

**Pakistan Air Force**

**Karachi Institute of Economics & Technology**

**Sign Language Recognition Using Image Processing**

Submitted by: Zain Baig, Hammad Abid

Advisor: Adnan Naseer

##### College of Computing & Information Sciences

A final year project report presented to the College of Computing & Information Sciences,

Pakistan Air Force – Karachi Institute of Economics & Technology in partial fulfilment of the requirements of the degree of Bachelor of Science

**2021**

**Table of Contents**

[**Abstract** 2](#_Toc69794384)

[**Introduction** 3](#_Toc69794385)

[**Literature Review** 4](#_Toc69794386)

[Paper No.1 4](#_Toc69794387)

[Paper No.2 5](#_Toc69794388)

[Paper No.3 6](#_Toc69794389)

[Paper No.4 7](#_Toc69794390)

[Paper No.5 8](#_Toc69794391)

[Paper No.6 10](#_Toc69794392)

[Paper No.7 10](#_Toc69794393)

[Paper No.8 12](#_Toc69794394)

[Paper No.9 13](#_Toc69794395)

[Paper No.10 14](#_Toc69794396)

[Paper No.11 15](#_Toc69794397)

[Paper No.12 16](#_Toc69794398)

[Paper No.13 17](#_Toc69794399)

[**Related Work** 19](#_Toc69794400)

[**Implementation** 19](#_Toc69794401)

[**Image collection** 20](#_Toc69794402)

[**Image pre-processing** 21](#_Toc69794403)

[**Feature extraction** 21](#_Toc69794404)

[**Classification** 22](#_Toc69794405)

[**Workflow** 24](#_Toc69794406)

[**Result** 25](#_Toc69794407)

[**Conclusion** 26](#_Toc69794408)

[**Reference** 26](#_Toc69794409)

# **Abstract**

Indian sign language (ISL) is one of the challenging topics as it is in rudimentary stage of its development, unlike American Sign Languages (ASL). This project aims at classification of Indian sign languages using machine learning models. There has been broad research on ASL and adequate data is available to analyze it. As, India is multi diverse country, there are several regions and cultures which results in different variations of languages for communication. So, there are very limited standard data sets, which has variations and noises. ISL uses both hands to make gestures instead of one hand unlike ASL. It leads to occlusion of features and this is a major barrier for the lack of development in this field. This project aims at helping in then research of this field further by providing a data set of ISL. A data of sign language was created by us for alphabets and numeric. Later, the features will be extracted from the collected segmented data using image pre-processing and Bag of words model. Histograms are generated to map the alphabets with images. In the final step, these features will be fed to supervised models for the classification.

# **Introduction**

A gesture is a pattern which may be static, dynamic or both, and is a form of nonverbal communication in which bodily motions convey information. Communication is an important aspect when it comes to share or express information, feelings, and it brings people closer to each other with better understanding. When it comes to disabled persons for example deaf and dumb people, it becomes tougher for them to communicate using natural language. So, they use sign language to communicate with themselves and with entire world. But normal people find it difficult to understand sign language as they do not have mostly any prior education or experience in this. Sign language is composed of visual gestures and signs, which are used by deaf and mute for their talking. It is a well-structured code gesture where every sign has a specific meaning allotted to it. These signs are not only used for alphabets or numeric but also for common expressions also for example greetings and sentences There are 143 existing different sign languages all over the world, mainly American Sign Language (ASL), British Sign Language, French Sign Language, Japanese Sign Language, and Indian Sign Language (ISL). Every country has its own language, similarly, sign language is not a universal language and differs from country to country. There has been a lot work already done on ASL recognition as it is widely learnt language all over the globe. ASL uses single hand in the gesture representation and it is simple comparing to ISL. ISL uses both the hands for gesture representation and it is complex comparing to ASL. Because of this reason, there is less research and development in this field. This project goal is to take the simple step in connecting the social and communication bridge between regular people and the disabled people with the help of Indian Sign Language. As our project only deals with alphabets and numeric in ISL, it can be extended to common expressions and also words which can be more effective for disabled and normal people in communication and understanding. As we live in a century where India is developing at a rapid pace in terms of digital and technological advances, this project could be one of the steppingstones where technology meets humanity and help the hearing impaired and mute community.

# **Literature Review**

### Paper No.1

**Title**: Comparative Analysis On Sign Language Recognition System

**Published**: INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 8, ISSUE 08, AUGUST 2019

**Publishers**: Swati Sharma, Varun Prakash Saxena, Kandukuru Satish.

**Problem Statement**: A process of exchanging information, expressions and views between two and more than two people is communication. It can be either Verbal or Non-Verbal communication. Non-Verbal communications make use of hand gestures for interaction. Hand Gestures can also be termed as sign language which works in a systematic way. With the help of this sign language, it is becoming easier for people with disability to interact with the normal people. The people who suffer from hearing disabilities makes them different from others as they lack behind due to their impaired hearing as there is a system which is used by them to interact with normal people is SLR. Sign Language Recognition (SLR) system basically used as an intermediate between normal people and hearing impaired. In this survey, various techniques that were implemented for sign lang.

**Introduction:** India is a country where more than 27 national languages are used, and as we know few peoples of this country cannot even able to speak single language at all. According to India’s National Association of Deaf, there are almost 18 million people roughly we can say 1 percent of the total Indian population is deaf. In our routine life, we need to interact with each other either by listening or talking whenever possible. But when some people of our society who are hearing impaired needs to interact with us then we find it as huge barrier as we are unable to understand their language properly and vice-versa. To remove that communication gap between the normal people and the hearing-impaired sign language is used. Sign language recognition system acts as an intermediate communication platform between the normal people and the hearing impaired. A gesture can be defined as a movement of any part of the body mainly the shapes made by hand and expression made by either the head or face that is used to express, a sign, emotion or sentiment

**Methodology:** The importance of Sign language can be explained from the fact that early humans used this language for communication between them because at that time there was no local language used by them. In this section, the details of previous framework is analyzed and compared on the basis of algorithms and methods used in them.A Methodology for the recognition of signs in Indian Sign Language in which the dataset has been divided into two parts one is training and other is testing. Images are then preprocessed and converted into binary format. Feature Extraction is applied where the HOG features and the geometric features are calculated for the image. Classification is performed by using support vector machine and K-Nearest Neighbor. Experimental Analysis shows that SVM provides more accuracy than KNN. SVM when used with HOG feature is able to provide an accuracy of 94.23%.

### Paper No.2

**Title**: Recognition of Indian Sign Language in Live Video

**Published**: International Journal of Computer Applications (0975 – 8887) Volume 70– No.19, May 2013

**Publishers**: Joyeeta Singha, Karen Das Department of Electronics and Communication DBCET, Assam Don Bosco University Guwahati, Assam

**Problem Statement**: Sign Language Recognition has emerged as one of the important areas of research in Computer Vision. The difficulty faced by the researchers is that the instances of signs vary with both motion and appearance. Thus, in this paper a novel approach for recognizing various alphabets of Indian Sign Language is proposed where continuous video sequences of the signs have been considered. The proposed system comprises of three stages: Preprocessing stage, Feature Extraction and Classification. Preprocessing stage includes skin filtering, histogram matching. Eigen values and Eigen Vectors were considered for feature extraction stage and finally Eigen value weighted Euclidean distance is used to recognize the sign. It deals with bare hands, thus allowing the user to interact with the system in natural way.

