Sign Language Translation Using 3D Sensor

[1]Sakshi Borgude, [2]Aditya Kanawade, [3]Sandesh Shardul, [4]Sneha Sonawane

Department of Computer Engineering
Savitribai Phule University, Pune

***Abstract*— Deaf people around the world don't have any other means of communication rather than sign language among themselves as well as normal peoples. They mainly rely on hand gestures and signs in shapes made with hands. Sign language is the way through which deaf and dumb people can communicate with each other. It has been observed that, impaired people find it very difficult to interact with the society. Normal individuals can't able to understand their sign language. To bridge this gap, the proposed system acts as the mediator between impaired and normal people. This System uses Kinect camera to capture the signs. The Kinect camera captures 3D dynamic gesture. Thus, the method is proposed for feature extraction of dynamic gesture of Indian Sign Language (ISL). As American Sign Language (ASL) is popularly used in the field of research and development, ISL on the other hand has been standardized recently and hence its ISL recognition is less explored. The propose method extracts feature from the sign and**
Converts it to the intended textual form. The method then integrates local as well as global information of the signs. This integrated feature improves the performance of the system, the system serves as an aid to disabled people. Its application includes hospitals, government sectors and some multinational companies. Concepts like machine learning and feature extraction will be used.

***Keywords*— Indian Sign language, American Sign Language, Kinect camera, Machine learning, Feature Extraction.**

1.1 INTRODUCTION

Humans have been endowed by nature with the voice capability that allows them to interact and communicate with each other. The spoken language becomes one of the main attributes of humanity. Unfortunately, not everybody possesses this capability due to the lack of sense, i.e., hearing or speaking. Loss of hearing or speech can cause people to become isolated
And lonely, having a tremendous effect on both their social and working life. To reduce this gap between the normal people and the impaired ones Sign language is introduced. Sign Language is the well-structured code gesture language, every gesture has meaning assigned to it. This is the most important communication way between impaired community and
Normal person. It is observed that ordinary people do not understand the sign language. So, to overcome this problem and make the communication possible this system is introduced. When the impaired person wants to communicate with ordinary person then at that time, he performs action in front of Kinect camera. The camera will recognize the actions being performed by the user, and gives the skeleton of human body. Kinect camera gives accuracy while performing actions, draws the skeleton of human body when user stand in front of Kinect sensor. These actions are then compared with actions stored in dictionary. A dictionary is maintained where all the actions and related text are stored. If the match is found then
Appropriate text is displayed on the screen. On the other side if the ordinary person wants to communicate then he will give audio input. This input is captured through Kinect sensor. Then the output will be displayed in the form of animated images and text. When the impaired user performs some actions which are not stored in the dictionary, then at that time we can dynamically add that action into the dictionary. So, there is no need to change in source code every time. Also, dynamic actions get stored along with its corresponding values in the dictionary. Thus, in the following way an interactive communication will take place. This system works with two-way communication. With this system impaired person
Can communicate with ordinary person and vice versa. It means it recognizes the actions performed by impaired people and converts into language understood by ordinary persons and vice versa. This system uses special hardware and maps that with software to produce the required result.

1.2 NEED

Communication is most important part of daily life. Through communication one can interact with the society. Communication with family, friends, at work and public place is vital. But if someone lack with this ability then it becomes serious issue. Deaf and Dumb people uses sign language for communication. It is difficult for normal user to understand the
Sign actions, so communication gap is increased. So in this situation a mediator is required to translate the languages. It becomes difficult to find a well experienced and educated translator for sign language every time and everywhere when needed which results in lack of communication between the ordinary and impaired people. So it becomes necessary to
Have such system which will help for impaired person to convey their messages in society. So there is a need of a System which acts as a mediator between impaired and normal people.

- The need of such system is in government sector where either the customer or the client May be impaired and wants to communicate with each other.
- Also in Hospitals where the impaired patient wants to communicate with the doctor regarding the disease.
- In Multinational companies this system is most useful where the impaired employee will get the platform to showcase his talent.
- For long distance education courses and to provide the Education material, this system can be useful.

Thus, the need of this Project is to make the communication easy for impaired people and Taking this project to the level of serving to the society.

