ISSN (Print): 0974-6846
ISSN (Online): 0974-5645
Indian Journal of Science and Technology, *Vol 11(29), DOI: 10.17485/ijst/2018/v11i29/128525, August 2018*

Investigating Hand Gestures for Interactivity in Legacy Notice Board System

Rabia Naz Khan*, Mohsin Ali Memon, Hira Abid Khan and Salahuddin Saddar

Department of Software Engineering, Mehran University of Engineering and Technology, Jamshoro, Pakistan; Rabia.naz33@gmail.com

Abstract

Objectives: This research has been carried out to add interactivity and information transfer in digital format by adding technology in legacy paper based notice board system. **Methods/Statistical Analysis**: To achieve the above-mentioned objective, we have used Hand gesture recognition Technology. A Microsoft Kinect sensor is placed in front of the notice board to detect hand gestures which serves as the medium of interactivity. Through specified American Sign Language Number gestures user can make selections and interact with the system. For the detection of gestures Visual gesture Builder has been used which implements AdaBoost Trigger Algorithm. This framework uses Data Driven Machine Learning Algorithm to detect gestures. For the analysis, training and testing of the framework we have collected Gestures data for each predefined American Sign Language Number gestures from 0−9, from both Left and right hands, from 49 people. The machine learning algorithm was trained by 80% of the gesture data and was tested by rest 20% gestures. The approach got varying Confidence value (accuracy values) for each gesture depending on varying hand space, hand size, person's height, clarity in gesture performance. The framework also tested based on both male and female candidates, the result for gender-based analysis is also formulated in graph. The confidence values vary from gesture to gesture for both male and female. **Findings:** The research come out with the results that this technique can be used to optimize and make static paper based notice boards system interactive. With this technique user is able to transfer the information or notice posted on the notice board to their digital platform by making selection and commanding using their hand gestures. This enabled the user to negotiate without changing the whole business model. **Application/Improvements:** This framework can be implemented on places where notice boards are used to deliver information to the users for example. Institutes, hospitals, stations etc. Using this system the user can easily interact with the notice boards and transfer information to their digital means.

Keywords: Hand Gesture Recognition, Static Hand gestures, Kinect, Notice boards, Information Displays, Human Computer Interaction, VGB, Machine learning.

1. Introduction

In this technological era, information is the most functional and important aspect of businesses, commercial and individual use. Information display and transfer is an expensive process requires infrastructure and technological means. Technically, Information is displayed via two mediums: Digital and Printed means.

Here our focus is on the legacy way of displaying information that is Paper/Poster based notice boards. This medium has been used since long. The problems in this type of system are: People try different ways to select and carry the information. For later use people either take pictures or write the information down on paper which is not an intuitive way of transferring the information and leads to erroneous and incomplete information.

1.1. Draw backs of Digital Signage System

The old-styled notice boards have now being replaced by digital screens at most of the places but not everywhere since digital screen is an expensive option and organizations like hospitals, universities, schools and public places where notices are posted for the audience might not

**Author for correspondence*

afford this option. Secondly, the digital system requires controller, media players (PC based or embedded), group of digital screens and Network[1]. Thirdly, the use of digital displays systems donates a plenty of amount of power consumption[2] from about 1000 W for a 14 inch × 48 inch display which will multiply by the number of screens implemented[3].

There are institutions that still work on printed ways of communication and notification instead of digital ways, so the use of digital screens will affect the business processes. There seems to be the need for the improvising such manual systems rather than replacing them completely with digital systems.

We've proposed a framework involving the use of Hand Gesture Recognition technology to solve the problems identified and targeted in this research. Hand Gesture recognition is mostly used to make it convenient to interact with the digital devices. The peripheral devices like mouse, Keyboard have shown their limitations. Whereas, Hand gesture recognition provides effective and natural way of communication and interaction. It also fulfils the need of silent communication[4].

To implement this natural communication, we have added Ms. Kinect sensor to detect the gestures performed by the user. Ms. Kinect sensor has a low resolution of 640x480, and the hand occupies the little portion of the body, it is still very difficult to identify fingers motion and hands gesture[5]. This opened up a lot of challenges.

We believe that this paper is the first to address the issue of automating the Static Notice board System. Our discourse is backed by an extensive literature survey, analysis and experimental work.

2. Literature Survey

Hand Gesture Recognition technology has been used in a variety of areas to make it interactive, including Digital signage[6]. Sign Language recognition[7] has been an active research area to use it as a means of communication.

Researchers have been working on the techniques to detect American Sign Language[8]. But the uniqueness of this research is that we are using American Sign Language (ASL) as a tool to manipulate a manual Paper based notice board system.

Microsoft Kinect has been used by various researchers to detect ASL because of various benefits like low cost, good performance in low light, and intuitiveness[9].