**Introduction:** A Sign Language is a language in which communication between people are made by visually transmitting the sign patterns to express the meaning. It is a replacement of speech for hearing and speech impaired people. Thus, because of which it has attracted many researchers in this field from long. Many researchers have been working in different sign languages like American Sign Language, British Sign Language, Taiwanese Sign Language, etc. but few works have been made progress on Indian Sign Language. The hearing-impaired people become neglected from the society because the normal people never try to learn ISL nor try to interact with the hearing-impaired people. This becomes a curse for them and so they mostly remain uneducated and isolated. Thus, recognition of sign language was introduced which has not only been important from engineering point of view but also for the impact on society.

**Methodology:** The proposed system which comprises of 3 major stages-preprocessing stage which includes the skin filtering and histogram matching to find out the similarity between frames, Feature Extraction stage in which the Eigen values and Eigen vector are being considered as features and finally Eigen value weighted Euclidean distance-based classification technique. The details of each stage will be discussed in the following sections.

The first step for our proposed system is the capturing of the video using webcam where different alphabets were taken into consideration. 24 different alphabets were considered for testing from 20 people. Some of the continuous video frames captured.

Skin Filtering was performed to the input video frames for detection of hand gestures. It was done so that the required hand could be extracted from the background. Skin Filtering is a technique used for separating the skin-colored regions from the non-skin-colored regions.

Feature Extraction stage is necessary because certain features has to be extracted so that they are unique for each gesture or sign. After the decision is made that a sign is present, then the last frame is taken into consideration and features like Eigen values and Eigen vectors are extracted from that frame. The procedure to calculate the features.

### Paper No.3

**Title**: Recent Developments in Indian Sign Language Recognition: An Analysis

**Published**: Daleesha M Viswanathan et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (1) , 2015, 289-293

**Publishers**: Daleesha M Viswanathan, Sumam Mary Idicula

**Problem Statement**: There exists significant variation between sign language recognition processes across the world, although there are many similarities. Pre-processing, feature extraction and classification are the three major steps involved in the sign language recognition process. An analysis of scientific literature indicates the potential of various methods in achieving significantly high accuracy in image recognition. Further examination of the literature indicates the voluminous works carried out in American Sign Language recognition systems and most of these works compare the potential of various methods and combination of methods for their accuracy. Although, the comparison using randomly selected gestures for their potential would result in realistic overall accuracy for ASL where the gestures are simple and distinct, the complete adoption of such methods for Indian Sign Language (ISL) recognition may not be ideal due to the complexity in ISL. Other than static gestures, the dynamic gestures, gestures including facial expression, similarity in gestures, all increase the complexity of ISL. Therefore, the potential of different methods and their combinations need to evaluate in the context of ISL. A preliminary study to analyze the potential of promising feature extraction methods indicated that the methods could vary significantly while handling gestures with resemblances. This clearly indicates the necessity to screen gesture recognition methods for their accuracy in handling gestures in the context of complex ISL.

**Introduction:** Sign languages have originated and evolved independently at different parts of the world. Sometimes, the connection with the native language and prevalent sign language are not very conspicuous. As an example, although British and Americans are predominantly native English speakers, their sign language differs significantly. American Sign Language (ASL) is single handed while handling alphabets, whereas British sign language (BSL) uses both hands except for the alphabet C (Parness et al., 2007). There exists significant variation between sign languages across the world although there are similarities. The ASL, for instance has a strong connection with French sign language, similar is the case for Arabic sign language (Mohandas, 2013), however, the ASL varies significantly with BSL and Australian sign languages (Parness et al., 2007). There exist enormous complexity in Indian sign Languages and the challenges are manifold. Morgan M. W (1998) reviewed the Indian sign languages and inferred that the prevalent sign language has a strong South Indian Connection. Attempts have been made to have regional sign languages based on the local dialect and this further enhance the challenge of having a common and uniform ISL (Rajma and Balakrishnan, 2012). Other than static gestures, the dynamic gestures, gestures including facial expression, similarity in gestures all increases the complexity of ISL (Nandy et al., 2010). For example, slight variation in positioning of index finger on nose would interpret as “Think” or “Woman” in ISL(Morgan, 2009).Also, at interpretation level the language follows Subject-Object-Verb pattern contrary to ASL where S-V-O pattern is followed.

**Methodology:** The Digital image processing techniques and artificial neural network were employed in recognizing finger spelling in ISL (Adithya et al., 2013). In an approach, Otsu's algorithm was employed in segmentation and Scale Invariant Feature Transform and Histogram of Oriented Gradient were combined to evolve the feature vector. An accuracy level of 93% was achieved by this approach (Agrawal et al., 2012).The Artificial intelligence possibilities were utilized to convert the clerk's speech to signs played by a virtual 3D animated human character playing the signs corresponds to the speech in Malayalam (language of Kerala state, India) to ISL (Anuja et al., 2009).The number of finger tips and their distance from centroid is utilized together with PCA for Indian sign language recognition and high accuracy is achieved by this approach (Deora et al., 2012). In an attempt, Indian Sign Language (ISL) is recognized with 96% accuracy and translated to normal text. The Hu invariant moment and a multi-class Support Vector Machine (MSVM) is employed in the recognition process (Dixit and Jalal, 2013). An attempt has been made by extracting Maximum Curvature Points (MCPs) as key frames resulted in high accuracy for Indian Sign Language (ISL) (Geetha and Aswathi, 2013).

The scalability problems of available recognition systems have minimized by segmentation based on Maximum Curvature Points (MCPs) and thus reduce the requirement of large training data set and reduce the complexity (Geetha et al., 2013). In an attempt, classification is performed using direction histogram due to high performance for illumination and invariance of orientation. The approaches based on K-nearest neighbor metrics and Euclidean distance resulted in high performance in recognizing ISL (Nandy et al., 2010). The Table 1 indicates the most dominating research works done on ISL.

### Paper No.4

**Title**: A Survey on Hand Gesture Recognition for Indian Sign Language

**Published**: International Research Journal of Engineering and Technology (IRJET), Volume: 03 Issue: 10 | Oct -2016

**Publishers**: Miss. Juhi Ekbote, Mrs. Mahasweta Joshi.

**Problem Statement**: Sign Language is most accepted and meaningful way of communication for deaf and dumb people of the society. Sign language uses gestures, head/body movements and facial expressions for communication. It is a powerful means of Non-Verbal communication among humans. Every country has its own developed Sign Language. The language which is used in India is called as “Indian Sign Language (ISL)”. Only little research work has been carried out in this area as ISL has got standardized recently. Currently many researchers have mainly focused on gesture recognition that has been recorded under static hand gesture. Only few works have been reported for recognition of dynamic hand gesture. Many methods have been developed to recognize alphabets and numerals of ISL. The major steps involved in designing such a system are: gesture acquisition, tracking and segmentation, feature extraction and gesture recognition. This paper presents a survey on various hand gesture recognition approaches for ISL.

