UDK 534.4:621.391

A.O. Chugay

Oles Honchar Dnipropetrovsk National University

SOFTWARE ENVIRONMENT OF ANALYSIS OF SPEECH WITH PURPOSE OF RECOGNITION OF THE FUNCTIONAL STATE OF MAN

Викладена послідовність обробки інформативних мовних сигналів і вилучення компонент, які характеризують відхилення функціонального стану людини: динаміка частоти основного тону голосу, темпоральні характеристики для розпізнавання стану.

Ключові слова: *мовний сигнал, вилучення компонент, основний тон мови, функціональний стан, темпоральні характеристики.*

Изложена последовательность обработки информативных речевых сигналов и выделения компонент, которые характеризуют отклонения функционального состояния человека: динамика частоты основного тона голоса, темпоральные характеристики для распознавания состояния.

Ключевые слова: *речевой сигнал, выделение компонент, основной тон речи, функциональное состояние, темпоральные характеристики.*

In the article is following Processing of the information about and speech Signal and the Extraction of the Components with typical Deviations of and functional State of the talking Man: and Dynamic of the Frequency of the ground Tone of the Speech, and temporal Characteristics considered and the Recognition of the State.

Keywords: *speech Signal, Extraction of the Components, ground Tone of the Speech, functional State, temporal Characteristics.*

Problem of operative estimation of the functional state of man in the process of activity by an incontact method and, in particular, on the dynamic parameters of acoustic characteristics of his spontaneous speech with every year becomes more topical. Especially it touches upon operator activity becoming more complicated automated human-machine systems. The fatigue and emotional tension belong to the number of the functional states of operator. These states appeare, as a rule, after the protracted work, as a result of operational tension, stressing influences, inadequate terms of activity and other.

We consider that the research of the functional state of a speaker in real time must be conducted with the use of wide enough set of informative parameters of acoustic characteristics of spontaneous speech with the

АКТУАЛЬНІ ПРОБЛЕМИ АВТОМАТИЗАЦІЇ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ. Том 14. 2010

purpose of subsequent combination of informative indexes of the functional state in general integral estimation. But for all that the main attention should be paid not so much to the absolute indexes of the selected informative parameters of acoustic characteristics as to their changes in time relative to some initial indexes of the functional state of the analyzed operator.

It is set by the researches that the acoustic parameters of complete fragment of the spontaneous speech are diagnostically informative [1]. In according to an acoustic theory of speechformation [2; 3] speech is the result of co-operation of resonance cavities of the vocal tract and cvaziperiodical impulses of the vocal copulas at formation of voiced sounds of the turbulent noise appearing in the places of narrowing in a vocal tract at unvoiced sounds.

Thus regenerated speech in a vocal signal on the output of the mouth can be represented as the furl of functions of signal of source $s(t)$ and impulsive characteristic of vocal tract $h(t)$: $f(t) = s(t) * h(t)$.

The spectrum of vocal signal is determined through transformation of the Fur'e function $f(t)$:

$$F(\omega) = F[f(t)] = S(\omega)H(\omega),$$

where $s(\omega)$ – is a spectrum of the source; $H(\omega)$ – transmission function of the vocal tract.

An aperiodic function $f(t)$ and its complex spectrum $S(\omega)$ is interconnected by the Fur'e transformation:

$$S(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt \ ;$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) e^{j\omega t} d\omega$$

However speech, being the product of psychical activity, is formed in accordance with the linguistic rules of the language by the speechform system guided by the sequence of commands of the cerebrum. The levels of speechform determine the blocks of psycho-lingual and psycho-acoustic presentations of the speech as some physical substance, which is characterized by the physical structure and bearing the complete psycho-linguistic reflection of vocal utterance. Accordingly, the parameters of speechform are formed at both psychical and psycho-acoustic level [4].

Thus, a vocal signal has the appearance of a complex multidimensional function $s(\Omega_1,...,\Omega_n,\omega,t)$, where $\Omega_1,...,\Omega_n$ – frequencies of the component of vocal signal with unknown operations of their co-operation, ω – frequency of the spectral parameters, t – current time.

