NEWS
OF THENATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 1, Number 335 (2021), 54 – 59 **https://doi.org/10.32014/2021.2518-1726.8**

UDC 44.01.11
IRSTI 004.5

Y.S. Nurakhov, A.E. Kami

Al-Farabi Kazakh National University, Almaty, Kazakhstan.
E-mail: y.nurakhov@gmail.com, qamiaiko@gmail.com

INFORMATION SYSTEM FOR PEOPLE WITH HEARING IMPAIRMENT

Abstract. The article presents the development of an information system for recognizing voice into text for people with hearing impairments, which makes it possible to improve the quality of life and interaction in society with other people. The device, software, functional blocks and subsystems of the information system are described. Examples of possible application and placement of the system in various spheres of public life are given. One of the types of implementation of the voice recognition information system is described. The development and creation of prototypes of a device for people with hearing impairments is considered. In the course of the research, the Google Speech Api technology was selected for speech recognition. In addition, this article presents a software and hardware complex that allows you to translate speech into text and then display it on the screen. Arduino UNO-based devices were chosen to achieve the goal. All information is processed on the smartphone of people with hearing impairments, which is sent to the device via Bluetooth with Arduino.

Keywords: information system, Google Speech Api, Arduino UNO, microcontroller, Ethernet, display, board, bluetooth, connector, port.

1. Introduction. Nowadays, it is impossible to develop various areas of human activity without the widespread use of computer technology and the creation of information systems. Information processing in such systems has become an independent scientific [1-2] and technical direction. In the modern world, information has become one of the most important resources, and information systems (IS) have become a necessary tool for virtually everyone, including people with disabilities.

Modern world information technologies have determined the direction of adaptation of people with disabilities. There are more than 360 million people with hearing impairments in the world community. In various countries, active research and development is underway to create information systems for the hearing impaired. Many scientific papers have focused on the use of neural networks to recognizing gestures. There are also gloves based on the Arduino Uno microcontroller for the hearing impaired. It is a bulky and inconvenient device, it is only at the design stage (no mass production), there is no text scoring, it is displayed only on the OLED screen.

The rapid process of miniaturization of electronics leads to the emergence of a new class of devices - wearable electronics, which is becoming the most relevant trend of our time. Mobile solutions can help transform speech and sound into visual representations for people with hearing impairments. However, where handheld phones pose problems, displays can facilitate further communication through privately written text, hands-free use, increased mobility, and socially acceptable interactions.

The progress in the field of augmented reality, which led to the creation of prototypes of augmented reality glasses - GoogleGlass and MicrosoftHololens, is especially bright and interesting [3-4]. Applications using these devices appear every day and are shown at the latest company exhibitions. The advent of augmented reality devices was a new step in the development of technology. Many companies managed to release analogs of these devices.

Another breakthrough was the emergence of an open unified universal digital platform Arduino, which made it possible to make a real prototype of the device. This platform includes a digital processor, a

programmer from a computer, and a development environment. The platform is open, which has allowed the emergence of a huge number of devices compatible with it - from the simplest sensors, screens to complex Ethernet boards, WiFi, etc.

Based on the knowledge gained, the capabilities of existing augmented reality devices were considered. The existing universal microcontrollers and devices interfaced with them, as well as writing programs for such devices, have been studied.

Using the Arduino [5-6] universal controller allows you to make a prototype of such devices. Using the open architecture of the controller allows you to write a universal platform for the device that makes life easier for people with hearing impairments.

2. Methods.

When researching this topic, the implementation of speech recognition and speech-to-text conversion, as well as the output of the received text to the device screen, were considered. The first part of the implementation consists of speech recognition and conversion to text.

The study examined some of the methods used in speech recognition.

Pre-emphasis. Different signals differ in volume. To bring audio to the same form, it is necessary to normalize the signals and filter them with a high-pass filter to reduce noise. Pre-emphasis is a filter for speech recognition tasks. It boosts high frequencies, which improves noise immunity and gives more information to the acoustic model.

