Information Technologies Applications for Sign Languages Investigations

Iurii Kryvonos
V.M.Glushkov Cybernetics Institute
Kyiv, Ukraine
e-mail: aik@incyb.kiev.ua

Iurii Krak
Taras Shevchenko National University
Kyiv, Ukraine
e-mail: krak@unicyb.kiev.ua

Waldemar Wojcik
Lublin University of Techology
Lublin, Poland
e-mail: waldemar.wojcik@pollub.pl

ABSTRACT
In the paper the problems of information technologies creation for sign language modeling and investigation are considered. An analysis of sign language information transfer which includes human hands, body, fingers movements, change of mimicry and emotions on human face is brought. Conception is developed and new information technology is proposed for sign language modeling on the base of human spatial model. For the transmission of movements of the real human-informant of sign language on a spatial model technology of motion capture is used. For dactyl alphabet modeling technology that uses a three-dimensional model of a hand based on an informational-parametric model has been developed. Efficiency of the developed information technology is shown on realization of Ukrainian sign language. The proposed approach carries a universal character and can be used for modeling of other sign languages

Keywords
Information technology, deaf people, design, fingerspelling alphabet, sign language

1. INTRODUCTION
Modern progress in computing and creation of new methods for data representation, storage and organization makes it possible to innovate and create new technologies for sign language learning [1]. As marked in [2]: "…Sign language interpretation services should also be provided to facilitate the communication between deaf persons and others…". From the practical point of view, teaching systems of sign language that use 3D human model are very promising. Since the information is transmitted by mean of arms movements, mimics and articulation it is necessary to research the process of construction of a sign language sentence as well as the synthesis of the elements in order to get a good understanding of the subject. The problem of description of human movements is complex enough with a high percentage of fuzzy knowledge about human body and its physiology [3]. Thus, it is important to conduct a research of the process of how the gestures are formed from the viewpoint of formalization for the problem of modeling using 3D model and for the problems of gestures analysis and synthesis.

The sign language is the primary mean for communication for deaf people, has national traits (e.g. English and French [4,5], Russian [6], Polish, Ukrainian [1], etc.). And there are two sign languages, different in grammar and sets of signs, used by the deaf:

a common sign language, which is typically used in an everyday communication; its grammar differs largely from the spoken language;

a calculative sign language, which is used for official communication; it has some aspects of the common one as well as fingerspelling; it does not have its own grammar and follows the natural spoken language.

Comprehension of a natural language by watching lips, "lipsreading", is an important skill for a deaf, because common people generally do not know the sign language. With regard to this fact, the solution capable of reading the text being pronounced by watching lips can be regarded as an alternative technology for facilitation of communication with people having difficulties in hearing or seeing. Besides, the visual-based speech recognition is an additional independent source of information for the problem of speech recognition itself, and can be used for improvement. The synthesized equivalent of the text received by lips reading will facilitate the communication for visually impaired people.

In the given paper authors argue on the necessity to create a complex informational technology aimed to facilitate the non-verbal communication between common people and the deaf, people having hard difficulties in hearing or having damaged hearing. It should be noted that on the basis of the offered technology can be realized.

Therefore, the following problems statement has been formed: to create computer-aided informational technologies for communication with deaf people, with intent for implementation, that would provide the following features:

1) creation of a system for sign language fingerspelling units (dactyls) modeling based on 3D model of human palm;
2) gesture synthesis for the common sign language and calculative sign language on a 3D human model;
3) representation on a spatial (3D) human model of the pronunciations process with regard to emotional components;
4) lipsreading analysis and modeling.

2. 3D HUMAN HAND MODELING AND ANIMATION OF FINGERSPELLING PROCESS
To teach fingerspelling, technology that uses a three-dimensional model of a hand based on a informational-parametric model has been developed. The technology allows observing hand from different viewpoint during the learning process, show sequence of letters etc. The main window is shown on Fig. 1.

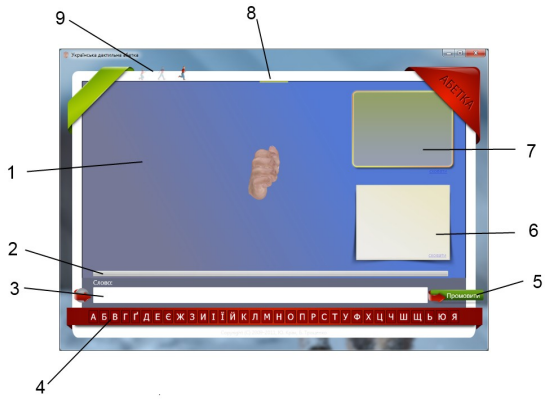

Figure 1. The main window of a program for fingerspelling alphabet modeling.

