2015

A Survey Paper about Image and Voice Processing

Submitted To,
Ms Jawata Afnan
Lecturer
Dept of Electrical & Electronic Engineering
Shahjalal University of Science & Technology, Sylhet

Submitted By,
Sarah Siddiqua Brishti
Reg no: 2011338025
Md. Mohiminul Islam
Reg no: 2011338023
Student
Dept of Electrical & Electronic Engineering
Shahjalal University of Science & Technology, Sylhet

Submission Date
12/19/2015

Image Processing

Hand Gesture Recognition(HGR)

Discussion: Hand gestures recognition (HGR) is one of the main areas of research for the engineers, scientists and bioinformatics. It is the natural to interact with vision enabled computers and other machines. There is some initial methods that is necessary for HGR method is given bellow with those reference papers that are using those methods.

Initial methods:

- Edge Detection. *ref* [4],[9]
- Skin Detection Algorithm. *ref* [4],[9],[10]
- Morphological Model. *ref* [6],[8]
- Fingertip Detection. *ref* [4],[7],[9]
- Curvature of Perimeter. *ref* [9]
- Automatic Gesture Area Segmentation. *ref* [10]
- Orientation Normalization. *ref* [10]
- Shape Recognition. *ref* [2], [4]
- Orientation Histogram. *ref* [2],
- Large Object Tracking. *ref* [2],
- Vectorization. *ref* [2],[4],[6]
- Hough Transform. *ref* [6]
- Gaussian smoothing. *ref* [6]
- Dilation & Erosion. *ref* [6]
- Pixel Count Algorithm. *ref* [8]
- Detection of Circles. *ref* [8]
- Scanning Method. *ref* [8]

In some paper, there is shown some new initial methods by which those papers become unique. But in HGR, researchers have some different approaches. They are given bellow. By creating new approach, there is some being published.

Approaches:

- Artificial Neural Network. *ref* [2], [4],[7],[10]
- Fuzzy Logic Based Approach. *ref* [7]
- Genetic Algorithm Based Approach. *ref* [7]
- Haar-like Feature. *ref* [1],
- AdaBoost Learning Algorithm. *ref* [1],
- Context-Free Grammar-Based Syntactic Analysis. *ref* [1],
- SIFT algorithm. *ref* [6]
- Finger-Earth Mover's Distance (FEMD). . *ref* [3]

Available Tools: Basically Matlab &

- Microsoft® Foundation Classes (MFC).
- OpenCV Library.
- SGONG network.

Working Possibility: It is kind of software based but if we use it to control robot than this can be hardware based also. In this field, people need bust knowledge on image processing. Resources are available in this field. So we can work in this field.

Ideas:

- Using HGR, we can convert bangla sign language into text and voice for dumb and deaf people in Bangladesh.(Now-a-days people work on Americal sign language(ASL) & Indian sign language(ISL) but do not work on Bangla sign language. In some cases, they are successful to convert ASL into text and voice. I read about briefly later.
- Now-a-days HGR is used to create virtual mouse, controlling robot , gaming etc. So using HGR, we can control computer from gesture just like SUST CSE build a computer bangla voice control system for blind people.
- Identifying body language in CCTV and give a basic idea of every person on the CCTV screen.

References:

Hand Gesture Recognition(HGR):

1) **2007** Real-time Vision-based Hand Gesture Recognition Using Haar-like Features.

   Abstract: This paper proposes a two level approach to solve the problem of real-time vision-based hand gesture classification. The lower level of the approach implements the posture recognition with Haar-like features and the AdaBoost learning algorithm. With this algorithm, real-time performance and high recognition accuracy can be obtained. The higher level implements the linguistic hand gesture recognition using a context-free grammar-based syntactic analysis. Given an input gesture, based on the extracted postures, the composite gestures can be parsed and recognized with a set of primitives and production rules.

2) **2009** Real-Time Hand Tracking and Gesture Recognition System Using Neural Networks

Abstract: This paper introduces a hand gesture recognition system to recognize real time gesture in unstrained environments. Efforts should be made to adapt computers to our natural means of communication: Speech and body language. A simple and fast algorithm using orientation histograms will be developed. It will recognize a subset of MAL static hand gestures. A pattern recognition system will be using a transforrn that converts an image into a feature vector, which will be compared with the feature vectors of a training set of gestures. The final system will be Perceptron implementation in MATLAB. This paper includes experiments of 33 hand postures and discusses the results. Experiments shows that the system can achieve a 90% recognition average rate and is suitable for real time applications.

3) **2011** Depth camera based hand gesture recognition and its applications in Human-Computer-Interaction

Abstract: Of various Human-Computer-Interactions (HCI), hand gesture based HCI might be the most natural and intuitive way to communicate between people and machines, since it closely mimics how human interact with each other. Its intuitiveness and naturalness have spawned many applications in exploring large and complex data, computer games, virtual reality, health care, etc. Although the market for hand gesture based HCI is huge, building a robust hand gesture recognition system remains a challenging problem for traditional vision-based approaches, which are greatly limited by the quality of the input from optical sensors. [16] proposed a novel dissimilarity distance metric for hand gesture recognition using Kinect sensor, called Finger-Earth Mover's Distance (FEMD). In this paper, we compare the performance in terms of speed and accuracy between FEMD and traditional corresponding-based shape matching algorithm, Shape Context. And then we introduce several HCI applications built on top of a accurate and robust hand gesture recognition system based on FEMD. This hand gesture recognition system performs robustly despite variations in hand orientation, scale or articulation. Moreover, it works well in uncontrolled environments with background clusters. We demonstrate that this robust hand gesture recognition system can be a key enabler for numerous hand gesture based HCI systems.

4) **2011** Hand gesture recognition using different algorithms based on artificial neural network

Abstract: Gesture is one of the most natural and expressive ways of communications between human and computer in a real system. We naturally use various gestures to express our own intentions in everyday life. Hand gesture is one of the important methods of non-verbal communication for human beings. Hand gesture recognition based man-machine interface is being developed vigorously in recent years. This paper gives an overview of different methods for recognizing the hand gestures using MATLAB. It also gives the working details of recognition process using Edge detection and Skin detection algorithms.

5) **2012** Vision based hand gesture recognition for human computer interaction: a survey

Abstract: As computers become more pervasive in society, facilitating natural human–computer interaction (HCI) will have a positive impact on their use. Hence, there has been growing interest in the development of new approaches and technologies for bridging the human–computer barrier. The ultimate aim is to bring HCI to a regime where interactions with computers will be as natural as an interaction between humans, and to this end, incorporating gestures in HCI is an important research area. Gestures have long been considered as an interaction technique that can potentially deliver more natural, creative and intuitive methods for communicating with our computers. This paper provides an analysis of comparative surveys done in this area. The use of hand gestures as a natural interface serves as a motivating force for research in gesture taxonomies, its representations and recognition techniques, software platforms and frameworks which is discussed briefly in this paper. It focuses on the three main phases of hand gesture recognition i.e. detection, tracking and recognition. Different application which employs hand gestures for efficient interaction has been discussed under core and advanced application domains. This paper also provides an analysis of existing literature related to gesture recognition systems for human computer interaction by categorizing it under different key parameters. It further discusses the advances that are needed to further improvise the present hand gesture recognition systems for future perspective that can be widely used for efficient human computer interaction. The main goal of this survey is to provide researchers in the field of gesture based HCI with a summary of progress achieved to date and to help identify areas where further research is needed.

