The visualization system of patrol squad coordinates with a voice user interface

M S Medvedev and S V Schukov
Institute of Space and Information Technology, Siberian Federal University, Kirenskogo 26, Krasnoyarsk, 660074, Russia

E-mail: 2149972@mail.ru

Abstract. The relevance of the topic is down to the need of developing a remote control system for operational groups of security companies with support for radio data speech recognition and location visualization. The developed system solves the problem of displaying information about the sectors of the orders on a map, which allows the operator (dispatcher) to visually monitor the situation in the given sector and ensures quick change of operators, eliminating errors under the influence of the human factor. For the convenience of managing the system, the voice command input subsystem was implemented and integrated.

1. Introduction

One of the factors of the stable work of enterprises engaged in the protection of real estate is the control of the location and the possibility of moving mobile transport orders. GPS/GLONASS systems are most commonly used to control the movement of vehicles. The location of the vehicles is received by the dispatcher who is monitoring. He, in turn, determines the sectors of the location of mobile outfits necessary to achieve the minimum time of arrival to the object of observation. The number of combinations of such sectors can reach several dozen, which in turn leads to actual problems, such as: increase of the terms of training the dispatcher and increase of the number of errors under the influence of the human factor.

And as a result, it is reflected in the quality of the security services provided by the enterprise. Therefore, there is an urgent problem in the need to expand the existing systems of remote control of transport by the radio data speech recognition subsystem on the coordinates of mobile groups and display the operational information on an electronic map. Such a subsystem will solve the problem of displaying information from all sectors on a common map with the required scale.

2. Analysis of existing systems

Today, there are several systems for remote online monitoring of transport, such as "Almaz" and "BARS-GLONASS" satellite systems for monitoring the transport.

"Almaz" system of satellite navigation and security control solves several problems at once, such as: determining the location of the nearest car; display of vehicles on an electronic map of the area.

“BARS-GLONASS” system is intended for satellite monitoring of mobile objects. With the help of the equipment installed at the facility, the system ensures its continuous tracking and displays information about its movement on the screen of the user's electronic device.

The considered systems do well with common tasks, such as taking into account the length of a route by a vehicle, fuel consumption, the location of a moving object, etc. However, developers do not complement the systems on the market, taking into account the employment specifications of the

client's enterprise. This led to the creation of a system that would allow to recognize speech commands about the coordinates of mobile groups and display operational information on an electronic map. This system is aimed at functioning as an addition to "Almaz" software.

3. System description

A schematic mapping of the developed system "Recognition of voice data on the coordinates of mobile groups and visualization of operational information on an electronic map" is shown in figure 1.