At this point the numbers mean: 1 – area of displaying fingerspelling alphabet; 2 – panel of displaying playback progress of letters or words; 3 – input panel for words; 4 – list of letters; 5 – button «spell», the process of fingerspelling of input word begins when the button is clicked; 6 – panel to display the verbal description of a hand configuration that corresponds to the currently displayed letter; 7 – panel to display a written letter and a picture that correspond to the currently displayed letter; 8 – indicator of a location of a hand rotation; 9 – defines the pace of fingerspelling.

The main features of the program:

- changing of a view angle. The use of a three-dimensional modeling enables the possibility to examine the hand model from different viewpoints. That would be impossible using video materials. The range of changing an angle of viewpoint vary to 80° right/left;
- presence of pictures which are associated with particular letter for whole alphabet (see Fig. 2, for example) (this panel can be hidden and shown back);
- verbal description of hand configuration which is shown;

Figure 2. Examples of fingerspelling letters “O” and “X”, respectively (for Ukrainian letters).

- presentation of dactyls is performed by selecting a particular letter from a list using the mouse or by pressing the letter-button on a keyboard. If user wants to repeat, he presses space. This feature allows implementing interactive learning process, when right arm (trained) is in the free position and the left one realizes interactions with the program;
- changing the pace of the animation. Three pace modes (slow, medium, and fast) are implemented in the program for different needs of learning process (repetition after the model, recognition of the foregoing, etc.);
- fingerspelling of a word. This feature allows entering words into the input panel and observing the process of fingerspelling of a word. That allows not only learn separate letters but also learn how to spell whole words;
- verification mode. The program has a feature to “hide” the panel of the verbal description of a hand configuration as well as the panel with a written letter and a picture. That allows conducting examination of knowledge displayed letter (hand configuration)(Fig.3).

Figure 3. Model and real letter of fingerspelling alphabet.

Based on this technology, training programs for any one-handed fingerspelling alphabet can be created. There are currently developed programs for Ukrainian, Russian, Polish, Azerbaijan and American fingerspelling alphabets.

3. INFORMATION TECHNOLOGY FOR NON-VERBAL COMMUNICATION FOR/WITH DEAF PEOPLE

The general approach for sign language using communication with deaf people is shown on Fig. 4.

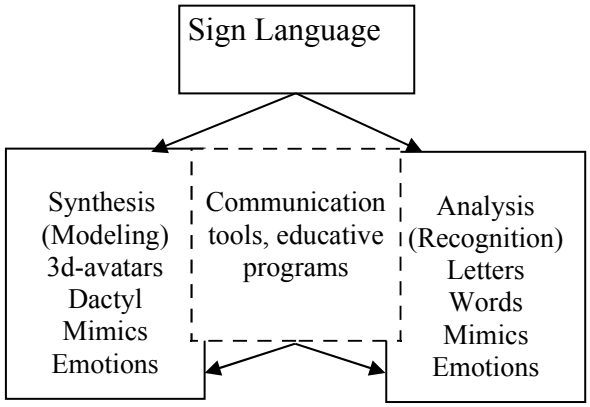

Figure 4. General scheme of the concept for communication with deaf people.

The complex information technology will provide the following features:

- a module for translation of the normal text into the sign language (text-to-gesture); the module will provide pronunciation animation of common and official sign languages by presenting the output on a 3D human model;
- mimics and animation (with regard to emotional components) during the pronunciation process;
- lipsreading module for recognition of the text being pronounced.

For the implementation of the suggested concept of computer-aided non-verbal communication, a series of research works has been made and the appropriate software has been developed. For the 3D sign language animation synthesis, the geometrical classes of vector-based gestures are described. These classes were formed using a motion capture technology [7]. Motion capture is a technology for retrieving real-world 3D coordinates using multiple video streams recorded from different viewpoints. Then the coordinates are used to determine values in the mathematical model. The key frames are determined by using a tracking technology [8].

