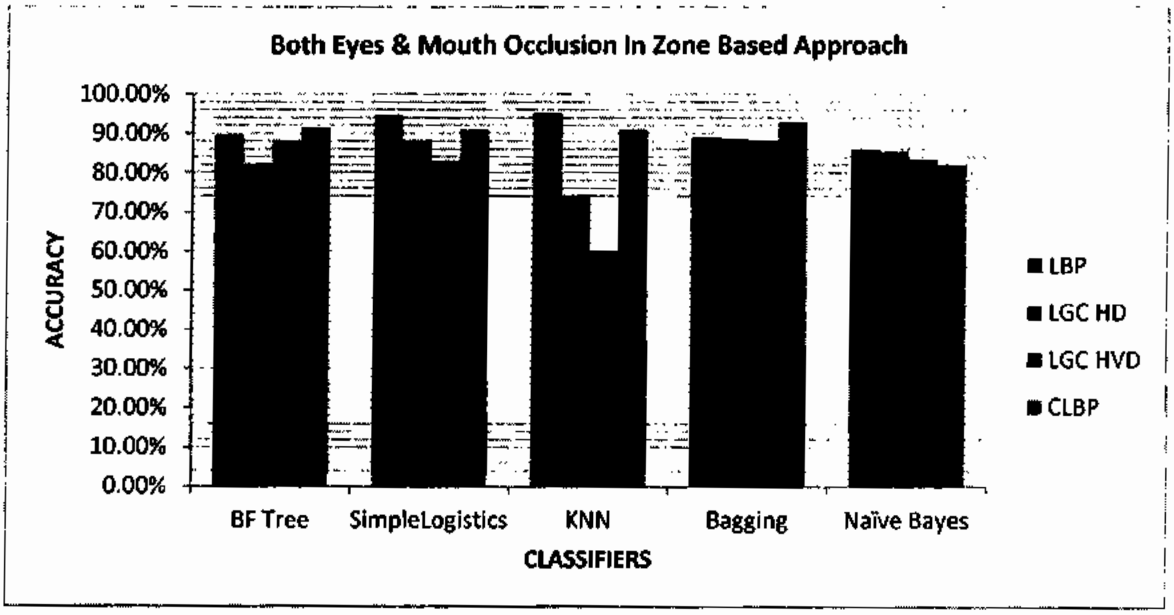


Figure 5. 32: Results of LBP, LGC-HD, LGC-HVD, MSBP and CLBP with Both Eyes & Mouth Occlusion for Zone based approach

Anger, fear and happy have showed 100% results. While surprise, neutral and sad shows some misclassification as sad, neutral and surprise.

Table 5. 28: Confusion Matrix of LBP classified with KNN

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	28	1	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	0	31	0	0	0
Neutral	0	0	0	0	29	1	0
Sad	0	0	0	0	3	25	3
Surprise	0	0	0	0	0	2	28

Both holistic and division based techniques when compared LBP generated maximum accuracy with both eyes and mouth occlusion.

5.1.5.8 Half Face Occlusion:

An occlusion is introduced, covering half face to identify expression for example in crowd where sometimes full face is not visible.

Figure 5. 33: Half face occlusion

a) Examining Feature Extraction Techniques on Full Face Image:

In order to get the optimum recognition rate, extraction technique is applied on whole image containing half face occlusion. A histogram of features is generated and is input to the classifier for classification. The procedure is repeated for all techniques.

Table 5. 29: Results of LBP, LGC-HD, LGC-HVD, and CLBP with Half Face Occlusion for Full Face Image

Technique	LBP	LGCHD	CLBP	LGCHVD
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	90.14%	96.71%	98.59%	95.77%
Simple Logistics	94.83%	93.89%	97.18%	90.14%
KNN	88.26%	84.97%	95.30%	78.87%
Bagging	92.95%	96.71%	98.12%	95.77%
Naïve Bayes	66.66%	78.87%	73.23%	81.22%

In comparison to other techniques CLBP attained maximum recognition i.e. 98.59%. LBP on the other hand achieved a maximum of 94.83% accuracy.

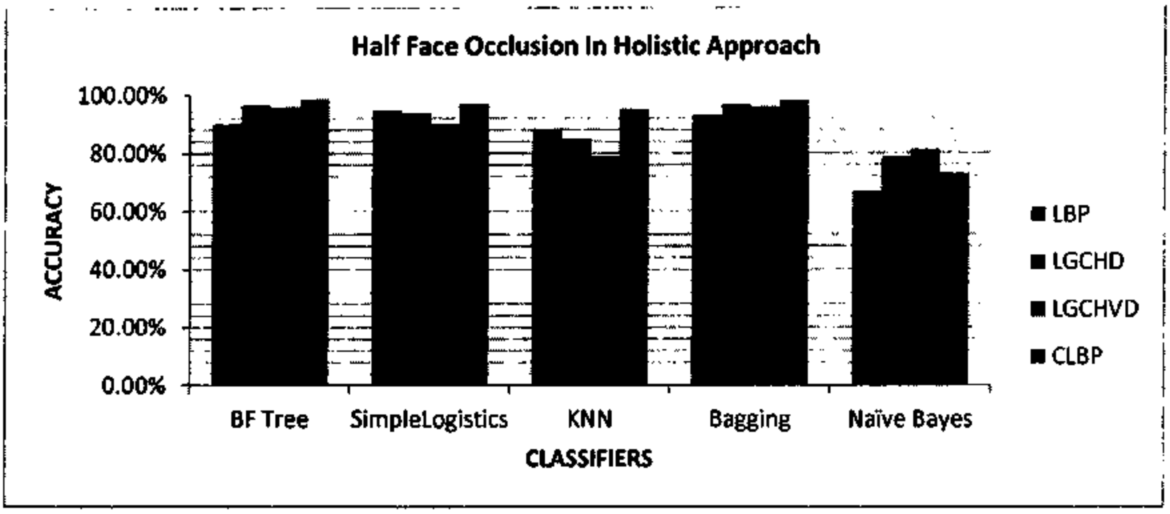


Figure 5. 34: Results of LBP, LGC-HD, LGC-HVD and CLBP with Half Face Occlusion for Full Face Image

Confusions matrix shows that only happy and anger showed minor misclassifications whereas; all other expressions are recognized to correctness.

Table 5. 30: Confusion Matrix of CLBP classified with BF Tree

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	29	1	0	0	0	0	0
Disgust	0	29	0	0	0	0	0
Fear	0	0	32	0	0	0	0
Happy	0	0	2	29	0	0	0
Neutral	0	0	0	0	30	0	0
Sad	0	0	0	0	0	31	0
Surprise	0	0	0	0	0	0	30

b) Examining Feature Extraction Techniques on Zone based Approach:

Image is divided in nine parts and features are extracted from each part. A histogram is generated to represent features for every part and resultant histogram after the concatenation of all nine histograms is advanced to the classifier to get the results.

