International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 5 | ISSN : 2456-3307

An Innovative Artificial Replacement to Facilitate Communication between Visually and Hearing- Impaired People

Sangeetha R*[1], Elakkiya S[2]
*[1]Research Scholar, VLSI, Vandayar Engineering College, Thanjavur, Tamilnadu, India
[2]Assistant Professor of ECE in Vandayar Engineering College, Thanjavur, Tamilnadu, India

ABSTRACT

In this work, the problem of the communication between visually- and hearing-impaired people is a special case of particular interest, where the aforementioned aids cannot be applied because these users do not share any common communication channel. A proposed scheme included algorithms for speech recognition and synthesis to aid communication between visually and hearing-impaired people. Our proposed framework consists of a situated communication environment designed to foster an immersive experience for the visually and hearing impaired. To achieve the desired result, our situated modality replacement framework combines a set of different modules, sign language analysis and synthesis, speech analysis and synthesis, into an innovative multimodal interface for disabled users.

Keywords: Sign Language Synthesis, ANN, Speech Synthesis, Feature Selection, Feature Extraction

I. INTRODUCTION

Hearing loss can range from mild to severe and can be attributed to anything including age, occupational hazards such as military combat and exposure to excessive noise at work, injury to the ear, infections and more. In addition, hearing loss does not just make things quieter, it can distort sounds and cause normal speech to sound jumbled and garbled.

The National Federation describes a person for the Blind as visually impaired if his or her sight is bad enough, even with corrective lenses, that the person must use alternative methods to engage in normal activity. Here are some modifications that can be made to effectively communicate with the visually impaired. The purpose of this document is to provide you with some guidelines. You are, however, encouraged to consult additional resources that assist you in writing a professional technical paper.

In recent years, there has been an increasing interest in human computer interaction (HCI) for multimodal interfaces. Since Sutherland’s SketchPad in 1961 and Xerox’s Alto in 1973, computer users have long been acquainted with technologies other than the traditional keyboard for interacting with a system. Recently, with the desire for increased productivity, seamless interaction, immersion, and e-inclusion of people with disabilities, along with progress in fields such as multimedia, multimodal signal analysis, and HCI, multimodal interaction has emerged as an active field of research. Multimodal interfaces are those encompassing more or other input and output devices than the traditional keyboard, mouse, and screen, and the modalities they enable.

A key aspect in many multimodal interfaces is the integration of information from several different modalities to extract high-level information nonverbally conveyed by users or to identify cross-correlation between the different modalities. This information can then be used to transform one modality to another. The potential benefits of such modality transformations to applications for disabled users are important because disabled users could access an alternate communication channel to perceive information that was not previously available.

However, despite the proliferation of research in this area, the problem of the communication between visually- and hearing-impaired people is a special case of particular interest, where the aforementioned aids cannot be applied because these users don't share any common communication channel. This problem has been only theoretically dealt with in other work, where a proposed scheme included vibrotactile sensors, cameras for 3D space perception, and algorithms for speech recognition and synthesis to aid communication between visually and hearing-impaired people.

II. EXISTING METHOD

Most existing aids for the visually impaired detect visual cues captured from cameras and present them to the visually impaired using symbolic information in alternate communication channels, such as audio or vibrotactile feedback. Concerning aids for the hearing impaired, [A. Caplier et al., ''Image and Video for Hearing Impaired People,'' Eurasip J. Image and Video Processing, vol. 2007, article ID 45641, 2007.]. Most developments focus on the recognition and synthesis of sign language [S.C.W. Ong and S. Ranganath, ''Automatic Sign Language Analysis: A Survey and the Future beyond Lexical Meaning,'' IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no. 6, 2005, pp. 873-891.]. A problem that stresses the high importance of modality replacement is that communication between visually- and hearing-impaired users is not possible using physical means.

To achieve the desired result, our situated modality replacement framework combines a set of different modules, such as gesture recognition, sign language analysis and synthesis, speech analysis and synthesis, and haptic interaction, into an innovative multimodal interface for disabled users. Unlike this method, existing approaches focus mainly on the nonsituated and symbolic presentation of information to the disabled.6,7,10 In this article, the descriptions of the framework and platform are mainly used to contextualize our contribution. The prospective value of such a system is huge due to the socioeconomic aspects of including people with disabilities in the information age, which modality-replacement technologies can make possible. It should be emphasized that the risk of exclusion is one of the largest threats for the disabled. The proposed system opens a new, previously unknown dimension for visually- and hearing-impaired people, and offers an interactive application where visually- and hearing-impaired users can entertain themselves while also getting familiar with novel interaction technologies. Potential applications include universally accessible phones, generic accessible computer interfaces, social networks, digital assistants, and so on.

