Technology and Disability 20 (2008) 51–53
IOS Press

Introduction to Special Issue

Electronic speech processing for persons with disabilities

This issue of *Technology and Disability* is dedicated to the phenomenon of speech, methods of electronic speech processing and the benefits of electronic speech processing for persons with disabilities.

It is well known that speech is most important form of communication among humans. Our daily life with its countless personal and professional contacts stands and falls with speech. If speech communication is disturbed or even impossible as is the case with deafness, then these persons are in the danger of loneliness and isolation.

Fortunately, modern technology has developed systems which are able to bridge some of the gaps in communication. This does not mean that technology can replace any human sense. The only possibilities are either to support or amplify the remaining capabilities of a sense (that's what hearing aids or glasses do) or, if a sense is damaged completely to transform an information into a sensory capability which is still functioning. Applications are the transformation of written text into synthetic speech for blind people or a speech-to-text transformation for a deaf person.

In this special issue, the attempt is made to summarize the state-of-the-art in communication technology for the disabled and elderly focused on electronic speech processing. Several researchers and developers present their work and show how far we are today and also at the same time the limitations and the remaining problems which are still unresolved.

The first paper from Fellbaum and Koroupetroglou is structured as an extended introduction to the principles and applications of electronic speech processing. In the first part the principles and the state-of-the-art of speech processing, and especially speech synthesis and recognition, are explained. Then, a speech-based human-computer dialogue system is discussed. The next section gives a brief overview of the available recommendations, guidelines and standards that are directly related with the application of speech technologies. The last part of the paper is dedicated to applications of speech technology for the disabled. The main focus is on blind and partially sighted people and those with hearing loss. Concerning the blind, many multilingual text-to-speech synthesis systems exist, and some polyglot ones, that can convert printed and electronic documents to audio, but further research is needed for structured text, tables and above all graphics to be efficiently transformed into speech. For the deaf persons, there are still big challenges in the development of adequate communication aids. Special applications discussed in this paper include speech-based cursor control for those with physical disabilities, transformation of dysarthric speech into intelligible speech, voice output communication aids for the language impaired and those without speech, and accessibility options for public terminals and Automated Teller Machines (ATMs) through the incorporation of speech technologies. The paper concludes with an outlook and recommendations for research areas that need further study. For further reading see [4,5,8,9].

The second paper, written by Kvale and Warakagoda, explores how multimodal interfaces make it easier for people with sensory impairments to interact with mobile terminals such as PDAs and 3rd generation mobile phones. The authors describe the development of a flexible speech centric composite multimodal interface to a map-based information service on a mobile terminal. The user interface has proven useful for different types of disabilities, from persons with muscular atrophy combined with some minor speaking problems to a severe dyslectic and an aphasic.

The third paper from Beskow, Engwall, Granström, Nordqvist and Wik deals with the visualization of speech and audio for hearing impaired persons. For

ISSN 1055-4181/08/$17.00

them, acoustic information must be supplemented or even replaced by cues using other senses. The authors argue that the best replacement of speech is the visual sense since these two modalities are complementary. In the contribution, a number of visualization techniques are proposed and exemplified with applications for hearing impaired persons, children as well as middle-aged and elderly with the aim to allow them equal participation in communication.

The next two papers are dedicated to people who are deaf and who communicate primarily by sign language. It must be clearly stated that sign language has a similar complexity and richness of articulateness as spoken language but both are very different in their linguistic structures. The paper of Vogler and Goldenstein discusses aspects of computational sign language with a focus on automatic sign language recognition and facial expression analysis. Despite the seeming similarities between computational spoken and sign language processing, signed languages have intrinsic properties that pose some very difficult problems, among others a high level of simultaneous actions, the intersection between signs and gestures, and the complexity of modelling grammatical processes. Additional problems are posed by difficulties that computers face in extracting reliable information on the hands and the face from video images [7].