Table 5. 31: Results of LBP, LGC-HD, LGC-HVD and CLBP with Half Face Occlusion for Zone based approach

Technique	LBP	LGCHD	CLBP	LGCHVD
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	88.26%	78.87%	91.54%	78.87%
Simple Logistics	86.38%	80.28%	85.91%	69.95%
KNN	80.75%	59.15%	84.03%	43.66%
Bagging	87.32%	84.50%	90.61%	85.44%
Naïve Bayes	72.30%	75.58%	76.52%	69.95%

According to the Table above it can be observed that CLBP gained the best result for half face occlusion dataset in 3x3 scenario.

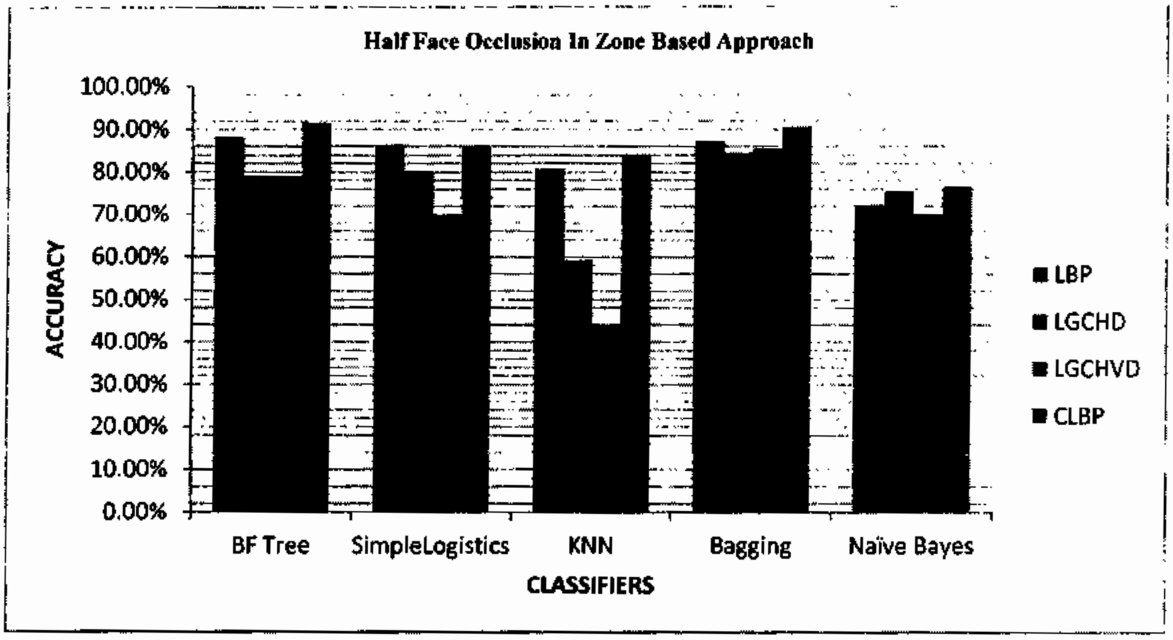


Figure 5. 35: Results of LBP, LGC-HD, LGC-HVD and CLBP with Half Face Occlusion for Zone based Approach

Major misclassifications can be seen in the confusion matrix representing CLBP features and their classification with BF Tree. Anger achieved 100% correctness.

Table 5. 32: Confusion Matrix of CLBP classified with BF Tree

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	1	28	0	0	0	0	0
Fear	0	1	30	1	0	0	0
Happy	0	0	1	29	1	0	0
Neutral	0	0	0	3	24	3	0
Sad	0	0	0	0	2	25	4
Surprise	0	0	0	0	0	1	29

For half face occlusion CLBP features give optimum results in holistic as well as in zone based method.

5.1.5.9 Miscellaneous Occlusion:

In order to get more close to the real world scenario, a dataset is generated including all types occlusion. Figure below illustrates some samples of miscellaneous occlusion dataset.

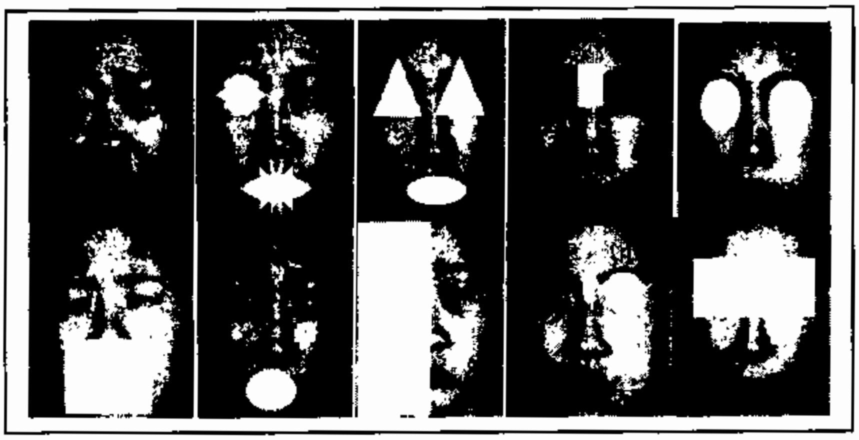


Figure 5. 36: Miscellaneous occlusion

a) Examining Feature Extraction Techniques on Full Face Image:

A histogram representing extracted features of full face is advanced towards the classifier for assessment of results. Images having different types of occlusion are processed to extract feature and their histograms are created and classified to give following findings.

Table 5. 33: Results of LBP, LGC-HD, LGC-HVD and CLBP with Miscellaneous Occlusion for Full Face Image

Technique	**LBP**	**LGC**	**LGCHVD**	**CLBP**
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	93.89%	81.69%	84.03%	92.95%
Simple Logistics	94.83%	91.07%	90.14%	95.77%
KNN	92.01%	92.01%	84.97%	94.36%
Bagging	92.95%	85.44%	85.44%	92.01%
Naïve Bayes	65.72%	69.95%	68.07%	69.48%

Table 5.33 portrays that in comparison to all the other feature extractor & classifier arrangement, highest classification rate i.e. 95.77% is achieved by CLBP features. Both LBP and LGC-HD attained 94.83 and 91.07%.