In another research Kinect Depth mapping capabilities are used for the recognition and verification of sign Language for deaf children to provide them a better way to play educational games[10]. In a research, the presentation system is made interactive by using Hand Gesture recognition technology. The presenter's hand is localized and the presentation works based on the commands given by Hand Gestures[11].

In another research, an algorithm is proposed for offline gesture detection and recognition in lecture videos. They defined gestures for deliberate interaction between a User and the objects[12].

The proposed research introduces a unique idea of automating the static notice boards by adding a depth sensor and gesture recognition algorithm and utilizing the ASL gesture as a tool of communication, instead of replacing them with digital screens. The benefits of using ASL is that it makes the user gestures independent of the background, other than that, it makes the notice board flexible to be organized in any possible way as well as portable.

3. Aim and Objectives

This research aims to develop a Hand Gesture Based solution in order to overcome the short comings of the Manual Paper based legacy information displays. The objectives of this research are,

1. To determine whether Hand Gestures can be helpful to make legacy System interactive.
2. To examine whether American Sign Language gestures are suitable for this purpose.
3. To measure the accuracy of each Gesture by varying hand shapes left and right hand.

4. Proposed Framework

In our framework the idea is to use the Natural User Interface to offer interactivity and medium to provide input. First, we needed to make the computer understand about the user's choice. Either we had to use any peripheral device to make selection or any natural user interface. Most of the advanced innovation driven applications use gestures as input[13]. Gestures are the most natural way of communication.

For gesture recognition we have used a data driven Microsoft Kinect sensor and Machine Learning Algorithm: AdaBoost Trigger comes bundled up with Microsoft Kinect Sensor. The launch of Microsoft Kinect with its capability of depth sensing[14], skeletal tracking, high performance, high resolution and cheap cost have open up the platform gesture detection easy to conquer[8].

We have also defined Gestures for this purpose. In this system Gestures representing numbers from American Sign Language has been used. The user will interact with the device using American Sign Language. The American Sign Language contains static gestures for numbers 0 to 9 (see Figure 1). Every notice on the notice board will be given a number. The user needs to perform the ASL gesture for any number, depending upon which notice the user wants to select. The gestures representing digits 0 to 9 in American Sign Language are static gestures.

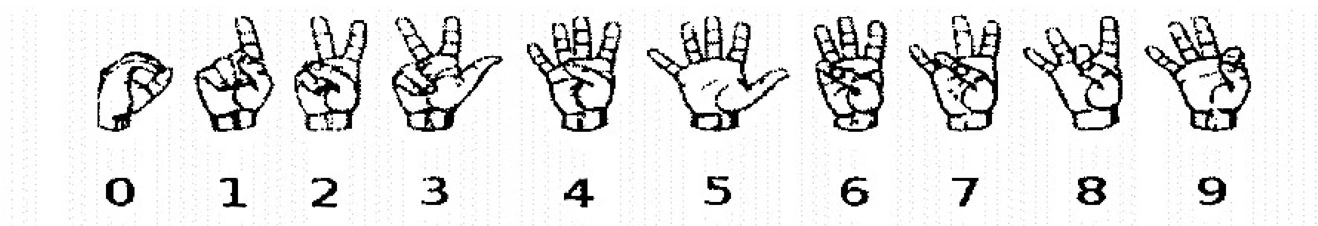

Figure 1. ASL Number gestures for the system.

The system includes following components:

1. Notice Board.
2. A depth Camera programmable for recognizing hand gestures, in this case Microsoft Kinect v1 is used.
3. A computer running the application for Hand Gesture Recognition and contains database.
4. Speakers.
5. System Feedback/Result (Optional)

The organization of the complete system can be seen in Figure 2 which is used for this research. The frame work is divided in to two parts: The User End and The Back End. The Notice board and Depth sensor is considered the Front end of the system, while a computer on which Hand Gesture Recognition algorithm and the application runs is considered the back end of the system. The user arrives at the notice board to retrieve any information, and then performs the particular ASL number gesture in front of the sensor. Subsequently, the gesture algorithm that is continuously polling the sensor receives the gesture event, identifies the gesture and sends the notice information via Email to the user and the user will be informed by the voice command. If the user's gesture doesn't match with any predefined gesture then the user will be informed with voice command to perform the gesture again (See Figure 3).

Figure 2. The System setup of the research.