Framing. The original signal is not stationary. It is divided into small intervals (frames), overlapping with each other, which are considered stationary. A Hann window function is applied to each frame to flatten the ends of the frames to zero.

Fourier. Fast discrete Fourier transform. The Fourier transform allows you to decompose the original stationary signal into a set of harmonics of different frequencies and amplitudes. This operation is applied to the frame, hence it transforms the frequency representation. We apply the Fourier transform to all frames, the spectral representation is formed and the power of the spectrum is calculated.

Log mel filterbank. Numerous scientific studies have shown that a person recognizes low frequencies better than high ones, and the dependence of his perception is logarithmic. Therefore, a convolution of N-triangular filters with one in the center is applied to the power spectrum. As the filter increases, the center shifts in frequency and increases logarithmically at the base. This allows you to capture more information in the lower frequencies and compress the representation of the high frequencies of the frame, subsequently the data is logarithm [7].

The second part of the implementation of the text output to the screen.

To study the principle of operation of the Arduino device with a universal control board, a number of experiments were considered to study the programming of universal controllers. For the hardware part, devices with certain functional properties were selected (table 1).

Table 1 - Functional components of the created device

№	Name	Main settings
1	Microcontroller board	Arduino Uno
2	Display	16×2, Integrated controller WS0010
3	Bluetooth	Module HC-05
4	Connecting wires / connectors	JST connector with wire (male-female kit), wire length 100mm
5	USB cable	A –B
6	Android device	Smartphone
7	IDE Arduino	Software environment

In the course of the research, the Google Speech Api technology was selected for speech recognition. Speech recognition works for 80 languages. Recognition of speech on the air through a microphone or audio recordings from files up to 2 minutes is possible. Numerous formats are supported, including FLAC, AMR and PCMU. Speech API technology can be embedded in any programming language. In this work, the Python language was chosen for this. The system provides the recognized text instantly in the process.

The recognition process is divided into several stages:

• Signal analysis. At this stage, the system sends the received request to the server. This is where the process of clearing interference and noise takes place. The next process is the compression of the received sound, it is divided into fragments, which is 25 milliseconds long. The divided fragments are passed through the acoustic model. For later recognition, this model determines which sounds have been pronounced.

• In the stage of signal recognition, the reference pronunciations stored in the acoustic model are compared with the speech fragments of the recording. Using machine learning, the system selects variants of spoken words and their context. As a result, it collects presumed words from sounds.

• The last step is to convert the signal to text. Using the language model, the system determines the word order and matches unrecognized words by context. The information obtained at these stages goes to the decoder. Therefore, with the most probable sequence of words, the decoder combines data from the acoustic and language models, converts them into text [8].

For programming the Arduino, the standard C ++ programming language was used (the AVRGCC compiler is used). Of course, there are several other development environments and several programming languages (Sketch, Basic), aimed mainly at beginners. We used the C ++ language and the ArduinoIDE device driver packages available on the Internet with an example of the program code [9-10].

To interact with the Arduino with a 16x02 LED display, the LiquidCrystal.h library was required, which includes a variety of commands to control the monitor.

On the Arduino Uno board, at the top are pins numbered from 0 to 13. These are digital pins, usually various LEDs are connected to them. However, pin number 13 is responsible for the operation of the built-in LED on the board and serves, most often, to check its operability. By specifying the pin number in the sketch, we can work with the device that is connected to it. For this we use the necessary functions: pinMode, DitalWrite [11-12]. The PinMode () function is used to customize this very pin. All pins on the board can work as inputs and outputs. By default, all pins are inputs, but the LED needs an output signal as it receives an operating command and voltage. The PinMode (pin, mode) function consists of two values. These are pin - pin number and mode - input / output mode (INPUT / OUTPUT). Further, the loop itself is written in the void loop () function. After determining the contact, as an output, you can set it to a high or low state, this is 0 or 5 volts of voltage [13].