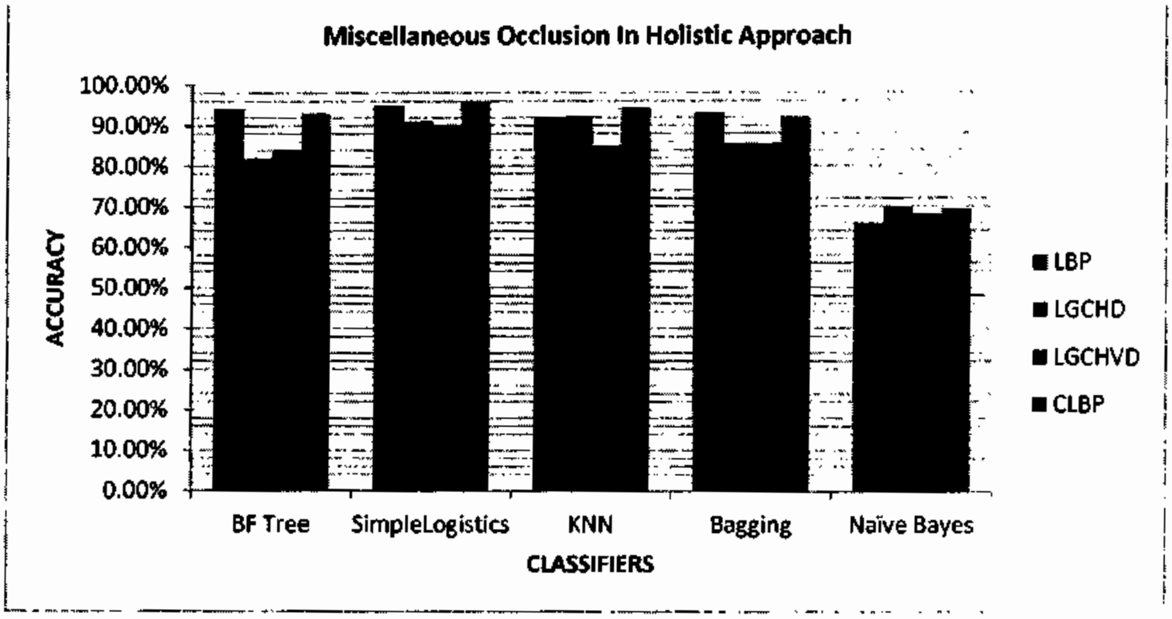


Figure 5. 37: Results of LBP, LGC-HD, LGC-HVD and CLBP with Miscellaneous Occlusion for Full Face Image

Confusion matrix explains results of CLBP features and their classification with NN. Anger contain 100% accuracy, disgust is misclassified as fear, happy is misclassified as fear, and fear is misclassified as happy and disgust etc.

Table 5. 34: Confusion Matrix of CLBP classified with KNN

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	30	0	0	0	0	0	0
Disgust	0	28	1	0	0	0	0
Fear	0	1	28	3	0	0	0
Happy	0	0	1	30	0	0	0
Neutral	0	0	0	1	28	1	0
Sad	0	0	0	0	1	28	2
Surprise	0	0	0	0	0	1	29

b) Examining Feature Extraction Techniques on Zone based approach:

A concatenated histogram of all the parts of an image is given to the classifier for expression recognition.

Table 5. 35: Results of LBP, LGC-HD, LGC-HVD, MSBP and CLBP with Miscellaneous Occlusion for Zone based approach

Technique	LBP	LGC	LGCHVD	MSBP	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	86.38%	74.17%	69.01%	62.91%	84.50%
Simple Logistics	85.91%	84.97%	75.58%	73.70%	83.56%
KNN	85.44%	73.23%	61.03%	67.13%	86.85%
Bagging	84.03%	76.05%	79.34%	76.05%	83.56%
Naïve Bayes	74.64%	67.60%	67.13%	68.54%	77.93%

A maximum accuracy of 86.85% is achieved with dataset including all types of occlusion. This highest accuracy is a result of features extracted by CLBP. Slightly declining LBP attained 85.44% accuracy.

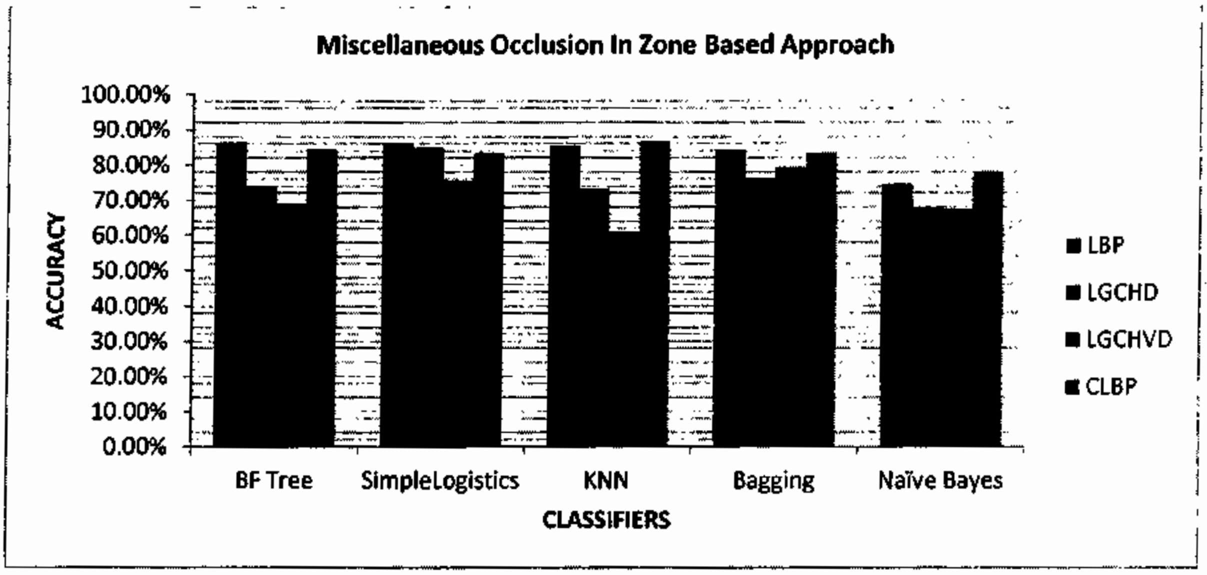


Figure 5. 38: Results of LBP, LGC-HD, LGC-HVD and CLBP with Miscellaneous Occlusion for Zone based Approach

Confusion matrix in Table exhibits no expression with miscellaneous dataset gives 100% accuracy.

Table 5. 36: Confusion Matrix of CLBP classified with KNN

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	29	1	0	0	0	0	0
Disgust	0	26	3	0	0	0	0
Fear	0	0	30	2	0	0	0
Happy	0	0	1	27	3	0	0
Neutral	0	0	0	1	24	4	1
Sad	0	0	0	0	3	24	4
Surprise	0	0	0	0	1	4	25

5.2: Crowd Emotion Classification by using FER:

Second part of this chapter discover a new dimension in analyzing crowd behavior through facial expression recognition is introduced. Crowd is generally gathering of people grouped together[86]. This is initial work in utilizing crowd behavior on the basis of FER. Section 5.3 includes experimental setup. As this is very initial effort so dataset is manually generated. Section 5.3.1 summarize methodology, section 5.3.2 comprises results and section 5.4 finalized chapter summary.