**Introduction:** Sign language provides a way for the deaf and dumb people to communicate with the outside world. Instead of voice, sign language uses gestures to communicate. Sign language is an organized way of communication in which every word or alphabet is assigned to a particular gesture [2]. Sign language is made up of a range of gestures produced by various facial expressions and movements of hands or head/body. In the last several years there has been an increased interest among the researchers in the field of sign language recognition to introduce a means of interaction from human –human to human – computer interaction. Various applications of Gesture Recognition System are: Human Computer Interface, Video gaming, Augmented reality, Home appliances, Robotics, Sign Language, etc. Hearing-impaired and speech impaired people depend on sign language interpreters for communication. But finding skilled and trained interpreters for their day-to-day interactions during life time is a very difficult job and also too expensive [7]. All over the world, a variety of sign languages exist. The sign language depends on the traditions and spoken language of that place. Indian sign language (ISL) is used by hearing and speech impaired people in India [3]. The gestures are primarily divided into two classes: Static gestures and Dynamic Gestures. Static gestures include only configurations and poses whereas dynamic gestures include postures, strokes, prestrikes, phases and also emotions [1]. ISL alphabets and numeric signs are represented.

**Methodology:** A webcam integrated in the laptop or an external webcam or a simple digital camera can be used to capture the images. Either an on-hand database can be used or it can be formed by the researchers themselves. To track the movement of hand, segmentation of hand is required. Segmentation separates ROI from background. Hand tracking is a technique used to know position of hand. Next step is feature extraction which is carried out to extract important features after successful completion of hand tracking and segmentation. The complete process of recognition can be separated into the following two stages - training and testing. Training is the initial stage where the classifier is trained using the training database. The key steps implicated in training stage are creation of database, pre-processing, extraction of features and training of the classifier. The key steps involved in the testing phase are gesture acquisition, preprocessing, feature extraction and classification.

### Paper No.5

**Title**: Perlustration of Deaf-And-Dumb Alphabet Detection and interpretation.

**Published**: 2017 IJEDR | Volume 5, Issue 2 | ISSN: 2321-9939.

**Publishers**: Ms.N.Priyadharsini, Mrs.N.Rajeswari, Department of Computer Science And Engineering, SVCE, Sriperumbudur, India.

**Abstract**: A sign language appertains to recognize the meaningful expressions which uses the gestures like combined hand-shapes, orientation and movement of the hands, arms or body, facial expressions and lip-patterns instead of sound to convey ideas. It is of uttermost importance in designing an intelligent and efficient human–computer interface. The applications of sign language ranges from medical rehabilitation to loud venues. However, successful recognition of lexical signs is not sufficient for understanding of proper Sign language communication. The Non-manual signs and the grammatical processes, results in the variation in the appearance of signs are integral as respects of interaction but have received comparatively less attention. In this survey, we examine the data acquisition, preprocessing, feature derivation and classification methods employed for scrutinize the sign language gestures. These are discussed with respect to the issues like modeling transitions between signs in continuous signing, signer independence, and adaptation, also with the overall progress towards a true test of sign recognition systems—dealing with natural signing by native signers. Every hand gesture recognition is depending on characters in different Sign Language (SL). The related research area of sign gesture recognition is Human Computer Interaction (HCI) and image processing which helps to solve this problem. Existing challenges and future research possibilities are also highlighted.

**Introduction:** Signs are very expressive, meaningful body motions involving physical changes of the body parts like fingers, hands, arms, head, face, or body with the intent of: 1) conveying the meaningful information or 2) interaction with the environment. They constitute one interesting subspace of possible human motion. A sign may also be recognized by the environment as a compression technique for the information to be transmitted elsewhere and subsequently reconstructed by the receiver. Sign recognition has wide-ranging applications such as the following:

• developing aids for hearing impaired people;

• enabling young children to interact with computers;

• designing techniques for finding forensics;

• recognizing the sign language;

• medically monitoring patients’ stress levels or emotional states;

• navigating and/or manipulating in virtual environments;

• communicating in the video conferencing;

• tele-teaching assistance/ distance learning

**Methodology:** In this section a survey of most recent works in Sign Language Recognition and their results are discussed. Hidden Markov Models (HMMs), Accelerometer and SEMG, Field Programmable Gate Arrays (FPGA) and Self Organizing Map (SOM), Histograms of Oriented Gradients (HOG) and Scale Invariant Feature Transform (SIFT), Bag of Features (BOF), Block Search Scheme and K-curvature algorithm, Labeled Graph SVM(LGSVM) and Labeled Graph Logistic Regressor (LGLR) are the most preferred methods used for Sign Language Recognition. Sign Language Recognition based on Hidden Markov Model is presented in[14]. Recognition system for understanding the words of home-service-related sign language. Data received from sensor are sequential HMM that has been successfully applied to the speech signals, entropy-based K-means algorithm is used to evaluate the number of states in the HMM model with the help of entropy diagram. Database contains 11 home-service-related Taiwan sign language words and each word is performed ten times, five males and five females are invited to perform such words with the average recognition rate of 91.3%.

### Paper No.6

**Title**: Hand Gesture Recognition Based on Karhunen-Loeve Transform

**Published**: Department of Electronics and Communication Engineering 1,2Assam. Don Bosco University, Guwahati, Assam, India.

**Publishers**: Joyeeta Singha , Karen Das.

**Abstract**: In this paper, we have proposed a system based on KL Transform to recognize different hand gestures. The system consists of five steps: skin filtering, palm cropping, edge detection, feature extraction, and classification. Firstly the hand is detected using skin filtering and palm cropping was performed to extract out only the palm portion of the hand. The extracted image was then processed using the Canny Edge Detection technique to extract the outline images of palm. After palm extraction, the features of hand were extracted using K-L Transform technique and finally the input gesture was recognized using proper classifier. In our system, we have tested for 10 different hand gestures, and recognizing rate obtained was 96%. Hence we propose an easy approach to recognize different hand gestures.

**Introduction:** Hand gesture recognition is one of the growing fields of research today which provides a natural way of human machine interaction. Gestures are some forms of actions which a person expresses in order to express information to others without saying it. In our daily life, we can observe few hand gestures frequently used for communication purpose like thumbs up, thumbs down, victory, directions etc. Some common examples are in cricket where the umpire uses different hand gestures to show different events that occurred at that instant on the match, hand gestures used by the traffic police, etc.

**Methodology:** In our paper, we have firstly used Skin filtering where the RGB image is converted to HSV image because this model is more sensitive to changes in lighting condition. And then K-L transform is performed. The advantage of K-L transform is it can eliminate the correlated data, reduces dimensions keeping average square error minimum, and gives excellent cluster character after the transform. Some applications in this field that has already been done, for example hand gesture recognition for sign language, hand gestures used for controlling robot’s motion, in video games, etc.

