Silent Speech Recognition by Surface Electromyography

Andrzej B. Dobrucki, Piotr Pruchnicki,
Przemysław Plaskota, Piotr Staroniewicz,
Stefan Brachmański and Maciej Walczyński

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/60467

Abstract

For some time, new methods based on a different than acoustic signal analysis are used for speech recognition. The purpose of nonacoustic signals is to allow silent communication. One of these methods based on the electromyography signal is generated by the human speech articulation system. This article presents a device for electromyographic (EMG) signal acquisition and the first measurements from its use.

Keywords: Electromyography, Speech recognition

1. Introduction

Speech recognition is a very important part of computer–human communication system. The most frequently used methods for speech recognition are based on the analysis of an acoustic signal. This signal can be represented as a real-time signal or previously recorded data with a human speech signal.

The use of the acoustic signal analysis is not always possible. In a noisy environment, it is not possible to separate the speech signal from background noise. In this case, the limitations of technical capabilities are present. On the other hand, in a very quiet environment, it is not possible to use voice communication because of the need to maintain silence. In both cases, other signals generated by a speech articulation system can be useful.

It is possible to use electromyographic (EMG) signals produced by muscles of the human speech system [1, 2]. These signals are characterized by low voltage levels and therefore are

difficult to record. Another problem is their low resistance to external interference due to the need for large gains [3].

For a practical use of EMG signals, it is necessary to record and gain EMG signals. In the next part of this article, the assumptions and implementation of recording devices are presented. Also, the results of the first attempts to use the device are described.

2. System concept

2.1. EMG signal sensor

The designed hardware part of the SVR (subvocal recognition) system should comply with some specific assumptions because of the specific characteristics of signals:

- high quality of the signal path due to work with low- level signals – characterized by low noise, high dynamics, and immunity to external interference,

- flexibility in the configuration of the signal path – variable usage of electrodes, different types of electrodes, and the possibility of simultaneous use of several signal paths,

- mobility – portable character of the system, small size, and weight.

Based on the assumptions outlined above, a block diagram of the system has been developed (**Figure 1**). Also the necessary requirements for its parts have been defined.
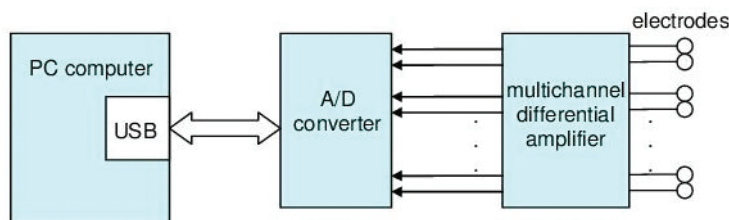

Figure 1. Block diagram of the EMG sensor.

Properties of electrodes that have been used in the system were selected primarily on the basis of SENIAM (Surface ElectroMyoGraphy for the Non-Invasive Assessment of Muscles) recommendations [4]. According to them, there are bipolar electrodes. Due to the small size of the analyzed muscle, the electrodes must be placed close to those muscles. Also, the electrodes must be small.

Electrical impulses collected by the electrodes need to be amplified around 500–1000 times. For this purpose, it is necessary to use specially designed low-noise amplifiers. Due to the use of bipolar electrodes, the amplifier operates in a differential regime that helps significantly to reduce the impact of external noise in the recorded signal.

For further analysis of EMG signals, it is necessary to replace electrical EMG signals with a digital form. This process allows to carry on digital processing of EMG signals on computer devices. This transformation process of analog signals into their digital form requires the use of a signal acquisition card (DAQ card). Due to the requirement of system mobility, the optimal solution is to use an external adapter with a USB interface. It enables the use of easy-to-carry laptop, and besides, there is no need to use additional power supplies since enough energy is supplied by the USB port. An important element of the system is a PC. On this device, the programs for SVR based on electromyography signals are installed.

2.2. Electrodes

The elements used to collect the electromyography signals are electrodes. The device uses various types of electrodes in order to allow finding the best solution. The self-adhesive electrodes (**Figure 2**) provide the most flexibility for an individual subject. The use of this kind of electrodes allows placing them almost anywhere. Initial experiments have shown that the adhesion of such electrodes is only possible on smooth and dry skin. The presence of hairs on the skin, sweat, and wrinkles often prevents the attachment of electrodes.