5.2.1 Motivation:

In video surveillance crowd behavior identification is an important aspect of research. But to analyze complete video requires capturing video data all the time, which serves as a bottle neck as capturing all the video requires not only a lot of memory but also is a time consuming task. More over in videos we mostly depends on optical flow helping us perceiving actions but this task becomes more difficult when the crowd is static or (static images). In this chapter our research focus will be on analyzing static images of crowd to determine their emotions. This research will be helpful in areas where static crowd is gathered e.g. customer care services, asylums, prison, schools, and orphanages etc. These are community services which are deprived of freedom of speech, so fair opinion cannot be taken by perceiving their actions so we have to judge their facial expressions to take a fair opinion of them. Further such technique can be used for surveillance of crowds to avoid any security risk.

5.2.1 Experimental Setup:

Crowd data in this research is collected manually as no such data set was available to us. M-FER (Multiple Face Expression Recognition Dataset) is a modeled expression dataset of 45 imageries in which partakers were requested to display four basic emotional expressions i.e. happy, anger, surprise, sadness and neutral. In our dataset 13 participants forming a group of four or five members modeled their expressions. Group of 4 to 5 people is considered as crowd. Images are labeled into five classes happy, anger, sad, surprise and neutral. Some examples of dataset hence collected are given in Figure 5.39.

Anger Sad Surprise Happy Neutral

Figure 5. 39: Examples of Crowd Dataset

5.2.3 Methodology:

Initial step in crowd behavior identification is data acquisition. Details of dataset are given in section 5.2.2.

To minimalize unwilling distortions or enhance some features significant for further processing, pre-processing is applied on image data. Following functions were performed step by step for pre-processing the image set available;

1. Data is cropped to remove unnecessary details from background.

Original Image Cropped Image

Figure 5. 40: Cropped Image

2. This cropped image is then converted to gray scale and normalized for making calculation easier.

Figure 5. 41: Normalized Data

After pre-processing local texture based techniques LBP, LGCHD, LGCHVD and CLBP were used for feature extraction of normalized data.

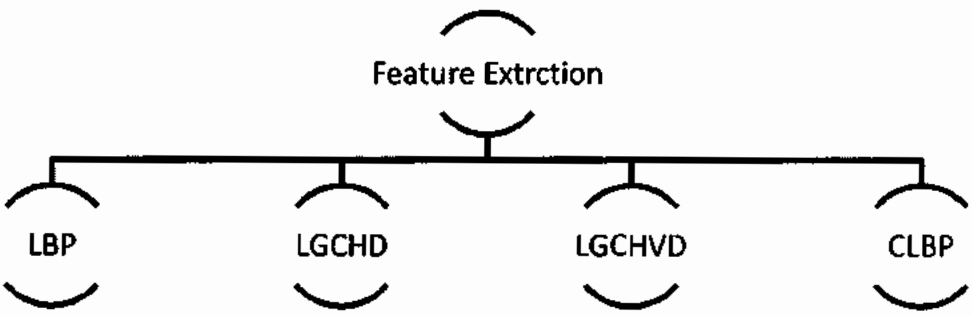


Figure 5. 42: Feature Extraction techniques in detection Crowd Emotion

After applying these local texture based techniques, histogram is generated which serve as feature vector.

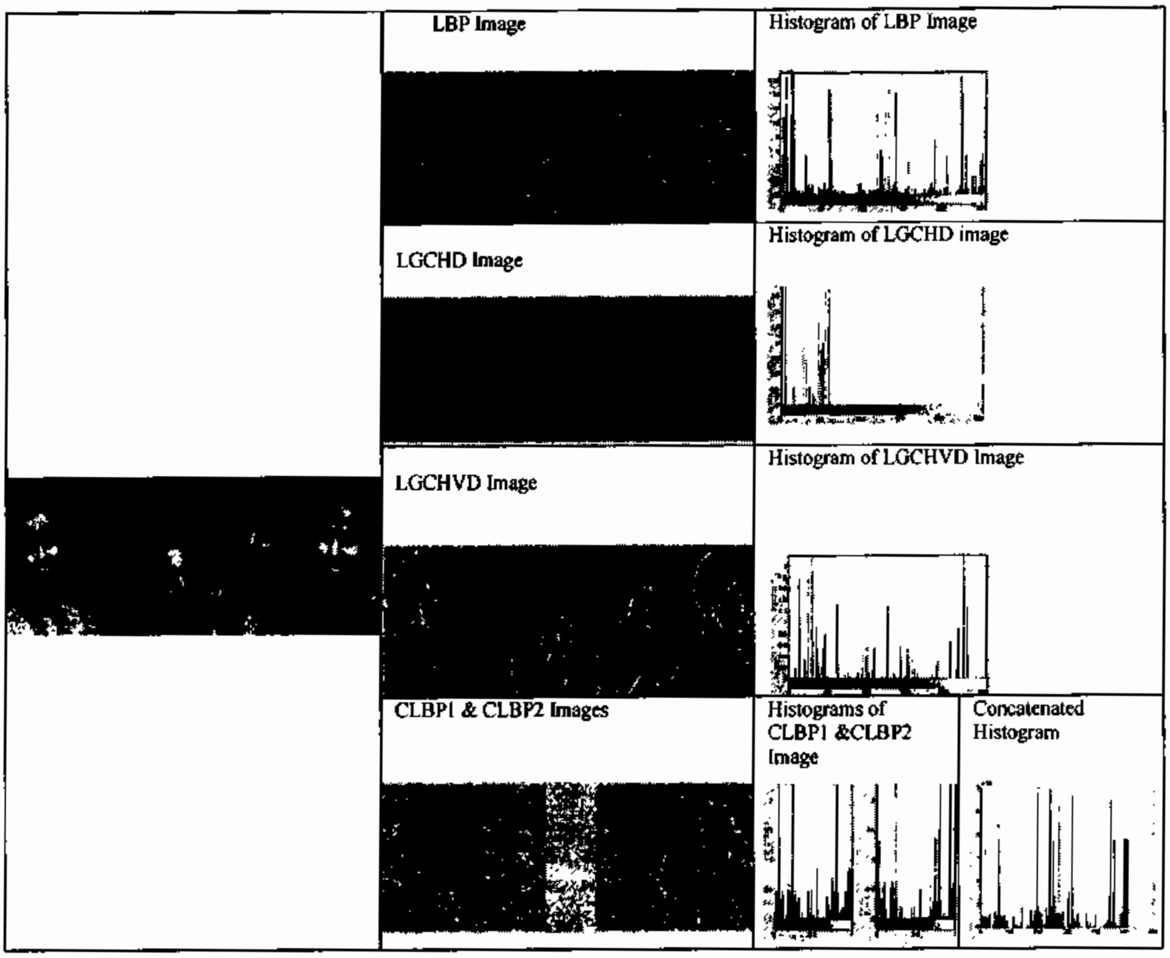


Figure 5. 43: Feature Extraction Techniques Applied on Crowd

6.2.3 Results:

Experiments are performed in grouped data to determine expression of crowd on the basis of facial expression recognition. In real crowd data challenges will increase than considered in this dataset. Though this data is collected in real environment with random light variations. But still this data does not hold occlusions. Further if size of crowd increases than results may vary. Being first effort in recognizing crowd behavior. Local texture based techniques like LBP, LGCHD, and LGCHVD and CLBP are used to conduct experiments. All these techniques are implemented and tested on the given MFER. Once technique is applied, its histogram is generated that is then provided to the classifier to generate results. Supervised learning is used for classification with ten cross validation. Table 5.37 shows the results of classifiers on MFER.