The other paper on sign language processing written by Dreuw, Stein, Deselaers, Rybach, Zahedi, Bungeroth and Ney, gives a deep insight into the problems of sign language and speech recognition. The authors describe a system which was created based on state-of-the-art techniques from statistical machine translation, speech recognition, and image processing research. As opposed to most current approaches, which focus on the recognition of isolated signs only, they present a system that recognizes complete sentences in sign language.

The next paper by Freitas and Koroupetroglou is dedicated to people who are blind. These persons obviously need a technology which transforms visual information into speech. There are many devices as well as application areas which are speech enabled, for example talking thermometers, weight scales, calculators, barcode and RFID tag readers, PDAs and so on. Today, there is in principle no problem to give these devices a voice or, with the aid of speech recognizers, a hearing capability for the execution of speech commands. But there are still serious difficulties to make these devices user-friendly, above all for blind persons. If we think of web pages, there are many graphic elements including icons, figures, video etc. which have to be 'translated' into a verbal description. In a scientific text, mathematical and chemical formulas appear and need to be expressed adequately. Using a PC, screen readers are required and the screen reader software must interpret the operating system messages to build an off-screen model, a substantially difficult task. If we think of voice browsers or voice portals, an interactive voice dialogue based on speech synthesis and recognition is probably the best and most efficient solution. Generally speaking, problems increase dramatically when the visual information is non-text. Further aspects discussed in the article are speech technologies in mobility systems, speech-based accessibility in the classroom and speech technologies in digital television.

The last paper written by Fellbaum is thought to be an outlook. It describes communication aspects in an Ambient Intelligence (AmI) environment. AmI provides a vision of the future Information Society where the emphasis is on greater user friendliness, more efficient services support, user empowerment, and support for human interactions. In an AmI environment people are surrounded by invisible, intelligent, and intuitive interfaces which act and react in a way of an attentive butler. AmI primarily aims at the private citizen and includes elderly persons as well as persons with disabilities. In order to illustrate the impact of AmI also for non-experts and in a thought-provoking way, the idea of scenarios is introduced which describe the all-day situation of people in the near future and being surrounded by AmI. Finally, examples of local and mobile applications of AmI with a focus on speech processing as well as applications for the disabled and elderly are given and commented. For further information about Ambient Intelligence see [1–3,6,10,11].

Klaus Fellbaum
Guest Editor

References

[1] E. Aarts and S. Marzano, *The New Everyday. Views on Ambient Intelligence*, Koninklijke Philips Electronics N.V. 0I0 Publishers, Rotterdam, The Netherlands, 2003.

[2] J. Ahola, Ambient Intelligence. *ERCIM News* No. 47, Oct. 2001. http://www.ercim.org/publication/Ercim_News/enw47/intro.html.

[3] P.L. Emiliani and C. Stephanidis, Universal access to ambient intelligence environments: Opportunities and challenges for people with disabilities, *IBM Systems Journal* **44**(3) (2005).

[4] K. Fellbaum, Speech Input and Output Technology – State of the Art and Selected Applications. *8th International Conference on Application of Natural Language to Information Systems*, Burg, Germany, June 2003.

[5] S. Furui, *Digital Speech Processing, Synthesis, and Recognition,* 2nd ed., New York: Marcel Dekker, 2001.
[6] ISTAG. Scenarios for Ambient Intelligence in 2010. www.cordis.lu/ist/istag.htm.
[7] K.F. Kraiss, ed., *Advanced Man-Machine Interaction,* Springer Berlin Heidelberg, New York, 2006.
[8] P.R.W. Roe, ed., *Towards an Inclusive Future,* A COST219 publication, Brussels, 2007.
[9] D. O'Shaughnessy, *Speech Communication Human and Machine,* 2nd ed., IEEE Press, 2000.
[10] Siemens Documentation: Horizons2020. A thought-provoking look at the future. http://w3.siemens.de/horizons2020/pdf/inhalt/presse/Horizons2020_szenarioreport_e.pdf.
[11] W. Weber, J.M. Rabaey and E. Aarts, eds, *Ambient Intelligence,* Springer: Berlin, Heidelberg, New York, 2005.