### Paper No.7

**Title**: Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique.

**Published**: (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 2, 2013.

**Publishers**: Joyeeta Singha, Karen Das Dept. of Electronics and Communication Engineering Assam Don Bosco University Guwahati, India

**Abstract**: In this paper, we have proposed a system based on KL Transform to recognize different hand gestures. The system consists of five steps: skin filtering, palm cropping, edge detection, feature extraction, and classification. Firstly the hand is detected using skin filtering and palm cropping was performed to extract out only the palm portion of the hand. The extracted image was then processed using the Canny Edge Detection technique to extract the outline images of palm. After palm extraction, the features of hand were extracted using K-L Transform technique and finally the input gesture was recognized using proper classifier. In our system, we have tested for 10 different hand gestures, and recognizing rate obtained was 96%. Hence we propose an easy approach to recognize different hand gestures.

**Introduction:** Sign Language Recognition is one of the most growing fields of research today. Many new techniques have been developed recently in these fields. Here in this paper, we have proposed a system using Eigen value weighted Euclidean distance as a classification technique for recognition of various Sign Languages of India. The system comprises of four parts: Skin Filtering, Hand Cropping, Feature Extraction and Classification. 24 signs were considered in this paper, each having 10 samples, thus a total of 240 images was considered for which recognition rate obtained was 97%.

**Methodology:** We have proposed a system which is able to recognize the various alphabets of Indian Sign Language for Human Computer interaction giving more accurate results at least possible time. It will not only benefit the deaf and dumb people of India but also could be used in various applications in the technology field. We have thus proposed a system that could overcome the difficulties faced by various. Our proposed system was able to recognize two hand gestures with an improved accuracy rate of 97%. Moreover, experiment was carried out with bare hands and computational time was also less thus removing the difficulties faced by use of the hand gloves with sensors. The input RGB image is first converted to the HSV image. The motive of performing this step is RGB image is very sensitive to change in illumination condition. The HSV color space separates three components: Hue which means the set of pure colors within a color space, Saturation describing the grade of purity of a color image and Value giving relative lightness or darkness of a color. The following Fig. 4 shows the different components of HSV color model.

Different images were tested and found that the new technique of classification was found to show 97% accuracy. Some images tested with other database images are given in the following table where 2 levels of classification were used to identify the gestures. Table I shows the Level 1 classification experimented for different test images and Table II shows the level 2 classification. A comparison between the first level and second level of classification is being made in Table III and it is seen that the success rate has improved from 87% to 97% with the use of the Eigen value weighted Euclidean distance between Eigen vectors as a classification technique.

### Paper No.8

**Title**: Sign Language Recognition System to aid Deaf-dumb People Using PCA.

**Published**: Shreyashi Narayan Sawant / International Journal of Computer Science & Engineering Technology (IJCSET).

**Publishers**: Shreyashi Narayan Sawant. Department of Electronics and Telecommunication Engineering Rajarambapu Institute of Technology Rajaramnagar, Islampur- 415 409.

**Abstract**: The Sign Language is a method of communication for deaf-dumb people. Here vision-based approach has been used. This paper presents design and implementation of real time Sign Language Recognition system to recognize 26 gestures from the Indian Sign Language using MATLAB. The signs are captured by using web cam. These signs are preprocessed for feature extraction using HSV color model. The obtained features are compared by using Principal Component Analysis (PCA) algorithm. After comparing features of captured sign with testing database minimum Euclidean distance is calculated for sign recognition. Finally, recognized gesture is converted into text and voice format. This system provides an opportunity for a deaf-dumb people to communicate with non-signing people without the need of an interpreter.

**Introduction:** The sign language is an important method of communication for deaf-dumb persons. As sign language is well structured code gesture, each gesture has a meaning assigned to it. In the last several years there has been an increased interest among the researchers in the field of sign language recognition to introduce means of interaction from human –human to human – computer interaction. Deaf and Dumb people rely on sign language interpreters for communications. However, finding experienced and qualified interpreters for their day-to-day affairs throughout life period is a very difficult task and also unaffordable [1]. The propose system is able to recognize single handed gestures accurately with a single normal webcam using bare human hands and convert it into text and voice message. The aim of this project is to recognize the gestures with highest accuracy and in least possible time and translate the alphabets of Indian Sign Language into corresponding text and voice in a vision-based setup.

**Methodology:** Here our system takes the input hand gestures through the web camera. In this proposed method, 26 combinations of Indian sign are developed by using right hand stored in training data base. Preprocessing is done on these captured input gestures. Then the Segmentation of hands is carried out to separate object and the background. The segmented hand image is represented using certain features. These features are used for gesture recognition using the PCA algorithm which gives optimized results. The final result obtained is converted into corresponding text and voice. The sign recognition procedure includes four major steps. They are a) Data Acquisition b) Preprocessing and segmentation c) Feature extraction d) Sign recognition and e) Sign to text, voice conversion.

### Paper No.9

**Title**: An Analysis of Features for Hand-Gesture Classification.

**Published**: IWSSIP 2010 - 17th International Conference on Systems, Signals and Image Processing.

**Publishers**: Thiago R. Trigo and Sergio Roberto M. Pellegrino Electronic Engineering and Computer Department Aeronautics Technological Institute (ITA) São José dos Campos, Brazil.

**Abstract**: The human-computer interaction, also known as HCI, depends mostly on physical devices. The goal of this work is the evaluation and analysis of methods which allows the user to interact to machines using a hand gesture based natural language. Here we present some approaches which are used in HCI systems based on hand gesture and a new proposal that uses geometric shape descriptors for hand gesture classification. The results analysis shows that this new proposal beats some limitations of other known HCI methods.

**Introduction:** The human-computer interaction, also known as HCI, depends mostly on physical devices. Normally, people use mouses and keyboards, but this kind of interaction can be unnatural for humans, who are used to express their ideas, their feelings and their wishes through voice, corporal gestures, facial expression, hands gestures, and so on. Therefore, an interface which allows one of those kinds of interaction between human and machines would be more natural and instinctive for users. If we consider interactions concerning to object manipulation, then the interaction based on hand gestures seems to be an attractive solution. In the most popular HCI systems based on hand gesture, the user needs to wear special gloves which measure the hand pose and the joint angles [1]. The problem of this kind of technique is that once the user has to wear a glove, the system becomes invasive, besides the fact of special gloves being expensive. Due to this, a hand-gestures interface based on computer vision appears as a reasonable option to reach a more natural human computer communication. Since the hands are capable to produce great number of gestures (thanks to its flexibility and fingers), we defined a gesture alphabet containing a reduced number of gestures to be recognized. The Figure 1 shows the defined alphabet. This paper presents a new technique to classify hand gestures which could be used in HCI systems. Next section shows some usual approaches for HCI systems based on hand gestures. In section three, we present some limitations of these approaches and we propose a new one. In section four we conclude this paper by analyzing the produced results and appointing to the future directions.

