

Jakub PEKSIŃSKI¹
Grzegorz MIKOŁAJCZAK²

VOICE SYSTEM FOR REGISTRATION OF GOODS DENOTED BY A BAR-CODE BASED ON THE SPECTRAL ANALYSIS OF THE HUMAN VOICE

The computer system for registration of goods denoted by means of the EAN-13 code is featured in the paper. Operation of the system is based on the spectral analysis of the human voice. Thereby, the system can be an alternative to traditional systems based on scanning a code by means of diode or laser scanners.

GŁOSOWY SYSTEM EWIDENCJONOWANIA TOWARÓW OZNACZONYCH KODEM KRESKOWYM OPARTY NA ANALIZIE SPEKTRALNEJ GŁOSU LUDZKIEGO

W artykule przedstawiono komputerowy system ewidencjonowania towarów oznaczonych przy pomocy kodu EAN-13. Działanie systemu opiera się na analizie spektralnej głosu ludzkiego. Dzięki temu system ten może być alternatywą dla tradycyjnych systemów opartych na odczytywaniu kodu przy pomocy skanerów diodowych czy laserowych.

1. REGISTRATION OF GOODS AT THE COMPANIES

EAN is an abbreviation for European Article Number. These numbers are marked almost all of the products sold, they form a bar code string 13, or 8 digits. Using a barcode scanner to be placed on computerized cash registers. In the case numbers 13-digit, first two digits indicate the country of origin (manufacturer), for example, Polish and 59 for 40, 41, 42, 43 or 44 for Germany. The next 5 digits means the same manufacturer, another group of 5 digits is characterized by a specific product. The last digit p is a check digit.

This figure we choose so that after multiplying successively, starting from the left, the first twelve digits of 1 or 3, etc., and then summing up of the products, adding to the result gave the number of digits of p is divisible by 10. Thus, if the number is a good " $ab cde hikmn fg p$ ", then p should satisfy the following congruence:

$$a+3b+c+3d+e+3f+g+3h+i+3k+m+3n+p\equiv 0(10)$$

Individual record numbers are encoded by the bar code shown in Table 1.

¹West Pomeranian University of Technology, Szczecin, Electric Faculty, POLAND, ul. 26 Kwietnia 10, 71-126 Szczecin Phone: +48 91 4495311, E-mail: kooba4@o2.pl

² West Pomeranian University of Technology, Szczecin, Electric Faculty, POLAND, ul. 26 Kwietnia 10, 71-126 Szczecin Phone: +48 91 4495311, E-mail: grzegorz.mikolajczak@zut.edu.pl

Tab. 1. Kod EAN

Digit Decimal	Code A	Code B	Code C
0	0001101	0100111	1110010
1	0011001	0110011	1100110
2	0010011	0011011	1101100
3	0111101	0100001	1000010
4	0100011	0011101	1011100
5	0110001	0111001	1001110
6	0101111	0000101	1010000
7	0111011	0010001	1000100
8	0110111	0001001	1001000
9	0001011	0010111	1110100

Using the above method of check digit, you can always reveal errors involving a change of one digit, and in most cases the switching of two adjacent digits. Often remain undetected switching two nonadjacent digits and the simultaneous changes in two digits.

Record the EAN code consists of dark lines of different thicknesses: a single (1), double (11), triple (111) and Quad (1111) and different intervals: single (0), double (00) and triple (000).

Code start's and end's with the two extremes (elongated) lines separated by a single thickness of a single spacing (101). In the midst of writing there are two lines separating the single-thickness and three single spacing (01010).



Fig.1. Sample code of product

Barcode of the article (Fig. 1) contains a total of 13 decimal digits of which 12 are encoded sequence of 7 binary digits, in the form of dashes.

2. METHODS OF REPRESENTATION OF THE HUMAN VOICE

The human speech is a basic tool used by a human for communication and transfer of information between humans. This process is so popular in our life that we practically do not analyse the complexity of the processes associated with it. As far as for the humans, the process of creation, reception and decoding of speech is a natural process, whereas for the

electronic devices the process is only an electrical signal that is interpreted and processed by means of the pre-defined algorithms.

The speech signal can be presented by means of various methods [1]. The standard representation is the time representation featured in the figure no.2.

