ARCHIVES OF ACOUSTICS
33, 1, 133–139 (2008)

THE ELECTROMAGNETIC MICROWAVE SENSORS IN IMPROVING SPEECH INTELLIGIBILITY AFTER LARYNGECTOMY

Mariusz MIKOŁOWICZ

Warsaw University of Technology
Institute of Radioelectronics
Nowowiejska 15/19, 00-665 Warszawa, Poland
e-mail: m.mikolowicz@olikon.com.pl

(*received June 15, 2007; accepted November 30, 2007*)

The paper contains a short review of analytic, constructional and development investigations of low power electromagnetic microwave sensors (EMS) destined for applying them in laryngectophone system. General procedure rule and construction examples are presented. The measurement stand used to research signals produced by sensors of articulators' motion are described. Analytic method used to examine the usefulness of these signals to identify the vibrant phonemes is reviewed.

Keywords: EMS (electromagnetic microwave sensors), Laryngectophone, Microwave sensor, speech.

1. Introduction

During scientific work on absent larynx people speech regeneration methods it has been established that solely acoustic identification methods of its lost vibrant fragments probably do not exist. Consequently, it is impossible to regenerate it without speaker active collaboration or tools showing articulators movements inside human voice tract. Medical, technical and ergonomic sense stays with uninvasive solutions harmless to health like using microphones and low power microwave sensors [3] within one measurement system. The system's task is to register exemplary signals produced by sensors to compare it with samples of various records arranged randomly. By building the whole phoneme, diphone, triphone and complete words exemplary matrix and measuring identification grade of spoken phonemes, the effectiveness grade of device so called laryngectophone [2] can be investigated.

2. EMS sensor's features

Conventional radar sends out short burst of single-frequency (narrow-band) electromagnetic energy in the microwave frequency range. Other radars step through multiple (wideband) frequencies to obtain more information about a scene. An impulse or ultrawide-band radar sends individual pulses that contain energy over a very wide band of frequencies. The shorter the pulse, the wider the band, thereby generating even greater information about reflected objects. Because pulse is so short, very little power is needed to generate the signal. EMS is unique because it inexpensively generates and detects very fast pulses – about nanoseconds. The drawback of using short and low-power pulses is that less energy can be measured on the radar returns. This problem was solved by transmitting many pulses rapidly and averaging all returns [1].

According to its physical properties, EMS has such features as:

- power consumption is about microamperes, because current is draw only during short pulses;
- the power of transmission is about tens of microwatts so keeps medical safety;
- the receiving reflections contain much information;
- because of sensor's operating across a wide band of frequencies it is little susceptible to interferences form other sensors and other electronic equipment.

3. Sensors used during measurements

3.1. Mini RF vibration sensor

The mini RF vibration sensor (MRFV) is a pulse-echo radar that can sense the speed and direction of motion (Doppler effect), mechanical vibrations (phase modulation), and electronic signatures (active reflections). The vibrations can be detected in "stereo" from bio-activity (e.g., vocal cords). By integrating inbound or outbound signatures (with external processing), a target's displacement or motion through distance can be sensed without false alarms from air conditioners, fluorescent lamps or loitering people. Inbound motion produces an upward Doppler shift (upper sideband or USB) and outbound motion produces a downward Doppler shift (lower sideband or

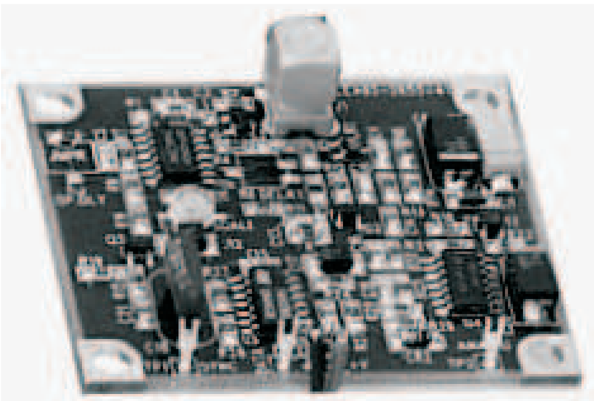

Fig. 1. Photo of vibration sensor.

