Human-Computer Interaction -- INTERACT'03
M. Rauterberg et al. (Eds.)
Published by IOS Press, (c) IFIP, 2003, pp. 825-828

The Ultrasound Image of the Tongue Surface as Input for Man/Machine Interface

Konrad Lukaszewicz

Institute of Biocybernetics and Biomedical Engineering PAS, ul. Ks. Trojdena 4, Warsaw, POLAND

konrad@ibib.waw.pl

Abstract: The paper presents a new concept of computer user interface dedicated for the disabled people. The concept is based on the recognition of ultrasound image of the selected region of tongue surface. Several ultrasound images of tongue surface has been stored and analyzed from the point of view to steering usability. The representative regions has been selected and the tongue surface movement boundary has been determined. As the input standard medical ultrasound sensor has been used. The digital sequence of image frames has been stored in the computer. The position of tongue surface has been determined by the distance between the main echo and the ultra sound sensor. The tongue position signals will be used to speech synthesis control.

Keywords: speech synthesis, ultrasound image, user interface, speech prosthesis, rehabilitation

1 Introduction

A typical computer user interface consists of a keyboard and a mouse as pointing device. However in many cases this solution can not be used by disabled people. For example, a blind user can not control the mouse without seeing the mouse cursor, while the people with ALS (whose hardly control their hands movement) can not even use standard keyboard. In this case the traditional input device has to be replaced by a special solution.

Another situation exists when the disabled people have problems in communication with other ones. For example the people after tracheotomy can not speak because the sound source (vocal cords) has been removed. The existing prosthesis available on the market gives poor speech quality. Better result can be obtained when an artificial speech is used. However, in this solution the user interface has to give a possibility to control the speech unit. In this case it is not the computer user interface but a man/machine interface which makes possible a fast and comfortable control of the speech. Described below is a new concept of freehand user interface which could be applied to help the people after tracheotomy to drive the speech unit and also could be used by the people which can not use their hands to work with a computer.

2 Artificial speech as a kind of prosthesis for mute people

The speech is the most natural way for conversation. The disabled people who can not speak can still provide conversation using the text-to-speech synthesis (TTS). The speech synthesizers available today are of so high quality that the artificial speech can be used for oral communication. The new methods of speech synthesis give possibility to obtain synthetic speech similar to a particular man voice. It is important when the synthetic speech is used for helping the disable people who loose his natural voice.

Using the TTS system as prosthesis for mute people the text for utterance has to be prepared very fast. In normal typical situation we can pronounce over 100 words per minute. When the speech synthesizer is used as the rehabilitation device the user interface should allow the user to prepare tens of words in few seconds. Because of that the typical keyboard without any special software can not be used. There are many different solutions which speed up the process of preparing the text (Sag 1999). Most of them use the predictive functions to help the user preparing the utterance. But also some of them (Wloskowicz, 2001) use limited dictionary with ready to use sentences. All of proposed methods used more or less standard keyboards but those solutions are not comfortable and are not hands free so they are difficult to use outside building.

3 The special computer input devices for disabled people

Many different solutions have been designed to allow the people with disabilities to use the computer. The special devices are replacing the standard keyboard or mouse.

There are many solutions for keyboards. Starting from different shape keyboards which consist of three separate units where each of them could be placed in different position satisfying the special needs of disabled user and ending on the touch sensing keyboard where the size of keyboard is so small that the user can use small stick placed at his mouth or between his fingers to easily press the keys pads just by touching it.

Also the computer mouse can be replaced by free hand devices monitoring the head movement or keep track of the eye movement and eye blinking. Some of them are supported by different switches which could simulate the left and right mouse button. Those switches could recognize any kind of action like suck or puff for example. When the switches can not be used special software can simulate the function of the left and right mouse button. Thanks to this software and the software which simulate keyboard on the screen the disabled people who can control only the mouse pointer movement can write a text. In some cases computers with this kind of devices establish only one way for disabled people to communicate with other people.

4 Ultrasound visualization

Looking for new possibility of steering the speech prosthesis we have paid attention for the method which could be used in device allowing the mute people to use the synthetic speech in normal oral communication in the natural people environment without engaging his hands. The devices should be also very small and easy to carry.

One muscle which is always under control and naturally used while we speak is the tongue. Because the tongue is one of the most important articulator in speech process using it for driving the speech prosthesis seems to be very natural. The ultrasound pictures obtained from the bottom part of the chin can be one of possibilities to keep track of the tongue movement (Yusuf 1998). Our experiment has been provided to prove if it is possible to make simple ultrasound device which will extract information from the simple ultrasound pictures of tongue position. If this information could be used directly to control the speech then the training duration should not be very long and difficult.

The observation of the tongue surface movement has been made using standard ultrasound equipment and the ultrasound sensor has been placed under the chin, Figure 1.

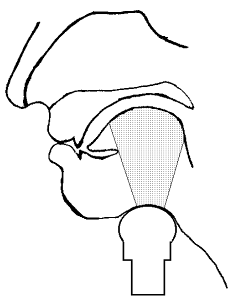

Figure 1: Ultrasound sensor has been positioned under chin. The ultrasound signal has been generated along the tongue muscle. Only a part of tongue surface has been seen which was reached through the tissue.

The ultrasound acoustic waves reflected off the tongue surface have been seen as bright regions on the gray scale images. Only a part of tongue surface has been seen on the ultrasound image which could be reach by the signal directly through tissue without any gap and the surface position was not parallel to signal direction, Figure 1. The sequences of images have been stored on the video tape and imported to the computer. Pictures of different tongue position have been selected manually.

The selected static images have been analyzed by the density histogram. Only the distance from the sensor to the brightest region on the selected direction has been considered.