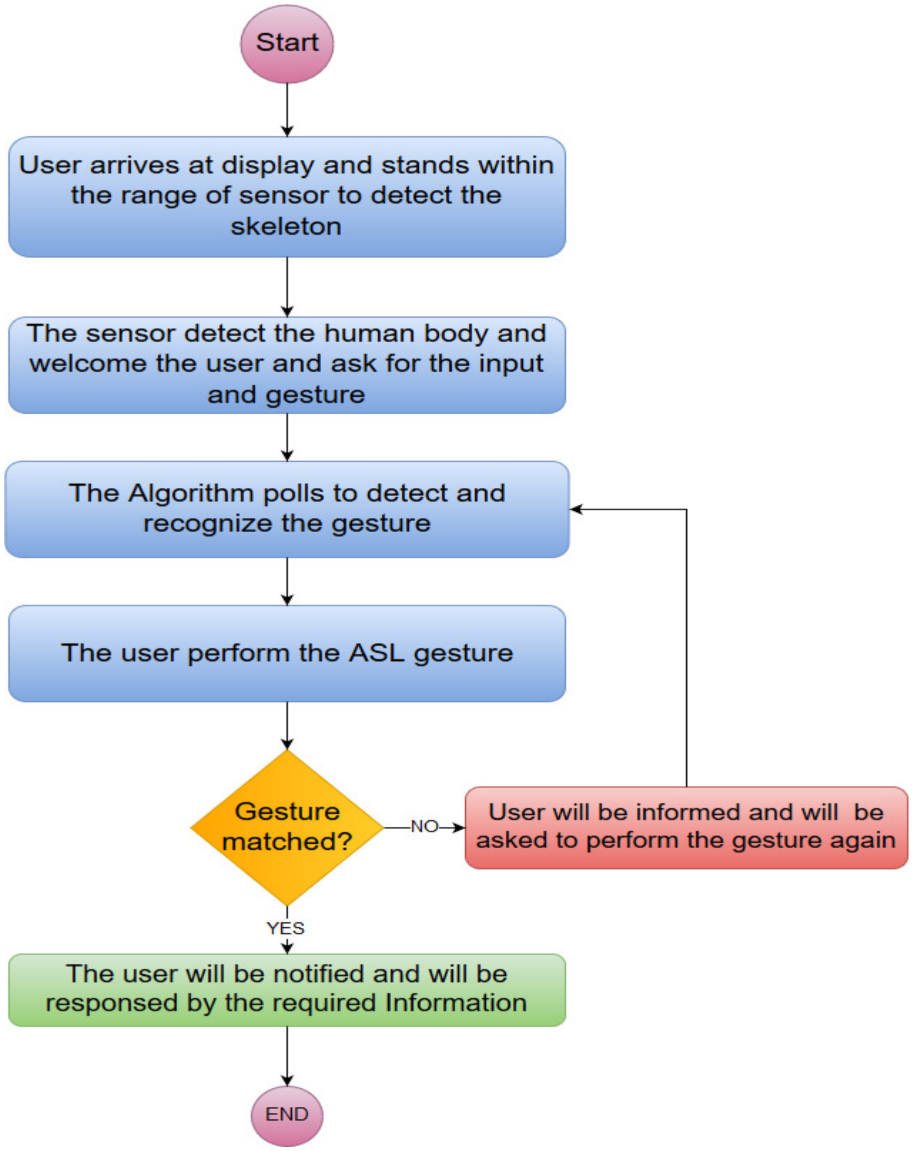

Figure 3. Workflow of the system.

4.1. Assumptions

The notice board should have maximum 10 notices. Notices can be of any size and color. Notices can be placed anywhere but they must be labelled by numbers between 0–9. Users can be of any height and wear any dressing. The User must be standing at any place within the range of Kinect sensor. Recommended distance between sensor and the user is 2 ft. The user must be facing the sensor. Height of sensor from the ground is 2 ft. Distance between the sensor and Notice board is 6 ft.

4.2. User End

The front end of the system is where user interacts with the display by performing ASL gestures in order retrieve the information. This portion of structure comprises of the Notice board, Microsoft Kinect sensor and the user.

4.2.1. Notice Board

The notice board considered here is the usual notice board. The notice board does not require any special organization, rather in this research; our motive was to

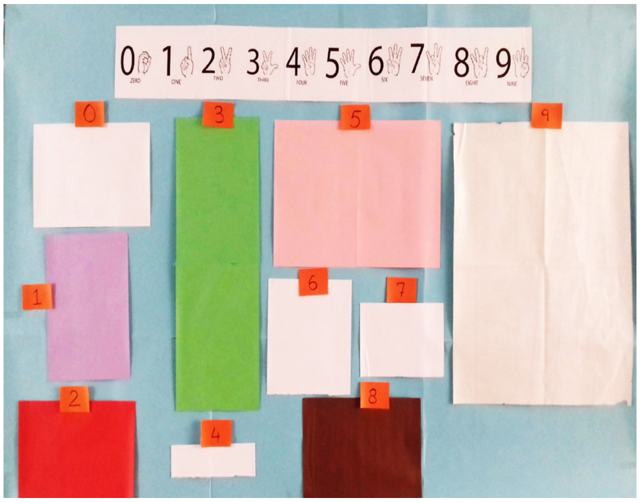

Figure 4. The Notice board organization used in this research.

keep the notice board simple with no specific changes. Only two things are added, first, all the notices are tagged by a number and second, the hints for the gestures are posted so that the user can see and learn (See Figure 4).