1.3 BASIC CONCEPT

Communication with deaf and dumb people is based on the sign language. Sign Language is the standard language designed particularly for impaired people. Basically this is based on movement of hands, which is taught to them in their school. Two or more impaired persons can communicate with each other through this language. But it has been observed that they find it very difficult to interact with the society, because not everyone in society is well familiar with the sign language also normal people find it difficult to understand their sign language. So to bridge the gap between normal people and impaired people this system is proposed. In the early days this gap is bridge by the mediator who converts the sign language into natural language and vice versa. It is not necessary that every time the mediator will be available 24X7. So to solve this problem and make ease of the new technology the proposed system is developed. The proposed System act as the mediator between impaired people and normal people.

In Mode 1 the system is initiated by the impaired person. When impaired person want to communicate with normal person then he gives the input as sign to System through Kinect camera. These actions are captured by the Kinect camera. System performs Processing on captured sign and produces text and audio as output. That is understood by normal people as shown in the Figure 1.1

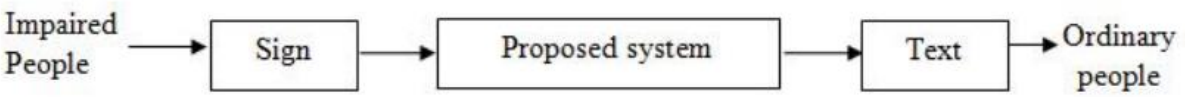
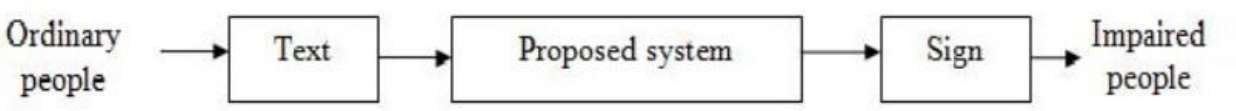


Figure 1.1: Sign to text conversion

In mode 2 the system is initiated by the normal person. When normal person wants to communicate with impaired person then he gives input as audio. This input is accepted by the microphone which is fitted on the base of Kinect camera. System performs processing on audio and converts it into text. Based on this text pattern matching is done. Once the match is found then the relevant images are displayed. This is shown in Figure 1.2. In this way the system converts sign to text and vice versa.



F

igure 1.2: Text to Sign conversion

2.1 CLASS DIAGRAM

Class diagram consist of number of classes, each has set of attributes and functions. Interactive System for Deaf and Dumb People, is the System helps to make communication between impaired and Ordinary User. In this System user can be impaired or ordinary user. Impaired user will perform actions in front of Kinect sensor. System will recognize the actions performed by the User and gives output in the form of audio. In other case, when ordinary user will give input in the form of voice Kinect will capture voice. Perform processing and give output in the form of text and animated image. Kinect camera is used to capture voice as well as actions of user, also recognizes the skeleton of human body as shown in Figure 2.1.