We change the state of the output signal using the digitalWrite (pin, level) function. It also has two meanings. This pin is the pin number (for example, our LED is connected to pin A0) and level is the output signal level: it can be low (HIGH) or high (LOW).

The use of the delay (1000) function serves to set the delay in the program code between the execution of various actions (since the controller cannot simultaneously perform several actions, but performs them sequentially). The function itself specifies the time of this very delay in milliseconds (1000ms = 1s).

Before assembling all the devices, a program code was entered on the Arduino Uno board that will translate speech into text and output it to a 16x2 Led Display. The program code was previously written in the Arduino IDE software (an overview of the program was described earlier). To enter the code, the board was connected to a computer via a USB port, which also serves as a power supply for the Arduino UNO.

To implement the device, the Arduino board must be connected to a power source. Here, the power source is a USB cable from the Arduino UNO coming from the computer, as well as from the wall power supply that is connected to the connector. I also uploaded the code to the Arduino board via USB. You should not use a power supply greater than 20 volts as this will overwhelm your Arduino. The recommended voltage for most Arduino models is 6 to 12 volts. The necessary ports were installed in the software environment, corresponding to the hardware, only after installing the input and output ports, the program code is loaded and launched.

Physically, it serves as a wireless intermediary, transferring data between the Arduino and any smartphone. The application is available for both Android and IOs operating systems. After connecting the two boards, the Led Display was connected to the Arduino, then the HC-05 Bluetooth module [14].

The connection diagram is shown in figure 1.