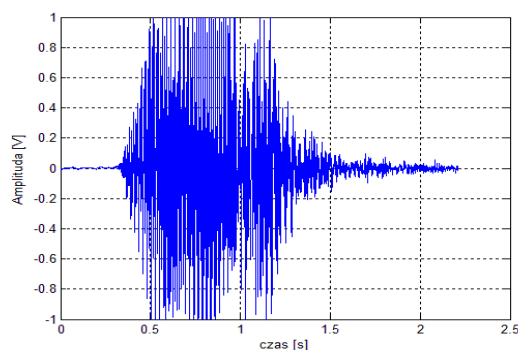


Fig.2. Time representation of the speech signal

However, the time representation of the signal does not transfer any information about frequencies in the speech signal. Therefore, the representation in the frequency domain is a more often used representation of the speech signal. It is possible to change from the time domain into the frequency domain through the Fourier transform [2] according to the formula no. 1.

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt \quad (1)$$

where: $x(t)$ - time domain signal

$X(\omega)$ - frequency domain signal

The figure no. 3 features representation of a frequency domain speech signal.

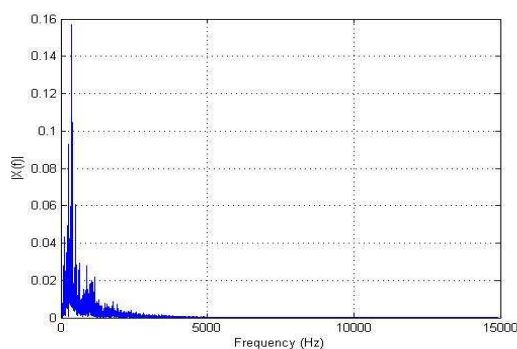


Fig.3. Representation of the frequency domain speech signal

The frequency domain signal transfers many useful information that can be used for the analysis but its drawback is that the information about the signal duration is lost. Thereby, the frequency domain speech signal cannot be associated with the signal duration.

There is, however, the representation of the speech signal that makes it possible to present the signal in the time, frequency and amplitude form. Such representation is based on calculation of the short-term Fourier transform featured by the formula no. 2.

$$STFT\{x(t, f)\} = \int_{-\infty}^{\infty} x(t)w(t - \tau)e^{-j\omega\tau} dt \quad (2)$$

where: $w(t)$ - time window function

The representation known as the spectrogram is obtained through squaring the $STFT\{x(t, f)\}$ module.

$$G(t, f) = |STFT\{x(t, f)\}|^2 \quad (3)$$

The figure no. 4 presents a speech signal spectrogram obtained as the result of application of the formula no. 2 and 3.

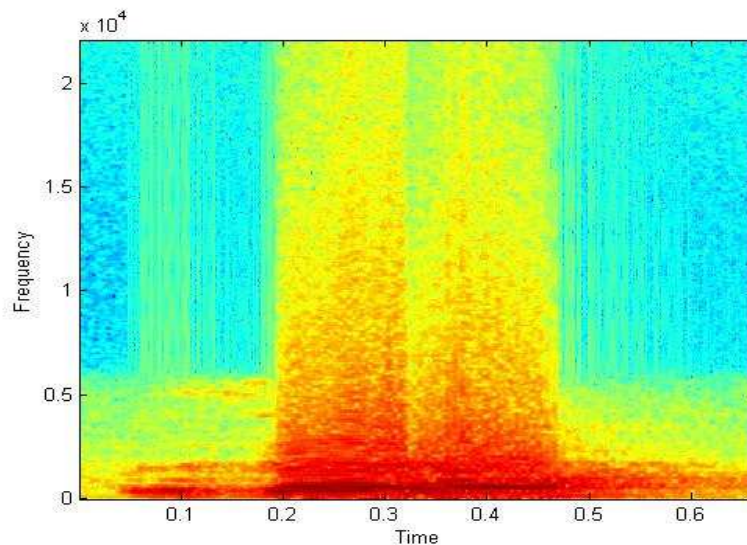


Fig.4. Representation of a speech signal by means of the spectrogram

The spectrogram in the computer screen is a colour image featuring the time and frequency axis, whereas the colours denote the signal amplitude. The spectrogram in this paper is a bit map, and the voice bar-code scanning system is based upon comparison of such map.

3. VOICE RECOGNITION SYSTEM FOR SCANNING A BAR-CODE

The bar-code scanning system proposed by the authors is based on comparison of the collected voice samples and the pre-recorded patterns, i.e. so called Speaker-Dependent analysis [3]. Therefore, it is necessary to collect the patterns corresponding to the numbers from 0 to 9 and to transform them into the spectrogram. The figure no. 5 presents spectrograms of the numbers obtained for a specific person (the 37 year old man).