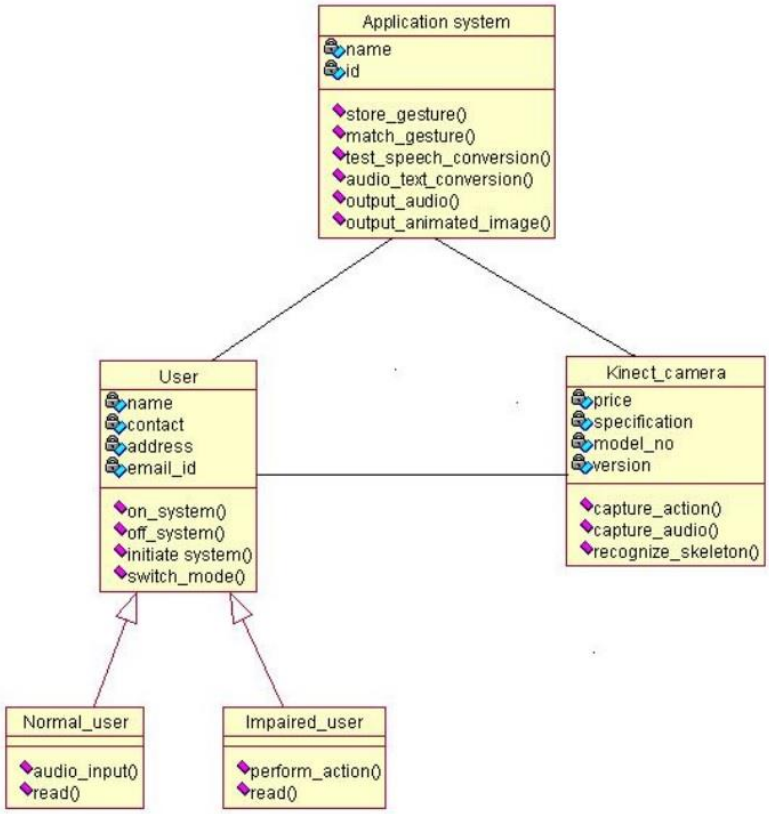


Figure 2.1: Class diagram of System

2.2 PACKAGE DIAGRAM

Interactive System for Deaf and Dumb User helps for the Interaction among Deaf dumb and normal people. This System acts as a mediator between impaired and ordinary user. Impaired user will interact with the Ordinary user by using this System as shown in Figure 4.6. Impaired and Ordinary user can give input to system through Kinect camera. Kinect camera captures the actions of user and draw the skeleton of human body, gives that skeleton as input to this System. When impaired user will give input in the form of actions, system will produce output in the form of audio. In other case, when Ordinary user will give voice as input, system will convert voice/audio into its corresponding Text. Then pattern matching is done on the converted Text. System gives output in the form of animated image as well as Text.

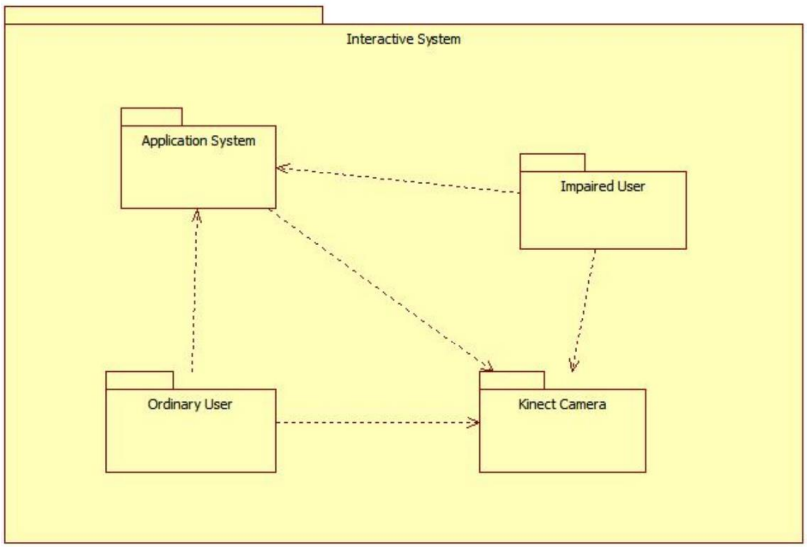


Figure 2.2 Package diagram of System

2.3 USE CASE DIAGRAM

Interactive System for Deaf and Dumb System used to make interaction among impaired and ordinary user as shown in Figure 2.3. There are two modes, in mode 1 impaired user will stand in front of Kinect sensor and perform actions. Sensor will capture the human body and draw skeleton of user. Then, system will give output in the form of audio. In mode 2, ordinary user will give reply in the form of voice. Kinect will capture the voice and give output in the form of text as well as animated image. System will dynamically record actions of impaired user into the dictionary along with its meaning.

A) Mode 1:

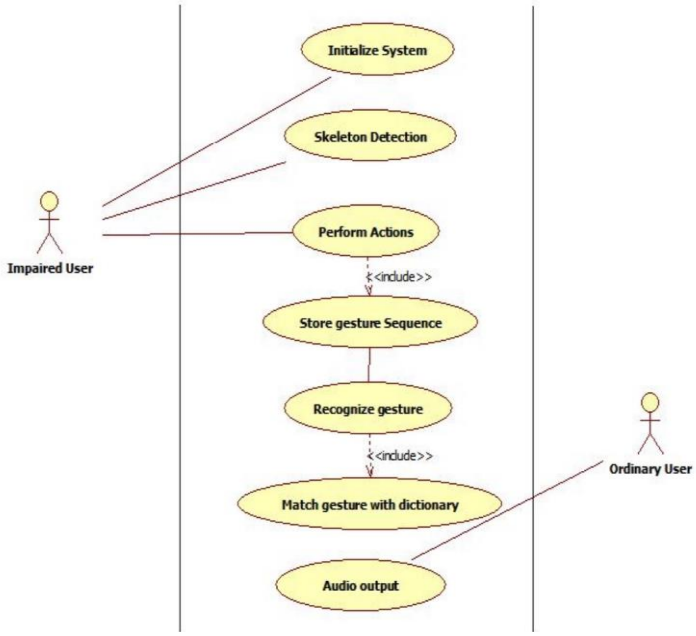


Figure 2.3: Use-case diagram for Mode 1

B) Mode 2:

In this mode, Different use cases are associated with the users of a System as shown in Figure 2.4. In Mode 2, Ordinary User initiates the System. Input for this mode will be in the form of Audio/ voice. Microphone of Kinect camera helps to capture the voice of Ordinary user. When Ordinary user gives voice/ Audio input, system converts Audio into its corresponding Text. System match that Text with Sign images, and displays Text as well as animated image on screen.

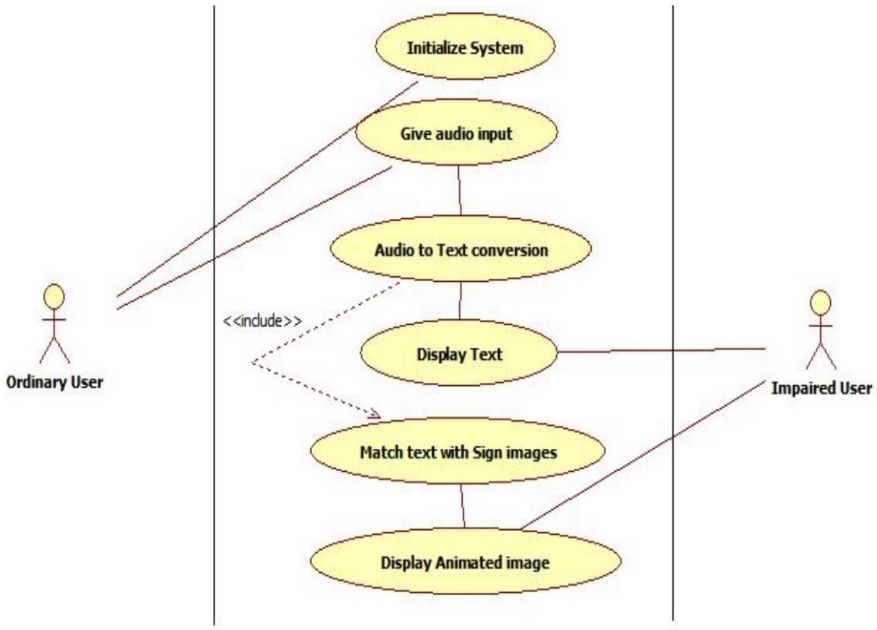


Figure 2.4: Use-case diagram for Mode 2

2.4 ACTIVITY DIAGRAM

1. Mode 1: This System performs the different activities in Mode 1, impaired user will perform the sign actions as input. Kinect sensor will capture the actions, system stores the gesture sequence and identify its meaning in the dictionary. System will give output in the form of Audio as shown in Figure 2.5.

