Smart Phone Based Indian Sign Language Maker for Aiding Deaf and Dumb People

R. Manoharan[#], R. Kavitha[*], A. Pavithra[*], Meenatchi[*]

[#] *Department of Computer Science and Engineering, Rajalakshmi Engineering College,*

[*] *Adhiparasakthi Engineering College*

***Abstract*— Sign language is the native language of deaf and dumb people which they prefer to use it on their daily life. This paper describes architecture of a smart phone based Indian sign language maker for aiding deaf and dumb people (special people). This system comprised of two features namely, speech to gesture (Indian Sign Language) and text to speech. The speech to gesture can be done with the help of Speech recognizer, semantic analyzer, gesture sequence generator and gesture player. The text to speech is done with the help of Google text to speech engine. The aim of the project is to aid special people by translating the speech to Indian Sign Language and converting text to voice. Speech recognition and semantic analysis is done with the help of the Google voice. Gesture sequence generation and gesture playing is the place where we focused our work. Gesture sequence is generated based on the semantic analyzer output in which the gesture sequence generator can recognize. Gesture sequence generator will produce sequences read by gesture player which is nothing but an animated human agent who will perform hand signs corresponding to the input given. This gesture sequence generator will also learn from the user if the speech is not properly recognized or the words are not found in the sequence generator dictionary. This paper is an initial work of the project in which, gesture generation is done for particular domains like railway, airport and bus transport kiosk where a deaf person can respond to the phone calls from various customers. The future work will be a complete system which can produce hand signs for any domain.**

Keywords-speech to gesture; voice recognition; gesture playing; indian sign language.

I. INTRODUCTION

A sign language (also signed language or simply signing) is a language which uses manual communication and body language to convey meaning, as opposed to acoustically conveyed sound patterns. This can involve simultaneously combining hand shapes, orientation and movement of the hands, arms or body, and facial expressions to fluidly express a speaker's thoughts. They share many similarities with spoken languages (sometimes called "oral languages", which depend primarily on sound), which is why linguists consider both to be natural languages, but there are also some significant differences between signed and spoken languages. A common misconception is that all sign languages are the same worldwide or that sign language is international. Aside from the pidgin International Sign, each country generally has its own, native sign language, and some have more than one, though sign languages may share similarities to each other, whether in the same country or another one. [1]

There are perhaps three hundred sign languages in use around the world today. The number is not known with any confidence; new sign languages emerge frequently through creolization and de novo (and occasionally through language planning). In some countries, such as Sri Lanka and Tanzania, each school for the deaf may have a separate language, known only to its students and sometimes denied by the school; on the other hand, countries may share sign languages, though sometimes under different names (Croatian and Serbian, Indian and Pakistani). Deaf sign languages also arise outside of educational institutions, especially in village communities with high levels of congenital deafness, but there are significant sign languages developed for the hearing as well, such as the speech-taboo languages used in aboriginal Australia. Scholars are doing field surveys to identify the world's sign languages. [2].

When a deaf or dumb person approaches a store or reception it will be hard to communication with the other persons through sign language which probably end up in ignorance. Since the evolution of mobile devices, these people always had some heavy feeling on their hear that they unable to use these mobile devices effectively. So to overcome these problems a system which can act as an intermediate between the special person and the normal person is built.

II. RELATED WORK

A lot of research has been carried on the development of systems that translate sign languages into spoken words and the reverse. At HTK in Stockholm, Joakim Gustafson [3], Bjorn Granstrom [4] and their colleagues have developed several multimodal dialogue systems where animated agents were incorporated to improve the interface. These include Waxholm [5] (a travel planning system for ferryboats in the Stockholm archipelago), August [6] (an information system at the Culture Center in Stockholm), and AdApt [7] (a mixed-initiative spoken dialogue system, in which users converse with a virtual real estate agent to locate apartments in Stockholm). Education is another domain in which language and animated agent technologies can be combined.

At this point, it is necessary to remark that there is a CSLU Toolkit which integrates an animated agent named

Baldi. This toolkit has been developed at CSLU (Center of Spoken Language and Understanding, Oregon Graduate Institute, OGI) [8,9] which is now being expanded at CSLR (Center for Spoken Language Research at University of lorado) [10]. This toolkit facilitates the speedy development of interactive books with multimedia resources and natural interaction. Nowadays, researchers have generated systems and architectures for representing and managing behaviors of animated agents. In [11, 12], the authors present a good overview of computational models for developing believable virtual humans.

Not only research centers but also companies like Microsoft and IBM are interested in animated agents. Microsoft has developed a software platform [13] where users can use several animated agents or create new ones (http://www.microsoft. com/msagent/). This platform began with the Persona Project [14]. IBM is also interested in technology which will be the future for human computer interfaces. In both the aforementioned systems, the synergy between language and virtual agent technologies is due to the fact that virtual humans offer a friendlier computer-user interface.