6) **2012** Real time robust hand gesture recognition and visual servoing

Abstract: The paper introduces hand gesture recognition and morphological model for visual servoing in real time unconstrained environments for positioning, docking or mobile target pursuing in robotics. Vision based hand tracking is an important problem in human computer interaction, since hand motions and gestures could potentially be used to interact with computers in more natural ways. A simple and fast process for gesture recognition and direction control is implemented. SIFT algorithm is used for hand gesture recognition by generating and comparing the feature vectors. Once the gesture is confirmed for particular system (e.g. motor) selection, further procedure would be to drive the system in required direction. For this purpose a morphological operation is performed and Hough transform is applied on the image of interest for angle measurement. This value is fed to the control system for further servo actions. The entire process is implemented in MATLAB. The paper includes experiments of 7 different hand postures and their results. Experiments show that the result is efficient and can be applied for real time applications.

7) **2013** Intelligent Approaches to interact with Machines using Hand Gesture Recognition in Natural way A Survey.

Abstract: Hand gestures recognition (HGR) is one of the main areas of research for the engineers, scientists and bioinformatics. HGR is the natural way of Human Machine interaction and today many researchers in the academia and industry are working on different application to make interactions more easy, natural and convenient without wearing any extra device. HGR can be applied from games control to vision enabled robot control, from virtual reality to smart home systems. In this paper we are discussing work done in the area of hand gesture recognition where focus is on the intelligent approaches including soft computing based methods like artificial neural network, fuzzy logic, genetic algorithms etc. The methods in the preprocessing of image for segmentation and hand image construction also taken into study. Most researchers used fingertips for hand detection in appearance based modeling. Finally the comparison of results given by different researchers is also presented.

8) **2014** Image processing algorithms for gesture recognition using MATLAB.

Abstract: Gesture recognition is the fast growing field in image processing and artificial technology. The gesture recognition is a process in which the gestures or postures of human body parts are identified and are used to control computers and

other electronic appliances. The most contributing reason for the emerging gesture recognition is that they can create a simple communication path between human and computer called HCI (Human Computer Interaction). This paper is confined to identification of hand postures and to establish a man-machine interaction. The hand region in the image is detected and the number of active fingers is determined. In this approach, the input which is an image or a frame from a video can be obtained from web camera or any other camera. This color image is converted into binary image and preprocessed and the number of fingers is counted using scanning method in MATLAB. This is a simple yet efficient approach. The main reason to employ scanning method is to make the code to recognize the finger count independent of size and rotation of the hand.

9) **2014** Real time hand gesture recognition for computer interaction.

Abstract: Hand gesture recognition is a natural and intuitive way to interact with the computer, since interactions with the computer can be increased through multidimensional use of hand gestures as compare to other input methods. The purpose of this paper is to explore three different techniques for HGR (hand gesture recognition) using finger tips detection. A new approach called "Curvature of Perimeter" is presented with its application as a virtual mouse. The system presented, uses only a webcam and algorithms which are developed using computer vision, image and the video processing toolboxes of Matlab.

10) **2015** Human Computer Interaction Using Face and Gesture.

Abstract: In this paper, we present a face and gesture recognition based human-computer interaction (HCI) system using a single video camera. Different from the conventional communication methods between users and machines, we combine head pose and hand gesture to control the equipment. We can identify the position of the eyes and mouth, and use the facial center to estimate the pose of the head. Two new methods are presented in this paper: automatic gesture area segmentation and orientation normalization of the hand gesture. It is not mandatory for the user to keep gestures in upright position, the system segments and normalizes the gestures automatically. The experiment shows this method is very accurate with gesture recognition rate of 93.6%. The user can control multiple devices, including robots simultaneously through a wireless network.

11) **2015** Tracking, Analysis, and Recognition of Human Gestures in Video.

Abstract: An overview of research in automated gesture spotting, tracking and recognition by the Image and Video Computing Group at Boston University is given. Approaches for localization and tracking human hands in video, estimation of hand shape and upper body pose, tracking head and facial motion, as well as efficient spotting and recognition of specific gestures in video streams are summarized. Methods for efficient dimensionality reduction of gesture time series, boosting of classifiers for nearest neighbor search in pose space, and model-based pruning of gesture alignment hypotheses are described. Algorithms are demonstrated in three domains: American Sign Language, hand signals like those employed by flight-directors on airport runways, and gesture-based interfaces for severely disabled users. The methods described are general and can be applied in other domains that require efficient detection and analysis of patterns in time-series, images or video.

Face Recognition(FR) & Facial Expression(FE)

Discussion: In between 2000 to 2012 year, Researcher create so many methods for FR but after that they focus on FE recognition and in between 2010 to 2015, they develop some FE recognition method also but now-a-days FR methods are kind of perfect but FE Recognition methods are not so perfect. So Facial Expression recognition is a developing field and working is going on in the field of artificial intelligence and human computer interaction. There is given some initial step or methods for FR.

FR Initial methods:

- C1-minimization. *ref* [6]
- Eigenface. *ref* [6]
- Laplacianface. *ref* [6]

In our survey, we can find some approaches for FR which are given bellow.

FR Approaches:

- Multiple Nose Region Matching. *ref* [1]
- Local Binary Patterns. *ref* [2]
- An Annotated Deformable Model Approach. *ref* [4]
- Sparse Representation. *ref* [6]
- Connected Component Analysis (YCbCr). *ref* [14]

Face expression recognition also have some initial steps or methods. And a interesting news is some that FR approaches are used as initial method for FE recognition. Those initial methods are given bellow.

FE Initial methods:

- Support Vector Machine (SVM). *ref* [5],[8],[10],[13]
- Multilayer Perceptron (MLP). *ref* [8]
- Principal Component Analysis (PCA). *ref* [8]
- Generalized Feed Forward Neural Network (GFFNN). *ref* [8]
- Patch matching operation. *ref* [9]
- GentleBoost. *ref* [10]
- Genetic Algorithm. *ref* [12]
- Connected Component Analysis (YCbCr). *ref* [15]
- Morphological operation. *ref* [15]
- AAM (Active Appearance Model) method. *ref* [15]

Researchers are developing some FE recognition approaches now-a-days. And latest approaches are most prominent than relatively oldest. In 2015, a researcher team is confident that their Viola and Jones algorithm is more prominent than others and they show comparison this and others in there paper. *ref* [19]. Those FE recognition approaches are given bellow.

FE Approaches:

- Facial-action-dynamics recognition. *ref* [3]
- Gabor wavelet. *ref* [7],[10],[11]
- Learning vector quantization. *ref* [7]
- Artificial Neural Network. *ref* [8]
- 3D Gabor features. *ref* [9]
- Human visual cortex. *ref* [11]
- Classifier synthesis. *ref* [11]
- Fuzzy Inference System. *ref* [12]
- Local Directional Number Pattern. *ref* [13]
- Euclidean Distance method. *ref* [15]
- Viola and Jones algorithm. *ref* [19]

Now-a-days FR & FE recognition is used for some application. There is some application given bellow.