Figure 2. Self-adhesive electrodes.

An alternative solution is to use cup-shaped metal electrodes (**Figure 3**). They also contain isolated electrodes, which make them any deployment. They require using a special adhesives conductive gel that is used for fixing the electrode and increases the conductivity of the connection. The electrodes may be made of different materials, such as silver, silver coated with silver chloride (Ag/AgCl), gold, iron, or tin. Due to the very good properties – low and stable in time voltage – the most commonly used are silver electrodes coated with silver chloride.

Figure 3. Cup-shaped electrodes.

The main advantage of individual electrodes is the possibility of putting them on any place, although in certain circumstances, it may be a significant drawback. That happens when the repeatable electrode placement is needed in order to couple each other. A good solution in such a situation is the use of bar-type bipolar electrodes (**Figure 4**). These electrodes are compact and are lead shielded. The distance between the electrodes is strongly determined by their structure.

Figure 4. Bar-type electrodes.

During the measurement process, in addition to EMG signal electrodes, it is necessary to use a reference electrode. The purpose of the additional electrode is to determine the reference potential of the speaker. Without this additional electrode, the amount of noise (especially from the power network) in the acquisition increases greatly, which prevents the measurement completely. During the measurements, the reference electrode is connected to the patient in a place where there is no muscular activity.

2.3. Amplifier

To amplify the electrical EMG signals from the electrodes, specially designed amplifiers are used [5]. The development of the designed systems allows for a perfect fit of the amplifier for the SVR purpose.

Amplifier design guidelines are defined on the basis of the solutions used in the world as follows:

- The frequency band should be in the range of 10–500 Hz; hence, the filters should be used in order to limit this frequency range.

- The total amplification factor has to be very high (up to 1000 V/V).

- The self-noise level in reference to the input signal is assumed less than 1 μV.

- In order to suppress interference from the environment, it is very important to get the largest possible value for CMRR (common-mode rejection ratio). It is assumed that CMRR is >95 dB value.

- Input impedance of EMG amplifiers is usually high and is at least of 10 MΩ. Because of high resistance of the human skin and nonideal coupling of the electrodes with the skin as well as adipose tissue, the impedance of EMG electrodes is significant and it can be as much as several dozens of kiloohms. Due to a high input impedance value of the amplifier, it is possible to obtain maximum values of EMG voltages.

- The need of decrease of external distortions induces the minimum length of a cable connecting electrodes and amplifiers; thus, the amplifiers should be placed as near as possible and their size must be also as small as possible.
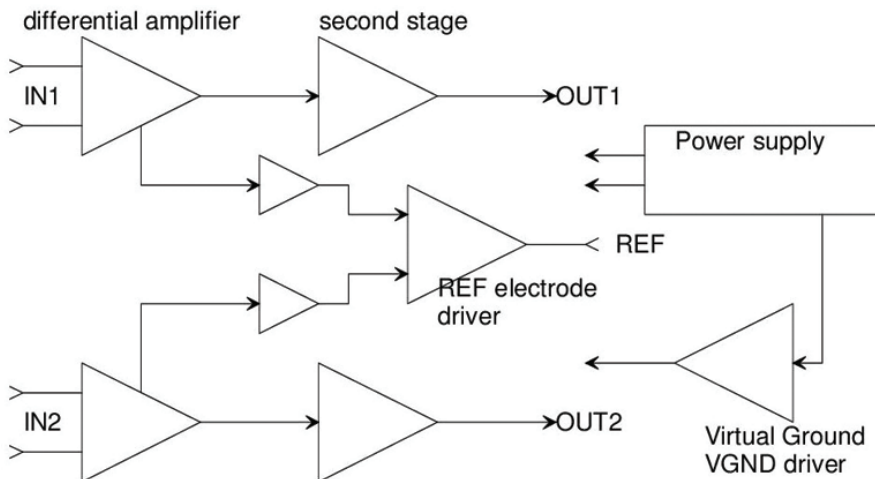

Figure 5. EMG amplifier.

- Power supply for the amplifier is on the board of the DAQ card; so the additional power unit is not necessary to use.

The scheme of the EMG amplifier is presented in **Figure 5**.