1. Components of unstationary of vocal signal.

Changeability of vocal signal is conditioned in a number of factors. The basic of them are following:

a) a vocal signal is a casual unstationary temporal process;

b) variety of speech, conditioned by the transformation (reduction), transcription;

c) individuality;

d) dialect;

e) state of announcer and changeability of the state.

Each of these factors is included in a vocal signal as component part with the unknown method of co-operation of these parts with one another.

We will assume that it is possible to find spectral–temporal description of vocal signal $S(\omega,t)$. All transferred factors will be present as

$$S(\omega,t)=S\,[S_r(\omega,t),S_a(\omega,t),S_c(\omega,t),$$
$$S_t(\omega,t),S_v(\omega,t),S_d(\omega,t),S_p(\omega,t),S_s(\omega,t)],$$

i.e.. $S(\omega,t)$ – complex frequency-temporal function, $s(t)$ – complex signal.

2. Algorithm of task solving.

For organization of decision initial function $s(t)$ digitizes [5,6] with frequency f_d, forming a signal $s(t_j)$, which is then divided into the intervals of analysis T_{lN} with the amount of the N reports on an interval at $l=\overline{0,n-1}$, n – amount of intervals T_{lN}, by submitting an initial signal as two-dimensional function of time as $s(\omega_k,T_l)$ (Table 2), at $j=\overline{lN,(l+1)N}$, $S(\omega_k,T_{lN})$ – spectral cut.

Table 2

T_0	$t_0 - t_{1(N-1)}$	→	$s(t_j,T_0)$	$S(\omega_k,T_0)$
T_{1N}	$t_{1N} - t_{2(N-1)}$	→	$s(t_j,T_{1N})$	$S(\omega_k,T_{1N})$
T_{lN}	$t_{lN} - t_{(l+1)(N-1)}$	→	$s(t_j,T_{lN})$	$S(\omega_k,T_{lN})$
$T_{(n-1)N}$	$t_{(n-1)N} - t_{n(N-1)}$	→	$s(t_j,T_{(n-1)N})$	$S(\omega_k,T_{(n-1)N})$

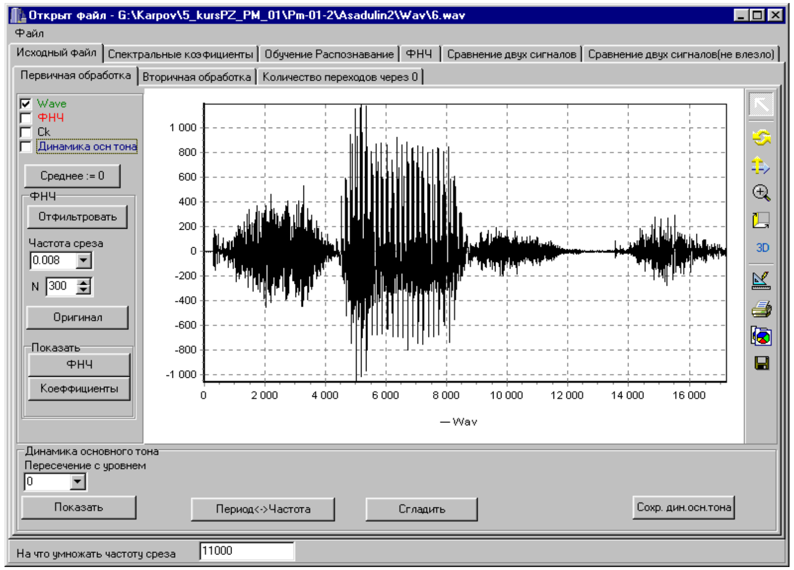

Fig. 1. Oscillogram of signal of the word «shest'»