FR & FE Applications:

- Facial expression recognition using PCA based interface for wheelchair. *ref* [16]
- Development of biometric security system using CBIR and EER. *ref*[17]
- Face Recognition in Surveillance System. *ref*[18]

Available Tools: Basically Matlab, Image processing tools and some mathematics algorithm

- Artificial Neuro-Fuzzy Inference System (ANFIS). *ref* [15]

Working Possibility: It is kind of software based but if we use it to control robot than this can be hardware based also. In this field, people need bust knowledge on image processing. Resources are available in this field. So we can work in this field. But if we compare this with HGR then working possibility of FR & FE is less than HGR because FE recognition is higher form of gesture recognition if we want to detect expression perfectly.

Ideas:

- Emotion Detection in CCTV.
- Allow personal computer to detect emotion from webcam HCI. By this information computer do some work relatively comparing his/her emotion.

References:

Face Recognition & Facial Expression:

1) **2006** Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression.

Abstract: An algorithm is proposed for 3D face recognition in the presence of varied facial expressions. It is based on combining the match scores from matching multiple overlapping regions around the nose. Experimental results are presented using the largest database employed to date in 3D face recognition studies, over 4,000 scans of 449 subjects. Results show substantial improvement

over matching the shape of a single larger frontal face region. This is the first approach to use multiple overlapping regions around the nose to handle the problem of expression variation.

2) **2006** Face Description with Local Binary Patterns: Application to Face Recognition.

Abstract: This paper presents a novel and efficient facial image representation based on local binary pattern (LBP) texture features. The face image is divided into several regions from which the LBP feature distributions are extracted and concatenated into an enhanced feature vector to be used as a face descriptor. The performance of the proposed method is assessed in the face recognition problem under different challenges. Other applications and several extensions are also discussed.

3) **2006** Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences.

Abstract: Automatic analysis of human facial expression is a challenging problem with many applications. Most of the existing automated systems for facial expression analysis attempt to recognize a few prototypic emotional expressions, such as anger and happiness. Instead of representing another approach to machine analysis of prototypic facial expressions of emotion, the method presented in this paper attempts to handle a large range of human facial behavior by recognizing facial muscle actions that produce expressions. Virtually all of the existing vision systems for facial muscle action detection deal only with frontal-view face images and cannot handle temporal dynamics of facial actions. In this paper, we present a system for automatic recognition of facial action units (AUs) and their temporal models from long, profile-view face image sequences. We exploit particle filtering to track 15 facial points in an input face-profile sequence, and we introduce facial-action-dynamics recognition from continuous video input using temporal rules. The algorithm performs both automatic segmentation of an input video into facial expressions pictured and recognition of temporal segments (i.e., onset, apex, offset) of 27 AUs occurring alone or in a combination in the input face-profile video. A recognition rate of 87% is achieved.

4) **2007** Three-Dimensional Face Recognition in the Presence of Facial Expressions: An Annotated Deformable Model Approach.

Abstract: In this paper, we present the computational tools and a hardware prototype for 3D face recognition. Full automation is provided through the use of advanced multistage alignment algorithms, resilience to facial expressions by employing a deformable model framework, and invariance to 3D capture devices through suitable preprocessing steps. In addition, scalability in both time and space is achieved by converting 3D facial scans into compact metadata. We present our results on the largest known, and now publicly available, face recognition grand challenge 3D facial database consisting of several thousand scans. To the best of our knowledge, this is the highest performance reported on the FRGC v2 database for the 3D modality.

5) **2007** Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines.

Abstract: In this paper, two novel methods for facial expression recognition in facial image sequences are presented. The user has to manually place some of Candide grid nodes to face landmarks depicted at the first frame of the image sequence under examination. The grid-tracking and deformation system used, based on deformable models, tracks the grid in consecutive video frames over time, as the facial expression evolves, until the frame that corresponds to the greatest facial expression intensity. The geometrical displacement of certain selected Candide nodes, defined as the difference of the node coordinates between the first and the greatest facial expression intensity frame, is used as an input to a novel multiclass Support Vector Machine (SVM) system of classifiers that are used to recognize either the six basic facial expressions or a set of chosen Facial Action Units (FAUs). The results on the Cohn-Kanade database show a recognition accuracy of 99.7% for facial expression recognition using the proposed multiclass SVMs and 95.1% for facial expression recognition based on FAU detection.

6) **2008** Robust Face Recognition via Sparse Representation.

Abstract: We consider the problem of automatically recognizing human faces from frontal views with varying expression and illumination, as well as occlusion and disguise. We cast the recognition problem as one of classifying among

multiple linear regression models and argue that new theory from sparse signal representation offers the key to addressing this problem. Based on a sparse representation computed by C1-minimization, we propose a general classification algorithm for (image-based) object recognition. This new framework provides new insights into two crucial issues in face recognition: feature extraction and robustness to occlusion. For feature extraction, we show that if sparsity in the recognition problem is properly harnessed, the choice of features is no longer critical. What is critical, however, is whether the number of features is sufficiently large and whether the sparse representation is correctly computed. Unconventional features such as downsampled images and random projections perform just as well as conventional features such as eigenfaces and Laplacianfaces, as long as the dimension of the feature space surpasses certain threshold, predicted by the theory of sparse representation. This framework can handle errors due to occlusion and corruption uniformly by exploiting the fact that these errors are often sparse with respect to the standard (pixel) basis. The theory of sparse representation helps predict how much occlusion the recognition algorithm can handle and how to choose the training images to maximize robustness to occlusion. We conduct extensive experiments on publicly available databases to verify the efficacy of the proposed algorithm and corroborate the above claims.

7) **2008** Recognition of facial expressions using Gabor wavelets and learning vector quantization.

Abstract: Facial expression recognition has potential applications in different aspects of day-to-day life not yet realized due to absence of effective expression recognition techniques. This paper discusses the application of Gabor filter based feature extraction in combination with learning vector quantization (LVQ) for recognition of seven different facial expressions from still pictures of the human face. The results presented here are better in several aspects from earlier work in facial expression recognition. Firstly, it is observed that LVQ based feature classification technique proposed in this study performs better in recognizing fear expressions than multilayer perceptron (MLP) based classification technique used in earlier work. Secondly, this study indicates that the Japanese Female Facial Expression (JAFFE) database contains expressers that expressed expressions incorrectly and these incorrect images adversely affect the development of a reliable facial expression recognition system. By excluding the two expressers from the data set, an improvement in recognition rate from 87.51% to 90.22% has

been achieved. The present study, therefore, proves the feasibility of computer vision based facial expression recognition for practical applications like surveillance and human computer interaction.

8) **2009** Emotion Recognition from Facial Expression Using Neural Networks.