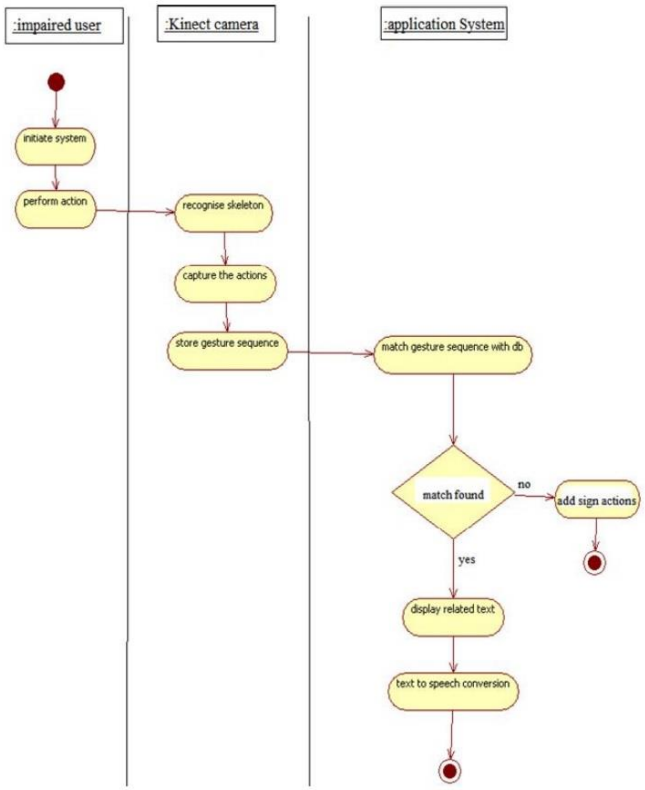


Figure 2.5: Activity diagram of Mode 1

2. Mode 2: When ordinary user will give input in the form of Voice/ Audio, Kinect captures the Voice. System will perform processing and convert Audio to Text. Then pattern matching is performed on converted Text, and gives output in the form of Text as well as animated image.

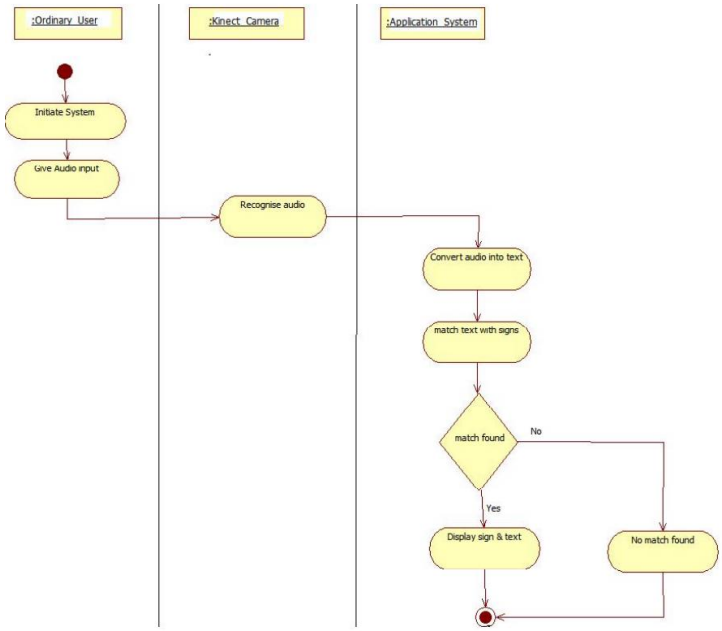


Figure 2.6: Activity diagram for Mode 2

2.5 SEQUENCE DIAGRAM

This System helps for the interaction among impaired and ordinary user. User will perform the sign actions, Kinect captures the action and gives as input to the system. System performs the conversion of the sign into Audio form. When ordinary user will give input as Voice, system will convert audio into text and search animated sign image for that text. System produces output in the form of Text and Animated image shown in Figure 2.7.

