iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42575

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Voice to Sign Language Converter

Akshay Kishore[1], Akshita Chauhan[2], Pooja Verma[3], Shivam Veraksatra[4]
[1, 2, 3, 4]*Department of Computer Science & Engineering, Meerut Institute of Engineering & Technology, 250005 Meerut, Uttar Pradesh India.*

Abstract: The aim of this paper is to design a convenient system that is helpful for the people who have hearing difficulties and in general who use very simple and effective method; sign language. This system can be used for converting sign language to voice and also voice to sign language. A motion capture system is used for sign language conversion and a voice recognition system for voice conversion. It captures the signs and dictates on the screen as writing. It also captures the voice and displays the sign language meaning on the screen as motioned image or video
Keywords: Motioned Image, Sign Language Converter, Voice Recognition.

I. INTRODUCTION

The point of this paper is to work on the correspondence with individuals who has hearingchallenges and utilizing any gesture based communication to put themselves out there. At the main sight, as a thought, how troublesome could make a communications via gestures converter. After nitty gritty exploration about gesture based communication phonetics, it is sorted out around 240 gesture based communications have exist for expressed dialectson the planet. To show how intense to functioning with any gesture based communication, the overall data about communications via gestures is given momentarily.

After have a thought regarding gesture based communication semantics, Microsoft Kinect Sensor XBOX 360 is chosen to use for catchingcapacities and specialized elements to the movement catch of sign to voice change.

Google Voice Recognition is utilized for the voice to sign transformation. Google Voice Recognition isaccessible just on android based projects

Eventually, the voice recognition program CMU Sphinxis chosen. This allows us to combine both components inJava. Conversion program is also designed and written in Java.

Finally, Java based program is produced which can make voice recognition, motion capture and convert both of them to each other. So a deaf person easily speaks to in sign language in front of motion sensor, theperson behind the screen can understand easily without ability to speak sign language and vice versa.

II. INFRASTRUCTURE ANDIMPLEMENTATION

Framework of a communication via gestures framework comprises of three principle branches as Sign Language, Speech Recognition and Implementation with MS Kinect XBOX 360TM. These are the principle inspirations of executing such a framework. The accompanying areas are depicting each term in subtleties and giving vital data.

A. Sign Language

It is not difficult to track down a wide number of communications via gestures all around the world and pretty much every communicated in language has its separate communication through signing, so there are about in excess of 200 dialects accessible. There are a few communications via gestures accessible like American, British, German, French, Italian, and Turkish Sign Language. American SignLanguage (ASL) is notable and the best concentrated on gesture based communication on the planet. The sentence structure of ASL has been applied to other communications through signing particularly as in British Sign Language (BSL).

BSL isn't firmly connected with ASL, so the distinctions among BSL and ASL are displayed in Figure 1. This part won't go further with subtleties of a solitary gesture based communication in light ofthe fact that each communication through signing has its own guidelines. The following segment will plan to give an overall depiction of the common or normal qualities between the different communications via gestures: beginning, phonology, and language structure. Plan a gesture based communication interpreter is certifiably not a simpleundertaking.

©IJRASET: All Rights are Reserved





Figure 1. Differences between American Alphabetand British Alphabet.

1) Origin of Sign Language

Hard of hearing individuals need communication through signing to speak with one another and other hard of hearing individuals. Also, a few ethnic gatherings that utilization totally various phonologies(for example Plain Indians Sign Language, Plateau Sign Language) have utilized gesture based communications to speak with other ethnicgatherings. The beginning of the gesture based communication is primarily connected with the startof the set of experiences.

The book of Juan Pablo Bonet called "Reduccion delas letras y Arte para ensenar a hablar los Mudos (Reduction of letters and craftsmanship forencouraging quiet individuals to talk) is distributed in Madrid in 1620 [1]. This is acknowledged as the main current composition of phonetics, organized a strategy for oral schooling for hard of hearing individuals by utilizing the manual signs, as displayed in Figure 2, of manual letter set to work ontheir correspondence. Nonetheless, this manual letterset was bad, but rather a method for making correspondence conceivable.



ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com



Figure 2. Manual Signs of Alphabet

The main genuine investigation of gesture based communications is accomplished in 1960s. Dr. William C. Stokoe distributed the monograph Sign Language Structure [2] in 1960. A portion of his hard of hearing understudies from the University of Gallaudet assist him with proposing the signs. Then, at that point, he distributed the main American Sign Language word reference [3]. In this first word reference, Dr. Stokoe coordinated the signs thinking about the place of the shape and movement. He didn't consider on its English interpretation. This is a foundation and give a beginning for research about the Sign Language semantics.

