

The Research of Correlation Between Pitch and Skin Galvanic Reaction at Change of Human Emotional State.

A.V.Agranovski, O.Y.Berg, D.A.Lednov

SPETSVUZAVTOMATIKA DESIGN BUREAU
51 Gazetny St., Rostov-on-Don, Russia,
e-mail: asni@ns.rnd.runnet.ru

1. Introduction and description of the measuring device

Reflection of the emotional state of humans in their speech was the subject of numerous publications, as can be seen from the review [1].

However, this issue remains open for discussion, since it is still unclear, how true is such reflection of the emotional state through the modification of the speech characteristics.

Earlier the most popular method consisted in stage acting, e.g., in [1,2], whereby the actor was the source of the preset emotional colouring of the voice. This method is intrinsically subjective and can hardly be relied upon.

The problem can be solved by a simultaneous use of a technique for speech characteristics measurement, and another, alternative, technique known valid.

As the latter technique, the authors chose the method for measuring the skin galvanic reaction (SGR), proven good in practice [3].

For measurements we used a standard Sound Blaster Pro.4.0 to make the hardware for input of the SGR and speech into the computer via two parallel channels. The block diagram of the device is shown in Fig.1. SGRs are recorded by means of direct

current based on the Farey's method.

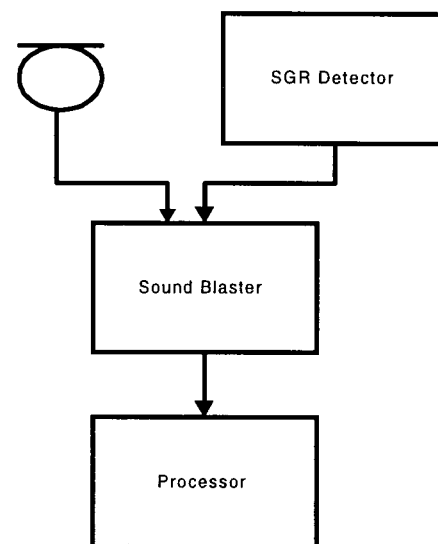


Fig.1

The block diagram of the SGR detector is shown in Fig.2.

The magnitude of the current passing through the electrodes is fixed at the level of 15 μA , since in the current range of 10-40 μA the skin resistance is not dependent on the current value, while the electrodes' polarization is insignificant [5]. The dynamic range of measurement of the SGR is preset within the limits of 10 $\text{k}\Omega$ - 1 $\text{M}\Omega$. To ensure the electrical safety of the subject, all circuits connected to the subject are galvanically isolated.

The differential amplifier of biosignals 2 (DAB) provides for amplification of bipolar biovoltages,

effective suppression of the common-mode interference, and galvanic isolation of the input and output circuits. To isolate the subject from the power circuits, the supply power transforming converter 1 is used.

From the DAB the voltage proportional to the SGR amplitude is supplied to the inverting input of the amplifier 3 to be amplified to a level required for coordinated operation with the automatic compensation circuit and convenient control of the SGR level.

The automatic correction circuit 5 is used for compensation of the constant component of the biovoltages. From the output of the amplifier 3 the signal is supplied to the comparators' unit 6, which enables, disables and determines the direction of count of the counter 7. The digital-to-analog converter 8 converts the readings of the counter into direct current voltage. This voltage is then supplied to the direct input of the amplifier 3. The level of the compensating voltage is controlled at the output 2.

The second-order low-frequency filter 4 with the cut-off frequency of 30 Hz limits the range of the dynamic component of the SGR and additionally suppresses the noise and industrial interference.

The sound card Sound Blaster Pro 4.0 for IBM PC/AT computers has a first-order top frequency filter, limiting the input signal band by the magnitude of 20 kHz, therefore, if the SB is to be used as a SGR input device, it must be transformed, e.g., by amplitude modulation. This is achieved through the use of the modulator 8 and the carrier generator 9.

Speech and SGR were each assigned 8 digits, and the digitization frequencies were respectively 5 kHz and 200 Hz.

The pitch interval duration was used as the characteristic index of the speech state.