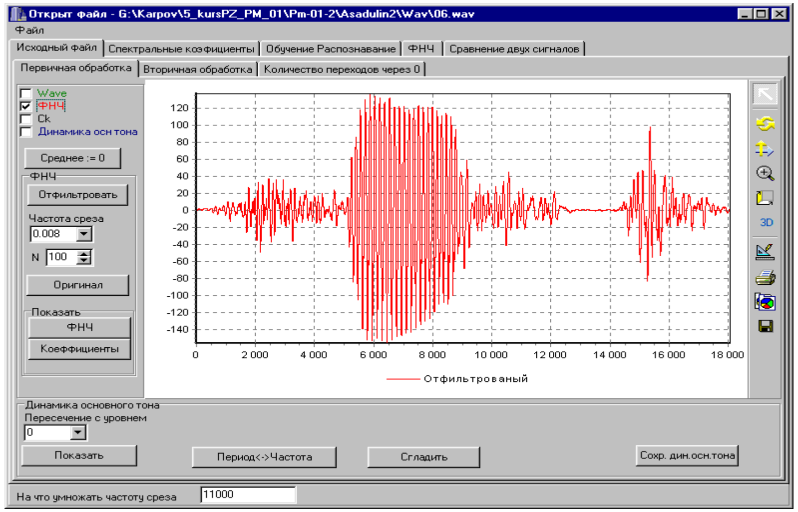

Fig. 2. Signal of frequency of the basic of the word «shest'»

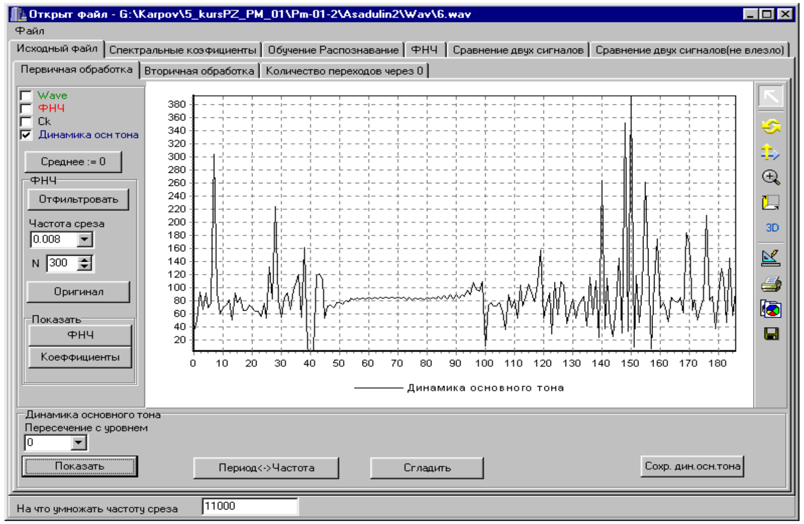

Fig. 3. Dynamic of frequency of the basic of the word «shest'»

The task solving of estimation of the functional state of a speaker by his speech can be realized as some hierarchy of levels different by the parameters attracted for the decision. It is possible to consider the zero level of recognition of words of speech (isolated words or phrases, for this variant of estimation of parameters of the functional state the problem of recognition of speech continuous together of arbitrary duration is not set).

The *first* level consists in comparison of parameters of dynamics of basic tone and temporal characteristics of speech.

The *second* level is comparison of spectral–temporal characteristics of speech utterances of the same format (words or phrases).

The first and second levels is solved for the recognized variant of format (words or phrases). Information technologies of construction of devices of speech recognition are put in [4; 6].

For the task solving of estimation of the functional state of the speaker of the first level from an initial digitazed signal $s(t_j)$ two groups of parameters are formed:

– signal with frequency of basic tone for a voice-frequency signal [7];

– temporal characteristics as amount of pauses (AP) in speeches, duration of speech (DR) on the interval of analysis, duration of pauses on the interval of analysis, the rate of articulation (RA) is attitude of duration of pauses toward duration of speech. If quantum of speech signal is the sign-variable sequence of integers with frequency of quantum