**Methodology:** The approaches described previously have some limitations. Hu invariant moments are sensitive to morphological deformations and it is inevitable produce hand gestures with variations in its morphology, moreover, hands of different people have different morphologies and it could affect directly the invariant moments performance. The k-curvature has two easily identifiable problems. First, the definition of constant k, a selected value that could work well in some cases, but it could fail in cases of scale variances. Second, the algorithm can indicate how many fingers there are in the produced gesture, but the method can produce false positives answers in the cases where different gestures has the similar number of curves e.g., gesture Victory and Gun in the template matching approach is invariant to rotation and translation, but it will misclassify if scale variations are introduced. Based on these observations, this article proposes the uses of different features to classify the hand gestures.

### Paper No.10

**Title**: A Survey on State-of-the-Art Methods for Indian Sign Language Recognition.

**Published**: JASC: Journal of Applied Science and Computations Volume VI, Issue II, February/2019 ISSN NO: 1076-5131

**Publishers**: Pradip Patel, Narendra Patel Research Scholar, Gujarat Technological University, Ahmedabad, Gujarat, India. \* Department of Computer Engineering, BVM Engineering College, Vallabh Vidyanagar, Gujarat, India.

**Abstract**: Sign language is the only way for the deaf people to communicate where, in place of voice, different gestures are used to represent meaning. Normal people do not understand of sign language. This results in lake of communication between normal people and deaf people. By developing computer based automatic system for Sign Language recognition, we can make this communication possible. Very little work has been done for recognition of Indian Sign Language due to lack of standardization and complexity. But recently, with the advancement of science and technology, many researchers are motivated to work in the field of Indian Sign Language recognition. It is emerging as current area of research in the field of human computer interaction. It will help deaf peoples to communicate and learn via use of image processing and computer vision techniques. Few researchers have recently worked in this area. This paper specifies a literature survey of the widely used techniques in sign language recognition system. Some of the renowned techniques for feature extraction and classification are discussed briefly. There are mainly two phases of sign language recognition, training and testing, each of which consists of Data acquisition, Preprocessing, Feature extraction and Classification.

**Introduction:** Sign language is the only way by which deaf people communicates. In sign language, in place of voice, different gestures are used to represent meaning. These gestures are combination of hand shapes, movement, position, palm orientation, arms or body, and facial expressions. There are many sign languages in existence all over the world - American Sign Language (ASL), British Sign Language (BSL), Indian Sign Language (ISL) etc. Figure 1 shows the sign language representation of ISL alphabets and numbers. As per the survey done by Indian government in year 2011, over 2.68 crore of people in India suffer from some form of the disability. Out of this 18.9% people have speech and 7.5% people have hearing disability [1]. It is difficult for them to communicate with the normal people because normal people do not understand sign language. If we develop a system that can recognize sign-language symbols then it can be used as a means of communication with deaf people. Using this system normal people can easily communicate with deaf people. This resulted in the requirement of development of automatic sign language recognition system that can automatically recognize the signs and convert them into corresponding text or sound. This will eliminate the help of sign language interpreters.

**Methodology:** During training phase, dataset is used by using which the classifier is trained. For this, we can create our own database or an existing database may be used. Any external or inbuilt web camera in the laptops can be used for the purpose of capturing images for training dataset. To train the system, these training images are presented to system along with class labels. Almost all the sign language recognition systems consider only hand gestures to classify signs and ignores facial expressions and body posture. Major steps of training phase include database creation, preprocessing, feature extraction and classifier training. These steps are briefly discussed in next sections of this paper. Once the system is fully trained, during testing phase an unknown gesture image is presented to system for classification. Testing also involves image acquisition, preprocessing, feature extraction, and sign recognition. A variety of techniques are used in various stages of recognition.

### Paper No.11

**Title**: Accessible System for Sign Language Computation.

**Published**: Wutan Huatan Jisuan Jishu Volume XVI, Issue XI, November/2020 ISSN:1001-1749.

**Publishers**: Patel Harshil1, Patel Jay2, Patel Smit3, Thakkar Mayur4. Scholar, Sigma Institute of Engineering, Vadodara, Gujarat, India Dr. Sheshang Degadwala.

**Abstract**: Sign Language is most accepted and meaningful way of communication for deaf and mute people of the society. Sign language uses gestures, head/body movements and facial expressions for communication. It is a powerful way of Non-Verbal communication among humans. Every country has its own developed Sign Language. The language which is used in India is called as “Indian Sign Language (ISL)”. Very less research work has been carried out in this area as ISL has got standardized recently. Currently many researchers have mainly focused on gesture recognition that has been recorded under static hand gesture. Only few works have been reported for recognition of dynamic hand gesture. Many methods have been developed to recognize alphabets and numbers of ISL. The major steps involved in designing such a system are gesture acquisition, tracking and segmentation, feature extraction and gesture recognition. This paper presents a survey on various hand gesture recognition approaches for ISL.

**Introduction:** Sign language provides a way for the deaf and mute people to communicate with the outside world. Instead of voice, sign language uses gestures to communicate. Sign language is an organized way of communication in which every word or alphabet is assigned to a particular gesture and meaning. Sign language is made up of a range of gestures produced by various facial expressions and movements of hands or head/body. In the last several years there has been an increased interest among the researchers in the field of sign language recognition to introduce a means of interaction from human –human to human – computer interaction. Various applications of Gesture Recognition System are: Human Computer Interface, Video gaming, Augmented reality, Home appliances, Robotics, Sign Language, etc. Hearing-impaired and speech impaired people depend on sign language interpreters for communication. But finding skilled and trained interpreters for their day-to-day interactions during life time is a very difficult job and also too expensive. All over the world, a variety of sign languages exist. The sign language depends on the traditions and spoken language of that place. Indian sign language (ISL) is used by hearing and speech impaired people in India. The gestures are primarily divided into two classes: Static gestures and Dynamic Gestures. Static gestures include only configurations and poses whereas dynamic gestures include postures, strokes, pre-strokes, phases and also emotions.

**Methodology:** The major steps involved in designing a sign language recognition system are: gesture acquisition, tracking and segmentation, feature extraction and gesture recognition. Primary step of gesture recognition system is to acquire gestural data. A webcam integrated in the laptop or an external webcam or a simple digital camera can be used to capture the images. Either an on-hand database can be used or it can be formed by the researchers themselves. To track the movement of hand, segmentation of hand is required. Segmentation separates ROI from background. Hand tracking is a technique used to know position of hand. Next step is feature extraction which is carried out to extract important features after successful completion of hand tracking and segmentation. The complete process of recognition can be separated into the following two stages - training and testing. Training is the initial stage where the classifier is trained using the training database. The key steps implicated in training stage are creation of database, pre-processing, extraction of features and training of the classifier. The key steps involved in the testing phase are gesture acquisition, preprocessing, feature extraction and classification of sign language recognition system.