5 Ultrasound images of the tongue

Below few examples of tongue positions are presented, figures 2 - 9. The teeth are positioned on the left side and the throat on the right side of each picture.

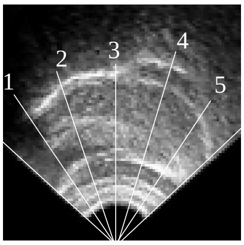

Figure 2: Tongue placed at rest position touching the hard palate. Mouth is closed.

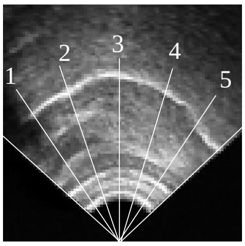

Figure 3: Tongue placed at rest position. Mouth is slightly opened.

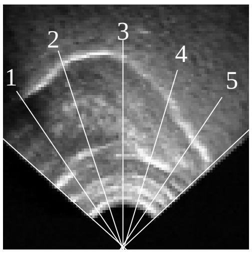

Figure 4: Tongue on the front up position. Mouths is slightly opened.

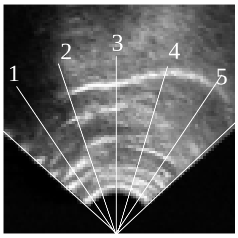

Figure 5: Tongue on slightly back position.

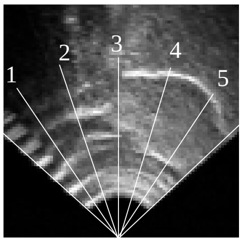

Figure 6: Tongue on the back position.

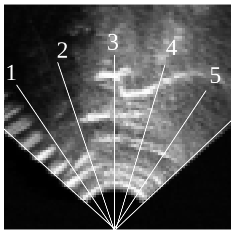

Figure 7: Tongue on the back position, tongue end is raised up.

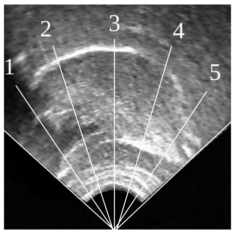

Figure 8: Tongue on the front position, back part of the tongue surface is touching the plate.

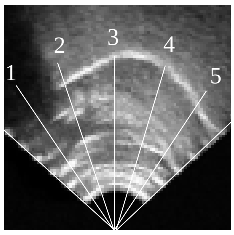

Figure 9: Tongue on the up position near the plate.

6 Determination of the tongue surface positions

To simplify the process of detection position of the tongue surface five directions have been defined. The value of the gray level has been considered only along one of lines (1-5) showed on the Figures (3-9). The line has been placed every 15° in the angle of ultrasound image. Each image has been described by one vector of 5 distances. If along desired a line the bright region has not been detected the 0 value has been placed for this direction.

The Table 1 below contains the distances (in cm) for presented 8 different positions of the tongue.

Each of the 8 tongue positions can be easily detected. Only two of them (Fig. 3 and 8) have the zero values on the same positions, but in this case the difference of distance in the direction 2 is 1cm.

Fig.	Direct. 1	Direct. 2	Direct. 3	Direct. 4	Direct. 5
1	0,0	8,2	8,4	9,2	0,0
2	7,4	7,6	8,0	7,6	6,4
3	0,0	9,0	8,9	7,2	6,2
4	0,0	0,0	7,4	8,4	8,8
5	0,0	0,0	0,0	8,6	8,4
6	0,0	0,0	8,0	7,8	0,0
7	0,0	9,2	9,2	0,0	0,0
8	0,0	8,0	8,8	8,8	7,4

Table 1: Distances for presented positions of the tongue.

7 Conclusion

The information which could be obtained from the position of tongue surface is not sufficient for direct steering the speech on the level of phonemes, but it could be used to select ready made utterance for simple oral conversation or to be used to control different devices like wheelchair. With special software drivers the ultrasound image of the tongue surface can be used by the disabled people to work with the computer simulating the mouse pointing device.

Simply ultrasound system can be built for monitoring the tongue movement. The system could consist of few ultrasound transmitters placed in collar which will touch the bottom part of the chin, Figure 10. The ultrasound one-direction fixed transmitters will take the information from 5 - 6 determined directions. The disabled people moving the tongue to defined region should keep it in this position until the detection will be made. This system will recognize few different static tongue positions and use it as a set of switches for the man/machine interface. However the tongue surface position can be also used dynamically to move (for example) the mouse pointer. Then one of the defined static tongue positions will select one direction (for example along the line 3) to move the mouse cursor up and down or left and right.

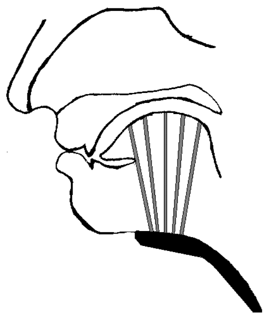

Figure 10: The ultrasound collar placed under chin.

References

Sag, I., H. Clark, A. Copestake, D. Flickinger, R. Malouf and J. Carroll (1999) `Natural language generation for a speech prosthesis'. Poster at *National Science Foundation Human-Computer Interaction Program Grantees' Workshop*, Orlando, Florida.

Wloskowicz, D. & Lukaszewicz, K. (2001), Phone Communication by Means of Synthetic Speech for Deaf-Mute People – ESEM 2001 Proceedings of the Sixth Biennial Conference of the European Society for Engineering and Medicine, Belfast, 3-5.-5.2001 Technology And Health Care, Vol. 0, No 1,2 pp. 58-60.

Yusuf, A. Kambhamettu, C. & Stone, M. Extraction and Tracking of The Tongue Surface from Ultrasound Image Sequences, *IEEE Conference on Computer Vision and Pattern Recognition*, Santa Barbara, CA, June 23-25 1998.