$f_{\hat{e}\hat{a}} = 22\hat{E}hz$ and digit capacity 16 bits (word structure) with the limits of values $\pm 32768 Khz$, the task solving of forming of signal of basic tone y_n (BT) the low frequency digital filter is used:

$$y_n = \sum_{k=-N}^{k=N} c_k x_{n-k}, \qquad ((c_k = c_{-k}),$$

$$c_k = \frac{1}{k\pi}\sin 2\pi k f_{cp}, \quad c_0 = 2 f_{cp}, \quad 0 < f_{cp} \le 0.5,$$

where c_k – coefficient of filter, $f_{cp} = 0.008$ – frequency of cut of filter, $x_{n-k} = s(t_j, T_{lN})$, y_n – signal with frequency of basic tone, $N = 100$, $0 < n \le N_1$, N_1 – length of vocal utterance of the set format. However, for those speech utterances which contain noisy sounds, a task becomes complicated because for most noisy sounds a voice-frequency signal does not exist (except for the mixed sounds «zh, z»). For a word «shest» the graphs have the appearance of the fig. 1– fig.3. On fig.3. by vertical lines the area of dynamics of duration of periods of basic tone for the phoneme «e» is limited. Outside this area the dynamics of low-frequencies component of loud sounds of «sh, s», unvoiced explosive sound of «t'» and unvoiced pause before him are reflected. For task solving of exposure the voice-frequency areas of speech utterance the analysis of transition through the zero of the signal got as a result of filtration of initial signal $s(t_n)$ with the help of the high-frequency filter is applied:

$$z(t_n) = s(t_n) - \sum_{k=-N}^{k=N} c_k s(t_{n-k}), \qquad ((c_k = c_{-k}),$$

$$c_k = \frac{1}{k\pi}\sin 2\pi k f_{cp}, \quad c_0 = 2 f_{cp}, \quad f_{cp} = 0.1$$

– lower frequency of cut of high-frequency filter (upper frequency of low-frequency (filter), corresponding to physical frequency $f_{cp} = 1100\,\Gamma\mu$, $N = 100$.

Graphs illustrating the process of selection of signal $z(t_n)$ and formings of function of frequency of transition through a zero $\rho(T_l)$ signal $z(t_n)$, is represented on a fig. 4–fig. 6.

As it is obvious from the graphs fig. 5-fig. 6. level of frequency of transitions through the zero of signal $z(t_n)$ for voice-frequency sounds does not exceed 15

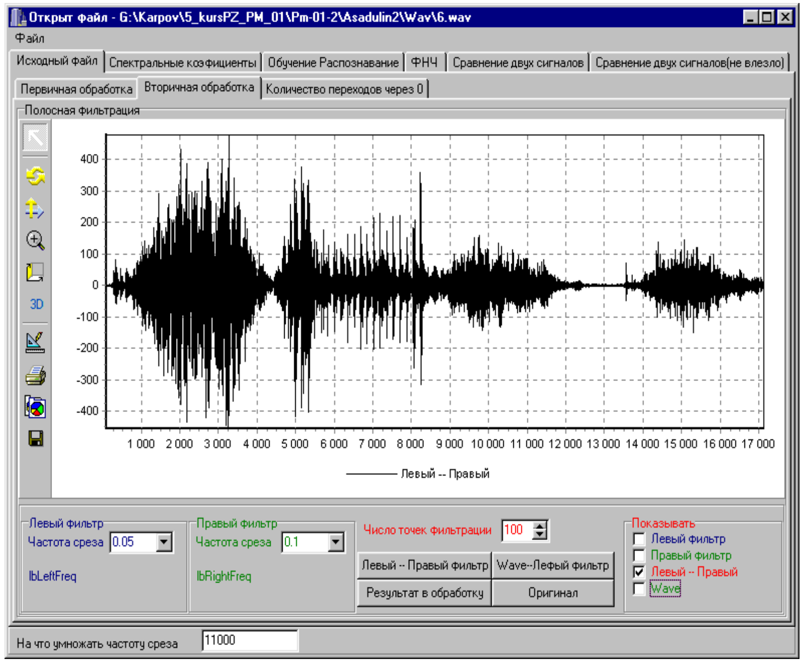

Fig. 4. Signal $z(t_n)$ of the word «shest'»