2.4. Driver of the reference electrode

The task of the reference electrode is fixing an adequate potential of the patient's body and ensuring returns to the bias currents of the inputs of a differential amplifier. Using the driver improves the damping of the common component for the amplifier (CMRR) of a value from 10 to 50 dB [6, 7]. The operation of this system consists of a continuous comparison of average voltage from measuring electrodes with the virtual ground's voltage. If a difference occurs, it is amplified with an opposite sign and put to the reference electrode. In this way, the system still maintains the proper value of the DC component of the measuring electrodes. The reference electrode driver (RLD) is used to raise the CMRR of the instrumentation amplifier. With this higher signal-to-noise ratio (SNR), the differential signal obtained is ensured to possess only relevant information and a minimum of interference currents or irrelevant data. The idea behind the RLD is to maintain a known voltage potential in the human subject that is directly related to the system board ground. This method then reduces the common-mode DC offset previously found in the system and thereby attempts to cancel any different DC offsets that individual channels or probes may experience. The actual method of the RLD is quite unique. A feedback network is created that depends on the averaged inputs from the combined instrumentation amplifier floating grounds and a GROUND signal originating from the human. This signal is then sent through an inverting gain stage that completes the feedback loop, which effectively counteracts any potential changes in the subject.

2.5. Virtual ground source

Because all electronic circuits of the power amplifier are supplied with a single voltage, it was necessary to produce an artificial "ground" potential. DC voltage of such a virtual ground should be equal to half the supply voltage. The easiest way to achieve this is to use a simple resistive divider (this version was used in the construction of the first amplifier). Unfortunately, the change of current from the divider causes changes of the voltage. Therefore, in the second version the buffer amplifier has been added, which holds the potential of virtual ground unchanged.

2.6. ADC card (DAQ)

Because of the narrow bandwidth of collected signals and a large level of signal after amplification, requirements for the DAQ card parameters are not critical.

Mobility of the whole system involves the use of an external card with a USB interface. A precise sampling runs with high dynamics and requires a resolution of not less than 16 bits. Because of the research reason, the originally formed multicard system must provide a sampling for at least eight independent channels. The sampling rate is not critical. At the top of the analyzed signal frequency of 500 Hz and the provided sample for the antialiasing filter

slope, there is quite an adequate sampling frequency equal to 4 kHz. Even while working with eight channels, the resultant frequency for the whole card will only be 32 kHz.

A desirable feature of the card used is switched ranges. They allow for a flexible adjustment range of the amplitude of the sampled signal and better use of dynamics DAC.

A software part of this project is written in C++ and runs on the Linux operating system. Due to the availability of drivers, the application was tested on Linux Debian, and due to the licensing requirements, it was decided that the program is a console application. The target of the software is data acquisition, signal parameterization, and classification.

3. EMG measurement by using cup-shaped and double (glued) electrodes

The measurements were made by using adhesive (**Figure 2**) cup-shaped (**Figure 3**) and double electrodes (Figure 4). The reference electrode was used as one of the glued electrodes (**Figure 2**). An example of the reference disc electrode placement is shown in **Figure 6**.

Figure 6. Example of the reference disc electrode placement.

3.1. Measurement system

The measurement system consists of electrodes, a preamplifier of an EMG signal, an AD converter, and a multichannel recorder. As an AD converter, a multichannel recorder has been used, or a sound card with a software for digital signal processing.

A few extra features have been checked before finalizing the configuration of the measurement system. To record the results of the measurements, five different measurement configurations

were used. Three of them have been reviewed negatively. The main problems concern the interference from the power supply, noise from FireWire (not always), and demodulation of the EMG signal cables from a nearby radio station transmitter. The use of battery power, the reduction of EMG wires length, and the use of the USB bus made it possible to eliminate the problems listed above.