Abstract: This research aims at developing "Humanoid Robots" that can carry out intellectual conversation with human beings. The first step of our research is to recognize human emotions by a computer using neural network. In this paper all six universally recognized principal emotions namely angry, disgust, fear, happy, sad and surprise along with neutral one are recognized. Various neural networks such as Support Vector Machine (SVM), Multilayer Perceptron (MLP), Principal Component Analysis (PCA), and Generalized Feed Forward Neural Network (GFFNN) are employed and their performance is compared. 100% recognition accuracy is achieved on training data set (seen examples) and test data set (unseen examples).

9) **2011** Facial Expression Recognition Using Facial Movement Features.

Abstract: Facial expression is an important channel for human communication and can be applied in many real applications. One critical step for facial expression recognition (FER) is to accurately extract emotional features. Current approaches on FER in static images have not fully considered and utilized the features of facial element and muscle movements, which represent static and dynamic, as well as geometric and appearance characteristics of facial expressions. This paper proposes an approach to solve this limitation using "salient" distance features, which are obtained by extracting patch-based 3D Gabor features, selecting the "salient" patches, and performing patch matching operations. The experimental results demonstrate high correct recognition rate (CRR), significant performance improvements due to the consideration of facial element and muscle movements, promising results under face registration errors, and fast processing time. Comparison with the state-of-the-art performance confirms that the proposed approach achieves the highest CRR on the JAFFE database and is among the top performers on the Cohn-Kanade (CK) database.

10) **2011** Fully Automatic Recognition of the Temporal Phases of Facial Actions.

Abstract: Past work on automatic analysis of facial expressions has focused mostly on detecting prototypic expressions of basic emotions like happiness and anger. The method proposed here enables the detection of a much larger range of facial behavior by recognizing facial muscle actions [action units (AUs)] that compound expressions. AUs are agnostic, leaving the inference about conveyed intent to higher order decision making (e.g., emotion recognition). The proposed fully automatic method not only allows the recognition of 22 AUs but also explicitly models their temporal characteristics (i.e., sequences of temporal segments: neutral, onset, apex, and offset). To do so, it uses a facial point detector based on Gabor-feature-based boosted classifiers to automatically localize 20 facial fiducial points. These points are tracked through a sequence of images using a method called particle filtering with factorized likelihoods. To encode AUs and their temporal activation models based on the tracking data, it applies a combination of GentleBoost, support vector machines, and hidden Markov models. We attain an average AU recognition rate of 95.3% when tested on a benchmark set of deliberately displayed facial expressions and 72% when tested on spontaneous expressions.

11) **2012** Facial expression recognition using radial encoding of local Gabor features and classifier synthesis

Abstract: Primarily motivated by some characteristics of the human visual cortex (HVC), we propose a new facial expression recognition scheme, involving a statistical synthesis of hierarchical classifiers. In this scheme, the input images of the database are first subjected to local, multi-scale Gabor-filter operations, and then the resulting Gabor decompositions are encoded using radial grids, imitating the topographical map-structure of the HVC. The codes are fed to local classifiers to produce global features, representing facial expressions. Experimental results show that such a hybrid combination of the HVC structure with a hierarchical classifier significantly improves expression recognition accuracy when applied to wide-ranging databases in comparison with the results in the literature. Furthermore, the proposed system is not only robust to corrupted data and missing information, but can also be generalized to cross-database expression recognition.

12) **2012** A novel fuzzy facial expression recognition system based on facial feature extraction from color face images.

Abstract: Emotion recognition plays an effective and important role in Human–Computer Interaction (HCI). Recently, various approaches to emotion recognition have been proposed in the literature, but they do not provide a powerful approach to recognize emotions from Partially Occluded Facial Images. In this paper, we propose a new method for Emotion Recognition from Facial Expression using Fuzzy Inference System (FIS). This novel method is even able to recognize emotions from Partially Occluded Facial Images. Moreover, this research describes new algorithms for facial feature extraction that demonstrate satisfactory performance and precision. In addition, one of the main factors that have an important influence on the final precision of fuzzy inference systems is the membership function parameters. Therefore, we use a Genetic Algorithm for parameter-tuning of the membership functions. Experimental results report an average precision rate of 93.96% for Emotion Recognition of six basic emotions, which is so promising.

13) **2012** Local Directional Number Pattern for Face Analysis: Face and Expression Recognition.

Abstract: This paper proposes a novel local feature descriptor, local directional number pattern (LDN), for face analysis, i.e., face and expression recognition. LDN encodes the directional information of the face's textures (i.e., the texture's structure) in a compact way, producing a more discriminative code than current methods. We compute the structure of each micro-pattern with the aid of a compass mask that extracts directional information, and we encode such information using the prominent direction indices (directional numbers) and sign-which allows us to distinguish among similar structural patterns that have different intensity transitions. We divide the face into several regions, and extract the distribution of the LDN features from them. Then, we concatenate these features into a feature vector, and we use it as a face descriptor. We perform several experiments in which our descriptor performs consistently under illumination, noise, expression, and time lapse variations. Moreover, we test our descriptor with different masks to analyze its performance in different face analysis tasks.

14) **2014** A Human Face Detection Method Based on Connected Component Analysis.

Abstract: This paper describes a human face detection algorithm where human skin colour is detected using 2D Colour Space Skin Clustering Method in YCbCr

color model. Then the face is detected using connected component analysis followed by a refining process containing a set of human face shape criteria.

15) **2014** Face Detection and Facial Expression.

Abstract: A human-computer interaction system for an automatic face recognition or facial expression recognition has attracted increasing attention from researchers in psychology, computer science, linguistics, neuroscience, and related disciplines. In this paper, an Automatic Facial Expression Recognition System (AFERS) has been proposed. The proposed method has three stages: (a) face detection, (b) feature extraction and (c) facial expression recognition. The first phase of face detection involves skin color detection using YCbCr color model, lighting compensation for getting uniformity on face and morphological operations for retaining the required face portion. The output of the first phase is used for extracting facial features like eyes, nose, and mouth using AAM (Active Appearance Model) method. The third stage, automatic facial expression recognition, involves simple Euclidean Distance method. In this method, the Euclidean distance between the feature points of the training images. and that of the query image is compared. Based on minimum Euclidean distance, output image expression is decided. True recognition rate for this method is around 90% - 95%. Further modification of this method is done using Artificial Neuro-Fuzzy Inference System (ANFIS). This non-linear recognition system gives recognition rate of around 100% which is acceptable compared to other methods.

16) **2014** Facial expression recognition using PCA based interface for wheelchair.

Abstract: Human beings transmit a lot of apparent information visually rather than vocally. This consistent expression recognition by machine remains a challenge as of today. To automate recognition of emotional state, machines must be taught to realize facial gestures. The wheelchair is one of the kinds of assistive tools for physically challenged people and elderly people to gain mobility and lead to self-governing life. The use of wheelchair with high maneuverability and navigational intelligence is one of the greatest steps towards the integration of severely disabled people. The aim of this work is to generate a model of a wheelchair command interface that does not require the other's hands. It includes 3 major modules. They are face detection, facial expression recognition and command generation. The software contains digital image processing for face

detection, principal component analysis for facial expression recognition and generating a command signals for interfacing the wheelchair. The algorithm is tested in MATLAB. The stagnant images represent different persons face and facial expressions of a subject, from Indian Face Database and Japanese Female Face Database are used to appraise the effectiveness of the algorithm respectively.