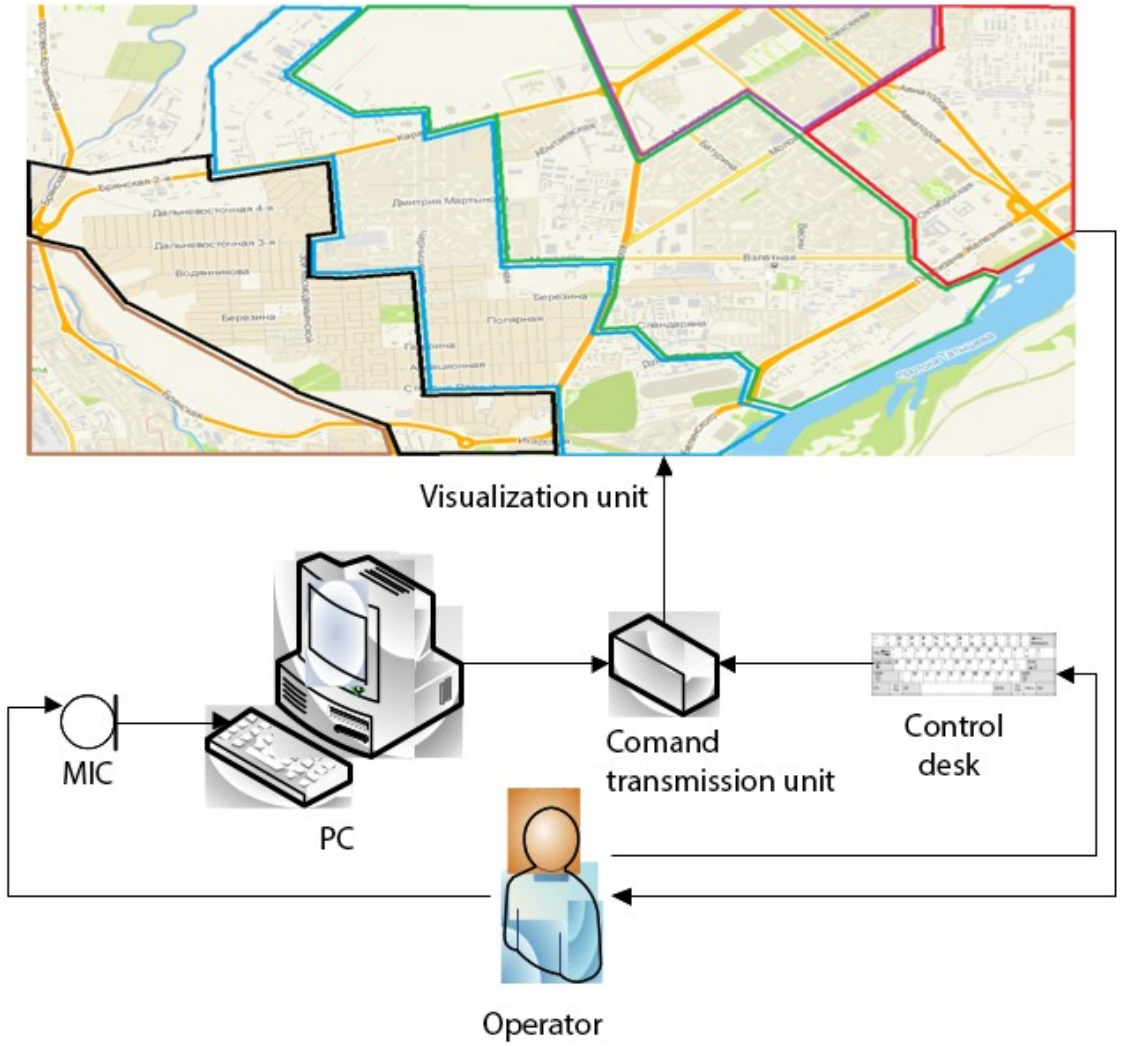

Figure 1. The visualization system of patrol squad coordinates with a voice user interface.

4. Voice-to-control module

Currently on the market of speech recognition systems, there are many solutions with high performance on the quality of recognition. However, most of them are presented in the form of online services, which limits their applicability to the tasks of security companies due to the confidentiality of the transmitted data. An important criterion of the speech module is also the presence of a data exchange protocol for integration with the visualization system used. Therefore, to recognize the speech commands of the dispatcher, a module for converting a speech signal into control commands was independently developed.

As a classifier of features of speech signals, a two-layer direct distribution neural network (a two-layer perceptron) was chosen, the weights of whose connections are formed in the learning process [1]. The input parameters of the recognizable speech signal are successively input to the network, i.e. separate command word. The number of neurons in the input layer is determined by the number of

components of the input feature vector. The number of neurons on the hidden layer is equal to the number of neurons in the input layer and equal to 500. The size of the output layer of the network depends on the number of classes of recognizable commands and changes dynamically. Sigmoid activation function is used.

To extract the parameters of speech commands, a method based on the wavelet transform was used [2].

The recognition quality of a developed speech signal conversion module into a control command depends on the quality of the training sample formation, the reliability of neural network command classification, as well as the chosen method for extracting command features. Experiments were conducted on the recognition of commands recorded by different speakers.

For the command speech signal digitized with a sampling frequency of 22050 Hz, the length and amplitude were normalized, then it was divided into frames. For each frame, the wavelet decomposition into 10 levels was applied, then the total energy value of the detail coefficients of these levels was calculated. Thus, vectors consisting of 500 elements, which were used as a training set, were formed. In the same way as the training set, vectors of parameters of the teams that did not participate in the network training were prepared. These examples were sequentially fed to the input of the neural network for classification. According to the results obtained in the course of the experiment, the recognition factor was determined for groups of commands from various speakers. The results of the experiments obtained on a training sample of 600 standards showed high value of the coefficient of correctly recognized words, which was 93%.