LSB). The MRFV resolves these Doppler components into quadrature channels. An external baseband quadrature combiner or the software equivalent can then resolve the lower and upper sidebands into separate channels.

3.2. Micropower motion sensor

The microPower Differential Pulse Doppler (μP-DPD) motion sensor uses the Doppler principle to detect motion, time gating to limit the detection range, and a special technique to improve sensitivity within the gated region. It does not respond to objects outside its range gate and does not cause false alarms on nearby small objects, such as insects.

As a target moves towards or away from the μP-DPD sensor, one sinewave cycle is generated for each half-wavelength the target moves radially to the antenna. If the target moves 254 cm/s, the analog output will produce a 100 Hz sinewave. The Doppler passband supports speed measurements from 4.1 to 406.4 cm/second, or 0.148 to 14.63 km/h, and a wider range at reduced sensitivity. The Doppler passband can be customized by changing RC components on the circuit board. Direction of motion, in-bound or outbound, cannot be detected with the μPDPD since the same sinewaves are produced in either direction.

Fig. 2. Photo of motion sensor.

4. Measurement stand

The measurements stand consists of [4]:

- EMS sensors installed on the support stand and directed to chin, mouth and larynx (auxiliary sensor during research with healthy patients); also there are microphone and loudspeaker performing communication during measurement procedure; patient opens and closes each registration procedure using computer keyboard.
- measurement card – card gathering signals from all sensor block elements; connected to the computer via USB;
- computer and program controlling measurement card and loudspeaker.

The range adjustment on jaw and lips sensors was set to 30 cm: there is no response to any motion at ranges > 30 cm. Range to lips and jaw (directly under chin) is about 10 cm.

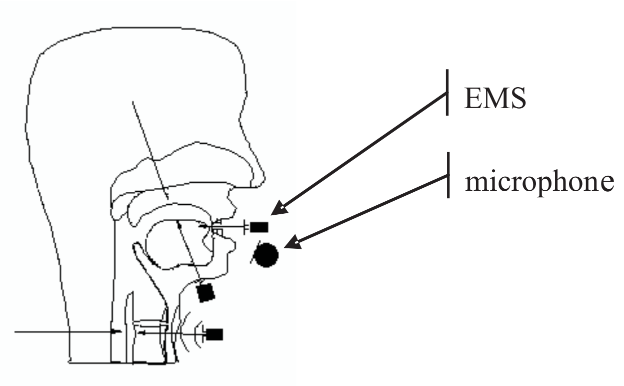

Fig. 3. Sensors' configuration in relation to a patient's head.

5. Methods of measurements and computation

5.1. Data acquisition

During measurement procedure patients speak indicated phonemes (or words) chosen from the collection of Polish phonemes.

The measured signals are recorded during sessions of pronunciation of whole phoneme series (or some isolated words) chosen from the collection of Polish phonemes arranged pseudo-randomly. The whole procedure is reviewed shortly below:

- Registration series process which consists in several stages:
  - computer serves randomly chosen phoneme to the loudspeaker,
  - patient pushes release button,
  - patient pronounces desired phoneme,

  - patient leaves release button,
  - if there is lack of some phoneme, the sequence is repeated.
- Series can be paused several times, last recording can be repeated.
- Every phoneme is registered only one time.