Table 5. 37: Accuracy Results of LBP, LGCHD, LGCHVD and CLBP for determining Emotion in Crowd

Technique	LBP	LGCHD	LGCHVD	CLBP
Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
BF Tree	80.43%	91.11%	93.33%	93.33%
Simple Logistics	89.13%	91.11%	93.33%	91.11%
KNN	93.47%	95.50%	91.11%	95.50%
Bagging	82.60%	93.33%	93.33%	95.55%
Naïve Bayes	86.95%	88.88%	86.66%	86.67%

It is obvious from the Table 5.37 that CLBP gives better result of determining emotion in crowd than rest of the techniques. It gives 95.5% accuracy when classified with KNN and bagging both. Simple reason of CLBP to outperform is that it takes into account sign and magnitude differences both. LGCHD and LGCHVD also performed well. Graphical representation of results is given in Figure 6.6.

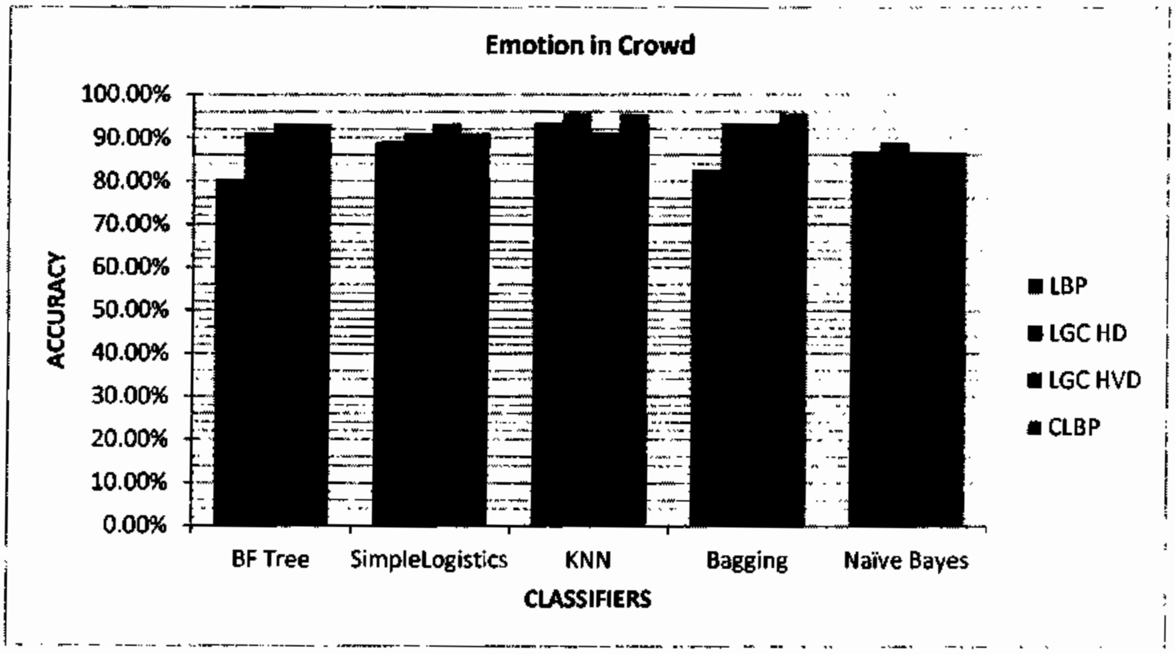


Figure 5. 44: Emotion in Crowd through Facial Expression Recognition

From graphical representation it is clear that CLBP outperforms rest of the techniques in determining emotion in crowd. LGCHD and LGCHVD also perform well specially with used with Simple Logistics by very small difference. Table 5.38 discussed confusion matrix of recognizing expressions of multiple faces. In confusion matrix results are shown in terms of percentage.

Table 5. 38: Confusion Matrix of CLBP in analyzing Emotion in Crowd

	Happy	Anger	Surprise	Sad	Neutral
Happy	88.8%	0	0	0	11.12%
Anger	0	100%	0	0	0
Surprise	0	0	100%	0	0
Sad	0	0	0	88.88%	11.12%
Neutral	0	0	0	0	100%

From confusion matrix in Table 5.38 it is clear that CLBP has little misclassification in case of happy and sad where both expressions are recognized false as neutral. In all rest of expressions it gives 100% true classifications.

6.4 Chapter Summary:

Novel dimension of crowd behavior recognizing through facial expression recognition is presented in this chapter. Being initial effort in this domain we have limited crowd size up to four to five persons. Five expressions in crowd are considered namely happy, sad, anger surprise and neutral. Local texture based techniques along with supervised learning are used to conduct experiments. Experimental results reveal that we can successfully determine emotion of crowd by using facial expression recognition. CLBP gives 95%accuracy in correctly identifying behavior of crowd. Time Complexity of these techniques for each image is same as size of the image i.e. O (mn) if image is of size m x n. If size of training set is T then this time complexity will be O (Tmn). Results show that CLBP and LGCHVD performs well in case of all expressions.

5.6 Chapter Summary:

In this chapter extensive types simulations of occlusions have been presented. Main contribution of this chapter is simulation of different type, shape and size of occlusion. Nine different types of occlusion have been introduced on JAFEE dataset. Further in order to validate local texture based technique same testing environments are provided. LBP, CLBP, LGCHVD and LGCHD have been compared on the same parameters i.e. type of occlusion and classifier have been kept same. Further for each feature extraction technique and occlusion type two set of experiments have been conducted holistic and division based. Results showed that CLBP performed better in maximum of the scenarios in holistic approach. Only in case of both eye occlusion LBP gave better results. In division based approach LBP achieved better results in most scenarios. In case of half face occlusion and miscellaneous occlusion CLBP outclassed all rest techniques in both holistic and division based case.

Chapter # 6

Conclusion and Future Enhancements

6. Conclusion and Future Enhancements:

Thesis is concluded as below.

6.1 Conclusion:

This research focuses on facial expression recognition through local texture based techniques in more efficient and effective way. Set of experiments have been conducted to explore different aspects such as recognizing emotions in real scenarios, analyzing impact of local and global illumination and its correction, occlusion handling and determining emotion of crowd.