III. PROPOSED METHOD

Figure 1 presents the architecture of the proposed system, including the communication between the various modules used for integration of the system as well as intermediate stages used for replacement between the various modalities. The left part of the figure refers to the visually impaired user's terminal, while the right refers to the hearing-impaired user's terminal.

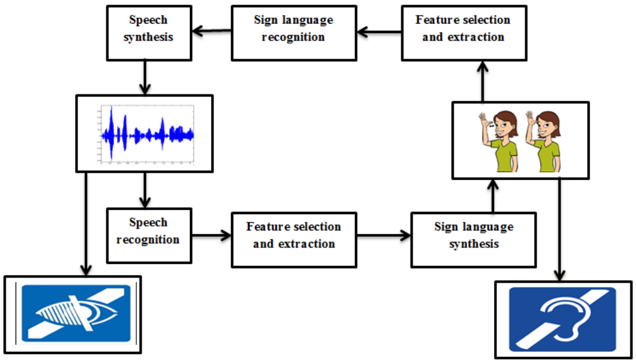


Figure 1. block diagram of the proposed method

The visually impaired user interacts with the computer using sign language and speech. Sign language is used to navigate in the virtual environment and to process and recognize the objects of the virtual scene. System tasks and verbal information are presented to the visually impaired user via speech synthesis, while the computer through speech recognition perceives verbal input from the visually impaired user.

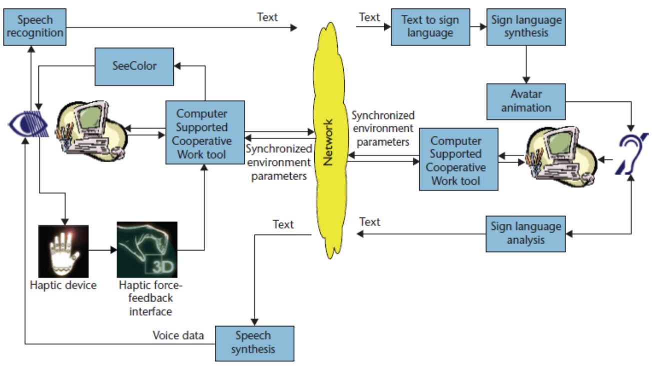


Figure 2. Real time architecture of the proposed method

At the hearing-impaired user's terminal, things are simpler because the user can use sight to navigate the virtual environment. To enable unobtrusive interaction, verbal information is presented to the user via sign language synthesis, while the user can provide input to the system through the sign language recognizer.

Visual information about the environment has to be conveyed to the visually-impaired user via the haptic and/or the auditory channel, while communication and the acquisition of various semantic information can be performed using natural language. The hearing-impaired user acquires visual information using vision and communicates with other people using sign language.

A problem that stresses the high importance of modality replacement is that communication between a visually- and hearing-impaired users is not possible using physical means. Ideally, as illustrated in Figure 1, a modality replacement system would be used to recognize all spoken language input of the visually impaired user, convert it into sign language, and present it to the hearing-impaired user with an animated avatar. Similarly, sign language gestures would be recognized and converted into text and would then be synthesized into speech using text-to-speech synthesis techniques. The present work takes the first step toward the development of such interfaces. It’s obvious that because multimodal signal processing is essential in such applications, specific issues such as modality replacement and enhancement should be addressed in detail. Figure 2 presents the architecture of the proposed system, including the communication between the various modules used for integration of the system as well as intermediate stages used for replacement between the various modalities. The left part of the figure refers to the visually impaired user’s terminal, while the right refers to the hearing-impaired user’s terminal. All actions are controlled by the Computer Supported Cooperative Work game system.

The visually impaired user interacts with the computer using sign language and speech. Haptic interaction is used to navigate in the virtual environment and to process and recognize the objects of the virtual scene. System tasks and verbal information are presented to the visually-impaired user via speech synthesis, while verbal input from the visually-impaired user is perceived by the computer through speech recognition. The SeeColor utility enables the perception of colors through the sounds of musical instruments. At the hearing-impaired user’s terminal, things are simpler because the user can use sight to navigate the virtual environment. To enable unobtrusive interaction, verbal information is presented to the user via sign language synthesis, while the user can provide input to the system through the sign language recognizer.