### Paper No.12

**Title**: A Real Time Hand Gesture Recognition Using Feature Extraction.

**Published**: International Journal of Technical Research and Applications e-ISSN: 2320-8163, [www.ijtra.com](http://www.ijtra.com/) Volume 1, Issue 5 (Nov-Dec 2013), PP. 41-45.

**Publishers**: Shruti Wale, Swapnil Patil, Prof. Neha Rai Assistant Professor, PG Student Electronics & Telecommunication Engineering PHCET, Rasayani, India.

**Abstract**: Gesture recognition is finding a meaning attached to some motions or actions. The main objective is of this paper is hand gesture recognition from the real time video. Gesture recognition is a topic in computer science and language technology with the goal of interpreting human gestures by using some mathematical algorithms. Gestures are commonly originated from the face or hand. Current focuses in the field include emotion recognition from face and hand gesture recognition. Many approaches have been made using cameras and computer vision algorithms to interpret sign language. However, the recognition of posture and human behaviors is also related to gesture recognition techniques. The major problem faced in the gesture recognition is that a person will appear at different scales in videos. The movement of camera is another problem as the person holding it may shake it while shooting the video. In certain cases, the camera may mount on something which moves with the person performing the action. Another problem faced is occlusions, background clutter, human variation and action variation. Gesture recognition also various with person to person because every person has different skin texture and color. The nonparametric histogram based on RGB model is used for the skin detection. In this system the gesture recognition is mainly divided into two phases: training phase and recognition phase. Feature extraction is the main function of both the phases. The fuzzy logic is used in this system for gesture recognition.

**Introduction:** The Gesture acknowledgment interface goes about as a correspondence channel amongst people and machines. The human-machine collaboration is like human-human connection, in which, the important data are conveyed utilizing the human organs like hand motion, head development, face demeanor, voice correspondence and general body stance. The configuration of a motion acknowledgment framework depends on basic equipment backing, for example, web-cams or portable incorporated cameras, to be pertinent to current PCs, cell phones, Digital Cameras, and so forth. While outlining the frameworks, certain parameters, for example, that it ought to be work under complex or non-uniform foundation, i.e., distinctive light force and loud environment, and so on are taken into contemplations. This anticipate presents a constant hand signal acknowledgment framework by utilizing webcam. This procedure includes recognizing the hand area, investigating the hand-position varieties. At that point the got movement data is been utilized as a part of the acknowledgment period of the motion.

**Methodology:** The system uses web cam for capturing the hand gestures and an algorithm that processes the acquired images and then classifies the hand gesture correctly. The work mainly highlights on the feature extraction from the hand gestures and use that features in the recognition algorithms. Firstly, features are extracted in the training phase by using the suitable algorithms for different hand gestures. Once this feature extraction is completed in training phase the system is able to classify the given hand gesture based on the gained knowledge during training phase. The efficiency of the recognition phase totally depends on the training phase. The better the task performed in training phase, better will be the performance of recognition phase. So, all the functions in training phase are carried out properly.

### Paper No.13

**Title**: Punjabi Sign Recognition and Education System for Differently-able Persons.

**Published**: (Special Issue) International Journal of Advanced Research in Computer Science. Volume 8, No. 4, May 2017.

**Publishers**: Kanika Goyal Chitkara University, Patiala) Rohit Sachdeva MM Modi College, Patiala Punjab (India).

**Abstract**: Sign languages commonly develop in deaf communities, that can include interpreters and friends and families of deaf people as well as people who are deaf or hard of hearing themselves. Sign Language Recognition is one of the most growing fields of research today. There are Many new techniques that have been developed recently in these fields. Here in this paper, we will propose a system for conversion of Indian sign language to text using LabVIEW Software. LabVIEW Software designed to generate motion template images that can be used to rapidly determine where that motion occurred, how that motion occurred, and in which direction it occurred. There is also support for static gesture recognition in LabVIEW Software which can locate hand position and define orientation (right or left) in image and create hand mask image. In this we will use image processing in which captured image will be processed which are digital in nature by the digital computer. By this we will enhance the quality of a picture so that it looks better. Our aim is to design a human computer interface system that can recognize language of the deaf and dumb accurately.

**Introduction:** In the era of advanced technology, where computers or laptops and other processor based devices are an integral part of day to day life, efforts are required to be done for making the disables more independent in life. In spite of that, there are people who are less fortunate than us and are physically i, may it be deafness or being aphonic. Such people lag behind impaired, their non-handicapped peers in using these technologies. These people have some expectations from the researchers and mostly from a computer scientist that we, computer scientists can provide some machine/model which help them to communicate and express their feelings with others. Very few researchers’ have them in mind and provide their continuous works for such people. Communication is the most important part of life. Around 1% of the total population of the world is suffering from hearing impairment, and their life is not as simple and easy as it is for human without limitations. Finding an experienced and qualified interpreter every time is a difficult work and also unaffordable. Moreover, people who are not impaired, never try to learn sign language for interacting with the impaired people. This becomes a cause of isolation of the impaired people. But if the system can be programmed in such a way that it can translate sign language to text , the difference between the normal people and the impaired people can be reduced. Closely related with image processing are computer graphics and computer vision. In computer graphics, images are processed from physical models of objects and lighting, instead of being acquired (via imaging devices such as video cam, cameras) from natural scenes, as like in most animated movies. Computer vision, on the other hand, is considered as the high-level image processing out of which a machine/computer/software intends to decipher the physical contents of an image or a sequence of images (e.g., videos or 3D full-body magnetic resonance scans).

**Methodology:** Many Different approaches have been used by different researchers for recognition of different hand gestures which were implemented in different fields. few of the approaches were vision-based approaches, soft computing approaches like Artificial Neural Network, Genetic Algorithm, Fuzzy logic, data glove-based approaches, and others like Canonical Analysis, PCA etc. All approaches could be divided into three categories- Feature extraction approaches, Hand segmentation approaches and Gesture recognition approaches. Few of the approaches have been discussed in this paper. Many researchers used skin filtering technique for segmentation of hand. This technique separates non-skin-colored pixels from the skin-colored pixels, thus extracting the hand from the background image. Fang used Adaptive Boost algorithm which could not only detect single hand but also the overlapped hands. In external aid like color gloves, data gloves were used by the researchers for segmentation purpose. In Principal Component Analysis (PCA) was used for extracting features for recognition of various hand gestures the system will provide an interface that can easily communicate with deaf people by Sign Language Recognition. The system is not only can apply in family environment, but also can apply in public. For the Social use this system is very helpful for deaf and dumb people.