The EMG signals were stored as a wave format file with a 48-kHz sampling frequency, with 16-bit resolution. The EMG signal dynamics reach about 20 dB, so that the signal parameters are sufficient. An easy analysis of the recorded signals is guaranteed by using a wave format file. During the measurements, the recorded signal can be observed, and these signals can be displayed in real time. It is also possible to insert some markers for easy highlighting in the next repetition of words. **Figure 7** shows some examples of the results of the measurements.
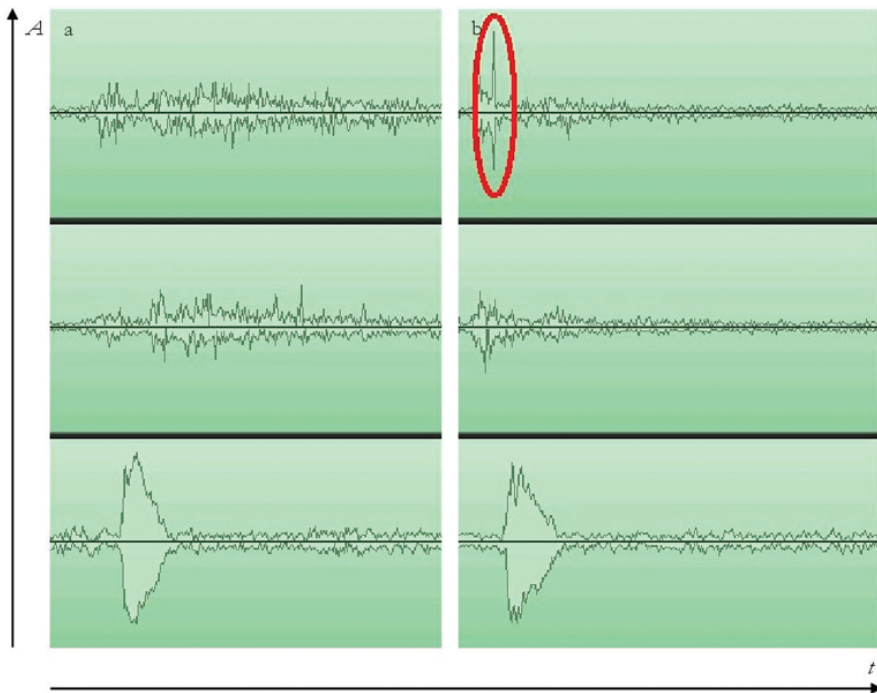

Figure 7. Sample measurement results for a person with a normal EMG signal (vowel "a," 1 s duration): a) a valid signal and b) a signal with interference (marked). From the top: an EMG signal from the bar-type electrode, an EMG signal from the cup-shaped electrode, an acoustic signal recorded by using a microphone. A – amplitude (arbitrary units), t – time.

3.2. Measurements

The measurements were provided for a group of 14 people (11 men and 3 women), aged 20–25 years. A typical location for electrodes is shown in **Figure 8**. The electrodes' location differed

slightly between individuals; the differences resulted primarily from the differences in the anatomy of the human body. In order to improve the conductivity between the skin and the electrodes, an electroconductive paste was used. The EMG signals were recorded with and without the conductive paste. The use of the paste in some cases slightly improved the conductivity (the recorded EMG signal was greater); the other did not affect the value of the EMG signal. There were no cases of deterioration of the EMG signals after applying the paste.

Figure 8. Example of cup-shaped electrodes' location.

To achieve the best repeatability of the measured signals, the research process was performed with the following procedure: 1) application of the sensor to the neck; 2) signal check: if bad – a change of sensor position is needed, if good – recording the signal. The electrodes were attached by using a hook and loop tape, ensuring the pressure stability over the entire process of measurements.

During the measurements, it was found that the EMG signals fundamentally differ within the group. For eight subjects, the correct values of the EMG signals were obtained immediately, while for six people, the correct values could not be obtained immediately. Three of these people learned to speak words in a short time, so that the EMG signal had the correct value. The other three were not able to obtain the EMG signal even after some time.

During the measurements, 10 repetitions of the following elements of the dictionary were recorded: the vowels a, e, i, o, u, y; the words: "stop", "start", "dalej" (further), "v levo" (left), "v pravo" (right), "pauza" (pause), "enter", "tak" (yes), "nie" (no).

To get the correct number of repetitions and the corresponding interval between different words, the graphical presentation on the screen was used. With this solution, the signals did not overlap, and it was easy to get the required number of repetitions.