17) **2015** Development of biometric security system using CBIR and EER.

Abstract: The paper discusses about the content based image retrieval systems for biometric security. As image databases are growing at a rapid rate, demand for efficient and effective tools for retrieval of images increased significantly. Among them, content-based image retrieval systems (CBIR) have become very popular for browsing, searching and retrieving images from a large database of digital images. Biometric security methods have been gaining importance increasingly in the recent years due to advances in biometrics technology and its reliability and efficiency in real world applications. Biometric security methods take into account human's unique physical or behavioural characteristics that help to identify them. The proposed approach is based on three main features of an image: colour, texture and shape which can be incorporated into biometric security for effective security systems. The inclusion of fuzzy heuristics presents a increased accuracy on image retrieval than the existing approaches.

18) **2015** Face Recognition in Surveillance System.

Abstract: Face identification has made its presence evident as an impact of advancement in technologies. Face identification in video surveillance has always been an challenging field where it needs to undergo a series of consideration to identify a particular face in the video In this paper initially a real time facial detection is performed using an open source environment that runs in processing 2,2.1 and later on captured faces are correlated withthe template faces that are stored in the database using Matlab once the faces are identified then the profile of the person is displayed to notify the behavioral status of the person who is under study.

19) **2015** Facial Expression Detection Techniques Based on Viola and Jones Algorithm and Principal Component Analysis.

Abstract: Facial expression is a prominent posture beneath the skin of the face. They are the way of communication in humans which convey many things non-verbally. During the past years face recognition has received significant attention as one of the most important applications of image understanding and analysis. Many algorithms have been implemented on different static and non-static conditions .Static conditions include static and uniform background, identical poses, similar illumination, neutral frontal face .Non static conditions include position, partial occlusion orientation; varying lightening conditions and facial hair which make recognition process a complex problem. All these factors influence face recognition process. The main stages for face recognition include face detection, feature representation and classifications. Researchers have described distinct approaches for face recognition. In this work we present a glimpse of face detection techniques, methods used, their performance & their limitations and proposed a new technique for Face Detection based on Viola and Jones algorithm and principal component analysis. At the end we have shown simulation results for the proposed technique and established that proposed technique is performing better than the existing one.

Sign Language

Discussion: Sign language converts to text & Voice is basically based on HGR. In this field, Researchers work on ASL & ISL but we can not find any sign of work in Bangla sign language. So it will be a new field in Bangladesh. In our world, Dump and deaf people face lots of problem in communication because maximum people do not know sign language. Microsoft and google are working on that. there is some youtube video link.

https://www.youtube.com/watch?v=YMbpnMXEq3E

https://www.youtube.com/watch?v=s-WGElbqw4s

The methods of this field are same as HGR that we are discussed. So it can be a great field for work. If we work hard on this than it can be possible for BSL

because this is fully software based and Matlab R2015a has video processing toolbox that can be so usefull.

In this link, we can know matlab features.

http://www.mathworks.com/products/matlab/whatsnew.html

References:

Sign Language:

1) **2010** Appearance Based Recognition of American Sign Language Using Gesture Segmentation.

   Abstract: The work presented in this paper goals to develop a system for automatic translation of static gestures of alphabets in American Sign Language. In doing so three feature extraction methods and neural network is used to recognize signs. The system deals with images of bare hands, which allows the user to interact with the system in a natural way. An image is processed and converted to a feature vector that will be compared with the feature vectors of a training set of signs. The system is rotation, scaling of translation variant of the gesture within the image, which makes the system more flexible. The system is implemented and tested using data sets of number of samples of hand images for each signs. Three feature extraction methods are tested and best one is suggested with results obtained from ANN. The system is able to recognize selected ASL signs with the accuracy of 92.33%

2) **2011** Real time Indian Sign Language Recognition System to aid deaf-dumb people.

   Abstract: The Sign Language is a method of communication for deaf - dumb people. This paper proposes a method that provides a basis for the development of Sign Language Recognition system for one of the south Indian languages. In the proposed method, a set of 32 signs, each representing the binary `UP' & `DOWN' positions of the five fingers is defined. The images are of the palm side of right hand and are loaded at runtime i.e. dynamic loading. The method has been developed with respect to single user both in training and testing phase. The static

images have been pre-processed using feature point extraction method and are trained with 10 numbers of images for each sign. The images are converted into text by identifying the finger tip position of static images using image processing techniques. The proposed method is able to identify the images of the signer which are captured dynamically during testing phase. The results with test images are presented, which show that the proposed Sign Language Recognition System is able to recognize images with 98.125% accuracy when trained with 320 images and tested with 160 images.

3) **2011** Shape, texture and local movement hand gesture features for Indian Sign Language recognition.

Abstract: This paper proposes an automatic gesture recognition approach for Indian Sign Language (ISL). Indian sign language uses both hands to represent each alphabet. We propose an approach which addresses local-global ambiguity identification, inter-class variability enhancement for each hand gesture. Hand region is segmented and detected by YCbCr skin color model reference. The shape, texture and finger features of each hand are extracted using Principle Curvature Based Region (PCBR) detector, Wavelet Packet Decomposition (WPD-2) and complexity defects algorithms respectively for hand posture recognition process. To classify each hand posture, multi class non linear support vector machines (SVM) is used, for which a recognition rate of 91.3% is achieved. Dynamic gestures are classified using Dynamic Time Warping (DTW) with the trajectory feature vector with 86.3% recognition rate. The performance of the proposed approach is analyzed with well known classifiers like SVM, KNN & DTW. Experimental results are compared with the conventional and existing algorithms to prove the better efficiency of the proposed approach.

4) **2014** An American Sign Language detection system using HSV color model and edge detection.

Abstract: The work presented in this paper is aimed to design an automatic vision based American Sign Language detection system and translation to text. To detect the human skin color from the image, HSV color model is used. Then edge detection is applied to detect the hand shape from the image. A set of morphological operation is applied to get a refined output for the sign language recognition.

Voice Processing

Discussion: Emotion recognition from Voice is a comprehensive framework for real-time recognition of emotions from acoustic properties of speech.

Initial Methods:

1. Audio characteristic detection(Formant/Pitch) *ref[1]*
2. Facial feature extraction *ref[1],[2]*
3. classification algorithm *ref[1] ref[2] ref[3]*
4. visual information detection ref[1]
5. Feature classification into different emotion classes *ref[2]*
6. Gaussian noise study *ref[2]*
7. Speech spectrogram *ref[6]*

Approaches:

1. Mel-frequency Cepstral Coefficient (MFCC)*ref[1],[5]*
2. HSV color model *ref[1]*
3. Mahalanobis distance *ref[1]*
4. multiclassifier scheme *ref[1]*
5. linear Support Vector Machine classifier *ref[2]*
6. Valence-Arousal approach *ref[3]*
7. Harmony parameters *ref [4],[5]*
8. Music theory *ref[4]*
9. Fourier parameter model *ref[5]*
10. Neural Network classifier *ref[6]*
11. Estimation of complex emotion as well as its dynamic changes in a three-dimensional PAD (Position–Arousal–Dominance) *ref[7]*

Voice-Emotion Recognition:

1) **2008** Recognizing Human Emotional State From Audiovisual Signals.