Artificial Neural Network (ANN):

ANN Classification is the process of learning to separate samples into different classes by finding common features between samples of known classes.

An Artificial Neural Network (ANN) is an information-processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true for ANNs as well.

Classifier Parameter Descriptions

Learners

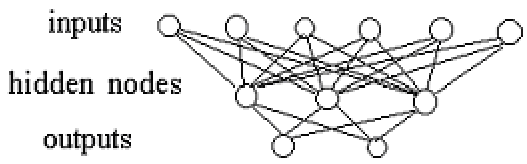
The number of learners to train. The samples are divided into N subsets. Each learner is trained on a different (N-1)/N samples and validated on the remaining 1/N samples. The default number is 10, corresponding to a conventional 10-fold cross-validation scheme. The number can be made as high as the number of samples (corresponding to leave-one-out cross-validation) or as low as 3. For most problems the default of 10 is fine.

Learner Votes Required

This is the number of learners which must vote for the same class in order for the Classifier to make a call (prediction) on a given sample. If fewer learners than this number agree, then the Classifier will make a class prediction of 'Unknown'. Raising this number may result in fewer misclassifications. Lowering it may lead to fewer 'Unknown' calls.

Hidden Units

This is the number of nodes in the hidden layer of each ANN. All ANNs have the same three-layer architecture: input nodes , hidden nodes, and output nodes. You can think of each node as corresponding to a neuron, and the interconnections between them as synapses, but this model should not be taken too literally.



There are as many nodes in the input layer as there are input features (genes) in the training dataset. There are as many nodes in the output layer as there are output classes. The number of hidden nodes in the middle layer is typically between these two numbers.

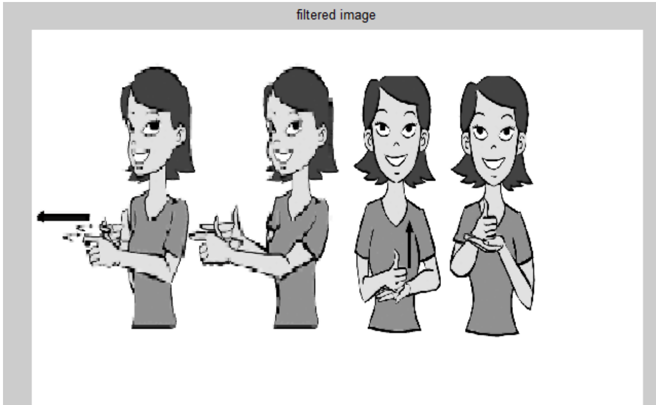
Setting the number of hidden nodes higher will usually result in overtraining, leading to poor results on test data. Setting the number of hidden nodes too low might result in an inability to learn even the training data, but this is easily detected by examining the results of the Create Classifier experiment. If the default number of hidden nodes yields good training results but poor test results, reduce the number of hidden nodes. If the default yields poor training results, try increasing the number of hidden nodes.

Conjugate Gradient Method

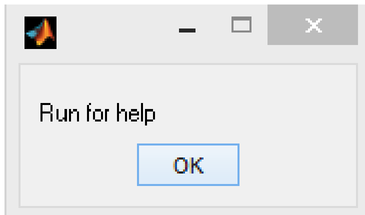
Polak-Ribiere and Fletcher-Reeves are two variants of the conjugate gradient algorithm used to optimize the neural network internal parameters during training. They differ in the formula used to update the search direction in internal parameter space.