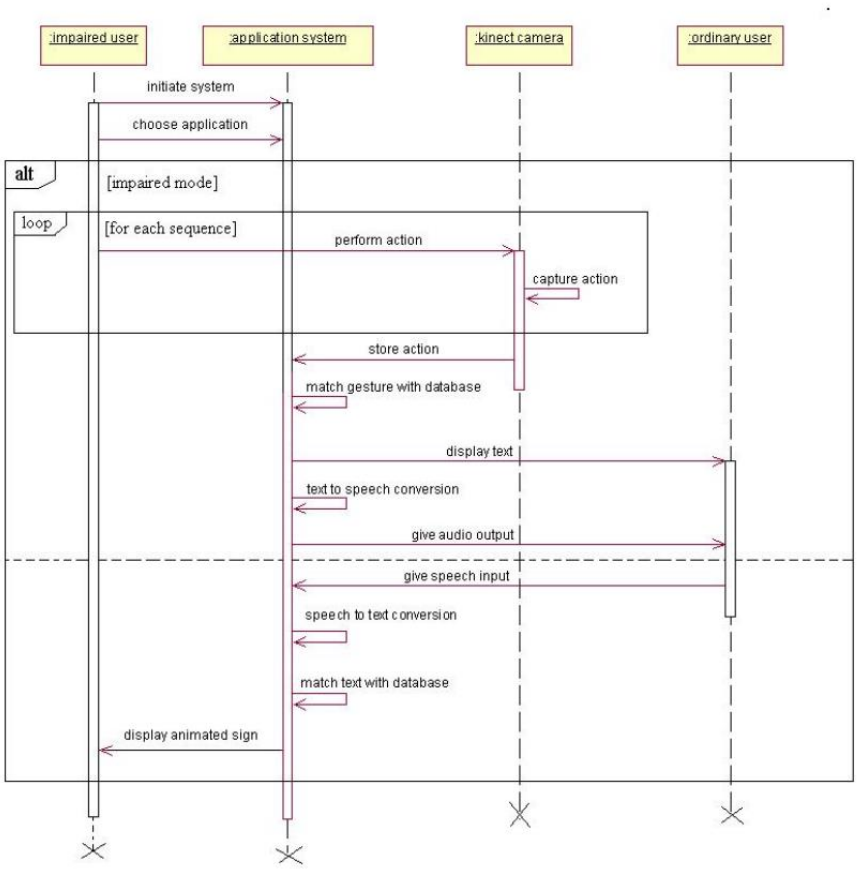


Figure 2.7: Sequence diagram of System

2.6 DATA DICTIONARY

When the Impaired User will perform the actions in front of Kinect camera, actions are captured and accordingly values are generated through distance calculation module. This numeric values get stored in the notepad file. When user add actions dynamically, actions are recorded and its values get stored in the dictionary. Whenever user perform any action, numeric values get appended into the dictionary. In this way dictionary is maintained, which helps to identify the actions.

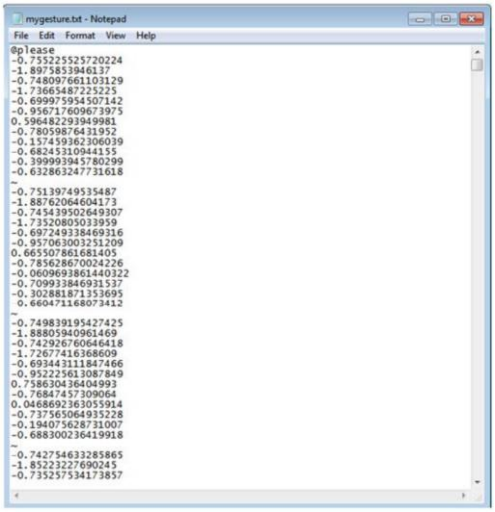


Figure 2.8: Data Dictionary

3. HARDWARE REQUIREMENTS

1. Kinect Camera:

(a) Kinect sensor captures the human body and draw the skeleton of it.

(b) Provide accuracy while performing Actions.

(c) Kinect XBOX 360 captures the Global points of human body.

2. Operating System

(a) Windows 7 and Windows 8 are supported.

4. SOFTWARE REQUIREMENTS

1. Kinect Studio v 1.8.0

This studio helps to make connectivity of Kinect camera with C sharp .net, this Studio also provides Developer Toolkit Browser v 1.8.0. In which sample template is provided, in this built-in packages and classes are given. This template can be used for Kinect learning.

2. C sharp .net

To design front end of this System WPF (Windows Presentation Foundation) is used, and Implementation is done in C sharp .net.

5. PERFORMANCE EVALUATION AND TESTING

This chapter focuses on different factors of performance evaluation and testing techniques, test cases that are successfully executed.

5.1 PERFORMANCE EVALUATION FACTORS

Performance evaluation factors is as follows:

1. Software Quality

- Software quality refers to two related but distinct notions that exist wherever quality is defined in a business context.
- Software functional quality reflects how well it complies with or conforms to a given design, based on functional requirements or specifications. That attribute can also be described as the fitness for purpose of a piece of software or how it compares to competitors in the market place as a worthwhile product
- Software structural quality refers to how it meets non-functional requirements that support the delivery of the functional requirements, such as robustness or maintainability, the degree to which the software was produced correctly.

2. Performance

Performance is characterized by the amount of useful work accomplished by a computer system compared to the time and resources used. Depending on the context, good computer performance may involve one or more of the following:

- Short response time for a given piece of work
- High throughput (rate of processing work).
- Low utilization of computing resources.

3. Scalability

- Scalability is the ability of a system, network, or process to handle a growing amount of work in a capable manner or its ability to be enlarged to accommodate that growth.
- Scalability, as a property of systems, is generally difficult to define and any particular case it is to define the specific requirements for scalability on those dimensions that are important. It is highly significant issue in electronics systems, databases, routers, and networking.
- A system, whose performance improves after adding hardware, proportionally to the capacity added, is said to be scalable system.