4.2.2. The User

To interact with the notice board the user must be standing in the range of the camera and perform the defined gesture with any hand. The gesture should be clear and should be same as it is defined by the system (See Figure 5).

Figure 5. User performing gesture in front of the sensor.

4.2.3. Microsoft Kinect Sensor for Xbox One

In the tech industry, for the development of IoT AR/VR there are various technologies that are available to use like HoloLens, but all these technologies are quite expensive. In contrast, the Microsoft Kinect sensor is a cheaper technology, with most accessibility and contains depth camera to draw the picture of the surroundings for computer to understand[1]. Microsoft Kinect uses data from many sensors it contains like RGB camera, and Infrared (IR) emitter, Multi-array microphone, 3 axis accelerometer and Tilt motor and sends this data to local or remote computers[1]. Kinect sensor has facial recognition and voice recognition capabilities. In our system, we have used the Kinect's depth sensing device and skeletal tracking capabilities to detect user and its gestures (See Figure 6). The Kinect stays active and keeps polling for a body to be tracked and as soon as the body is detected, and the gesture is recognized, the sensor fires an event to the application program.

Figure 6. Microsoft Kinect Sensor for Xbox One.

4.3. Back End

The gesture recognition technology is referred here as the back end of the System. The mechanism that is used in this research to detect and recognize the gestures is the AdaBoostTrigger Algorithm. It is a data-driven Machine Learning solution for the purpose of gesture recognition. This mechanism offers high accuracy for even very difficult gestures[13].

4.3.1. Tools And Technology

4.3.1.1. Kinect Studio

Kinect Studio (KS) is a tool that is used to record and play back the color and depth data streams from Kinect sensor. This tool is used to read and write data streams to help create repeatable scenarios for testing, debugging functionality, and analyzing performance[3].

The Kinect Studio requires sufficient computer resources for recording and playing back Kinect data. Data frames drop if the throughput of the data is not enough. The system should have plenty of space in hard drive, Fast CPU, extra RAM and run few other applications to get good performance while recording data[2].

4.3.1.2. Visual Gesture Builder

Visual Gesture Builder (VGB) is a tool developed by Microsoft and comes bundled up with Microsoft Kinect for Windows SDK 2.0. This tool incorporates within itself the algorithms and mechanism for data-driven machine learning solution for gesture recognition. The algorithm for gesture recognition needs to work for a variety of people with various physical characteristics in diverse environments like changes in lightning, physical data, heights and angles of sensor. To handle all these factors and environmental characteristics, the code becomes very much complex as more and more code is being added[13]. VGB simplifies this task and let the developers to work creatively and raise the quality of their applications with reduced latency and better gesture recognition[13]. Microsoft Kinect has been used in various researches to detect gestures using depth sensors[7, 15-20].

4.3.1.3. Machine learning technologies

There are numerous technologies in VGB to comprehend gestures[13]. All these technologies are combined into two groups: Discrete and Continuous.

Since in this research, static gestures are used therefore discrete indicators will be used for gesture detection. The discrete indicator of VGB i.e. AdaBoost Trigger determines following:

a) If the Person is performing a gesture or not.
b) Confidence value of the system in that gesture.

The AdaBoost Trigger is a detection technology that produces a binary or discrete result. During training time it accepts input tags which mark the occurrence of a gesture.

4.3.2. Methodology

VGB uses the clips recorded by Kinect studio. In the clips, the gestures are recorded by the people in raw format then converted using Kinect for windows SDK 2.0 tool. Later on, from the clips, all the frames that have the intended

