System of Finger Movement Identification for Sign Language Recognition

The sign language (language of people with hearing disabilities), as a special communication system used by specific groups of people in definite situations and problem areas, is an interesting and promising object for linguistics.

In many situations and cases the sign language is the only possible opportunity of communication realization. Until recently the sign language was used only in the context of human communication. But now with the development and wide implementation of computer information technology, the issue of translation from signs (gestures) to regular text language following its transformation into sound form without a human translator became a point of active research interest.

Our research is devoted to creation of program and technological computerized application that would allow significantly improve the situation concerning sign language, especially in the cases when no alternative communication is available. The situation when one of the communication sides does not know the sign language, but the communication is to be effectively conducted.

The characteristic features of sign language communication from the process and technology point of view are: its social direction and social meaning; technical and technological convenience and ease of use; problematic direction of application areas.

Each of the listed peculiarities allows to strictly define problematic communication situations, which can be served by this specific linguistic application. We should say that this kind of task can be solved only by using wide variety of methods, tools and technology from different sciences and arts. This research can be called a boundary one, and it's conducted by the team of linguists, teachers, mathematicians and computer scientists.

This paper gives the results of the first phase of linguistic-computing project on computer support of communication when one of the sides is a person who can use only the sign language. The paper looks into one of the scenario plans of usage of the sign language, in particular use of modern multimedia technology when the management of presentation is carried out by the linguistic analysis of the speaker's sign language.

We would like to offer the results of research carried out to create a prototype of the interactive presentation control system. A speaker can apply a wide spectrum of facilities of accompaniment of the presentation, including multimedia projector, overheads, slides projector, control system of room illumination, window shades, microphones, facilities of recording and presentation of information: by a dictaphone, tape recorder, video camera, digital camera and others. An operator can manage all these devices from a central stand. In such a case the presentation must take place in the specially equipped room with a complete set of indicated equipment, and operator is to have the scenario of presentation.

In the offered system the simplified chart of which is shown on Fig.1, necessity in this equipment is absent. The system can be collected from the present devices in free configuration with the minimum requirements to the cost and compatibility of their work, in particular, we do not necessarily need a central stand and specially trained operator. All devices are connected to a central server which manages their work. A speaker is constantly recorded by a video camera, a computer processes video signal, selects the gestures of speaker and passes a managing signal to the proper device. A speaker, screen, video camera, hardware and software complex and devices of presentation of material (multimedia projector, overheads, slides projector) are the constituents of this system.

We will consider the functions of constituents of the offered system.

1. A speaker comes forward and manages the complex set of devices by the certain set of gestures.

2. Presentation materials are projected on a screen. It is also possible to project gestures with the purpose of recognition of their projections. By other method the management of the

devices can be done by a set of marks which are projected on a screen during the presentation. In such case management of the devices is possible to be executed not by a set of gestures, but touch of finger, pointer or light mark on the proper section of screen.

3. A video camera records the speaker and the screen and passes the received image for processing by a hardware and software complex on the server.

4. The devices of materials presentation pass these materials on a screen.

5. A hardware and software complex consists of server and software installed on it. A hardware and software complex receives a signal from a video camera, works it over in the real time and passes managing signals on the devices connected to it.
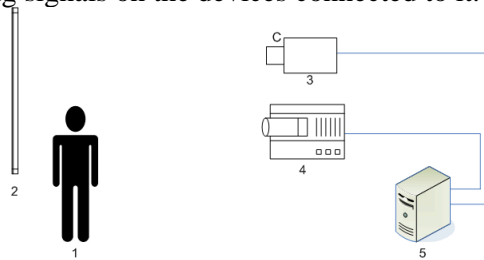

Fig. 1. The structure of the hardware and software complex

Basic tasks that are solved during creation of the hardware and software complex are related to processing of the image of hand of the speaker, recognition the ends of fingers on this image and classification of gestures. The images which come from a video camera and pass the number of stages of processing are the data for this task. The task of recognition of ends of fingers on the image is formulated as a task of classification.

The system of recognition of gestures consists of two basic parts. First of them is recognition of the ends of fingers on the image and selection them among the other elements, second - actually classification of the recognized combinations of fingers which are given the according words or letters of Ukrainian alphabet.

Basic tasks which are solved during creation of the system are related to the processing of the image and recognition of a gesture. It is carried out by the selection on the image the position of fingers of a hand and construction of a gestures classifier. The task of recognition of fingers configurations and gestures on the image is solved as a task of classification.

Construction of the translation system of sing language to a text consists of such stages:

1) forming of basic set of images from the videorow in the real time;

2) previous image processing with the purpose of removing noise on the image, that appears as a result of compression or low-quality digitalization of the image;

3) reducing the number of information dimensions by the removing of the constituents of image, which do not influence the process of recognition;

4) classification of elements of the videorow.

Training examples and standards were created from the shots of the videorow. The first step in their creation involved previous procession of videoshots for reducing of volume of information. The hands, images of which subject to processing, are in a constant motion. For the functioning of the system of recognition it is not necessary to fix hands in the special static positions.

Fig 2. Examples of the recognized gestures

The offered prototype of the system of recognition of the sing language does not require specialized video equipment. For the functioning of the system of recognition it is enough to use a web camera with frequency of 30 shots per second with resolution of 320x200 pixels. Time of procession of one shot on a computer with the processor of Celeron 1,2 Mhz is 0,15c. On Fig. 2 the examples of gestures recognition results are shown.

Used literature

1. Белозерский Л.А. Основы построения систем распознавания образов. – Донецкий Государственный институт искусственного интеллекта, 1997. 2. Русин Б.П. Системи синтезу, обробки та розпізнавання складноструктурованих зображень. – Львів: Вертикаль, 1997. – 264 с. 3. Форсайт Д.А., Понс Ж. Компьютерное зрение, современный подход. – М.: Издательский дом Вильямс", 2004. – 928с. 4. Mitchell T. Machine Learning. McGraw-Hill Companies, Inc. 1997. 5. Горбань А.Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей. – Сибирский журнал вычислительной математики, 1998, Т.1, №1. С.12-24. 6. Круглов В. В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – М.: Горячая линия – Телеком, 2001. – 382 с. 7. Fravolini M. L., Campa G., Napolitano M., La Cava M. Comparison of Different Growing Radial Basis Functions Algorithms for Control Systems Applications. 2000. http://www2.cemr.wvu.edu/~gcampa/pdf/acc02f8.pdf. 8. Bishop C. Neural Networks for Pattern Recognition. Oxford University Press. 1995. 9. LeCun Y., Matan O., Boser B., Denker J.S., Henderson D., Howard R.E., Hubbard W., Jackel L.D., Baird H.S. Handwritten Zipcode Recognition With Multilayer Networks, Proc. of International Conference on Pattern Recognition, Atlantic City, 1990. 10. Rowley H.A. Baluja S., Kanade T. Rotation invariant neural-network based face detection. – Proceedings, Computer Vision and Pattern Recognition, 1998. 11. Мисюрев А. В. Практика використання штучних нейронних мереж для розпізнавання рукодрукованих символів. http://www.ocrai.narod.ru/. 12. Пшеничный Б.Н., Данилин Ю.М. Численные методы в экстремальных задачах. – М.: Наука, 1975. 13. Jain A.K., Jianchang Mao, Mohiuddin K.M. Artificial Neural Networks: A Tutorial. Computer. Vol.29, No.3, March 1996, pp. 31-44.