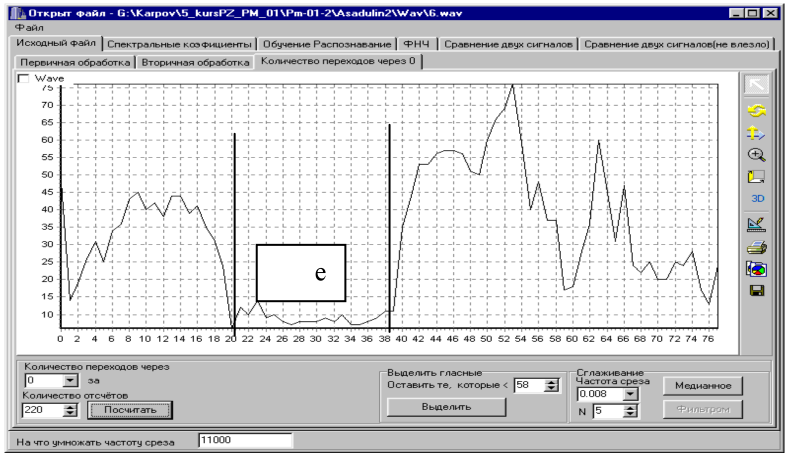

Fig. 5. Function of frequency of passing zero level $\rho(T_l)$

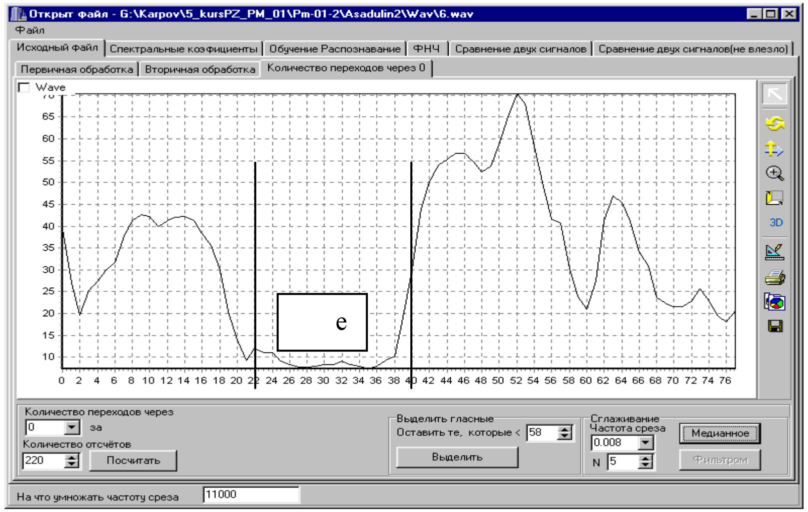

Fig. 6. Flat function of frequencies of passing zero level $\rho(T_l)$ of signal $z(t_n)$

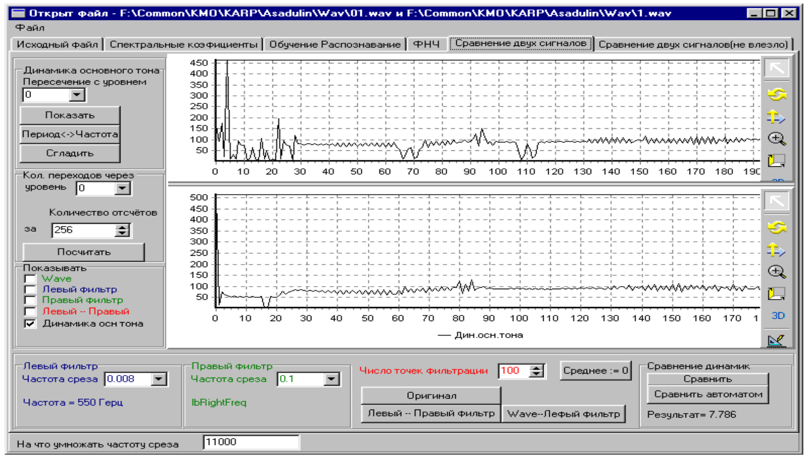

Fig. 7. Dynamic of frequency of the main tone of the word "adin" in two realizations

АКТУАЛЬНІ ПРОБЛЕМИ АВТОМАТИЗАЦІЇ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ. Том 14. 2010

At the same time for noisy sounds depending on a phoneme can reach into 60 for the interval of analysis $T_{l\,N} = 220$ counting out of signal $z(t_n)$. The task of comparison solves for the function of dynamics of frequency of signal of basic tone.