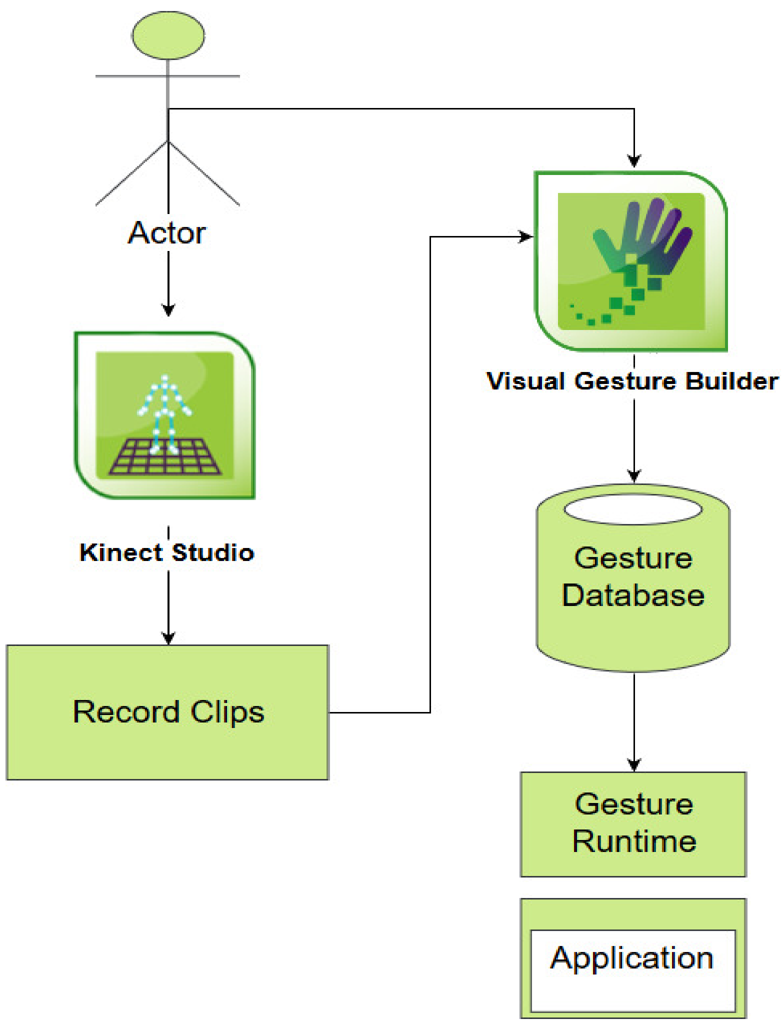

Figure 7. Backend Methodology of the architecture.

gestures are identified and then tagged or given label as positive or negative using Kinect VGB. Tagging is the process of identifying gesture frames in the clips. The machine learning algorithm builds a gesture database that can be run as gesture detector at runtime (See Figure 7). Two projects are created for this research. 1) Building/Training the gestures, 2) Analyzing/Testing the gestures.

4.3.3. Training Gestures

The clips used to build the database used as the training data are added in build project. The number of clips required to train varies from gesture to gesture, like if gesture requires hundreds of clips the other clip may require thousands of clip to train[13].

Training data should contain all the gestures that are defined for the system and should be recorded with all the considered factors for example: light, clothing, environment, background, people classification, tilt angles etc. Training gesture could take a lot of time[13] as it did in our research and the clips also occupy a lot of space. In our case the number of 840 clips each of 5 seconds occupied 53 GB of memory space.

4.3.4. Building And Analyzing Gestures

For testing and analyzing the gestures there are two ways, either use live preview or add test clips in the project. In live preview, user can directly perform gestures at runtime and measure confidence value. In the second case, clips are recorded for testing purpose involving various considered factors and are added in Analyze Project. The Analyze Project runs against the Build Project and the algorithm at the back end recognizes the gestures from it. Finally, it measures the confidence for each gesture. A very detailed log is also generated in the log tabs where all the classifiers and parameters are displayed that were generated during the project build.

5. Results and Discussion

The system under consideration is tested by the gestures collected from 49 people. The people were 99% students. Out of them 33 were male and 16 were female. The gestures for all 0–9 ASL gestures were collected for both left and right hands from each person. 80% of the gestures were used to train the database and 20% of the gestures were used to test the trained data set in order to find if the gestures have been detected or not and by what confidence value. Each recorded clip was of length 5 seconds. The confidence value varied based on factors like hand space, hand size, person's height, clarity in gesture performance. Figure 8 shows the confidence values of the set of testing data runs against the set of training data for both left and right hand. It can be seen that gestures which are clearer in terms of palm exposure and thumb joint have high confidence values for both hands. For example: 3, 4 5, 7, 8 & 9. In these gestures the palm and thumb is quite visible. The gestures which are complex in terms of clarity of the palm and thumb have low confidence values. For Example: 1 and 2. In these gestures only one or two fingers are quite visible and rest of the palm is not which seems to affect the detection of the hand or gesture.