2. Description of the algorithm

For preliminary definition of the pitch duration we used the technique described in [4]. Briefly, this technique consists in that the speech is filtered by means of five Chebyshev bandpass filters. The bandwidths of the filters evenly cover the frequency range from 100 to 400 Hz. The instant of synchronical appearance of the minimums of the filters' output functions is the reference point for the pitch. The interval between two nearest counts is used as the preliminary estimate of the pitch duration. Twenty duration values are then accumulated and used to draw a bar graph of the occurrence rate for the interval with a certain duration. This bar graph is used to calculate the mean duration value T .

The preliminarily analysed section of speech is then analysed again to obtain more accurate results. Based on the repeat analysis, an envelope is constructed, the value of which at each discrete moment of time t will be designated as $Y(t)$. The following conditions are imposed on the reference points and the pitch pulse lengths.

The moment of time t is the reference point for the pitch, provided

$$a) t = \operatorname{argmax} M(t)$$

where $M(t)$ - the maximum value of the speech discrete function $f(t)$, at $f(t) > 0$;

$$b) | \max M(t) - Y(t) | < Q$$

here Q is the threshold, $Y(t)$ is the speech envelope.

The pitch duration L_i is taken account of for subsequent analysis (i - pitch ordinal number), provided

$$a) | L_i - L_j | < 2,$$

where L_i is measured as the quantity of the numerical speech counts;

$$b) T - d < L_i < T + d,$$

d is the threshold.

The pitch duration values thus obtained were then used to draw the bar graph for the whole phrase.

3. Experimental data

It is known that the SGR amplitude values at false answers exceed the amplitude values at true answers [3]. In this connection we developed the following technique.

Ten announcers (male) were asked to answer six questions. Each of them had to provide five true answers and one false answer, under the requirement to answer in full. Each of the announcers was subjected to five tests, whereby the sets of questions were changed.

For each answer we constructed the diagram of the occurrence rate of the preset-value pitch interval, and recorded the maximum amplitude value of the SGR.

Here we present the most typical bar graphs for one person based on a sequence of tests. The "Parameters" box of the Fig.3 shows the frequencies of the corresponding mode maximums in hertz.

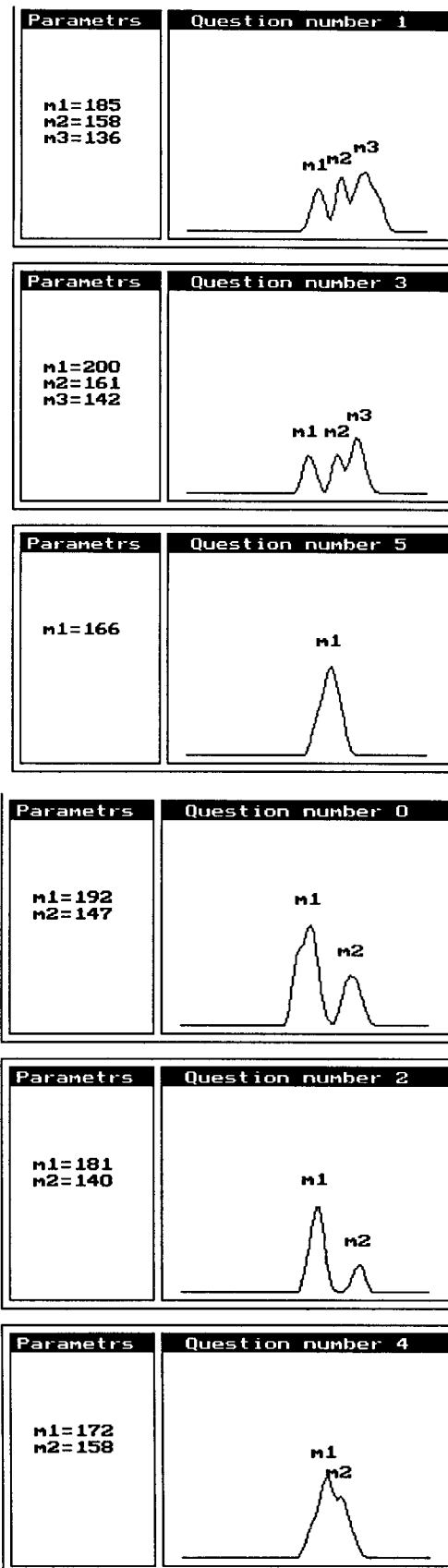


Fig.3

It was noted that the diagrams of the pitch intervals for true answers supported by SGR show marked polymodality, see Fig.3 (questions from 0 to 4).