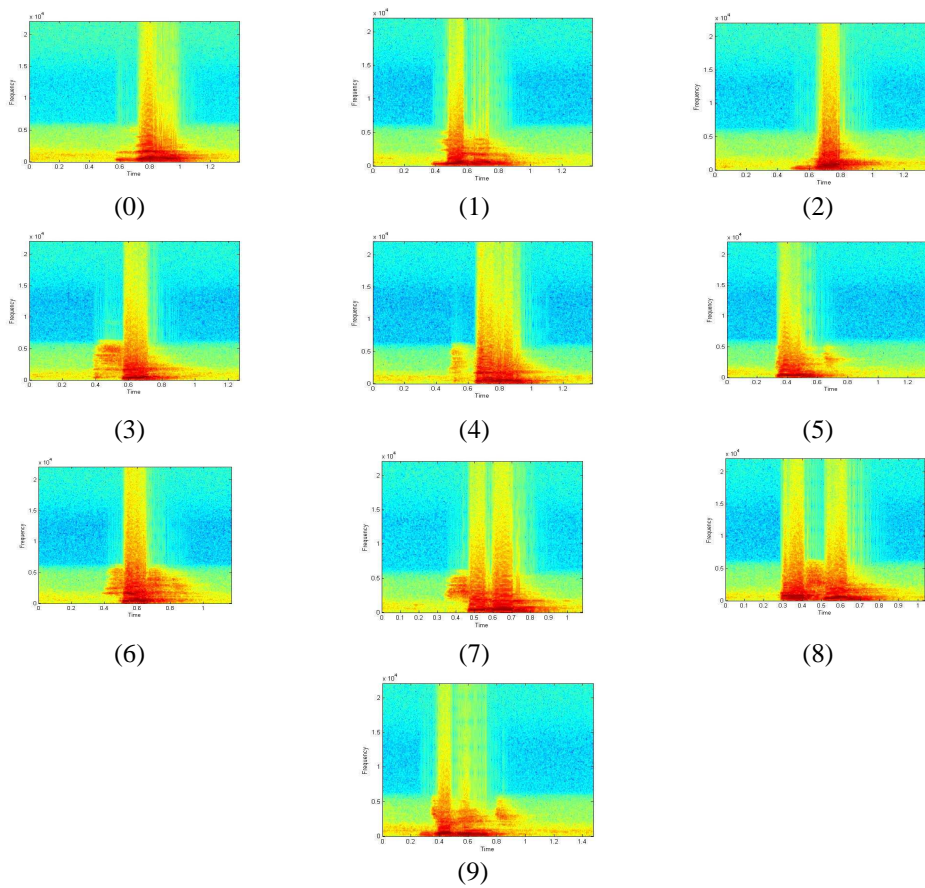


Fig.5. Spectrograms of the numbers from 0 to 9

Such base of spectrograms representing the numbers from 0 to 9 can be used as the reference pattern. The cross-correlation function featured by the formula number 4 [4] is used as a comparison criterion of spectrograms.

$$r = \frac{\sum_i \sum_j f_{i,j} f'_{i,j}}{\sqrt{\left(\sum_i \sum_j (f_{i,j})^2\right) \cdot \left(\sum_i \sum_j (f'_{i,j})^2\right)}} \tag{4}$$

where: $f(i, j)$ – pattern spectrogram; $f'(i, j)$ – tested spectrogram

$$f_{i,j} = f_{i,j} - \bar{f}; f'_{i,j} = f'_{i,j} - \bar{f}'; \bar{f}, \bar{f}' - \text{average values}$$

The process of scanning and recognition of a bar-code by means of voice is performed as follows:

- System user logging.
- Reading out consecutive code numbers by a user;
- Transformation of the recorded signals representing the code numbers into the spectral form;
 - Comparison of the scanned spectrogram and the pattern located in the base for a specific user by means of the criterion featured by the formula no. 4;
- Bar-code interpretation;
- Registration of goods denoted by a code.

The block diagram of the voice system for registration of goods is presented in the figure no. 6.