MSBP is developed for handling challenges of real world data. It is tested on one of most challenging dataset SFEW. This algorithm generates two binary patterns one on the basis of sign difference and second on the basis of gradient difference for each neighbor. These two codes are then combined on the basis of 4- neighbor and 8-neighbours to generate MSBP1 and MSBP2 codes. This algorithm works in the time complexity of $O(mn)$, where $m \times n$ is size of the image. Gradient difference along with sign difference makes these patterns powerful enough to recognize expression in SFEW. SFEW is dataset having real expressions as data is taken from the movies. Results shows that MSBP outperforms than LBP, LGC and CLBP on such challenging dataset. We also explored that for expression recognition holistic approach generates better results than zone based approach. It gives 96%result son SFEW in holistic approach and 60% results in SFEW in zone based approach. Thus two strong contributions of this chapter are MSBP provide discriminant gradient and sign differences which preserves feature information in challenging environment of SFEW. Further holistic approach give better results in texture based techniques than zone based techniques

RGPs are novel and are generated to resolve issues originated from local illuminations. .RGP works on comparison of central pixel with gradient differences. We have used standard dataset JAFEE and SFEW for experimentation. JAFFE Dataset is distorted with global and local illuminations and all three scenarios i.e. images without any illumination change, with global illumination change and with local illumination change are tested. Results show that in case of global illumination LBP, LGCHD and RGP works equally good but in case of local illuminations RGP performs best.in case of no illumination changes RGP gives 96% results, in case of global illumination changes it gives 95% and incase of local illumination it gives 97%

results. RGP is also tested with SFEW, it gives 94% results. Thus in challenging light variations gradient difference give better results.

Occlusion is one of the major hindrances in FER systems. To investigate impact of occlusion nine different types of occlusion are simulated and its results are studied for FER. Main contribution in this chapter is simulation of various types and styles of occlusions and providing same test environment to local texture based techniques in order to compare their results. As in literature every paper and technique have its own size and form of occlusion so it is hard to conclude which technique perform better. Further both holistic and zone based approaches are used to verify results. It is concluded that holistic approaches work good in case of occlusion as well. Among texture based techniques results of CLBP are impressive than rest of the techniques. JAFEE dataset is used for conducted experiments. Four categories of occlusion is simulated on JAFEE dataset namely eye occlusion, mouth occlusion, half face occlusion and miscellaneous occlusion. In eye occlusion single eye, both eye and eye with eye brows are occluded. In mouth occlusion lips region and lips along with side wrinkles are occluded. In half face occlusion pose variations are accommodated. In miscellaneous occlusion eye lip, both eye and lip etc. is simulated. Results shows that CLBP outperforms than all rest of techniques in both holistic and zone based approach.

Last module of the research is about either crowd behavior can be determined on the basis of facial expression recognition. For this purpose we create dataset having five main expressions happy, sad anger, and surprise. All these texture based techniques are applied and it shows that CLBP excels in determining emotion of crowd with 915 accuracy.

In this research we used all efficient algorithms which time complexity does not extend more than the size of the image. Efforts are made to standardize the methods as different texture based techniques are tested in same scenarios to determine their accuracies.

6.2 Future Enhancements:

In future we are intending to extend number of emotions in crowd i.e. emotions other than happy, sad, anger and surprise. Further Crowd emotion can be recognized in videos as real data with each frame of the video results of emotion in crowd may vary. We can determine results by increasing crowd size that to which extent of crowd we can determine its emotions via expression recognition. Moreover accurate face detection in crowd may lead to better results in Crowd behavior identification. Efforts should be made to identify crowd expression where people might exhibiting different expressions.

We are also concerned that either we can reconstruct occluded feature in the effective manner rather accommodating them. As reconstruction of occluded features will improve accuracy and intimation of expression.

We are also interested in finding out utilization of these algorithms in other areas of research in computer vision such as medical imaging, image retrieval and crowd activity and interaction modeling etc.

References

References:

[1] R. W. Picard, "Affective Computing for HCI.," in *HCI (1)*, 1999, pp. 829–833.

[2] R. W. Picard and R. Picard, *Affective computing*, vol. 252. MIT press Cambridge, 1997.

[3] Y. Gao, M. K. Leung, S. C. Hui, and M. W. Tananda, "Facial expression recognition from line-based caricatures," *IEEE Trans. Syst. Man Cybern.-Part Syst. Hum.*, vol. 33, no. 3, pp. 407–412, 2003.

[4] C. Darwin, P. Ekman, and P. Prodger, *The expression of the emotions in man and animals*. Oxford University Press, USA, 1998.

[5] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, "Web-based database for facial expression analysis," in *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, , p. 5 pp., 2005.

[6] O. Langner, R. Dotsch, G. Bijlstra, D. H. Wigboldus, S. T. Hawk, and A. van Knippenberg, "Presentation and validation of the Radboud Faces Database," *Cogn. Emot.*, vol. 24, no. 8, pp. 1377–1388, 2010.

[7] M. P. Brown *et al.*, "Knowledge-based analysis of microarray gene expression data by using support vector machines," *Proc. Natl. Acad. Sci.*, vol. 97, no. 1, pp. 262–267, 2000.

[8] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, 2006.

[9] B. Moghaddam, T. Jebara, and A. Pentland, "Bayesian face recognition," *Pattern Recognit.*, vol. 33, no. 11, pp. 1771–1782, 2000.

[10] M.-H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods.," in *Fgr*, , vol. 2, p. 215, 2002

[11] J. Yang, D. Zhang, A. F. Frangi, and J. Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131–137, 2004.

[12] H.-B. Deng, L.-W. Jin, L.-X. Zhen, and J.-C. Huang, "A new facial expression recognition method based on local gabor filter bank and pca plus lda," *Int. J. Inf. Technol.*, vol. 11, no. 11, pp. 86–96, 2005.

[13] L. Ma and K. Khorasani, "Facial expression recognition using constructive feedforward neural networks," *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 34, no. 3, pp. 1588–1595, 2004.

[14] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 98–113, 1997.

[15] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image Vis. Comput.*, vol. 27, no. 6, pp. 803–816, 2009.

[16] I. Kotsia and I. Pitas, "Facial expression recognition in image sequences using geometric deformation features and support vector machines," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 172–187, 2007.

[17] I. Cohen, N. Sebe, F. G. Gozman, M. C. Cirelo, and T. S. Huang, "Learning Bayesian network classifiers for facial expression recognition both labeled and unlabeled data," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, , vol. 1, pp. I–I, 2003.

[18] J. Ou, X.-B. Bai, Y. Pei, L. Ma, and W. Liu, "Automatic facial expression recognition using Gabor filter and expression analysis," in *Computer Modeling and Simulation, 2010. ICCMS'10. Second International Conference on*, , vol. 2, pp. 215–218. 2010

[19] T. Jabid, M. H. Kabir, and O. Chae, "Robust facial expression recognition based on local directional pattern," *ETRI J.*, vol. 32, no. 5, pp. 784–794, 2010.

[20] L. Zhang, S. Chen, T. Wang, and Z. Liu, "Automatic facial expression recognition based on hybrid features," *Energy Procedia*, vol. 17, pp. 1817–1823, 2012.