Our studies showed that such appearance of the diagrams is due to that the pitch duration in stressed syllables differs from that in unstressed syllables (at the stressed syllable pitch frequency of 125 Hz, this difference was 16%).

Diagrams of the false answers, see Fig.3 (question 5) are identified by merging of the modes and a tendency to monomodality.

In 70% of false answers supported by SGR data we observed monomodality.

References

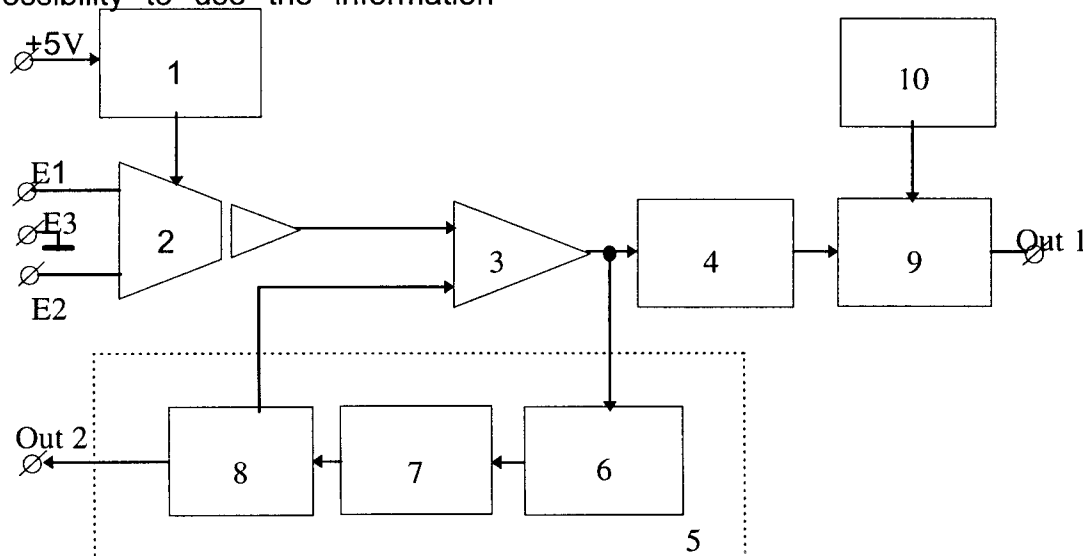
1. Frolov M.V. Control of the functional state in a human operator. M.:Nauka Publ., 1988.
2. Blokhina L.P. On the possibility to use the information

about the pitch frequency modulations for automatic recognition of emotional states. Proc.,rep. and pres. at the 5th All-Union Conference/Symposium on the Speech and Hearing Acoustics Cycle. Emotions and automatic recognition of speech, Odessa, 1989, p.41-42.

3. Ivanov Y.K. History, theory and practice of studies of skin galvanic reactions in man. Preprint 74-9, Kiev, 1974.

4. Agranovsky A.B., Lednov D.A. Variation of the Fundamental Tone Characteristics with the Emotional Changes in Man. Text Proc.of the SPECOM'96. St.Petersburg, Oct.1996.

5. Okushko V.R., Shugailov I.A., Salabai V.D. A device for recording the psychoemotional state and the pain sensitivity levels in a patient. Med.tekhnika, 1992, No.3.



1. voltage converter 2.biosignal amplifier with galvanic isolation 3.correction amplifier 4.low-frequency filter 5.correction circuit 6.comparators' unit 7.counter 8.digital-to-analog converter 9.modulator 10.carrier generator E1, E2 pickup electrodes E3 neutral electrode Out 1 output of modulated signal Out 2 output of the compensating voltage level

Fig.2 Block diagram of the device for recording of the skin galvanic reaction