This synergy becomes stronger in our case where we want to develop a system to translate speech into gestures for Deaf people. In the recent years several groups have shown interest in machine translation for Sign Languages, developing several prototypes based on different language translation techniques: example-based [15], rule-based [16], full sentence [17] or statistical [18] approaches. In a speech to sign language translation system, the virtual agent appears as an essential part of the system. It represents the gestures obtained from the semantic analysis of the recognized words. All of the fore mentioned agent platforms suffer the inconvenience of the great effort needed to develop the agent animations. This is one of the problems we focus on in this paper: the development of a platform where minimal effort is required to create a new agent animation. This is an important aspect because the amount of gestures required by our system is higher than those of the aforementioned systems. With the help of the android based smart phones it is easy to retrieve the agent animations from the database and play it effectively [19]. Recently a project called MiMix which is been released for American Sign Language [20], have similarities with the system proposed here. But if any words are not recognized the system fails to learn from the user.

III. SYSTEM DESIGN

Figure 1 Presents the Architecture of Speech to Gesture Conversation. There are four main modules, which carry out the four steps needed in the translation process: speech recognition, semantic analysis, gesture sequence generation and gesture playing.

At first the voice is taken as input for the Speech Recognizer which will convert the voice in to text. The text is checked with semantics using the semantic analyzer. The speech recognizer and semantic analysis will be acting together to form the concepts used for the gesture sequence generation.

Gesture sequence Generation is done based on the concepts provided by the semantic analyzer. If the recognized word cannot find the proper sign then it will ask the user to provide the proper sign for the unrecognized word. Reinforcement Learning is used to learn the unrecognized sign from the user. By this way the user defined gestures also added with the system. Finally the gesture sequence which is used for the gesture playing is generated. The gesture play is nothing but the human model which will be doing hand signs in Indian Sign Language for the corresponding voice signal.

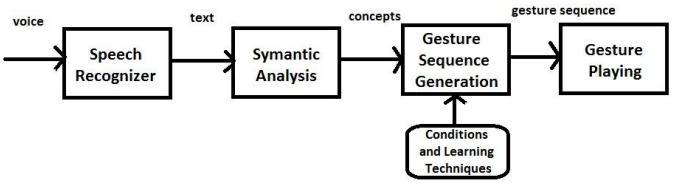

Figure 1. Architecture of Speech to Gesture Conversion

IV. IMPLEMENTATION IN SMART PHONE

Mobile Phones are improved in a way that any complex tasked can be performed easily in it. Even an octa core processor based smart phones are available in market with affordable costs. For deaf persons it's always been a problem on using a mobile phone. They cannot respond to a call without an aid. So to overcome this problem the system of speech to gesture is proposed. By using the proposed system, when a Deaf person uses a mobile phone, the voice from the other end (normal user) is given as the input to the speech to gesture conversion system. This System will play the Corresponding sign language on the mobile screen instantly. The voice is recognized with the help of the Google voice which will convert the given voice to text. This text is used to generate the Indian Sign Language through Gesture Sequence generation and Gesture Playing.

After seeing the gesture the deaf person will understand what the person on call says and reply correspondingly. As specified, this system is not only for aiding deaf people but also dumb people. The system has an additional feature to translate text to voice using Google Text to Speech Engine. For an example when a deaf and dumb person try to buy a good from a provisional store it is hard to specify why he/she needs. So through the proposed system the special person can easily communicate with the provisional shopkeeper and buy what he/she needs easily. Figure 2 shows how the Architecture of implementing speech to gesture conversion in smart phone in which the Google speech to text engine is used for speech recognition and semantic analysis. Gesture sequence generation and gesture playing are same as the actual system.

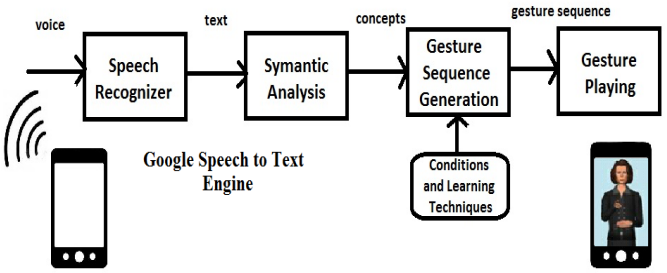

Figure 2. Architecture of Speech to Gesture Conversion in Smart Phone

V. CONCLUSION AND FUTURE WORK

In this paper we proposed a smart phone based Indian sign language maker for aiding deaf and dumb people. It will be useful for all the special people in India as well as the special people around the world. Since Indian Sign Language is easy to learn, it can be easily learnt and understood. Soon the working model of the proposed system will be released in Google Play store.