Abstract: Machine recognition of human emotional state is an important component for efficient human-computer interaction. The majority of existing works address this problem by utilizing audio signals alone, or visual information

only. In this paper, we explore a systematic approach for recognition of human emotional state from audiovisual signals. The audio characteristics of emotional speech are represented by the extracted prosodic, Mel-frequency Cepstral Coefficient (MFCC), and formant frequency features. A face detection scheme based on HSV color model is used to detect the face from the background. The visual information is represented by visual information We perform feature selection by using a stepwise method based on Mahalanobis distance. The selected audiovisual features are used to classify the data into their corresponding emotions. Based on a comparative study of different classification algorithms and specific characteristics of individual emotion, a novel multiclassifier scheme is proposed to boost the recognition performance. The feasibility of the proposed system is tested over a database that incorporates human subjects from different languages and cultural backgrounds. Experimental results demonstrate the effectiveness of the proposed system. The multiclassifier scheme achieves the best overall recognition rate of 82.14%.

2) **2009** Voice and Facial Expression Based Classification of Emotion Using Linear Support Vector Machine.

Abstract: The paper provides a novel approach to emotion recognition from facial expression and voice of subjects. The subjects are asked to manifest their emotional exposure in both facial expression and voice, while uttering a given sentence. Facial features including mouth-opening, eye-opening, eyebrow-constriction, and voice features including, first three formants: F1, F2, and F3, and respective powers at those formants, and pitch are extracted for 7 different emotional expressions of each subject. A linear Support Vector Machine classifier is used to classify the extracted feature vectors into different emotion classes. Sensitivity of the classifier to Gaussian noise is studied, and experimental results confirm that the recognition accuracy of emotion up to a level of 95% is maintained, even when the mean and standard deviation of noise are as high as 5% and 20% respectively over the individual features. A further analysis to identify the importance of individual features reveals that mouth-opening and eye-opening are primary features, in absence of which classification accuracy falls off by a large margin of more than 22%.

3) **2013** Valence-arousal approach for speech emotion recognition system.

Abstract: The current state-of-the-art speech emotion recognition approaches focus on discrete emotion classification. In practical perspective, emotion is deemed complex to be individually segregated and it is a continuous process that will change dynamically over time. In this paper, an alternative approach of Valence-Arousal is introduced based on the psychologists' understanding that emotion can be represented in two emotion primitives of Valence and Arousa. Results show better speech emotion understanding and analysis.

4) **2015** Emotion recognition from speech signals using new harmony features.

Abstract: In this paper we propose a new set of harmony features for automatic emotion recognition from speech signals. They are based on the psychoacoustic harmony perception known from music theory. Starting from the estimated pitch contour of an utterance, we calculate the circular autocorrelation of the pitch histogram on the logarithmic semitone scale. It measures the occurrence of different two-pitch intervals which cause a consonant or dissonant impression. Experiments of emotion recognition using these harmony parameters in addition to state of the art features show an improved recognition performance.

5) **2015** Speech Emotion Recognition using Fourier Parameter.

Abstract: Recently, studies have been performed on harmony features for speech emotion recognition. It is found in our study that the first- and second-order differences of harmony features also play an important role in speech emotion recognition. Therefore, we propose a new Fourier parameter model using the perceptual content of voice quality and the first- and second-order differences for speaker-independent speech emotion recognition. Experimental results show that the proposed Fourier parameter (FP) features are effective in identifying various emotional states in speech signals. They improve the recognition rates over the methods using Mel frequency cepstral coefficient (MFCC) features by 16.2, 6.8 and 16.6 points on the German database (EMODB), Chinese language database (CASIA) and Chinese elderly emotion database (EESDB). In particular, when combining FP with MFCC, the recognition rates can be further improved on the aforementioned databases by 17.5, 10 and 10.5 points, respectively.

6) **2015** Improvement of speech emotion recognition with neural network classifier by using speech spectrogram.

Abstract: This research presents a novel algorithm for detecting human emotion via speech recognition by using speech spectrogram. The proposed algorithm aims to detect the emotional by using information inside the spectrogram. Neural network was used for being the classifier. A new approach to feature extraction based on analysis of two dimensions time-frequency representation of a speech signal have been presented. The algorithm was tested with EMO-Database. The experimental results show that the proposed framework can efficiently find the correct speech emotion compared to using the comparing method.

7) **2015** Emotion recognition and affective computing on vocal social media.

Abstract: Vocal media has become a popular method of communication in today's social networks. While conveying semantic information, vocal messages usually also contain abundant emotional information; this emotional information represents a new focus for data mining in social media analytics. This paper proposes a computational method for emotion recognition and affective computing on vocal social media to estimate complex emotion as well as its dynamic changes in a three-dimensional PAD (Position–Arousal–Dominance) space; furthermore, this paper analyzes the propagation characteristics of emotions on the vocal social media site WeChat.

Voice Recognition

Initial Method:

1. Phoneme recognization *ref[1]*
2. Word recognization *ref[1]*
3. Verification algorithm *ref[2]*
4. Acoustic feature extraction *ref[3]*
5. MFCC *ref[3]*
6. Feature Vectorization
7. Vector Quantization *ref[9]*

Approaches:

1. Two stage multilayer neural networks. *ref[3]*
2. Dynamic Time Warping (DTW). *ref[4]*
3. Voice Activity Detection (VAD) *ref[5]*

4. Speech Enhancement Algorithm (SEA) *ref[5]*
5. Hidden Markov models (HMMs). ref[3] *ref[5]*
6. Mel frequency cepstrum Coefficient (MFCC). *ref[6]*
7. Linear prediction coefficients (LPC) *ref[7]*
8. Artificial neural network *ref[7]*
9. Voice Print (VP) *ref[8]*
10. Gaussian Mixtures Models (GMM) *ref[8]*
11. Neural and fuzzy logic implimentation *ref[9]*
12. Hamming windowing *ref[10]*
13. discrete fractional Fourier transform *ref[11]*

Tools:

Mainly Matlab

References:

1) **2004** A voice recognition system for speech impaired people.

   Abstract: We show how MATLAB can be used to realize voice recognition for people with cerebral paralysis. The first part of this paper is about what cerebral paralysis is, and its consequences on the way people with this illness talk. We then mention the mathematical background needed to realize this project. We explain the first stage of the system in which some phonemes are recognized. Next, we describe the functioning of the second stage of the project in which words are recognized. Finally, we give some results and we discuss them.

2) **2008** Security system using biometric technology: Design and implementation of Voice Recognition System (VRS).