# **Related Work**

As discussed above, sign languages are a key of communication for disabled people. Mrs. Daleesha M Viswanathan [3] presents a technique which perceives the Indian Sign Language (ISL) and converts into a normal text. This process comprises of three phases, to be specific a preparation stage, a testing stage and a classification stage. Combinational parameters of Hu invariant second and basic shape descriptors are made to frame another component vector for sign recognition. Classification is done using a multi-class Support Vector Machine (MSVM) model. The viability of the proposed technique is approved on a data set having 720 pictures. Trial results exhibit that the proposed framework can effectively perceive hand signal with 96% accuracy rate. Subhash Chand Agrawal et al [1] presents the recognition of two-handed Indian sign language (ISL). Their method consists of three stages mainly, Image segmentation, Feature extraction, Recognition. They used Otsu algorithm which is a automatic image thresholding method for segmentation step. In the next step, they used SIFT key points extraction algorithm and HOG descriptors. To classify all gestures, they used multi-class Support Vector Machine (MSVM) model and trained it with all features. Using this method, their model scored 96% accuracy rate in recognizing gestures. Our method is also similar to them, but we use canny edge and speeded up robust features (SURF)key points algorithm instead of SIFT in segmentation phase. The model that has been proposed by Mrs.N.Rajeswari [5] perceives static pictures of the marked letters in order in the Indian Sign Language. Dissimilar to the letter sets in other sign languages via gestures like the American Sign Language and the Chinese Sign language, the ISL letter uses both the double hands and as well single hand. Hence, it makes ease the recognition of the gestures by categorizing them into single hand and doublehanded signature. For the two classes two different features have been used, namely HOG and SIFT which are separated for a set of training images and are consolidated into a single matrix. After which, the features of HOG and SIFT for the input test images are joined with the feature matrices of HOG and SIFT. The resultant classification of the test image is obtained by feeding k-Nearest neighbor classifier with computed correlation for the matrices.

# **Implementation**

The algorithm implemented in this recognizes Indian Sign Language gestures taken from static pictures. The system comprises of several steps which are Image collection, Image pre-processing (segmentation), Feature extraction, Classification as shown in the figure 2. Bag of visual words (BoW) model has been implemented to classify the images. Idea of BoW is adapted from Natural language processing (NLP). In image processing, BoW model concept can be called as “histogram-based representation of independent features”. So, an image can be viewed as a document in order to depict any gesture using the BoW model. Likewise, it is important to describe “words” in images too. To accomplish this, the following three steps are normally included: feature description, and generation of codebooks (visual words). Using these codebooks, histograms can be generated for all the images. Further, classification of images can be done using Support Vector Machine (SVM) model.

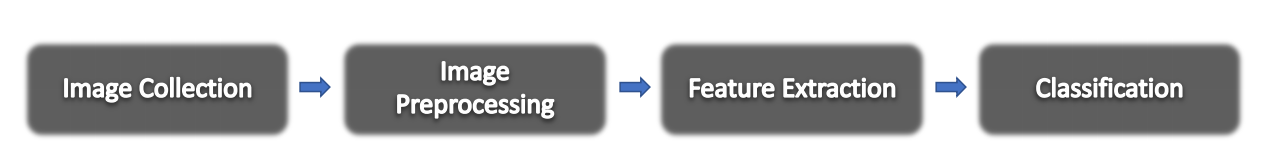
# Figure 1. ISL gestures for numeric and alphabets



# **Image collection**

As there is lack of research in this field, proper data set for ISL is not available currently. So, we gathered data set of ISL which contains 35 classes, each class with 1200images. Total 35 classes comprise of alphabets (A-Z) and numeric (1-9) with 42000images. All images for each class are captured by certain environment. Each frame of the video is saved as an image. To reduce the noise, gestures were captured with a black background. All gestures with classes can be seen in the figure 1.

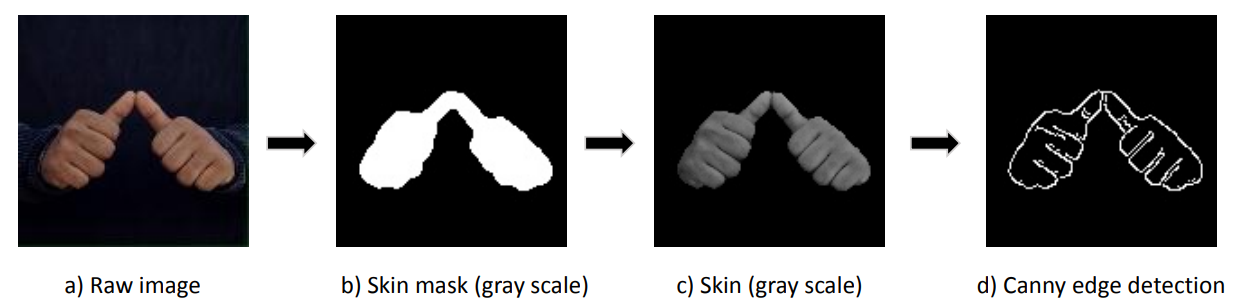
# Figure 2. Overview of the system

****

# **Image pre-processing**

In this phase, all images are pre-processed so that they can be used for feature extraction. This phase contains three steps which are image segmentation (skin masking), skin detection, edge detection. From the raw image (figure 3a) skin mask is generated by converting the image to HSV color space. The (H, S, V) range of all pixels from (0,40,30) to (43,255,254) are treated as skin pixels. Using the skin mask, skin can be segmented (see figure 3c). Finally, the Canny Edge technique [3] is used to detect and recognize the presence of sharp discontinuities in an image, thus detecting the edges of the image (see figure 3d).