2) Phonology

The phonology alludes to the investigation of actual sounds present in human discourse. The phonology of communication via gestures can be characterized. Rather than sounds, the phonemes are considered as the various signs present in succession of hand signs.They are considering the accompanying boundaries:

a) *Arrangement:* Hand shape while doing the sign.
b) *Direction of the Hand:* Where the palm ishighlighting.
c) *Position:* Where the sign is finished.
d) *Movement:* Movement of the hand while doingthe sign (straight, influencing, circularly) [2].
e) *Contact Point:* Which portion of the hand contactthe body.
f) *Plane:* The sign is relying upon the distance to thebody.
g) *Non-manual Parts:* Information given by the body.For instance, when the body inclines front, it communicates future tense.

©IJRASET: All Rights are Reserved

3) Morphology

Spoken languages have inflectional morphology and also derivational morphology. The inflectional morphology refers to the modification of words. Thederivational morphology is the process of forming a new word on the basis of an existing word.
Sign languages have only derivational morphologybecause there are no injections for tense, number orperson. The most important parameters regarding morphology are represented as:

a) Degree: Mouthing.
b) Reduplication: Repeating the same sign severaltimes.
c) Compounds: Fusion of two different words.
d) Verbal Aspect: Expressing verbs in different ways.Several of these involve reduplication [2].
e) Verbal Number: To express plural or singularverbs. Reduplication is also used to express it.

4) Syntax

It is principally brought through a blend of word request and non-manual highlights. It is depicted by:

a) Word Request: A full construction as [topic][subject] [verb] [object] [subject-pronoun-tag].
b) Point and Principle Provisions: Background datasets.
c) Nullification: Negated provisos can be referenced by shaking the head during the wholestatement.
d) Questions: The inquiries are referenced bybringing down the eyebrows.
e) Conjunctions: Separate sign in ASL is a briefdelay.

5) Concepts of Sign Languages

Momentarily the fundamental information about gesture based communications has been addressed. The super phonetic attributes utilizedby the framework are important for the Phonology segment. The accompanying boundaries are thought of:

a) Position: The place that the sign ishappened.
b) Motion: Movement of the hand while doing the sign (straight, influencing, circularly).
c) Plane: The distance regarding the body.

The different fundamental signs from the ASL word reference have been taken and the boundaries doled out to these qualities as regardto the position, movement and plane.

III. IMPLEMENTATION

In this part, the execution steps of communications via gestures converter are addressed. Initially, the sensors are thought ofand afterward programming and middleware determinations are referenced

A. Speech Recognition

1) Types of Automatic Speech Recognition (ASR): ASR items have existed since the 1970s. Be that as it may, early frameworks were over the top expensive equipment gadgets. These gadgets were not entirely solid and easy to understand that could perceive a couple of separated words implies that client stops each word and should have been prepared by clients rehashing every one of the words a few times. The 1980s and 90s had a significant improvement in ASR calculations and items and furthermore the innovation created. In the last part of the 1990s, programming fordesktop correspondence openedup for two or three many dollars. Disposal with amechanical viewpoint it is feasible to describe between two expansive sorts of ASR: "direct voice input" (DVI) and "huge jargon nonstop discourse acknowledgment" (LVCSR). DVI gadgets are essentially focused on voice order and-control, while LVCSR frameworks are utilized for filling the structures or voice-based archive creation. DVI frameworks are ordinarily designed for little to medium measured vocabularies and might utilize word or expression spotting strategies. In the two cases the fundamental innovation is pretty much something similar.

2) *Speech Recognition Techniques:* Discourse acknowledgment strategies are summedup in subtleties in this segment.

a) Template Based ApproachesMatching: Obscure discourse is thought about against a bunchof pre-recorded words (layouts) to track down the best match.
b) Knowledge Based Approaches: An expert knowledge about variations in speechis hand coded into a system
c) Statistical Based Approaches: In which variations in speech are modelled statistically, using automatic, statistical learning procedure, typically the Hidden Markov Models,or HMM.

IV. METHODOLOGY

There are two parts of methodology:

1) Database
2) Voice Recognition Procedure

A. Database

Words for Speech Recognition, .gif images andMotions together create the database.

1) Words for Speech Recognition

For the speech recognition [8-13], there are fifty words has chosen as shown in Table 2. There are13 personal pronouns, 14 verbs, 5 adjectives, 12 nouns, 3 question words, and 3 yes, no and not for yes/no statements. With the flexibility of Sphinx it is possible to add new words to the system.