3.3. Measurements summary

For higher values of EMG signals during speaking, it is necessary to use more sensitive sensors. The use of the conductive paste did not increase the value of the signal. The EMG signal has the highest value depending on the subjects and positions of the electrode. Finding the best location of electrodes requires several attempts. For some people, the correct EMG signal could not be obtained, but some of these people could learn to speak such words, and thus the EMG signals were correct. The use of rigid location of the sensor helps to improve the correct EMG signal quality.

The first experiments related to the registration of the EMG signals were conducted with the use of the first introductory version of the electronic circuit. The characteristics of this device were wires of 1.5 m, connecting the amplifier with the surface electrodes, which were placed on the skin of the investigated person.

In the process of measuring, some problems were observed, as follows:

- high level hum of 50 Hz – the level of this noise was strongly dependent on the wires position,

- change of a noise level during the measurement due to changing of the wires position, and

- signal from the FM transmitter (analog radio program), depending on the position in the building.

All the presented problems are connected with a large length of wires connecting the electrodes with the amplifier. The EMG signals amplifier has a very large amplification and input impedance. It is very sensitive to violations from the electric–magnetic fields around the device.

The best method to minimize these type of violations is the reduction of the wire length from the electrode to the amplifier. To realize this aim, the miniaturization of the device is necessary. The first step is to fasten the electrodes directly to the printed circuit board, and the second step is to place the whole device on the neck of the measured person.

4. State of the art of EMG speech recognition

The early research in the field of electromyography in speech technology was basically carried out only for proper speech prosthetic (parameters comparison between healthy people with prosthesis).

First significant trials to use EMG for automatic speech recognition were done in 1985–1986 in the USA and Japan. Morse et al. [8] applied four-channel surface electrodes with a sampling frequency of 5120 samples/channel/s. After the limitation of amplitude and frequency band to the range of 100–1000 Hz, 20 repetitions of isolated words (digits in English) for two voices were examined. After the use of ML (maximum likelihood), the scores at the level of 60% were

obtained. In a later research, the same team obtained similar results with the use of neural networks [9].

In a parallel study, Japanese scientists Sugie et al. [10] with three-channel surface electrodes with a sampling frequency of 1250 samples/channel/s succeeded in real-time phone recognition experiments. For three voices and 50 Japanese syllables, they reached a proper classification at the level of 64%.

Among research centers where most advanced works on subvocal speech (based on EMG) are done, the NASA Ames Research Center should be mentioned [11, 12]. Jorgensen et al. use wavelet transform or linear predictive coding as the parameterization method, and they use artificial neural networks, SVM (support vector machines), and HMM (hidden Markov models) as classifiers. For the limited set of five voice commands, they obtained the scores of the order of 90%. For the set of English phones and female voices, the scores were lower – at the level of 50%. Among applications designed by NASA, beside voice commands controls (e.g., for remote steering of application or robot), also speech recognition in difficult conditions should be mentioned (e.g., acoustic – with high noise level or with speech pronounced when the gas mask is on).

Research carried out by Schultz et al. [13–15] at Karlsruhe University can be considered as the most crucial in the last decade because they tried to overcome the speaker dependence on one hand and the limitation of the dataset on the other. The collected signal database (EMG-PIT) was recorded for 78 speakers with the use of six electrodes placed on the speaker's face and neck. For the signal parameterization, the parameters of time domain were applied (mean value in frame, energy, crossing zero density). For the sample of 14 speakers (10 sentences each), they obtained 47.15% of WER (word error rate).

Manabe et al. [16, 17] carrying out research at the Japanese NTT DoCoMo used a novel technique of setting of the EMG electrodes. In the proposed technique, three electrodes are set on the speaker's fingers which he keeps on his face. In their preliminary tests with the neural network and signals energy, they reached around a 90% recognition score for five vowels. In their later works where more sophisticated signal parameterization techniques (MFCC, LPC, etc.) and HMM as classifier were used, the score at the level of 60% for several comments in Japanese was obtained.

The research on EMG speech signal recognition was also carried out by Instituto de Investigacion en Ingenieria de Aragon and Dpt. De Informatica e Ingenieria de Sistemas [18]. The prepared system uses eight electrodes detecting signals of face muscles. For three voices and 30 syllables of Spanish, they reached a proper classification at the level of 69%.