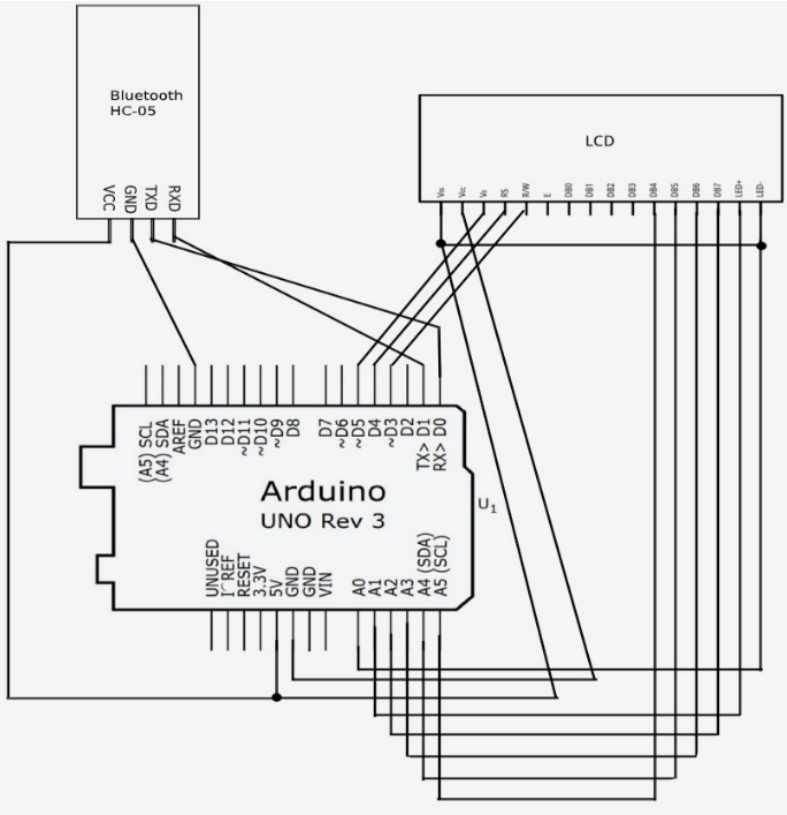

Figure 1 – Layout diagram of the device

Pin A0 is connected to pin 16 of Led Display.
16th pin of LED -> 1st pin of LED Display
2nd pin Led Display -> + 5V Arduino Uno
3rd pin of Led Display -> GND Arduino Uno
4th Pin Led Display -> 5th Pin Arduino Uno
5th Pin Led Display -> 4th Pin Arduino Uno
6th Pin Led Display -> 3rd Pin 1Sheeld
Connection diagram of the Bluetooth module HC-05 to 1Sheeld
RXD pin Bluetooth HC-05 -> TXD pin Arduino Uno
TXD pin Bluetooth HC-05 -> RXD pin Arduino Uno
GND pin Bluetooth HC-05 -> GND pin Arduino Uno
VCC pin Bluetooth HC-05 -> 5V pin Arduino Uno

Therefore, when the required boards have been connected, the device is operational as the program code was previously entered. In this project, an Android smartphone was selected and the OneSheeld application was downloaded. The device and the application communicate through the HC-05 Bluetooth module [15].

4. Conclusions. Given that at the current stage of the project we decided to focus on the implementation of the application and its interaction with the interface, new implementations were inserted into it at a later stage, but we cannot fail to mention this due to the usefulness of these resources. We will refer to the final stage of the project implementation: changes in the application interface, improvement of the usability of the Android device by the end user.

As a result, there were studies of ways to implement speech recognition developed on the Google Speech API technology. In the course of the research, the methods and stages of speech-to-text reproduction were considered.

This article is the result of research that studies and explores the possibilities of automation using Arduino[18]. Built with inexpensive resources such as Arduino and other devices, and an Android app to facilitate socialization of people with disabilities, this app will certainly make its own difference in their lives. However, this prototype is not intended for commercial use. The main goal is to help people with hearing impairment, however, and in the future this solution can be easily implemented.

Е.С. Нурахов, А.Е. Қами

[1]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан

ЕСТУ ҚАБІЛЕТІ НАШАР АДАМДАРҒА АРНАЛҒАН АҚПАРАТТЫҚ ЖҮЙЕ

Аннотация. Мақалада есту қабілеті бұзылған адамдарға мәтінді дауыстан тануға арналған ақпараттық жүйенің даму жағдайы ұсынылған, бұл өмір сүру сапасын жақсартуға және қоғамдағы басқа адамдармен өзара қарым-қатынас жасауға мүмкіндік береді. Ақпараттық жүйе құрылғысы, бағдарламалық жасақтама, функционалды блоктары мен ішкі жүйелері сипатталған. Жүйені қоғамдық өмірдің түрлі саласында қолдануға және орналастыруға мысалдар келтірілген. Дауысты танитын ақпараттық жүйені енгізу түрлерінің бірі сипатталған. Есту қабілеті төмен адамдарға арналған құрылғы прототиптерін жасау және құру жолдары қарастырылады. Зерттеу барысында сөйлеуді тану үшін Google Speech Api технологиясы таңдалды. Сонымен қатар, бұл мақалада сөйлеуді мәтінге аударуға, содан кейін оны экранға шығаруға мүмкіндік беретін бағдарламалық-аппараттық кешен ұсынылған. Осы мақсатқа жету үшін Arduino UNO негізіндегі құрылғылар таңдалды. Барлық ақпарат есту қабілеті бұзылған адамдардың смартфонында өңделеді, ол Arduino көмегімен Bluetooth арқылы құрылғыға беріледі.

Түйін сөздер: ақпараттық жүйе, Google Speech Api, Arduino UNO, микроконтроллер, Ethernet, дисплей, тақта, Bluetooth, коннектор, порт.