PERSONAL PRON.
I She TheyYou He His It We Me HerHim My Us

VERBS	ADJECTIVES
Am/Is/Are Like Thank Feel See Make Go Come Can Help Do Eat Drink Love	Good Bad Sick Fine Great

NOUNS
Tea Sign Student Teacher Doctor Language Father Mother School Brother Fruit Sister

QUESTIONS	YES AND NO
How Where Who	Yes No Not

2) Images (.gif format)

In the Speech Recognition part, the program utilizes the .gif pictures to show the appropriate significance for the perceived discourse. Each words or word bunches have an importance on the Sign Language. For instance, on the off chance that the speaker says: "I'm a specialist", the program will show the accompanying .gif pictures continuously for client to comprehend the Sign Language significance of the sentence.

3) Procedures

As the undertaking has two sections with the communication via gestures to voice, and voiceto communication through signing interpretations, there are two systems in the task. These systems are working with deferencethe decision of the client. When a strategy begins to work, it completes its part and afterward give the client one more decision for program. Principle Menu, Voice to Sign and Sign to Voice Menu of the Program are displayed in Figure. Sign to Voice system can be said as working inverse as Voice to Sign Procedure. Sign to Voice technique has two attempts to do, as to record the move, then, at that point, observe the appropriate text importance for the move, and inside the program changes it over to the Acoustic Signal.While, Voice to Sign methodology is, recordingthe acoustic sign and converts it to the text and afterward to the .gif documents.

B. Voice Recognition Procedure

Discourse handling is the field which deals withthe discourse signals and the handling of them. The signs are normally handled in an advanced portrayal, albeit the signs are simple. Discourse handling is keen on to accumulate, store, control, move discourse signals.
It is quicker to speak with the voice than text, hence with the interpretation of voice to the picture will give sound individuals to speak withindividuals with the consultation issues.
When the client press the button to record the discourse, PC's mouthpiece begins to tune in, and in the wake of getting the voice with the assistance of CMU Sphinx, it tracks down the importance as the text. Then, at that point, in Java it is coordinated with the legitimate .gif picture, so the other client will comprehend. Thegraph of the Voice Recognition Procedure is given in Figure 03.