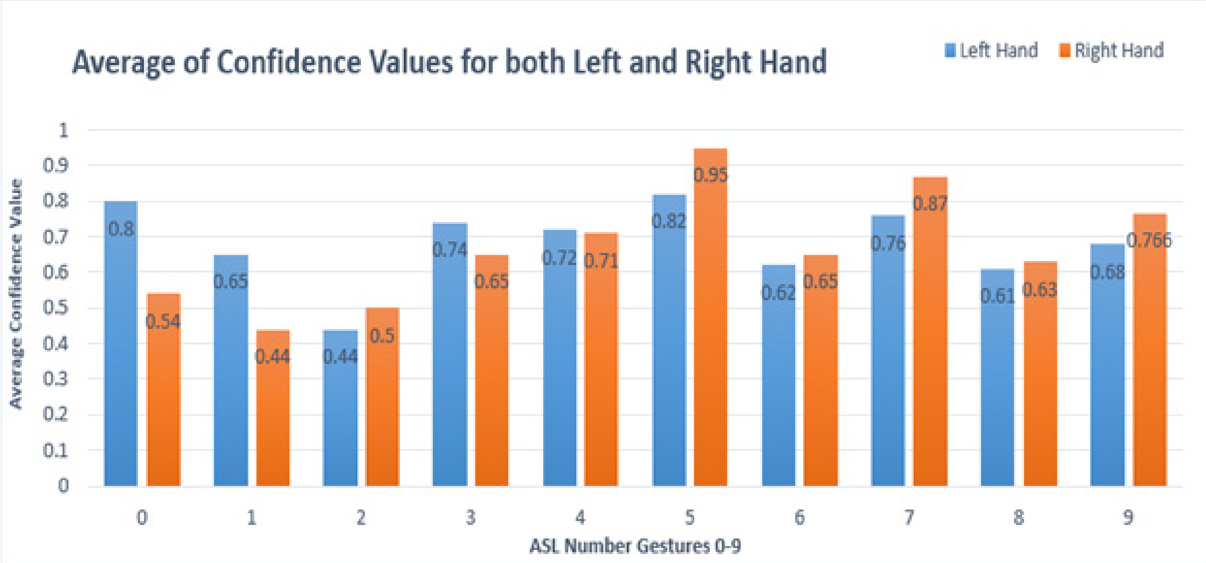

Figure 8. The average values of the Confidence values for each gesture performed by Left and Right hand.

5.1. Gender Based Study

We have also done the study based on gender to find out the confidence values for males and female gestures. Since the characteristics of hands of male and female are different. Usually the hands of female are smaller and less wide but the hands of men are wider. Hence, the results vary due to the shape of the hand, depending on the clarity of the performed gestures. We compared the gesture data of both Right and Left hands of 4 male and 4 females. Each clip comprises of 5 seconds length. The confidence values for all the gestures were calculated and represented in the form of graph for each gesture (See Figures 9-10). The confidence values vary based on gesture complexity against the variation in hand shape. It has been concluded from this gender based study that using machine learning algorithm the results detection of hand and recognition of gestures may vary based on gender depending on the complexity of the gesture and palm exposure. The Training database should be rich enough to train the algorithm for more complex gestures.

6. Survey

During the process of System testing, 49 users tested the system by performing all 0–9 gestures from both the hands. After the testing process a survey was conducted to record their feedback regarding the system. The survey was conducted to find out the general understanding of the people about the system, usability and the easiness to perform and remember the gestures to interact with the notice boards. The survey paper consists of following questions:

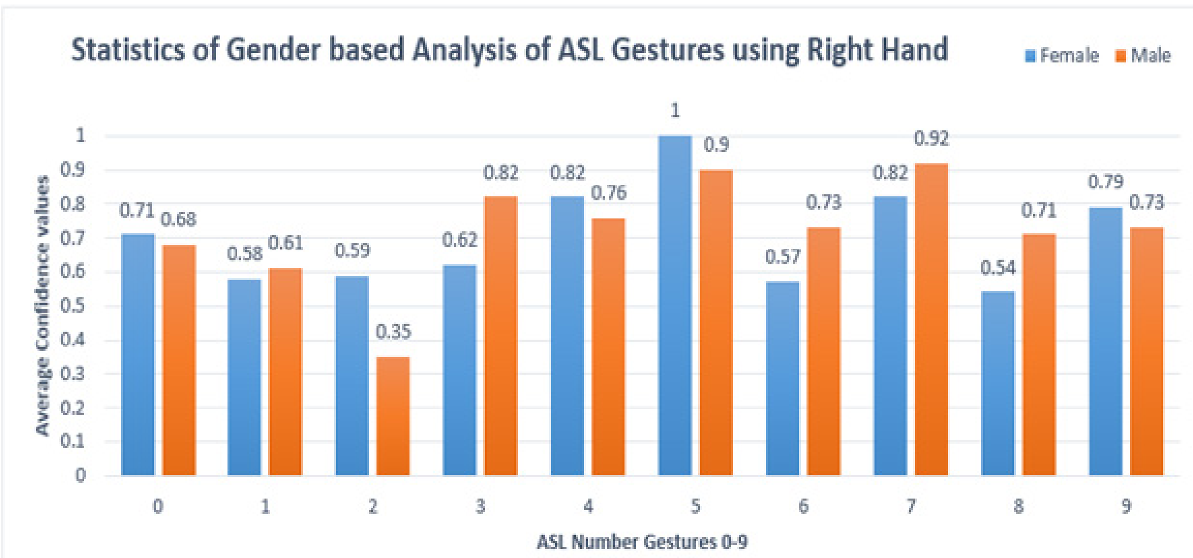

Figure 9. The average values of the Confidence values for each gesture of Left hand based on genders.