Our future work will be based on the Indian Sign Language maker for Tamil Language. Google has developing Tamil voice recognition and soon it will be online. Through Google Translate it is feasible to convert it in to English and implement the sign language maker using the above mentioned system.

REFERENCES

[1] http://en.wikipedia.org/wiki/Sign_language.

[2] http://en.wikipedia.org/wiki/List_of_sign_languages.

[3] J. Gustafson, Developing multimodal spoken dialogue systems-Empirical studies of spoken human-computer interactions, PhD. Dissertation, Department of Speech, Music and Hearing, Royal Institute of Technology, Stockholm, Sweden, 2002.

[4] B. Granstrom, D. House, J. Beskow, Speech and Gestures for Talking Faces in Conversational Dialogue Systems, Multimodality in Language and Speech Systems, Kluwer Academic Publishers, Donrecht, 2002 pp 209-241.

[5] J. Bertenstam, et al., The Waxholm system-A progress report, in: Proceedings on Spoken Dialogue Systems, Vigso, Denmark, 1995.

[6] M. Lundeberg, J. Beskow, Developing a 3D-agent for the August dialogue system, in: Proceedings on Audio-Visual Speech Processing, Santa Cruz, CA, 1999.

[7] J. Gustafson, L. Bell, Speech technology on trial: experiences from the August system, Journal of Natural Language Engineering: Special Issue on Best Practice in Spoken Dialogue Systems (2003) 273-286.

[8] S. Sutton, R. Cole, Universal speech tools: the CSLU toolkit, in: Proceedings of the International Conference on Spoken Language Processing, Sydney, Australia, 1998, pp. 3221-3224.

[9] R. Cole, et al., New tools for interactive speech and language training: using animated conversational agents in the classrooms of profoundly deaf children, in: Proceedings ESCA/ SOCRATES Workshop on Method and Tool Innovations for Speech Science Education, London, 1999, pp. 45-52.

[10] R. Cole, S. Van Vuuren, B. Pellom, K. Hacioglu, J. Ma, J. Movellan, S. Schwartz, D. Wade-Stein, W. Ward, J. Yan, Perceptive animated interfaces: first steps toward a new paradigm for human computer interaction, IEEE Transactions on Multimedia: Special Issue on Human Computer Interaction 91 (9) (2003) 1391-1405.

[11] W.L. Johnson, J.W. Rickel, J.C. Lester, Animated pedagogical agents: face-to-face interaction in interactive learning environments, International Journal of artificial Intelligence in Education 11 (2000) 47-78.

[12] J. Gratch, J. Rickel, E. Andre, N. Badler, J. Cassell, E. Petajan, Creating interactive virtual humans: some assembly required, IEEE Intelligent Systems 17 (4) (2002) 54-63. Microsoft Agent web. < http://www.microsoft.com/msagent/ index.html >.

[13] G Ball, D. Ling, D. Kurlander, J. Miller, D. Pugh, T. Skelly, A. Stankosky, D. Thiel, M. Van Dantzich, T. Wax, Lifelike Computer Characters: the Persona Project at Microsoft Research, 1999, <www.microsoft.com>.

[14] S. Morrissey, A. Way, An example-based approach to translating sign language, in: Workshop Example-Based Machine Translation (MT X-05), Phuket, Thailand, September, 2005 pp. 109-116.

[15] M. Huenerfauth, A multi-path Architecture for Machine Translation of English Text into American Sign language animation, HLT-NAACL, Boston, MA, USA, 2004.

[16] S.J. Cox, M. Lincoln, J. Tryggvason, M. Nakisa, M. Wells, M. Tutt, S. Abbott, TESSA, A System to Aid Communication with Deaf People, ASSETS, Edinburgh, Scotland, 2002, p. 205-212.

[17] J. Bungeroth, H. Ney, Statistical sign language translation, in: Workshop on Representation and Processing of Sign Languages, LREC, 2004, pp. 105-108.

[18] W. Stokoe, Sign Language structure: an outline of the visual communication systems of the American deaf, Studies in Linguistics, Buffalo University Paper 8, 1960.

[19] R. San-Segundo, J.M. Montero , J. Macías-Guarasa , R. Córdoba , J. Ferreiros , J.M. Pardo in : Proposing a speech to gesture translation architecture for Spanish deaf people, Science Direct,2008.

[20] https://play.google.com/store/apps/details?id=me.mimix.roid&hl=en

Manoharan Ramachandran Received the B.Tech Degree in Information Technology from Adhiparasakthi Engineering College, India in 2013 and currently pursuing his M.E Computer Science and Engineering Degree in Rajalakshmi Engineering College, India. He presented various research papers in National and International Conferences. His research interests include Vision Based Intelligence, Machine Learning, Advanced gadgets and safety in vehicles and Psychology. For his work he has received several awards including IRNET young investigator award in 2012 and ISTE best student award in 2013. He is a registered member of the IEEE.