[21] D. Ghimire and J. Lee, "Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines," *Sensors*, vol. 13, no. 6, pp. 7714–7734, 2013.

[22] V. J. Mistry and M. M. Goyani, "A literature survey on facial expression recognition using global features," *Int J Eng Adv Technol*, vol. 2, no. 4, pp. 653–657, 2013.

[23] M. Madhu and R. Amutha, "A combinatorial approach to human face recognition and expression identification," *Sci. Res. Essays*, vol. 8, no. 31, pp. 1464–1470, 2013.

[24] A. V. Surbhi, "The Facial expression detection from Human Facial Image by using neural network," *Int. J. Appl. Innov. Eng. Manag. IJAIEM*, vol. 2, no. 6, 2013.

[25] N. Shukla and A. Kumar, "Using Back-Propagation Recognition of Facial Expression," *J. Environ. Sci. Comput. Sci. Eng. Technol. JECET*, vol. 2, no. 1, pp. 39–45, 2013.

[26] S. T. Roohi, D. S. Sachin, K. Manikantan, and S. Ramachandran, "Feature Accentuation Using Uniform Morphological Correction as Pre-processing Technique for DWT based Face Recognition," *Procedia Technol.*, vol. 10, pp. 690–698, 2013.

[27] P. S. Subramanyam and S. A. Fegade, "Face and Facial Expression recognition-A Comparative study," 2013.

[28] J. Kalita and K. Das, "Recognition of facial expression using eigenvector based distributed features and euclidean distance based decision making technique," *ArXiv Prepr. ArXiv13030635*, 2013.

[29] J. Chen, Z. Chen, Z. Chi, and H. Fu, "Emotion recognition in the wild with feature fusion and multiple kernel learning," in *Proceedings of the 16th International Conference on Multimodal Interaction*, , pp. 508–513, 2014.

[31] W.-L. Chao, J.-Z. Liu, J.-J. Ding, and P.-H. Wu, "Facial expression recognition using expression-specific local binary patterns and layer denoising mechanism," in *Information, Communications and Signal Processing (ICICS) 2013 9th International Conference on*, , pp. 1–5, 2013.

[32] L. Zhang, D. Tjondronegoro, and V. Chandran, "Facial expression recognition experiments with data from television broadcasts and the World Wide Web," *Image Vis. Comput.*, vol. 32, no. 2, pp. 107–119, 2014.

[33] A. Majumder, L. Behera, and V. K. Subramanian, "Local binary pattern based facial expression recognition using Self-organizing Map," in *Neural Networks (IJCNN), 2014 International Joint Conference on*, pp. 2375–2382, 2014.

[34] F. Ahmed, H. Bari, and E. Hossain, "Person-independent facial expression recognition based on compound local binary pattern (CLBP).," *Int Arab J Inf Technol*, vol. 11, no. 2, pp. 195–203, 2014.

[35] Y. Tong, R. Chen, and Y. Cheng, "Facial expression recognition algorithm using LGC based on horizontal and diagonal prior principle," *Opt.-Int. J. Light Electron Opt.*, vol. 125, no. 16, pp. 4186–4189, 2014.

[36] V. Ramachandran and E. S. Reddy, "Facial expression recognition with enhanced feature extraction using PSO & EBPNN," *Int. J. Appl. Eng. Res.*, vol. 11, no. 10, pp. 6911–6915, 2016.

[37] M. H. Siddiqi *et al.*, "Human facial expression recognition using curvelet feature extraction and normalized mutual information feature selection," *Multimed. Tools Appl.*, vol. 75, no. 2, pp. 935–959, 2016.

[38] H. Qayyum, M. Majid, S. M. Anwar, and B. Khan, "Facial Expression Recognition Using Stationary Wavelet Transform Features," *Math. Probl. Eng.*, vol. 2017, 2017.

[40] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. Oliveira-Santos, "Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order," *Pattern Recognit.*, vol. 61, pp. 610–628, 2017.

[41] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image Vis. Comput.*, vol. 27, no. 6, pp. 803–816, 2009.

[42] M. Emadi, M. Khalid, R. Yusof, and F. Navabifar, "Illumination normalization using 2D wavelet," *Procedia Eng.*, vol. 41, pp. 854–859, 2012.

[43] X. Yuan, Y. Meng, and X. Wei, "Illumination normalization based on homomorphic wavelet filtering for face recognition," *J. Inf. Sci. Eng.*, vol. 29, no. 3, pp. 579–594, 2013.

[44] V. Vidya, N. Farheen, K. Manikantan, and S. Ramachandran, "Face recognition using threshold based DWT feature extraction and selective illumination enhancement technique," *Procedia Technol.*, vol. 6, pp. 334–343, 2012.

[45] Y. Luo, C. Wu, and Y. Zhang, "Facial expression recognition based on fusion feature of PCA and LBP with SVM," *Opt.-Int. J. Light Electron Opt.*, vol. 124, no. 17, pp. 2767–2770, 2013.

[46] S. T. Roohi, D. S. Sachin, K. Manikantan, and S. Ramachandran, "Feature Accentuation Using Uniform Morphological Correction as Pre-processing Technique for DWT based Face Recognition," *Procedia Technol.*, vol. 10, pp. 690–698, 2013.

[47] Y. Liu, L. De Dominicis, R. R. Martin, and B. Wei, "Illumination normalization using weighted gradient integral images," *Lect. Notes Softw. Eng.*, vol. 1, no. 4, p. 329, 2013.

[48] C. JayaMohan, M. S. Deepak, M. A. E. Manuel, and D. J. W. Wise, "Face Recognition under Expressions and Lighting Variations using Artificial Intelligence and Image Synthesizing," *Int. J. Comput. Sci. Netw. Secur. IJCSNS*, vol. 15, no. 9, p. 59, 2015.

[49] C. JayaMohan, M. S. Deepak, M. A. E. Manuel, and D. J. W. Wise, "Face Recognition under Expressions and Lighting Variations using Artificial Intelligence and Image Synthesizing," *Int. J. Comput. Sci. Netw. Secur. IJCSNS*, vol. 15, no. 9, p. 59, 2015.

[50] S. Lee and C. Lee, "Multiscale morphology based illumination normalization with enhanced local textures for face recognition," *Expert Syst. Appl.*, vol. 62, pp. 347–357, 2016.

[51] J.-Y. Zhu, W.-S. Zheng, F. Lu, and J.-H. Lai, "Illumination Invariant Single Face Image Recognition under Heterogeneous Lighting Condition," *Pattern Recognit.*, 2017.

[52] V. Vidya, N. Farheen, K. Manikantan, and S. Ramachandran, "Face recognition using threshold based DWT feature extraction and selective illumination enhancement technique," *Procedia Technol.*, vol. 6, pp. 334–343, 2012.

[53] R. Azmi and S. Yegane, "Facial expression recognition in the presence of occlusion using local Gabor binary patterns," in *Electrical Engineering (ICEE), 2012 20th Iranian Conference on*, pp. 742–747, 2012.