In the research of Bu et al. [19] instead of standard bipolar electrodes put on a singular muscle, they applied a differential signal between two unipolar electrodes put on different muscles, which let them use less number of electrodes to obtain similar recognition results. In the classification tests of several Japanese phones, they reached the score of 90.6%.

The parameterization and classification techniques used for the EMG speech signal recognition are presented in **Table 1**. Despite the fact that there are crucial differences between acoustic

and EMG signals, most of them are used also in automatic acoustic speech and speaker recognition systems.

Parameterization	Positions	Classification	Positions
WT (Wavelet Transform)	[11, 12]	HMM (Hidden Markov Models)	[11, 13, 14, 16]
STFT (short time Fourier transform)	[12, 13, 15, 18]	ANN (Artificial Neural Network)	[12, 19]
LPC (Linear Predictive Coding)	[12, 13]	SVM (Support Vector Machines)	[12]
Time-domain and zero-crossing parameters	[14, 15, 18,]	DTC (Decision Trees Classifier)	[14, 18]
Cepstral parameters	[13, 16–18]	LDA (Linear Discriminant Analysis)	[15]
Phonetic parameters	[14]		

Table 1. Parameterization and classification techniques used for the EMG speech signal recognition described in the literature.

5. Automatic recognition tests

The block diagram of general speech recognition procedures is presented in **Figure 9**. The digital EMG signal is given to the front end of automatic speech recognition system. The system also uses the signals stored before in the EMG signals database. The task of the front-end procedure is to obtain the parametric picture of the registered digital EMG signal that can be later recognized in the classification process. The classification consists of two stages: the training and recognition.

The preliminary tests were carried out for the EMG recordings of six Polish vowels [a, e, i, o, u, y] and eight isolated words [stop, start, dalej, v levo, v pravo, pauza, enter, tak] [vowels and words are given in the SAMPA (Speech Assessment Methods Phonetic Alphabet) notation]. The tested material was recorded with two types of electrodes (cup-shaped and self-adhesive) during one recording session of one speaker. Since the problem of speech recognition from the EMG signal is still very new, in the literature there is no commonly accepted set of best features for that kind of task. Most features applied and described in the literature are copied directly from nowadays well-developed techniques of speech recognition based on the acoustic signal. The EMG signal which has completely different physical origins than the acoustic one has also a completely different nature both in frequency (limited to substantially lower frequency range) and time domain (longer time intervals caused by preparation and relaxation of muscles for each utterance than for acoustic signal). Since the selection of most efficient features set is a big task, it demands longer studies and tests. In the preliminary tests, the feature vector was the combination of spectral MFCC (mel frequency cepstral coefficients) and LPC (linear predictive coefficients) features. In the following test results, the 76 parameters were applied (average and deviation for each): 13 MFCC, 10 LPC, 5 spectral moments, and 10 area method of moments of MFCC. MFCC and LPC are the most common features applied in speech technology; additionally the spectral variability was accounted for by using the FFT (fast Fourier transform) moment and its modification for MFCC.
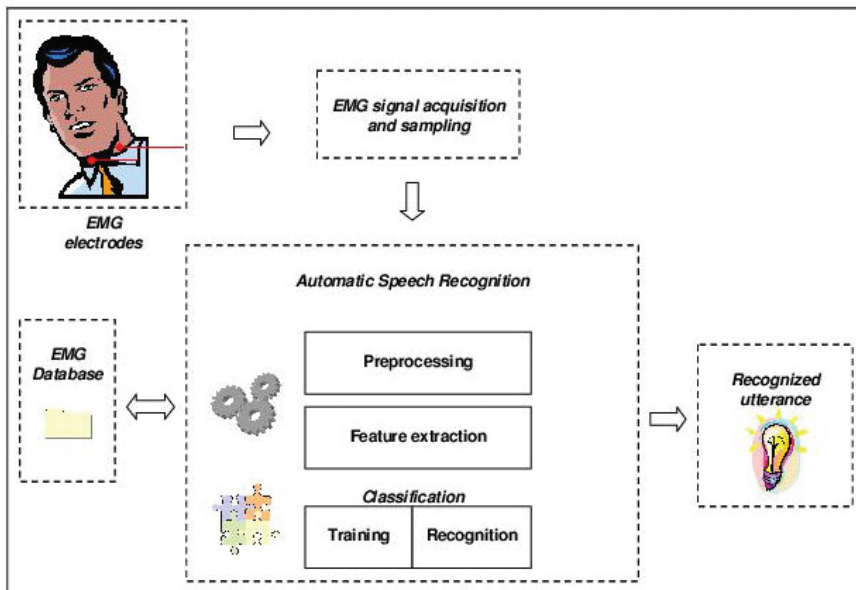
Figure 9. Block diagram of speech recognition procedures.