# Figure 3. Image Pre-processing

****

# **Feature extraction**

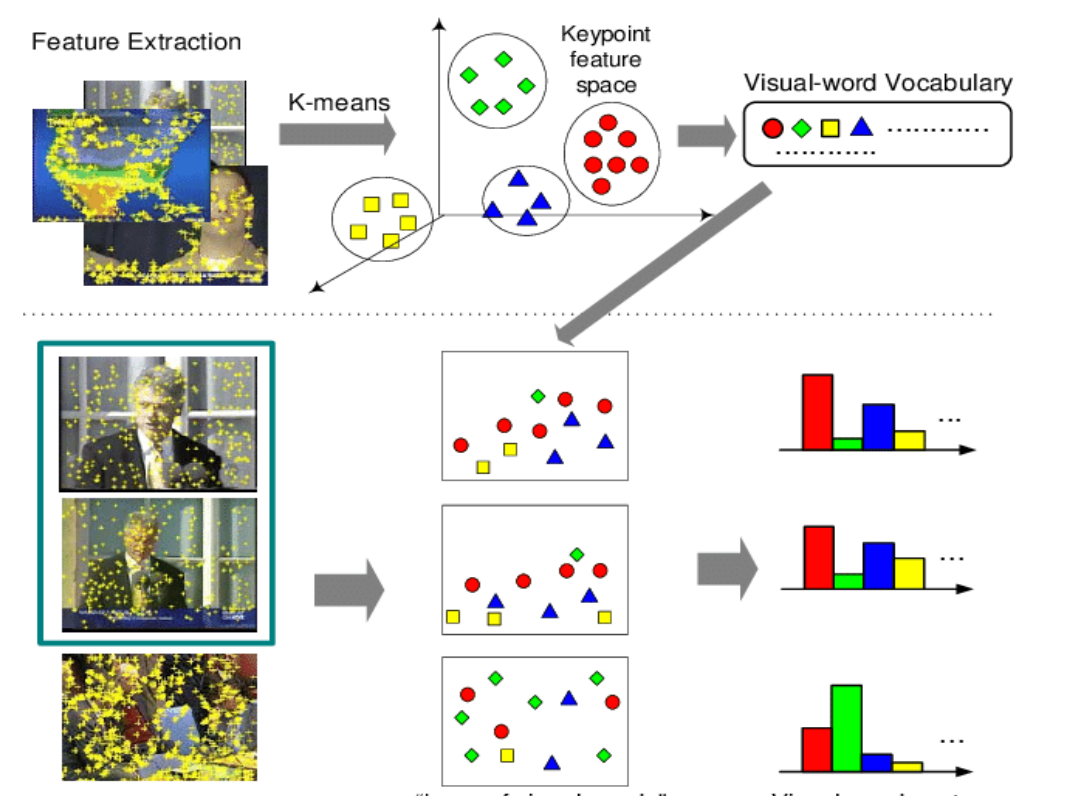
Next phase is feature extraction. This phase contains three steps which are image feature detection, Clustering, code book generation for Bow model. Initially, forfeiture detection we discussed to use Scale invariant feature transform (SIFT)algorithm to identify key features in the images. But we moved to speeded up robust features (SURF) as SURF is a novel feature extraction algorithm and robust against scaling, rotation, variation and occlusion in viewpoint. SURF features of an image can be seen in the figure 4. Next step is to cluster all these SURF features which are similar to make a visual vocabulary. It’s unlikely possible to human to find all similar feature descriptors. So, K-Means clustering algorithm can be useful in this step. K-means algorithm is a clustering un-supervised algorithm that is widely used. It aims at partitioning n features to k number of clusters and predicts a new feature belonging cluster by mean (centroid) of the clusters. As, we are using large data of SURF features of 42000 images, K-Means clustering takes much processing time and memory. So, we used mini batch K-Means which is similar to K-means but has advantages in terms of processing time and memory consumption. Once all SURF features are trained to mini batch K-means, it clusters all the similar features into a bag and number of clusters (visual words) are equal to k. We used k value 280 as there are 35 classes to classify. Using this model, visual words for each image can be predicted. Next step in this phase to calculate histograms using these predicted visual words. Histogram can be calculated by finding the frequency of occurrence of each visual word that belonging to image in total visual words. The whole phase can be seen in the figure5.

# **Classification**

Once all histograms are generated for total data set, the next step is classification. Before going to classification, it is important to divide the data for training and testing. So, we split the total data in 80:20 ratio such that each class has 960 images for training and 240 images for testing. Once the data is ready, the next step is to feed the training data to machine learning model. We used Support Vector Machine to train and predict the data with linear kernel. Similarly, other models can also be used for predictions such as K-Nearest Neighbors, Convolution Neural Network, Logistic Regression classifier etc. Additionally, real time recognition also developed so that a user can predict the gestures using video feed in real time.

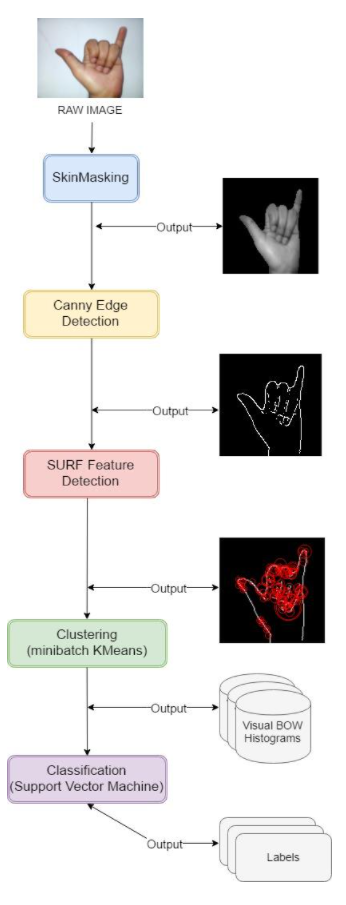
# Figure 4. SURF key points and feature descriptors for an image.

****

Figure 5. Bag of visual words representation. Above: Training for the K-means model using all features (SURF features) Below: Prediction of visual words and representing histograms for each image.

# **Workflow**

# Figure 6 . whole workflow of our project

****

# **Result**

Using Bow, integrated with robust SURF feature descriptors, the model scored 99% accuracy. The confusion matrix for the model can be seen in the figure 6. All labels have been predicted correctly by the SVM except label 2. Real time recognition prediction results can be seen in the figure 7. The precision, recall, f1 score was also 99.98%. But there could be slight biasing in the model prediction as data set has much similar images without variations for example in light and skin tone. So, using a large and variety of images in data set, this approach can be more robust for real world applications.

# **Figure 6.** Confusion matrix using SVM model

****

# **Figure 7.** Real time recognition of gestures results

a) Alphabet A b) Alphabet H c) Alphabet X d) Alphabet I

****

# **Conclusion**

ISL is a key for communication for deaf and dumb people in India. This report gives a detailed implementation for Indian sign language recognition using Bag of words model. In section 3, step wise implementation has been discussed which are image collection, image pre-processing, feature extraction (using K-means clustering, visual words collection) and Classification. Finally, results were presented of Bow model in section 4. Recognition for not only static images, but also real time recognition of gestures also developed. This project can also be extended for simple expressions and words in ISL including alphabets and numeric.

# **Reference**

[1]. S. C. Agrawal, A. S. Jalal, and C. Bhatnagar. Recognition of indian sign language using feature fusion. In 2012 4th International Conference on Intelligent Human Computer Interaction (IHCI), pages 1–5, 2012.

[2]. Herbert Bay, TinneTuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In AleˇsLeonardis, Horst Bischof, and Axel Pinz, editors, Computer Vision – ECCV 2006, pages 404–417, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[3]. J. Canny. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6):679–698, 1986.

[4]. K. Dixit and A. S. Jalal. Automatic indian sign language recognition system. In 2013 3rd IEEE International Advance Computing Conference (IACC), pages 883–887, 2013.

[5]. B. Gupta, P. Shukla, and A. Mittal. K-nearest correlated neighbor classification for Indian sign language gesture recognition using feature fusion. In 2016 International Conference on Computer Communication and Informatics (ICCCI), pages 1–5, 2016.

[6]. Yu-Gang Jiang, Jun Yang, Chong-Wah Ngo, and Alexander Hauptmann. Representations of keypoint-based semantic concept detection: A comprehensive study. Multimedia, IEEE Transactions on, 12:42 – 53, 02 2010.

[7]. David G. Lowe. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision, 60(2):91–110, November 2004.

[8]. Dutta, Kusumika& K, Satheesh& S, Anil & Sunny, Breeze. (2015). Double handed Indian Sign Language to speech and text. 374-377. 10.1109/ICIIP.2015.7414799.