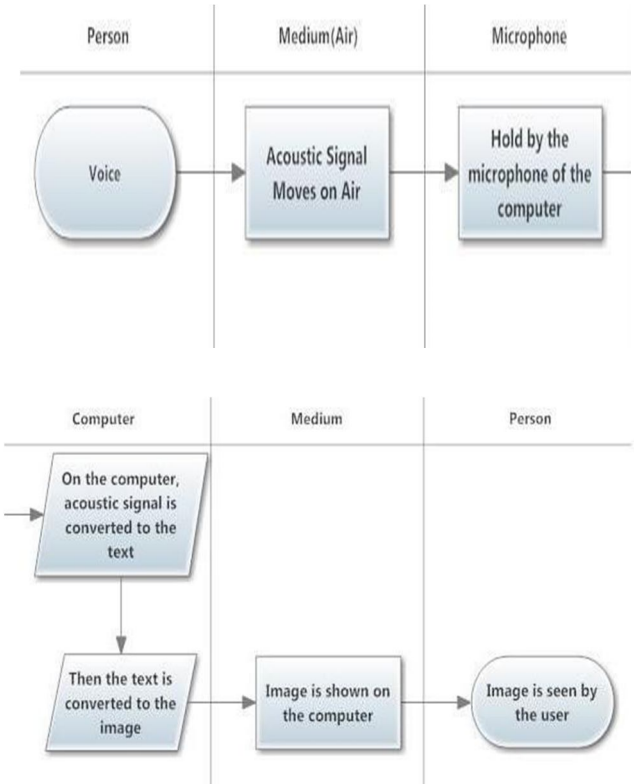


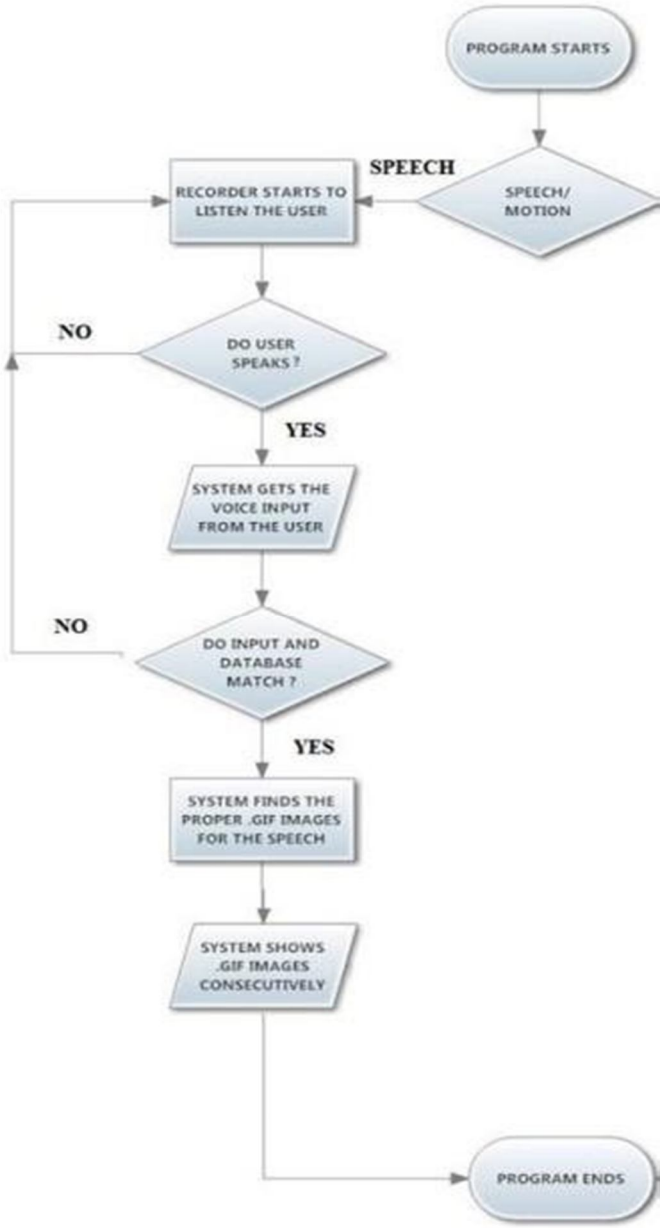
Figure 03. The diagram of the Voice Recognition Procedure

V. CONCLUSIONS

This paper is about a framework can uphold the correspondence among hard of hearing andnormal individuals. The point of the review is to give a total exchange without knowing gesture based communication. The program has two sections. First and foremost, the voice acknowledgment part utilizes discourse handling techniques. It takes the acoustic voice sign and converts it to an advanced sign in PC and afterward show to the client the .gif pictures as result. Also, the movement acknowledgment part utilizes picture handlingstrategies. It utilizes Microsoft Kinect sensor and afterward provide for the client the result as voice.

The task provides us with the many benefits ofuse area of communication through signing. After this framework, it is a potential chance to involve this kind of framework in any spots,for example, schools, specialist workplaces, schools, colleges, air terminals, social administrations offices, local area administration organizations and courts, momentarily all over the place.