The window size for feature vector was set up experimentally. The tests were carried out for recognition of six Polish vowels with the Bayes network classification. Due to a small number of instances, instead of dividing them into training and testing sets, a cross-validation was applied. **Table 2** presents the results of correctly classified instances in the function of window size (represented in samples and corresponding time intervals). Kappa is a chance-corrected measure of agreement between the classifications and the true classes. It is calculated by taking the agreement expected by chance away from the observed agreement and dividing by the maximum possible agreement. The windows overlap was set to 0.5.

Window size [samples]	Window size [ms]	Correctly classified instances	Kappa
128	8	90.70%	0.89
256	16	93.02%	0.92
512	32	97.67%	0.97
1024	64	93.02%	0.92

Table 2. Results of correctly classified instances in the function of window size.

Since the best results were obtained for the window size of 32 ms, the following tests were carried out for this value.

Table 3 presents the overall results of correctly classified utterances for vowels, words, and three tested classifiers (Bayes network, Naive Bayes, and multilayer perceptron). Some conclusions from the obtained results can be drawn.

Classifier	Correctly classified instances			
	Vowels		Words	
	Cup-shaped electrodes	Self-adhesive electrodes	Cup-shaped electrodes	Self-adhesive electrodes
Bayes network	97.67%	90.70%	97.40%	96.10%
Naive Bayes	95.35%	83.72%	93.51%	92.21%
Multilayer perceptron	93.02%	79.07%	80.52%	84.42%

Table 3. Overall results for vowel and word recognition.

6. Conclusions

The level of the EMG signals obtained from the applied electrodes is sufficient for further analysis for most cases.

Subjects, who during the first trial did not have the correct EMG signal, may learn to speak such words to be able to record the signal. After a few minutes of training, it was possible to speak the words correctly, so the EMG signals could be recorded. It should be noted that the way of speaking words was associated with an unnatural speaking out, with very expressive facial expressions.

Subjects involved in the research can be divided into three groups:

- People who were possible to speak in such the way that the registered EMG signals could be done at once. The expressive facial expressions allow to record signals with a higher amplitude.

- People who were able to produce the sufficient for recording EMG signals after training. The expressive facial expressions allow to capture data with normal values.

- People for whom EMG signals could not be registered even after training.

The great difficulty in the measurement process was to obtain the undisturbed EMG signals. Each movement of the speaker was reflected in the EMG signal parameters. Particularly, a strong interference was generated by the reflex of swallowing saliva – the signal is two to three times greater than the speech signal. It is important to ensure the adequate measurement conditions – separate from the power supply, and other signals that can be picked up by the electrodes.

The best scores during automatic recognition tests were obtained for the Bayes network, but the differences between chosen classifiers are considerably insignificant. Very similar results

were obtained for words and vowels (in vocal speech signal recognition, the vowel recognition would gain better results than word recognition).

The difference between electrodes indicates that the influence of electrodes type or their position can have a significant impact on final scores (recording for both types of electrodes were carried out during the same recording session). The difference was quite big in the case of vowels recognition (difference around 10%) and considerably insignificant in case of words recognition. Considerably high results of correctly classified instances are partly caused by a session-dependent and a speaker-dependent case.

Author details

Andrzej B. Dobrucki*, Piotr Pruchnicki, Przemysław Plaskota, Piotr Staroniewicz, Stefan Brachmański and Maciej Walczyński

*Address all correspondence to: andrzej.dobrucki@pwr.edu.pl

Department of Acoustics and Multimedia, Faculty of Electronics, Wrocław University of Technology, Wrocław, Poland

References

[1] Basmajian JV, De Luca CJ. Muscles Alive, Their Function Revealed by Electromyography. Baltimore: Williams Wilkins; 1985.

[2] De Luca CJ, Knaflitz M. Surface Electromyography: What's New?. Torino: C.L.U.T.; 1992.