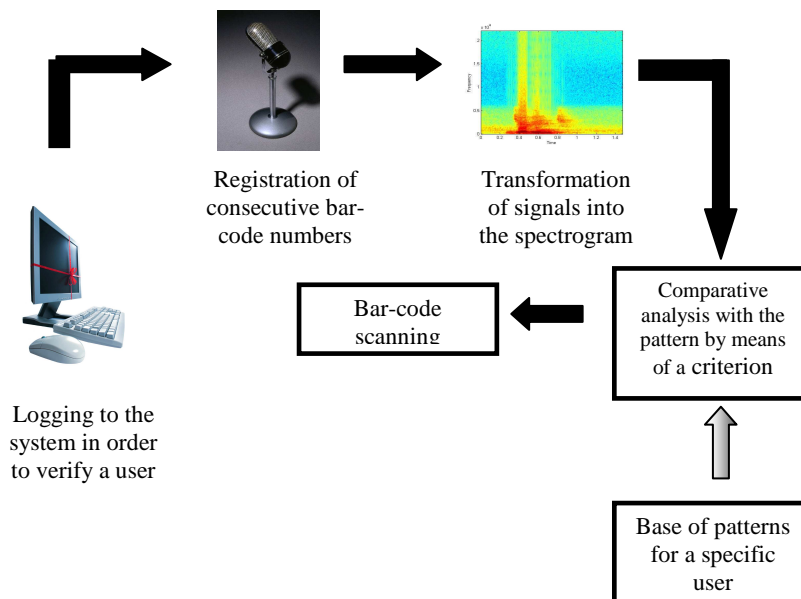


Fig.6. Block diagram of the voice system for registration of goods

4. EXPERIMENTAL TEST RESULTS

In order to verify the efficiency of operation of the voice bar-code scanning system, the authors of the paper performed an experimental test. A group of 10 people participated in the test. The samples of voice signals corresponding to the numbers from 0 to 9 were collected from these people. The samples were transformed into the spectrograms by means

of the formula 2 and 3 and recorded into separate databases. Then, every person read out consecutive numbers of a bar-code of any product after logging to the system. The results of the test are presented in the table number 2

Tab. 2. The results of the test

User 1				
	Read-out 1	Identification 1	Read-out 2	Identification 2
Code 1	5903199200212	5903199200212	-	-
Code 2	5904172189671	590417X189671	904172189671	904172189671
Code 3	9788320434095	9788320434095	-	-
Code 4	9788373614864	9788373614864	-	-
Code 5	9040172279612	904017227961X	9040172279612	9040172279612
User 2				
	Read-out 1	Identification 1	Read-out 2	Identification 2
Code 1	5903199200212	59X3199200212	5903199200212	5903199200212
Code 2	5904172189671	5904172189671	-	-
Code 3	9788320434095	9788320434095	-	-
Code 4	9788373614864	97X8373614864	9788373614864	9788373614864
Code 5	9040172279612	9040172279612	-	-
User 3				
	Read-out 1	Identification 1	Read-out 2	Identification 2
Code 1	5903199200212	5903199200212	-	-
Code 2	5904172189671	5904172189671	-	-
Code 3	9788320434095	97883X043409X	9788320434095	9788320434095
Code 4	9788373614864	9788373614864	-	-
Code 5	9040172279612	9040172279612	-	-
User 4				
	Read-out 1	Identification 1	Read-out 2	Identification 2
Code 1	5903199200212	5903199200212	-	-
Code 2	5904172189671	5904172189671	-	-
Code 3	9788320434095	9788320434095	-	-
Code 4	9788373614864	9788373614864	-	-
Code 5	9040172279612	9040172279612	-	-

The analysis of the results in the table no. 2 proves the efficiency of the voice bar-code scanning system based on the spectral analysis. Of course, sometimes it is necessary to read out a code again but the same problems can occur during scanning by means of the bar-code scanners.

The efficiency of the first read-out is about 80 %, thereby being consistent with the *Speaker-Dependent* system specification. The second read-out is, however, 100 % efficient.

It is obvious that the system does not replace a traditional system based on optical bar-code scanners but it can be successfully used as support system. It can be especially useful if a bar-code is damaged and it is not possible to read out the code by means of an optical

scanner. The voice bar-code scanning system can be also used at the supported employment enterprises where the employees can have problems with activities involving hands.

5. REFERENCES

- [1] Tadeusiewicz R.: *Sygnal mowy*; Wydawnictwo Komunikacji i Łączności, Warszawa 1988.
- [2] Szabatin J.: *Podstawy teorii sygnałów*, Wydawnictwo Komunikacji i Łączności, Warszawa 2000.
- [3] Shalk T.: *The Desing and Use of Speech Recognition Data Bases*, National Bureau of Standards 1982.
- [4] Zieliński T.: *Cyfrowe Przetwarzanie Sygnałów*, Wydawnictwo Komunikacji I Łączności, Warszawa 2005.
- [5] Kłaczyński M.: *Zjawiska wibroakustyczne w kanale głosowym człowieka*, Rozprawa Doktorska, AGH, Kraków 2007.
- [6] Janiszewski J.: *System analizy i rozpoznawania mowy w oparciu o procesor sygnałowy TMS320C6711*, Praca magisterska, Politechnika Śląska 2003.