   Abstract: In this paper, we present an implementation of a security system based on voice identification as the access control key. Verification algorithm is developed using MATLAB (SIMULINK) function blocks which is capable of authenticating a person identity by his or her voice pattern. A voice match will produce logic dasia1psila while a mismatch, logic dasia0psila. A microcontroller circuit controlling access to a door is built to test the reliability of this voice controlled security system. It is found out that the developed voice recognition software has successfully activated the door opening mechanism using a voice

command that ONLY works for the authenticated individual. The system is proven to be able to provide medium-security access control and also has an adjustable security level setting to account for the variations in onepsilas voice each time a voice identification occurs.

3) **2010** Bangla speech recognition using two stage multilayer neural networks.

Abstract: This paper describes a Bangla phoneme recognition method for Automatic Speech Recognition (ASR). The method consists of two stages: i) a multilayer neural network (MLN), which converts acoustic features, mel frequency cepstral coefficients (MFCCs), into phoneme probabilities and ii) the phoneme probabilities obtained from the first stage and corresponding Δ and ΔΔ are inserted into another MLN to improve the phoneme probabilities for the hidden Markov models (HMMs). it is observed that the proposed method provides higher phoneme recognition performance than the existing method.

4) **2011** Based on Artificial Neural Networks for voice recognition word segment.

Abstract: Speech recognition is to study so that the machine can accurately hear the voice of the contents of people's problems, that is, accurate identification of words. Since the 90's it has been the focus of speech recognition has shifted to large vocabulary, continuous speech, speaker-up, man-machine voice communication has become a research focus. Use one kinds of arithmetic to study speech recognition: Dynamic Time Warping (DTW). Build a basic test platform of speech recognition using DTW.

5) **2012** Performance analysis of hybrid robust automatic speech recognition system.

Abstract: In this paper, we evaluate the performance of several objective measures in terms of predicting the quality of noisy input speech signal through the Hybrid method using Voice Activity Detection (VAD) and Speech Enhancement Algorithm (SEA). Demand for Speech Recognition technology is expected to rise dramatically over the next few years as people use their mobile phones and voice recognition system everywhere. This paper enlighten the implementation process which includes a speech-to-text system using isolated word recognition with a vocabulary of ten words (digits 0 to 9). In the training period, the uttered digits are recorded using 8-bit Pulse Code Modulation (PCM) with a sampling rate of 8 KHz and save as a wave format file using sound

recorder software. For a given word in the vocabulary, the system builds an Hidden Markov Model (HMM) model and trains the model during the training phase. The training steps, from VAD, Speech Enhancement to HMM model building, are performed using PC-based Matlab programs.

6) **2012** The extraction and simulation of Mel frequency cepstrum speech parameters.

   Abstract: This paper takes consideration of (voices of) the characteristics of voice processing by the human auditory system, adopts triangle filter to do signal preprocessing, and uses logarithm operations of all filter output for extracting Mel frequency cepstrum Coefficient (MFCC). By Matlab simulation of MFCC vectors of typical signal of male and female, an analyses is given of the probability to be applied into speech recognition systems.

7) **2012** Security enhancement of voice over Internet protocol using speaker recognition technique.

   Abstract: Speaker recognition (SR) technique has proven to be an effective tool in many applications. In this study, SR was used to enhance the security of voice over Internet protocol system. This work focused on authentication issues and user identification detection depending on the voice signature of the incoming call. The authors suggest the use of SR concepts to support user identification process. First, a complete SR system based on linear prediction coefficients (LPC) feature extraction and decision based on artificial neural network was described and built using MATLAB package. Then, MATLAB package was used to develop a prototype version of the modified session initiation protocol (SIP), which is used to manage calls establishment among the network nodes. MATLAB codes were written for both SIP server and client nodes. The modified SIP protocol was implemented on an experimental network which is built to demonstrate the operation of the enhanced authentication procedure.

8) **2013** Automatic speaker recognition using a unique personal feature vector and Gaussian Mixture Models.

   Abstract: This article presents an automatic speaker recognition system implemented in Matlab, which uses a unique feature vector, the so-called “Voice Print” (VP), to describe the voice. The system uses Gaussian Mixtures Models

(GMM) in the classification process. The final part of the paper presents research on the efficiency of speaker recognition for different variants of the system, as well as the results of optimisation of the system.

9) **2013** Design and implementation of an Automatic Speaker recognition system using neural and fuzzy logic in Matlab.

Abstract: This paper presents a Text Dependent Automatic Speaker recognition system developed and simulated using Matlab. For a better computational efficiency, the system is trained to store voice of the same person under various physiological conditions such as coughing, shouting, during chewing, mouth covered etc. A dictionary is created to store the signature features of each user's voice. A neural networks is then trained using back propagation and accordingly weights are obtained to recognize voice in the testing phase. The efficiency of the proposed system is then compared to the system implemented using vector quantization.

10) **2014** Evaluation of MFCC for speaker verification on various windows.

Abstract: Speech identification has been an imperative part of human life acting as one of the five senses of individual body, due to which application developed on the basis of speech identification has high degree of approval. Different steps are carried out in order to analyze the speaker verification using small programs with the help of MATLAB 10.0 and for this paper; the system will recognize continuous words or sentences. This paper discusses a modified approach to implement a text-independent speaker verification system and to evaluate the recognition rate using different windows. The experiment has been performed on English language speech database for speaker recognition(ELSDSR) and Indian voice database (IVD) and has been found that as the number of training voice samples increases the recognition rate of the system increases and out all of the windows used hamming window gives the best result. Further result show that highest recognition rate can be achieved by other windows in case if test voice samples persists in one train voice samples.

11) **2014** Discrete Fractional Fourier Transform and Vector Quantization Based Speaker Identification System.

Abstract: In the study of speaker identification, Mel Frequency Cepstral Coefficient (MFCC) method is the best and most popular which is used to feature extraction. Further Vector Quantization (VQ) technique is used to minimize the amount of data to be handled and mapping vectors from a large vector space to a finite number of regions in that space in recent years. Voice, like other biometrics, cannot be forgotten or misplaced, unlike knowledge-based (e.g., password) or possession-based (e.g., key) access control methods. In the present work, modified Mel frequency cepstral coefficients using discrete fractional Fourier transform and vector quantization is obtained. The experimental results are analyzed with the help.

12) **2015** Voice control for a gripper using Mel-Frequency Cepstral Coefficients and Gaussian Mixture Models.

Abstract: This work presents an implementation of a speaker-dependent speech recognition system used to control a gripper. The application was made using MATLAB and the gripper was assembled using the Lego Mindstorm NXT robotic kit. Four commands are implemented for controlling the gripper: Open, close, rotate left and rotate right. The development was divided into two stages. In training stage, weuse Mel Frequency Cepstral Coefficients (MFCCs) and Gaussian Mixture Models (GMMs) to generate a representation of each defined command. Then, in testing stage, those models are used to identify the speaker's utterance and send the command to the actuator. Finally, we present test results that show a performance of 95.09% for our system, and then we compare it with similar works.

Application of Voice Recognition system:

1) **2008** Voice controlled automation system.