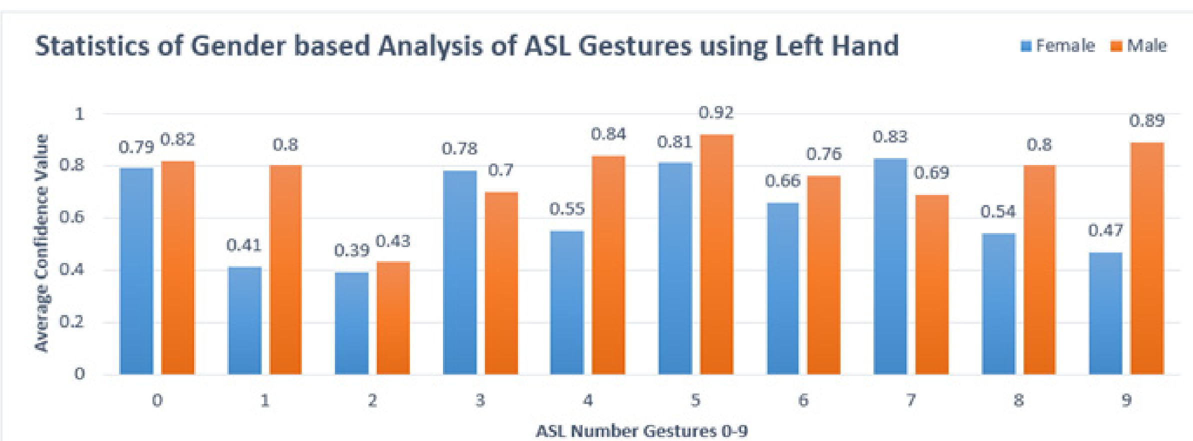

Figure 10. The average values of the Confidence values for each gesture of Left hand based on genders.

1. How much easy it is to learn ASL gestures?
2. How likely it is to remember ASL gestures?
3. For How Long the gestures can be remembered? (very short time period, short time period, long
4. How difficult it is to perform ASL gestures?
5. How much useful it is to embed gesture technology with displays?

The results for the above given questions are given in Table 1. People were given options from 'very easy' to 'very difficult' to select the option best suited for them. This survey was performed by all the people who tested the system. Almost 48.91% of the students found it very easy to learn ASL gestures. And rest of the student found it easy, not any of the users found it difficult to learn ASL gestures. 32.6% of the user said it is very easy to remember the gestures where as other 67.4% said it is easy to remember the gestures. 44.89% of the people said they can remember the gestures for short time period and 38.77% of user said they can remember it for long time period. Few user said they can remember for very short time period, i.e., 6.12% and few said they can remember for very long time period, i.e. 10.22%. 4.09% of user found it difficult to perform the ASL gestures rest found them either very easy (36.73%) or easy (59.18%). All of the users have opinion that the idea is very useful (55.10%) or useful (44.9) none said that it a useless system (See Table 1).

Table 1. Result of the Survey

How much easy it is to learn ASL gestures?	Very easy 48.91%	Easy 51.09%	Difficult 0%	Very difficult 0%
How likely it is to remember ASL gestures?	Very easy 32.6%	Easy 67.4%	Difficult 0%	Very difficult 0%
For How Long the gestures can be remembered?	Very short time period 6.12%	Short time period 44.89%	Very long time period 38.77%	Long time period 10.22%
How difficult it is to perform ASL gestures?	Very easy 36.73%	Easy 59.18%	Difficult 4.09%	Very difficult 0%
How much useful it is to embed gesture technology with notice boards?	Very Useful 55.10%	Useful 44.9%	Useless 0%	

7. Conclusion

This research aimed to make the static notice boards interactive so that a user could make choices and could transfer information from those paper based notice boards into their digital devices. With the framework proposed in this research, this objective is achieved with greater accuracy. The frame work showed promising results with varying hand shapes and notice board organizations. The results were satisfactory with up to 90% accuracy and confidence. This system cost lesser than digital solutions. Using this approach, the customers do not need to change the business processes.

8. Future Works

We are looking forward to improvise the gestures from being limited to American Sign Language only and go beyond 10 notices only. In future, we are going to analyze the system based on many other parameters and factors in addition to make this system easier to learn and use.