5. Executive transmission module

The array of executive commands is packaged in a matrix, which reduces the number of additional elements at the stage of hardware design. The rows of the matrix are alternately transmitted to the rendering unit via the SPI interface. Each line contains execution commands. In the first column of the row there is information about the line of the matrix, to which we transmit the value of the voltage level, in the second and subsequent ones - the corresponding values at a given moment in time to each element.

During the experiments, it turned out that the transfer of one matrix row takes 0.64 ms. which is seen from the oscilloscope readings (figure 2).

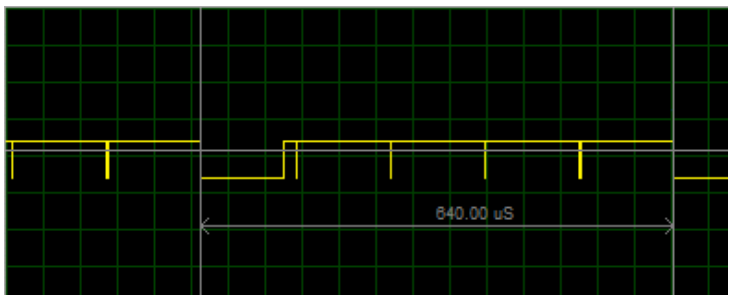

Figure 2. Control command signal.

That is, for a serial transfer of the entire byte with the data of the voltage level on all eight lines, 40.96 ms will be required, which is 24.4 Hz. With such signal, LED blinks will be visible. To prevent this, it is necessary to increase the transmission frequency to 50 Hz. In order to solve this emerging problem, such load bytes are selected for the voltage level values, in which the first four bits are respectively equal to the second four. However, this reduced the number of possible simultaneously displayed sectors to 125, which thus gives a substantial reserve of this system.

6. Operational information visualization unit

The reception of packets of executive commands and the output of control pulses are carried out by 74HC595 eight-bit shift registers connected in series. The control pulses after processing at the hardware level are transmitted to the information output unit, implemented in the form of a stretched canvas of a card with LED illumination from the inside. Visualization takes place by including the desired segments of LED strips.

The developed system requires a microcontroller, the capacity of the internal memory of which is capable of preserving the software implementation of control vector processing algorithms and the formation of output signal vectors. Requires SPI and USART interfaces; the first, to transfer the vector of the output signal to the unit of visualization of operational information; second, to receive data from the speech signal conversion module into the control command.

The widespread microcontroller from Atmel Atmega16 [3-4] has satisfactory characteristics.

7. Software implementation

The speech signal conversion module in the control command is implemented as an application for Windows OS. This module is an addition to the hardware modules of the developed system.

The following functions are implemented in the software model of the speech-to-text conversion module: training sample formation, neural network training for command classification, saving and loading the training sample in the form of a connected database of wavelet signs of commands, exporting and importing the command table as a connected database, saving and importing the trained neural network The program is controlled by the user through a graphical interface created in the Matlab GUIDE user interface development environment.

8. Conclusion

The developed system of visualization of patrol squad coordinates with voice interface support will allow to display overlapping mobile groups on sectors quickly, which will reduce the dispatcher's error when relocating moving crews. The display on the map takes place automatically, by means of recognition of the voice commands of the dispatcher, or by using a special 16-button keypad. During the implementation of the system, the possibility of applying the wavelet transform as a method for extracting features of speech commands and neural networks for their classification was investigated. Tests that showed a high rate of recognition and conversion of the speech signal in the control command were carried out.

The processing and visualization system is implemented at the hardware level. During its implementation, the task of managing electrical networks with high voltage and current through a combination of electronic components was solved, which significantly reduced: the potential costs of acquiring them; the total number of electronic components; PCB area when designing it.

References

[1] Mirkes E M 2002 Neuroinformatics (Krasnoyarsk)

[2] Medvedev M S and Dashkevich I V 2016 Using the wavelet transform in the speaker's voice identification task nnovations in the field of technical sciences *Proceedings of international. scientific-practical conference Tyumen, 2016* pp 67-70

[3] Prokopenko V S 2012 *Programming ATMEL microcontrollers in C language* (Kiev: MK-Press)

[4] Transistor switch 2014 (Admin) Available from http://www.doneathome.ru/archives/1152