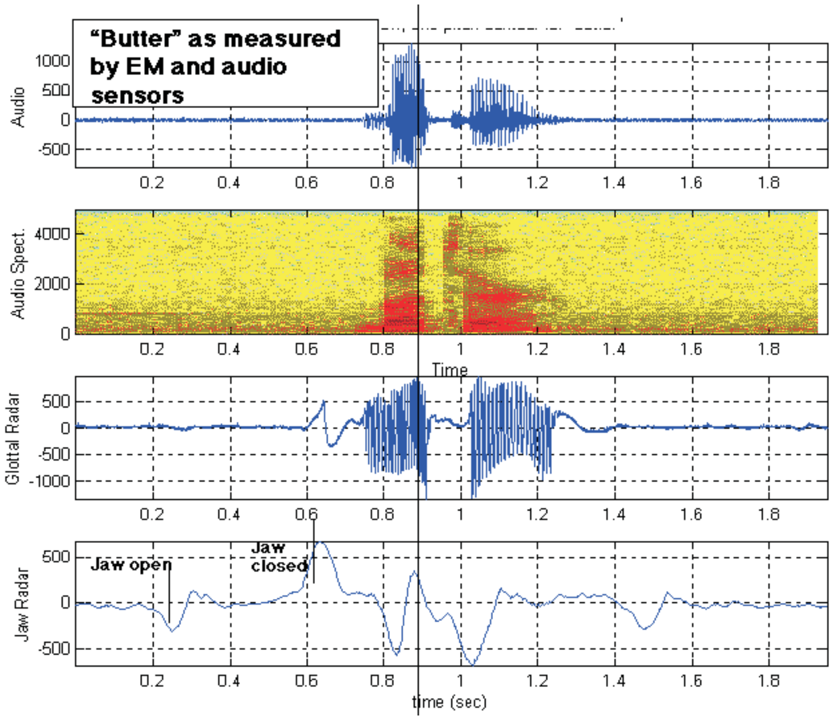

Fig. 4. Example of signals from glottal and jaw sensor.

5.2. Data analysis

The purpose of result analysis is to prove, that each phoneme and word spoken by patients are represented by different signals gathered from microphone and EMS sensors. Comparison of all gathered samples, establishment of convergence grade between samples of the same phonemes as well as divergence grade of different phonemes needs to be made to attain the objective [4].

The registered signals (enclosing separated phonemes) are ordered into 20 ms time interval vectors. For example lips vector:

$$L_n = \begin{bmatrix} l_1 \\ \vdots \\ l_x \end{bmatrix},$$

where n – index of time interval; l_1, l_2, ... – samples from 20 ms interval, gathered from lips EMS; x – index of last sample in current time interval.

Vectors ordered by sensors' index and phonemes' index create arrays indexed by successive 20 ms intervals and phonemes:

$$\sum_{n=1\ldots k}^{m} L_n = L^m = \begin{bmatrix} l_{11} & \cdots & l_{1n} & \cdots & l_{1k} \\ & \ddots & & & \\ \vdots & & \iddots & & \vdots \\ & & & \ddots & \\ l_{x1} & \cdots & l_{xn} & \cdots & l_{xk} \end{bmatrix}^m ,$$

where m – index of phoneme; k – index of last time interval of m-phoneme; l_{1n} – first sample from n-interval, gathered from lips EMS; x – index of last sample in time interval.

Every phoneme (and di-, tri-, word) is represented by set of such arrays: glottal, jaw, lips, microphone:

$$P^m = \{G^m, J^m, L^m, M^m\} ,$$

where m – index of phoneme; G^m – glottal array; J^m – jaw array; L^m – lips array; M^m – microphone array.

Genetic algorithms are used to searching out phonemes that are the most similar to gathered phoneme.

6. Conclusions

The constructions of EMS and method used to examine them has been presented. There are two main problems waiting for solution: search algorithms' performance and right number of sensors used to watch out human articulators. The high effectiveness of calculations is expected due to not high performance of mobile CPUs, which are naturally dedicated for laryngectophone's construction. The EMS' number is related to number of articulators which are observing.

References

[1] Azevedo S., McEwan T. E., *Micropower impulse radar*, Science & Technology Review, January/February (1996).

[2] Holzrichter J. F., Burnett G. C., Ng L. C., Lea W. A., *Speech articulator measurements using low power EM-wave sensors*, Journal of Acoustical Society of America, vol. 103, January (1998).

[3] Mikołowicz M., *Concept of the system for speech improvement after laryngectomy – laryngectophone* [in Polish], Proc. of IX$^{\text{th}}$ Symposium "New Trends in Audio and Video", pp. 137–145, Warszawa 2002.

[4] Mikołowicz M., *Implementation and planning test of microwave electromagnetic sensor designed for use in the laryngectophone system* [in Polish], Proc. of X$^{\text{th}}$ International Symposium of Sound Engineering and Tonmeistering, pp. 159-163, Wrocław 2003.

[5] Mikołowicz M., *Experimental stand for testing the use of the microwave electromagnetic sensors applied in the laryngectophone system* [in Polish], Proc. of X$^{\text{th}}$ Symposium "New Trends in Audio and Video", pp. 119-126, Wroclaw 2004.