Е.С. Нурахов, А.Е. Қами

Казахский национальный университет имени аль-Фараби, Алматы, Казахстан

ИНФОРМАЦИОННАЯ СИСТЕМА ДЛЯ ЛЮДЕЙ С НАРУШЕНИЕМ СЛУХА

Аннотация. В статье представлена разработка информационной системы распознавания голоса в текст для людей с нарушениями слуха, позволяющей повысить качество жизни и взаимодействия в обществе с другими людьми. Описаны устройство, программное обеспечение, функциональные блоки и подсистемы информационной системы. Приведены примеры возможного применения и размещения системы в различных сферах общественной жизни. Описан один из видов реализации информационной системы распознавания голоса. Рассматривается разработка и создание опытных образцов устройства для людей с нарушениями слуха.

В ходе исследования была выбрана технология Google Speech Api для распознавания речи. Кроме того, в данной статье представлен программно-аппаратный комплекс, позволяющий переводить речь в текст и затем выводить ее на экран. Для достижения этой цели были выбраны устройства на базе Arduino UNO. Вся информация обрабатывается на смартфоне людей с нарушениями слуха, который передается на устройство по Bluetooth с помощью Arduino.

Ключевые слова: информационная система, Google Speech Api, Arduino UNO, микроконтроллер, Ethernet, дисплей, плата, Bluetooth, разъем, порт.

Information about authors:

Yedil Sergazievich Nurakhov, lecturer of the Department "Informatics", Al-Farabi Kazakh National University, Almaty, Kazakhstan. E-mail: y.nurakhov@gmail.com, https://orcid.org/0000-0003-0799-7555;

Ayaulym Yerbolkyzy Kami, 2nd year master's degree in “Computer engineering", Al-Farabi Kazakh National University, Almaty, Kazakhstan. E-mail: qamiaiko@gmail.com, https://orcid.org/0000-0001-9840-9178

REFERENCES

[1] Weiss S.M., Indurkhya N. and Zhang T. Fundamentals of predictive text mining, Springer Science & Business Media, 2010.

[2] Voice Recognition // Human Interface Technology Laboratory URL: http://www.hitl.washington.edu/research/ knowledge_base / virtual-worlds / EVE / I.D.2.d.VoiceRecognition.html (date accessed: 09.06.2017).

[3] Albrecht U.V., von Jan U. and Kuebler J. et al., (2014) 'Google Glass for documentation of medical findings: evaluation in forensic medicine', Journal of Medical Internet Research, Vol. 16, No. 2.).

[4] CNN (2014) Movie Theaters Ban Google Glass and Other Wearables [online] http://money.cnn.com/2014/10/30/technology/google-glass-movies/ (accessed 22 June 2015)

[5] McRoberts, Michael. Arduino Básico. 456. Editora Novatec, Brazil, 2011

[6] Arduino project site. Electronic publication. Access mode https://www.arduino.cc

[7] How Neural Networks Recognize Speech-to-Text [Electronic resource] - URL: https://dzone.com/articles/how-to-train-a-neural-network-to-recognize-speech/

[8] Kim Y., Lee H., and Provost E. M. "Deep learning for robust feature generation in audiovisual emotion recognition," 2013, pp. 3687-3691, doi: 10.1109/ICASSP.2013.6638346. [Online]. Available:

http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6638346 Series of articles Arduino programming from scratch. Electronic publication. Access mode http://Techy-news.com

[9] Wen G., Li H., Huang J., Li D., and Xun E. "Random Deep Belief Networks for Recognizing Emotions from Speech Signals," Computational intelligence and neuroscience, vol. 2017, pp. 1-9, 2017

[10] OLED I2C display with Arduino Tutorial. Electronic publication. Access mode https://www.instructables.com/OLED-I2C-DISPLAY-WITH-ARDUINO-Tutorial/

[11] Arduino Programming Notebook. Brian W. Evans. 2007

[12] Simon Monk Programming the Arduino. Professional work with sketches

[13] Massimo Banzi Getting Started with Arduino, O'REILY 2009

[14] https://www.instructables.com/Arduino-Bluetooth-remote-lcd-display

[15] https://startingelectronics.org/tutorials/arduino/modules/OLED-128x64-I2C-display/