Abstract: In this era of technology, rapid advancements are being made in the field of automation and signal processing. The developments made in digital signal processing are being applied in the field of automation, communication systems and biomedical engineering. Controlling through human speech is one of the fascinating applications of digital signal processing (DSP) and Automation Systems. This paper discusses the speech recognition and its application in control mechanism. Speech recognition can be used to automate many tasks that usually

require hands-on human interaction, such as recognizing simple spoken commands to perform something like turning on lights or shutting a door or driving a motor. After speech recognition, a particular code related to spoken word is transferred through wireless communication system to 8051 microcontroller and it works accordingly. Many techniques have been used to compare the pattern of the speech signals and recent technological advancements have made recognition of more complex speech patterns possible. Despite these breakthroughs, however, current efforts are still far away from 100% recognition of natural human speech. Therefore, the project is considered involving processing of a speech signal in any form as a challenging and rewarding one.

2) **2010** Application of Speech Emotion Recognition in Intelligent Household Robot.

Abstract: This thesis aims to perform categorized recognition of 5 speech emotions as represented by joy, grief, anger, fear and surprise by means of algorithm with the combination of HMM and SOFMNN models. In this way, robot can recognize emotional information as contained in the human speech signals for friendly interaction with human beings, and eventually satisfactory performance.

3) **2011** Voice command interpretation for robot control.

Abstract: This paper presents some initial results from an analysis of performance of a voice command interpretation and authorisation system using voiceprint to identify the human-commander. Two approaches based on human voice related algorithms are proposed. Mel-frequency cepstral coefficient (MFCC) and perceptual linear predictive (PLP) are two feature extraction methods that are closely mimic the human auditory system. The two methods were applied to the proposed system to determine their suitability for use in a commander recognition system. Vector Quantization (VQ) with Linde-Buzo-Gray (LBG) iterative algorithm was used for clustering for the classification of commanders. The performance of the algorithms was evaluated to compare between two methods in MATLAB simulation environment based on, false rejection rate (rejecting an authorised commander), false acceptance rate (accepting unauthorised commander) and the execution time. Based on the initial results, both methods achieved accurate classification and PLP method has shown better execution time and lower false-acceptance rate compared to the MFCC. The combined approach (MFCC-PLP) did not show considerable improved performances to the individual feature models PLP and MFCC without incurring high computational costs that will compromise the performance of the speaker recognition tasks. Therefore, PLP method is the best candidate for command-recognition system to be developed in the second phase of this research.

4) **2012** An interactive and efficient voice processing system for embedded applications.

   Abstract: This paper proposes a novel speech recognition and colour sensing technique based on Formant frequency and Euclidean distance analysis for embedded systems. In this paper Formant frequencies are used as acoustic features for speech recognition. The objective of this paper is to present the design of an embedded system that will be helpful for the physically impaired individuals sspeech recognition system, a central controller and the robotic arm. In the time to come, it is anticipated that speech recognition systems will play a major role as man-machine interface to control robots in manufacturing industries, healthcare, entertainment etc. Formal modeling and Matlab simulation approach is used as a robust foundation for developing real-time and embedded applications. The experimental and simulation results show that the proposed algorithm makes a good balance between the computational complexity and recognitionaccuracy, and thus is more useful for embedded systems.

5) **2013** Voice Recognition and Visualization Mobile Apps Game for Training and Teaching Hearing Handicaps Children.

   Abstract: This paper introduces a mobile application (app) game for teaching hearing handicapped children, Then, we discuss the definitions of hearing handicapped, hearing impairment children, their study and learning problems in the literature review. We also proposed a design model for the development of mobile application game that can teach and train the hearing handicapped children to learn new knowledge and for speaking purposes. We also propose the block diagram of the voice recognition mobile apps game design for hearing handicapped children.

6) **2014** Hybrid brain computer interface in wheelchair using voice recognition sensors.

   Abstract: Brain computer interface is a direct technological interface between a brain and the computer. It is a fast growing emergent technology in which researches try to build a direct channel between a brain and a computer. In this paper we use a Hybrid BCI in order to provide large number of commands with increased accuracy to the BCI user. BCI can serve as powerful aids for severely

disabled people in their daily life. This technology is especially used to help the disabled people to move voluntarily. In BCI wheelchair the movement of the wheel chair is going to be controlled by adding voice recognition sensors in order to give an effective result with less effort. The voice and the brain signals when synchronises with each other then it is converted as a movement in the wheelchair by moving it towards left, right, accelerating, decelerating etc,. In this paper we propose a hybrid brain computer interface in wheelchair using voice recognition sensors which can be used by physically challenged persons by reducing the effort for high concentration power.

7) **2015** An integrated system for voice command recognition and emergency detection based on audio signals.

Abstract: this paper presents a complete solution for voice command recognition and emergency detection based on audio signals entirely integrated in a low-consuming embedded platform. The system combines an active operation mode were distress calls are captured and a vocal interface is enabled for controlling the home automation subsystem, and a pro-active mode, were a novelty detection algorithm detects abnormal acoustic events to alert the user of a possible emergency. In the first operation mode, a Voice Activity Detector captures voice segments of the audio signal, and a speech recogniser detects commands and distress calls. In the pro-active mode, an acoustic novelty detector is employed in order to be able to deal with unknown sounds, thus not requiring an explicit modelling of emergency sounds. In addition, the system integrates a VoIP infrastructure so that emergencies can be communicated to relatives or care centres. The monitoring unit is equipped with multiple microphones and it is connected to the home local area network to communicate with the home automation subsystem. The algorithms have been implemented in a low-consuming embedded platform based on a ARM Cortex-A8 CPU. The effectiveness of the adopted algorithms has been tested on two different databases: ITAAL and A3Novelty. The obtained results show that the adopted solutions are suitable for speech and audio event monitoring in a realistic scenario.

8) **2015** Preprocessing for elderly speech recognition of smart devices.

Abstract: In general, speech recognition systems are optimized to an average adult's voice and tend to exhibit a lower accuracy rate when recognizing an

elderly person's voice due to the effects of speech articulation and speaking style. Thus, using a preprocessing application on a smart device can not only deliver better speech recognition but also substantially reduce any added costs. The speech recognition rate for elderly adults could be improved by means of increasing the speech rate, eliminating silence periods and boosting the energy of the formant frequency bands. After all the preprocessing, a 12% increase in the accuracy of elderly speech recognition was achieved.

9) 2015 Development of a voice-controlled home automation using Zigbee module.

Abstract: In this study, we aimed speech recognition-based remote control of home devices. The system was designed at two stages: speech recognition module and transmitter module that sends commands including fewer than two main headings. In the first stage speech recognition algorithm is implemented in MATLAB. Mel Frequency Cepstrum Coefficients (MFCCs) are used as a feature extraction method. Dynamic Timing Warping (DTW) is used as a feature matching method. After the recognition of speech commands entered, Arduino pin situation is changed via USB. The voltage change in pin is detected by the same pin connected to the transmitter Zigbee module. When the change is detected, Zigbee transmitter sends data to connecting on the desired device to be controlled the receiver Zigbee module via wireless network. When sending data detected by the receiver Zigbee module, 5V relay is triggered connecting on the device pins of the module. Thus, it was possible to successfully remotely control any device with speech commands using wireless network.