IV. RESULTS

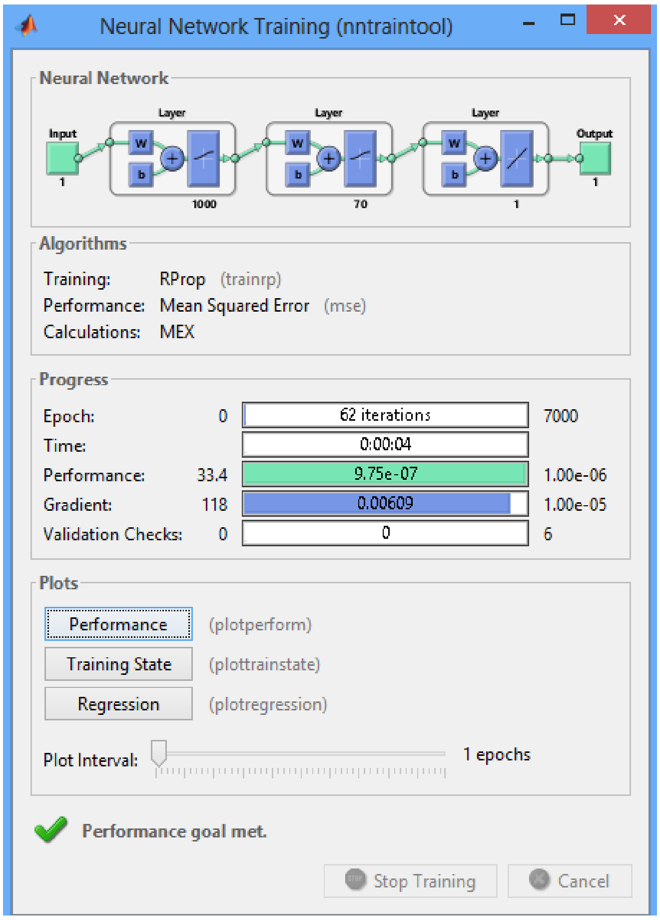
Input sign language after filtered:



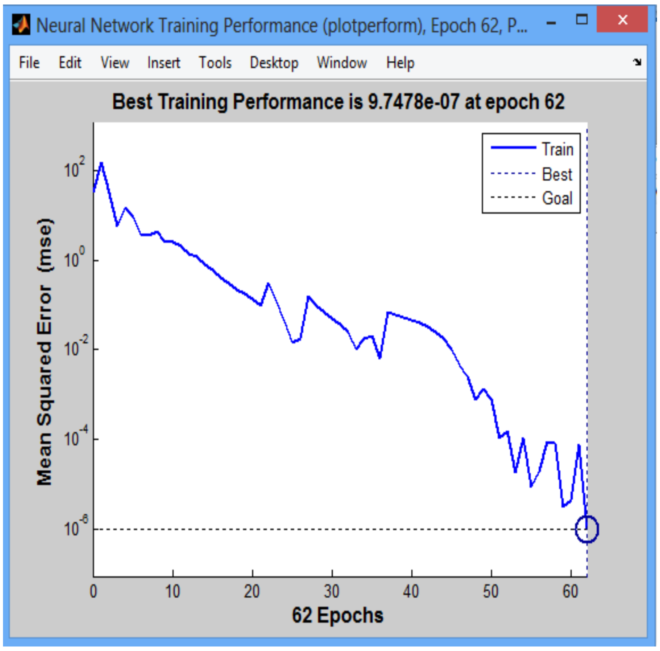
Sign language recognition



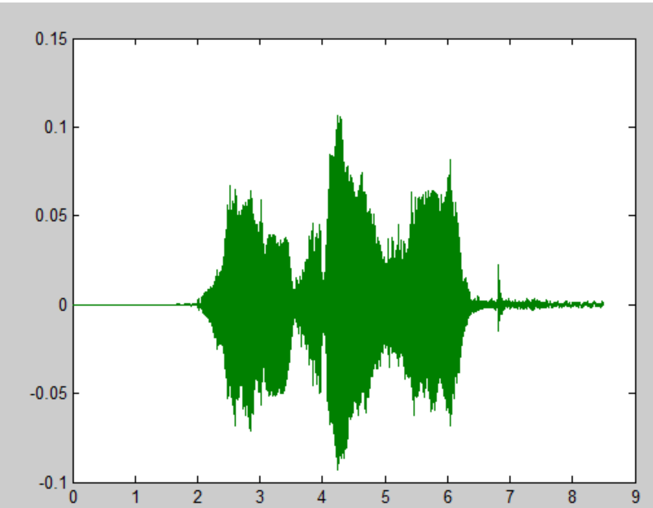
Ann training output



Training performance



Speech signal generation based on sign language



V. CONCLUSION

We consider it important that even if the application scenario (game) can be further enriched, the interaction framework and all its novel technologies was well received by all users and could be used as the basis for further development of such games or interaction systems. Concerning the applicability and future potential of the proposed system, the users were asked about their impression of communicating with other users to elicit information. Users were asked how it was to communicate with their partner and how they imagined that communication could be extended. All users said they found the advancements fostered by this proposed framework exciting. In particular, a visually-impaired user mentioned that the blind and the deaf tend to socially interact with similarly impaired people, while further development of the proposed system could open entire new possibilities for the visually impaired to have social interaction with the hearing impaired.