5.2 TESTING TECHNIQUES

Software testing is an investigation to provide stake holders with information about the quality of the product or service under test. Software testing can also provide an objective, independent view of software to allow the business to appreciate and understand the risk of software implementation. Test techniques include the process of executing a program with intent of finding bugs. In this chapter it is intended to cover the objective of testing, making sure that the developed system is bug free and free from functional errors. The study of test plan, the formal review and the test cases is conducted.

5.2.1 FORMAL TECHNICAL REVIEWS

The formal technical review is conducted by all the software engineers. The primary objective is to find errors during the process so that they do not become defects after release of software as they uncover errors in function, logic design or implementation.

5.2.2 TEST PLAN

The plan act as the anchor for execution, tracking and reporting of the testing projects. It describes the overall strategy that the project will follow for testing the final application. It is document describing the scope, approach, resources and schedule of intended testing activities. Test plan consist of following components.

1. Introduction

We applied unit testing and integration testing on the front end manually. The test cases are created and tabulated in the following section

2. Test items

(a) GUI

(b) Kinect

(c) Operation results

3. Features to be tested

(a) Skeleton recognition through Kinect.

(b) Record actions dynamically.

(c) Play audio as per the actions performed.

(d) Display animated image and text corresponding to the audio input.

4. Test Team

Each member of our project group performed the allocated part of testing of modules they have implemented.

5. Test Strategy
We performed following Testing:
(a) Unit Testing
(b) Integration Testing
(c) System Testing
(d) Black-box Testing, White-box Testing

5.3 TEST CASES AND RESULTS

A test case in software is a set of condition or variable under which a tester will determine whether an application or software system is working correctly or not. The mechanisms to determine whether a software program or system has passed or failed such a test known as test oracle. Test cases are often referred as test script particularly when written.

Test Id	Test Case	Expected Result	Actual Result	Pass / Fail
1	Display skeleton on canvas	Skeleton should be displayed on the canvas and joints should be displayed properly	Through joint points the skeleton is drawn on the canvas.	Pass
2	Play audio as per the actions performed	System should give correct audio output as per the actions performed	Audio output is given as per the actions performed.	Pass
3	Display animated images and text corresponding to audio input	Relevant animated image and text should be displayed.	Animated images and text is displayed.	Pass
4	Record dynamically actions	Dynamically new actions should get added into the dictionary.	Actions are recorded and added dynamically	Pass
5	Handle normalization	Same action performed by different users in different manner should be handled properly	Actions are identified properly	Pass

6. CONCLUSION AND FUTURE SCOPE

The proposed Interactive System can handle different types of words. Also it is suitable for dynamic signs. This system helps for easier interaction and communication with impaired people. It acts as mediator between impaired User and Ordinary User. They can easily convey the messages to each other by this System. While communicating, User can also add as many signs into the dictionary along with its corresponding meaning. The experimental results show that the system is working system for native Indian sign language recognition. The system is designed to support recognition of words in ISL only. This proposed system can be enhanced to recognize for continuous sentences also.
Examples: Stories, News.

REFERENCES:

1. Geetha M, Manjusha Cy, Unnikrishnan PZ and Harikrishnan Rz, "A Vision Dynamic Gesture Recognition of Indian Sign Language on Kinect based Depth Images, Dept. of Computer Science and Engg. Amrita School of Engineering Amritapuri, Kollam, Kerala, India, vol.58, No.20, pp-2059-2062, June 2012.
2. D.Martinez, "Sign Language Translator using Microsoft Kinect XBOX 360TM, V IBOT 5, Department of Electrical Engineering and Computer Science, Computer Vision Lab, University of Tennessee, vol.23, No.20, pp-2059-2062, June 2011.
3. Geetha M and Manjusha U C, "A vision based Recognition of Indian Sign language Alphabets and Numerals using Bspline approximation," INTERNATIONAL JOURNAL ON COMPUTER SCIENCE AND ENGINEERING(IJCSE), VOL. 4, NO. 3, MARCH 2012.
4. Matthew Tang, "Recognising hand gestures with Microsoft Kinect," STANFORD UNIVERSITY, vol.38, No.20, pp-2059-2062, June 2013.
5. Xiaodong Yang and YingLi Tian, Eigen Joints Based Action Recognition Using Nave Bayes Nearest Neighbour, THE CITY COLLEGE OF NEW YORK, NEW YORK.