[54] J. Shermina and V. Vasudevan, "Face recognition system with various expression and occlusion based on a novel block matching algorithm and PCA," *Int. J. Comput. Appl.*, vol. 38, no. 11, 2012.

[55] X. P. Burgos-Artizzu, P. Perona, and P. Dollár, "Robust face landmark estimation under occlusion," in *Proceedings of the IEEE International Conference on Computer Vision* , pp. 1513–1520, 2013.

[56] J. Yan, Z. Lei, D. Yi, and S. Li, "Learn to combine multiple hypotheses for accurate face alignment," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 392–396,2013.

[57] P. Vandana and C. N. Manjunath, "Face Recognition In Presence Of Occlusion Using Machine Learning Classifier." *Expert Syst. Appl.*, vol. 62, pp. 327–333, 2016.

[58] G. Ghiasi and C. C. Fowlkes, "Occlusion coherence: Localizing occluded faces with a hierarchical deformable part model," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2385–2392,2014.

[59] R. Min, A. Hadid, and J.-L. Dugelay, "Efficient detection of occlusion prior to robust face recognition," *Sci. World J.*, vol. 2014, 2014.

[60] L. Zhang, D. Tjondronegoro, and V. Chandran, "Random Gabor based templates for facial expression recognition in images with facial occlusion," *Neurocomputing*, vol. 145, pp. 451–464, 2014.

[61] A. Vyas and R. Hablani, "Evaluating Facial Expressions with Different Occlusion around Image Sequence." *Sensors*, vol. 13, no. 6, pp. 7754–7760, 2013.

[62] V. P. Viswanath, N. K. Ragesh, and M. S. Nair, "ACM Based ROI Extraction for Pedestrian Detection with Partial Occlusion Handling," *Procedia Comput. Sci.*, vol. 46, pp. 45–52, 2015.

[63] R. Li, P. Liu, K. Jia, and Q. Wu, "Facial Expression Recognition under Partial Occlusion Based on Gabor Filter and Gray-Level Cooccurrence Matrix," in *Computational Intelligence and Communication Networks (CICN), 2015 International Conference on*, pp. 347–351, 2015.

[64] E. J. He, J. A. Fernandez, B. V. Kumar, and M. Alkanhal, "Masked correlation filters for partially occluded face recognition," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pp. 1293–1297, 2016.

[65] P. Ramalingam and S. Dhanushkodi, "Recognizing Expression Variant and Occluded Face Images Based on Nested HMM and Fuzzy Rule Based Approach," *Circuits Syst.*, vol. 7, no. 06, p. 983, 2016.

[66] P. Moallem, S. Shokrani, and M. Habibi, "Optimized PHOG and KNN for Robust Frontal Facial Expression Recognition against Occlusion," *Int. J. Tomogr. Simulation*^{TM}, vol. 29, no. 3, pp. 32–47, 2016.

[67] P. Ji, Y. Kim, Y. Yang, and Y.-S. Kim, "Face occlusion detection using skin color ratio and LBP features for intelligent video surveillance systems," in *Computer Science and Information Systems (FedCSIS), 2016 Federated Conference on*, pp. 253–259, 2016.

[68] W.-J. Li, "The Study of Anomaly Detection in Crowded Scenes using a Subspace Approach," 2017.

[69] S. Mohammadi, A. Perina, H. Kiani, and V. Murino, "Angry Crowds: Detecting Violent Events in Videos," in *European Conference on Computer Vision*, pp. 3–18., 2016

[70] J. M. Grant and P. J. Flynn, "Crowd Scene Understanding from Video: A Survey," *ACM Trans. Multimed. Comput. Commun. Appl. TOMM*, vol. 13, no. 2, p. 19, 2017.

[71] A. Das, R. K. Kumar, D. R. Kisku, and G. Sanyal, "Attention Identification via Relative Saliency of Localized Crowd Faces," in *Proceedings of the 10th International Conference on Informatics and Systems*, pp. 101–106, 2016.

[72] A. Patwardhan, "Edge Based Grid Super-Imposition for Crowd Emotion Recognition," *ArXiv Prepr. ArXiv161005566*, 2016.

[73] A. Das, R. K. Kumar, D. R. Kisku, and G. Sanyal, "A novel approach to attend faces in the crowd through relative visual saliency," in *Region 10 Conference (TENCON), 2016 IEEE*, pp. 3778–3781,2016.

[74] D.-Y. Chen and P.-C. Huang, "Visual-based human crowds behavior analysis based on graph modeling and matching," *IEEE Sens. J.*, vol. 13, no. 6, pp. 2129–2138, 2013.

[75] B. Solmaz, B. E. Moore, and M. Shah, "Identifying behaviors in crowd scenes using stability analysis for dynamical systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 2064–2070, 2012.

[76] J. Foytik, P. Sankaran, and V. Asari, "Tracking and recognizing multiple faces using kalman filter and modularpca," *Procedia Comput. Sci.*, vol. 6, pp. 256–261, 2011.

[77] P. Ekman, "An argument for basic emotions," *Cogn. Emot.*, vol. 6, no. 3–4, pp. 169–200, 1992.

[78] S. T. Roohi, D. S. Sachin, K. Manikantan, and S. Ramachandran, "Feature Accentuation Using Uniform Morphological Correction as Pre-processing Technique for DWT based Face Recognition," *Procedia Technol.*, vol. 10, pp. 690–698, 2013.

[79] J. Anil and L. P. Suresh, "Literature survey on face and face expression recognition," in *Circuit, Power and Computing Technologies (ICCPCT), 2016 International Conference on* pp. 1–6, 2016.

[80] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image Vis. Comput.*, vol. 27, no. 6, pp. 803–816, 2009.

[81] B. Jun and D. Kim, "Robust face detection using local gradient patterns and evidence accumulation," *Pattern Recognit.*, vol. 45, no. 9, pp. 3304–3316, 2012.

[82] A. Dhall, R. Goecke, S. Lucey, and T. Gedeon, "Acted facial expressions in the wild database," *Aust. Natl. Univ. Canberra Aust. Tech. Rep. TR-CS-11*, vol. 2, 2011.

[83] M. J. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba, and J. Budynek, *The Japanese female facial expression (JAFFE) database*. 1998.

[84] R. Azmi and S. Yegane, "Facial expression recognition in the presence of occlusion using local Gabor binary patterns," in *20th Iranian Conference on Electrical Engineering (ICEE2012)*, pp. 742–747, 2012.

[85] S. V. Khamele and S. G. Mundada, "An Approach for Restoring Occluded Images for Face-Recognition." *Artificial Intelligence Review*, pp.1-31,2017.

[86] B. F. Kingshott, "Crowd management: Understanding attitudes and behaviors," *J. Appl. Secur. Res.*, vol. 9, no. 3, pp. 273–289, 2014.