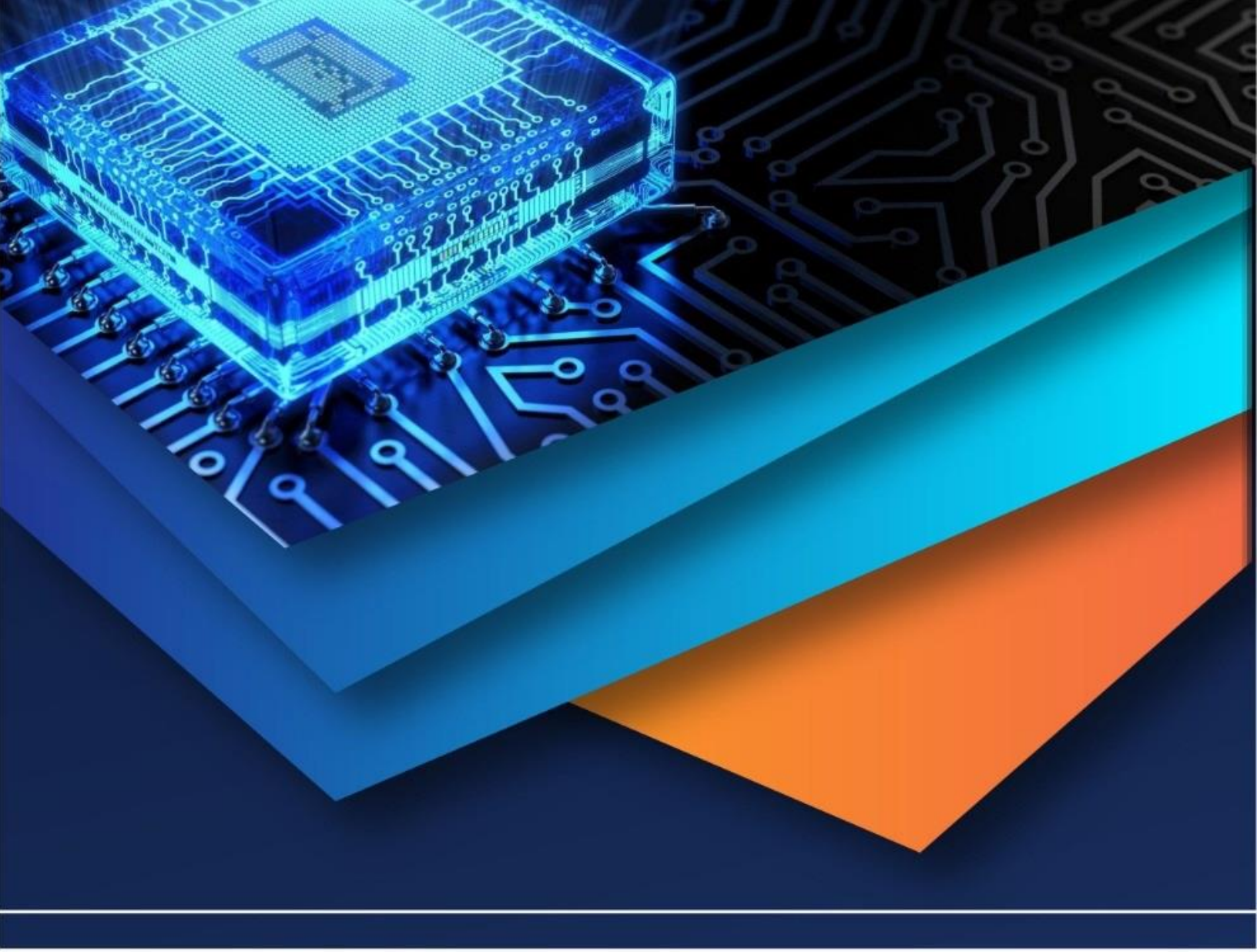
One of the main showings of the capacity for correspondence to assist with marking language clients speak with one another happened. Communications through signing can be utilized wherever when it is required and it would arrive at different neighborhoods.The future works are tied in with creating versatile use of such framework that empowers everybody have the option to talk with hard of hearing individuals.



REFERENCES

[1] J.P. Bonet. “Reducci_on de las letras yarte para ense~nar a hablar a los mudos”, Coleccion Cl_asicos Pepe. C.E.P.E., 1992.
[2] William C. Stokoe. Sign Language Structure [microform] / William C. Stokoe.Distributed by ERIC Clearinghouse, [Washington, D.C.], 1978.
[3] William C. Stokoe, Dorothy C Casterline,and Carl G Croneberg. “A Dictionary of American Sign Language on Linguistic Principles” Linstok Press, [Silver Spring, Md.],New Edition, 1976.
[4] Code Laboratories. CL NUI Platform.http://codelaboratories.com/ kb/nui
[5] The Robot Operating System (ROS),http://www.ros.org/wiki/ kinect.
[6] Open Kinect Project, http://openkinect.org/wiki/Main_Page.
[7] Open NI API Reference. http://openni.org/Documentation/Reference/index.html.
[8] Bridle, J., Deng, L., Picone, J., Richards, H., Ma, J., Kamm, T., Schuster, M., Pike, S., Reagan, R., “An Investigation of Segmental Hidden Dynamic Models of Speech co- articulation for Automatic Speech Recognition.”,Final Report for the 1998 .
[9] Ma, J., Deng, L., “Target-directed MixtureLinear Dynamic Models for Spontaneous Speech Recognition”, IEEE Transactions on Speech and Audio Processing, Vol. 12, No. 1, January 2004.
[10] Ma, J., Deng, L., “A Mixed-level SwitchingDynamic System for Continuous Speech Recognition”, Elsevier Computer Speech andLanguage 18 (2004) 4965, 2004.
[11] Mori R.D, Lam L., Gilloux M., “Learning & Plan Refinement in a Knowledge Based System fo r Automatic Speech Recognition”, IEEE Tra. on Pattern Analysis Machine Int., 9(2):289-305, 1987.
[12] Rabiner, L., R., and Wilpon, J. G., “Considerations in Applying Clustering Techniques to Speaker- independent Word Recognition”, Journal of Acoustic Society of America, 66(3):663-673, 1979.
[13] Tolba, H., and O’Shaughnessy, D., “SpeechRecognition by Intelligent Machines”, IEEE Canadian Review (38), 2001.
[14] Kathryn LaBelle, “Kinect RehabilitationProject”, http://netscale.cse.nd.edu/twiki/bin/ view/Edu/Kinect Rehabilitation, June 2009.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)