9. Acknowledgment

This research is funded by the Administration of Mehran University of Engineering and Technology, Jamshoro. This research required Microsoft Kinect sensor which is funded by the University.

10. References

1. Chen Q, Malric F, Zhang Y, Abid M, Cordeiro A, Petriu EM, Georganas ND. Interacting with digital signage using hand gestures. International Conference Image Analysis and Recognition. Springer, Berlin, Heidelberg. 2009 Jul; p.347–58. https://doi.org/10.1007/978-3-642-02611-9_35.
2. Davies N, Clinch S, Mikusz M, Bates O, Turner H, Friday A. Better off: when should pervasive displays be powered down? In Proceedings of the 6th ACM International Symposium on Pervasive Displays. 2017 Jun; p.1–19. https://doi.org/10.1145/3078810.3078821.
3. Toha TR, Ishmam AM, Islam MH, Al Maruf MA, Nandi SS, Chakraborty A, Estyak S, Al Alamin MA, Al Islam AA. An approach towards greening the digital display system. IEEE 4th International Conference on Networking, Systems and Security (NSysS). 2017 Dec; p.1–6. https://doi.org/10.1109/NSYSS2.2017.8267794.

4. Smari K, salim Bouhlel M. Gesture recognition system and finger tracking with kinect: Steps. 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT). 2016 Jul; p.544–8.
5. Kulshreshth A, Zorn C, La Viola JJ. Poster: Real-time markerless kinect based finger tracking and hand gesture recognition for HCI. IEEE Symposium on 3D User Interfaces (3DUI). 2013 Mar; p.187–8.
6. Hardy J, Rukzio E, Davies N. Real world responses to interactive gesture based public displays. In Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia. 2011 Dec; p.33–9. https://doi.org/10.1145/2107596.2107600.
7. Li Y. Hand gesture recognition using Kinect. IEEE International Conference on Computer Science and Automation Engineering. 2012 Jun; p.196–9. https://doi.org/10.1109/ICSESS.2012.6269439.
8. Dong C, Leu MC, Yin Z. American Sign Language alphabet recognition using Microsoft Kinect. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2015; p.44–52.
9. Huang F, Huang S. Interpreting American sign language with Kinect. Journal of Deaf Studies and Deaf Education. 2011 Dec; p.1–5.
10. Zafrulla Z, Brashear H, Starner T, Hamilton H, Presti P. American sign language recognition with the kinect. In Proceedings of the 13th international conference on multimodal interfaces. 2011 Nov; p.279–86. https://doi.org/10.1145/2070481.2070532.
11. Zeng B, Wang G, Lin X. A hand gesture based interactive presentation system utilizing heterogeneous cameras. Tsinghua Science and Technology. 2012 Jun; 17(3): p.329–36. https://doi.org/10.1109/TST.2012.6216765.
12. Wang F, Ngo CW, Pong TC. Simulating a smart board by real-time gesture detection in lecture videos. IEEE transactions on multimedia. 2008 Aug; 10(5):926–35. https://doi.org/10.1109/TMM.2008.922871.
13. Xdk XO, Windows KF. Visual Gesture Builder : A Data-Driven Solution to Gesture Detection. 2014; p.1–17.
14. Suarez J, Murphy RR. Hand gesture recognition with depth images: A review. IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication. 2012 Sep; p.411–7. https://doi.org/10.1109/ROMAN.2012.6343787.
15. Biswas KK, Basu SK. Gesture recognition using Microsoft kinect®. The 5th International Conference on Automation, Robotics and Applications. 2011 Dec; p.100–3.
16. Du H, To T. Hand gesture recognition using Kinect. Technical Report. Boston University. 2011 Dec.
17. Yang C, Jang Y, Beh J, Han D, Ko H. Gesture recognition using depth-based hand tracking for contactless controller application. IEEE International Conference on Consumer Electronics (ICCE). 2012 Jan; p.297–8. https://doi.org/10.1109/ICCE.2012.6161876.
18. Chang YJ, Chen SF, Huang JD. A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. Research in developmental disabilities. 2011 Nov; 32(6):2566–70. https://doi.org/10.1016/j.ridd.2011.07.002 PMid:21784612
19. Ronchetti M, Avancini M. Using kinect to emulate an interactive whiteboard. Final thesis, Corso di Laurea Magistrale in Informatica. Università Degli Studi Di Trento. 2011; p.1–129.
20. Ren Z, Yuan J, Meng J, Zhang Z. Robust part-based hand gesture recognition using kinect sensor. IEEE transactions on multimedia. 2013 Aug; 15(5):1110–20. https://doi.org/10.1109/TMM.2013.2246148.