VI.REFERENCES

[1] J. Lumsden and S.A. Brewster, ''A Paradigm Shift: Alternative Interaction Techniques for Use with Mobile & Wearable Devices,'' Proc. 13th Ann. IBM Centers for Advanced Studies Conf., IBM Press, 2003, pp. 97-100.

[2] O. Lahav and D. Mioduser, ''Exploration of Unknown Spaces by People Who Are Blind, Using a Multisensory Virtual Environment (MVE),'' J. Special Education Technology, vol. 19, no. 3, 2004, pp. 15-24.

[3] T. Pun et al., ''Image and Video Processing for Visually Handicapped People,'' Eurasip J. Image and Video Processing, vol. 2007, article ID 25214, 2007.

[4] A. Caplier et al., ''Image and Video for Hearing Impaired People,'' Eurasip J. Image and Video Processing, vol. 2007, article ID 45641, 2007.

[5] S.C.W. Ong and S. Ranganath, ''Automatic Sign Language Analysis: A Survey and the Future beyond Lexical Meaning,'' IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no. 6, 2005, pp. 873-891.

[6] N. Bourbakis, A. Exposito, and D. Kobraki, ''MultiModal Interfaces for Interaction-Communication Between Hearing and Visually Impaired Individuals: Problems and Issues,'' Proc. 19th IEEE Int'l Conf. Tools with Artificial Intelligence, IEEE Press, 2007, pp. 522-530.

[7] D. Tzovaras et al., ''Design and Implementation of Haptic Virtual Environments for the Training of Visually Impaired,'' IEEE Trans. Neural Systems and Rehabilitation Eng., vol. 12, no. 2, 2004, pp. 266-278.

[8] H. Petrie et al., ''Universal Interfaces to Multimedia Documents,'' Proc. 4th IEEE Int'l Conf. Multimodal Interfaces, IEEE CS Press, 2002, pp. 319-324.

[9] N.O. Bernsen and L. Dybkjaer, Multimodal Usability, HumanComputer Interaction Series, Springer, 2010.

[10] O. Aran et al., ‘‘Signtutor: An Interactive System for Sign Language Tutoring,’’ IEEE MultiMedia, vol. 16, no. 1, 2009, pp. 81-93.

[11] G. Bologna et al., ‘‘Transforming 3D Coloured Pixels into Musical Instrument Notes for Vision Substitution Applications,’’ Eurasip J. Image and Video Processing, special issue on image and video processing for disability, vol. 2007, article ID 76204, 2007.

[12] M. Papadogiorgaki et al., ‘‘Gesture Synthesis from Sign Language Notation Using MPEG-4 Humanoid Animation Parameters and Inverse Kinematics,’’ Proc. 2nd Int’l Conf. Intelligent Environments, IET, 2006.

[13] G.C. Burdea, Force and Touch Feedback for Virtual Reality, Wiley-Interscience, 1996.

[14] K. Moustakas, D. Tzovaras, and M.G. Strintzis ‘‘Sq-map: Efficient Layered Collision Detection and Haptic Rendering,’’ IEEE Trans. Visualization and Computer Graphics, vol. 13, no. 1, 2007, pp. 80-93.

[15] K. Moustakas et al., ‘‘Haptic Rendering of Visual Data for the Visually Impaired,’’ IEEE MultiMedia, vol. 14, no. 1, 2007, pp. 62-72.

[16] R. Ramloll et al., ‘‘Constructing Sonified Haptic Line Graphs for the Blind Student: First Steps,’’ Proc. ACM Conf. Assistive Technologies, ACM Press, 2000.

[17] S. Young, The HTK Hidden Markov Model Toolkit: Design and Philosophy, tech. report CUED/ F-INFENG/TR152, Cambridge Univ. Engineering Dept., Sept. 1994.

[18] B. Bozkurt et al. ‘‘Improving Quality of Mbrola Synthesis for Non-Uniform Units Synthesis,’’ Proc. IEEE Workshop Speech Synthesis, IEEE Press, 2002, pp. 7-10.