[3] Konrad P. The ABC of EMG. A Practical Introduction to Kinesiological Electromyography. USA: Noraxon U.S.A. Inc. ; 2005.

[4] Hermens HJ, Freriks B. SENIAM 5. The State of the Art on Sensors and Sensor Placement Procedures for Surface ElectroMyoGraphy: A Proposal for Sensor Placement Procedures. The Netherlands; 1997.

[5] Winter DA. Biomechanics and Motor Control of Human Movement. New York: John Wiley & Sons; 1990.

[6] MettingVanRijn AC, Peper A, Grimbergen CA. Amplifiers for bioelectric events: a design with a minimal numbers of parts. Medical & Biological Engineering & Computing. 1994, 32: pp. 305–310.

[7] Benning M, Boyd S, Cochrane A, Uddenberg D. The Experimental Portable EEG/EMG Amplifier, University of Victoria Faculty of Engineering, ELEC 499A Report; 2003.

[8]   Morse MS, O'Brien EM. Research summary of a scheme to ascertain the availability of speech information in the myoelectric signals of neck and head muscles using surface electrodes. Computers in Biology and Medicine. 1986, 16, no. 6: pp. 399–410.

[9]   Jorgensen C, Dusan S. Speech interfaces based upon surface electromyography. Speech Communication. 2010, 52: pp. 354–366.

[10] Sugie N, Tsunoda K. A speech prosthesis employing a speech synthesizer – vowel discrimination from perioral muscle activities and vowel production. IEEE Transactions on Biomedical Engineering. 1985, 32, no. 7: pp. 485–490.

[11] Betts BJ, Binsted K, Jorgensen C. Small-vocabulary speech recognition using surface electromyography. Interacting with Computers. 2006, 18: pp. 1242–1259.

[12] Jorgensen C, Binsted K. Web browser control using EMG based sub vocal speech recognition. In: Proc. of the 38th Hawaii International Conference on System Sciences; 2005.

[13] Maier-Hein L, Metze F, Schultz T, Waibel A. Session independent non-audible speech recognition using surface electromyography. In: IEEE Workshop on Automatic Speech Recognition and Understanding, San Juan, 27–27 November, pp. 331–336; 2005.

[14] Schultz T, Wand M. Modeling coarticulation in EMG-based continuous speech recognition. Speech Communication. 2010, 52: pp. 341–353.

[15] Szu-Chen J, Schultz T, Walliczek M, Kraft F, WaibelA. Towards continuous speech recognition using surface electromyography. In: Proc. Interspeech 2006, pp. 573–576, Pittsburgh, Pennsylvania, September 17–21; 2006.

[16] Manabe H, Zhang Z. Multi-stream HMM for EMG-based speech recognition. In: Proc. of the 26th Annual International Conference of the IEEE EMBS, San Francisco, CA, USA, September 1–5; 2004.

[17] Zhang Z, Manabe H, Horikoshi T, Ohya T. Robust methods for EMG signal processing for audio-EMG-based multi-modal speech recognition. In: COST278 and ISCA Tutorial and Research Workshop on Robustness Issues in Conversational Interaction, University of East Anglia, Norwich, UK, August 30–31; 2004.

[18] Lopez-Larraz E, MozosO M, Antelis JM, Minguez J. Syllable-based speech recognition using EMG. Conference Proceedings of the IEEE Engineering in Medicine and Biology Society. 2010, 1: pp. 4699–4702.

[19] Bu N, Tsuji T, Arita J, Ohga M. Phoneme Classification for Speech Synthesizer using Differential EMG Signals between Muscles. In: Proc. of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, September 1–4; 2005.

[20] Staroniewicz P, Brachmański S, Dobrucki AB. Subvocal speech recognition based on electromyographic signal [in Polish: Rozpoznawanie mowy subwokalnej w oparciu o

sygnał elektromiograficzny], Przegląd Telekomunikacyjny, Wiadomości Telekomuni-
kacyjne No. 6; 2013, pp. 571–574.

[21] Mendes J, Robson RR, Labidi S, Barros AK. Subvocal speech recognition based on EMG
signal using independent component analysis and neural network. In: Congress on
Image and Signal Processing; 2008.