The task of comparison solves by the method of the dynamic programming only for voice-frequency part of signal, got in the temporal window on the threshold level of parameter $\rho(T_l)$.

Conclusion. By the stated algorithm the programmatic system which can be applied for estimation of declinations of parameters of vocal signals as function of the psycho physiological state of a speaker is realized.

Bibliographical references

1. **Chugay A. A.** Elaboration and research of hardwares of analysis of speech for diagnostics of the emotional state of man-operator. Abstract of the dissertation doc. – Dnepropetrovsk,1988.
2. **Fant G.** Acoustic theory of speechformation. – M. Science,1964.-284p.
3. **Sapojcov M. A.** Vocal signal in cybernetics and communication. – M.Svjazizdat.1963. – 472 p.
4. **Carpov O. N.** Technology of construction of devices of speech recognition. Monograph. – Д, – 2001. – 190 p.
5. **Carpov O. N.** Calculable charts of presentation of functions of many variables in the classes of functions of less number of variables. Monograph. – D., 2003. – 120 p.
6. **Carpov O.N., Gabovich A.G., Marchenko V.G., Choroshko V.A., Tcserbak L.N.** Computer technologies of recognition of vocal signals. Monograph. – K., 2005.– 138 p.
7. Pat. 773689, USSR, MKY G 10 L 1/02. Device for the selection of frequency of the basic tone /O. N. Carpov, E. L. Nosenko, A. A. Chugay / Openings. Inventions. –1980.– № 39.– 247 p.

Надійшла до редколегії 25.05.10

UDC 543.318.3+515.124

V. V. Romanuke

Khmelnytskyy National University

OPTIMAL WLF EQUATION OVER EXPERIMENTAL VISCOSITY MEASUREMENTS BY FINITE SET OF FIXED TEMPERATURES

За критерієм мінімальної відстані у функціональному просторі $L_p[T_1; T_N]$ визначено оптимальну температурну залежність в'язкості полімеру у формі ВЛФ-рівняння за N вимірюваннями його в'язкості при N фіксованих температурах зі сегмента $[T_1; T_N]$. Повністю представлено розроблене програмне забезпечення у формі MATLAB-функції для знаходження оптимальної температурної залежності згідно з відповідними вхідними даними.

Ключові слова: *полімер, в'язкість, температурна залежність, температура переходу у скловидну речовину, ВЛФ-рівняння, функціональний простір, мінімальна відстань, поліетилентерефталат.*

По критерию минимального расстояния в функциональном пространстве $L_p[T_1; T_N]$ определена оптимальная температурная зависимость вязкости полимера в форме ВЛФ-уравнения по N измерениям его вязкости при N фиксированных температурах из сегмента $[T_1; T_N]$. Полностью представлено разработанное программное обеспечение в форме MATLAB-функции для нахождения оптимальной температурной зависимости согласно соответствующим входным данным.

Ключевые слова: *полимер, вязкость, температурная зависимость, температура стеклования, ВЛФ-уравнение, функциональное пространство, минимальное расстояние, полиэтилентерефталат.*

By the minimal distance criterion in the functional space $L_p[T_1; T_N]$ there has been determined the optimal temperature dependence of polymer viscosity in the WLF equation form over N measurements of its viscosity by N fixed temperatures from the segment $[T_1; T_N]$. There has been fully represented the designed software in the MATLAB-function form for finding the optimal temperature dependence in compliance with the corresponding input data.

Key words: *polymer, viscosity, temperature dependence, glass transition temperature, WLF equation, functional space, minimal distance, polyethylene terephthalate.*