For the storage of a gesture, the BVH file format was used. It allows the gesture to be applied on a virtual human (e.g. using Character Studio module in 3D studio MAX or using Poser software). The suggested implementation of motion capture technology contains the following stages:

1) A person showing a particular gesture is recorded full-face, from the left and right side views;

2) The video streams are processed: arms coordinates are detected and the motion is tracked;

3) Based on the position of arms obtained on the previous stage, the BVH is formed for further synthesis of 3D animation;

4) The BVH is applied on a virtual human for creation of the animation process (using Character Studio in 3D studio MAX or Poser).

For the input text preprocessing, the appropriate informational technology was created, which considers the stress location for each word, specifies its normalized word form; contains synonyms and idioms. The model is represented as a set of tables in a relational database along with a set of stored procedures which implement all the required functionality. For the implementation of visualization and pronunciation feature of a custom text, the appropriate synthesizer has been created. It allows creating the voice equivalent of a custom text using different voices and voice characteristics (volume, distance). The synthesizer allows to visualize the process of pronunciation by showing 2-dimensional views as well as 3D ones.

For the complex verification of the suggested technology the appropriate software has been created (Fig. 5). It is used for translation of a custom text into the calculative sign language.

Figure 5. Computer system for sign language modeling and learning.

The software uses the following algorithm for the sign language synthesis: 1) a speech equivalent is synthesized for the input text; 2) the input text is parsed into words; 3) a speech equivalent is synthesized for the input text; 4) the input text is parsed into words; 5) for each word its normalized form (infinitive) is found by performing a look-up in the database; 6) for each normalized word form a gesture is looked up (represented as a sequence of movements); 7) in case the gesture is not found, the word will be shown using the dactyl alphabet.

The 3D model displays the gesture accompanied by the speech synthesized.
Further studies were aimed at the 3D model improving. The real sign language interpreter was taken as a basis for building models (see Fig.6)

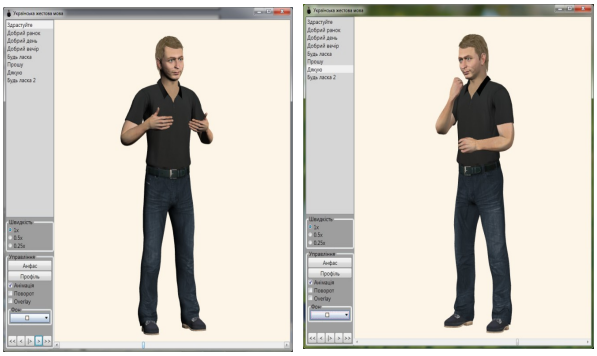

Figure 6. 3D model of real sign language interpreter.

4. CONCLUSIONS

The work suggests a complex of informational technologies for non-verbal communication with deaf people using sign language. Further development will be targeted at supporting the full set of signs in the Ukrainian sign language. Great attention will be paid to the consideration on the use of the created systems for modeling Polish, German, Kazakh and other sign languages.

REFERENCES

[1] Iu.G.Kryvonos, Iu.V.Krak, O.V.Barmak, ets. "Information technology for Ukrainian sign language modeling", Artificial Intelligence, No 3, pp. 186-197, 2009. (on Ukrainian).

[2] Standard Rules on the Equalization of Opportunities for Persons with Disabilities. Appendix N 49 (A/48/49), article 14, pp. 292-306. Document 995_306, edited on 20th Dec 1993. http://www.un.org/documents/ga/res/48/a48r096.htm.

[3] William C. Stokoe, Jr. Sign language structure: An outline of the visual communication systems of the american deaf. – Univ. of Buffalo, 1960.

[4] C. Neidle, S. Sclaroff, V. Athitsos. SignStream™: A Tool for Linguistic and Computer Vision Research on Visual-Gestural Language Data, Boston University, Boston, Massachusetts, Vol.33, No.3, pp.311-320, 2001

[5] IBM Research Demonstrates Innovative Speech to Sign Language Translation System, Extreme Blue programme HURSLEY, UK - 12 Sep 2007, IBM Media Relations http://www-03.ibm.com/press/us/en/pressrelease/22316.wss

[6] A.L. Voskresenskii, G.K. Khakhalin. Multimediinyi tolkovyi slovar russkogo gestovogo yazuka. Komputernaya lindvistika I Intellektualny tehnologii: Trudy mesgdunarodnoi konferencii «Dialog 2007» . - M.: Izd-vo RGGU, 2007. p.658. (in Russian).

[7] A.Menache. Understanding Motion Capture for Computer Animation and Video Games, Morgan Kaufmann, 2000.

[8] S. Avidan, "Support vector tracking" / Proc. IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii, 8-14 December, vol. I, pp. 84–191, 2001