ARCHIVES OF ACOUSTICS
23, 4, 463–479 (1998)

PARAMETRICAL PATTERNS OF VOICES IN FORENSIC APPLICATIONS

W. MAJEWSKI and CZ. BASZTURA

Institute of Telecommunications and Acoustics
Wrocław University of Technology
(50-370 Wrocław, Wybrzeże Wyspiańskiego 27, Poland)

A method of speaker recognition based on visual comparisons of sets of voice patterns is presented. These sets contain graphical patterns obtained from parameter extraction in the time-domain, frequency-domain and LPC (Linear Predictive Coding). In comparison to classical spectrograms, the parametrical patterns permit a better adaptation to specific speech signals that are available in forensic applications. Since parametrical patterns are more explicit than spectrograms, the decision reached by an expert may be more objective. The laboratory experiments on visual comparison of selected sets of patterns permitted to observe that in parametrical patterns of voices the intraspeaker variations are generally smaller than the interspeaker ones. The presented method of speaker recognition has been verified in a real-life case yielding valuable information on particular speakers. The results of laboratory experiments and real-life applications indicate that the visual method of speaker recognition based on various parametrical patterns of selected speech segments constitutes a very useful tool supporting the evidence proceeding with a much higher confidence level than it could be obtained on the basis of the classical spectrograms.

1. Introduction

Individual voice features contained in a speech signal permit to recognise speakers from their utterances. The speaker recognition may be carried out by means of subjective or objective methods. The subjective methods, that may be either aural-perceptual or visual, are based on subjective judgements, whereas the objective methods encompass all automatic approaches to the speaker recognition where the decisions are made by a computer.

Recognition of speakers, based on their utterances, have many different applications, from among which the most important ones are the forensic applications. Forensic applications of speaker recognition require a high level of reliability that is difficult to achieve by means of a single method of speaker recognition because of the unfavourable conditions (different environmental and transmission conditions for the evidence and reference speech samples, uncooperative speakers, voice disguise, etc.) typical of this type of applications. To solve this problem a system consisting of a parallel application of an aural-perceptual procedure and a visual and automatic one has been proposed [3]. The present paper is focused on the visual procedure of the speaker recognition.

The visual method of speaker recognition is based on comparisons of sets of patterns found in the utterances produced by the examined speakers. Usually spectrograms (sonagrams) are utilised for this purpose. Since a speech pattern depends on both the speaker and the speech content, the patterns under comparison have to be obtained from the same utterance for all the examined speakers. It has to be remembered that in forensic applications the patterns are influenced also by different transmission and recording conditions for the evidence and reference speech samples and by the voice disguise.

The effectiveness of this method of speaker recognition is strongly influenced by the qualification and experience of the expert. Moreover, a considerable confusion exists with respect to the effectiveness of visual speaker recognition based on spectrograms [1, 2, 4]. One of the possible reasons of this controversy is the fact that the spectrograms contain too many details that are hard to follow. Our idea is to include to the visual inspections sets of patterns obtained from parameter extraction in the frequency-domain, time-domain or LPC (Linear Predictive Coding) and presented as parameters averaged for particular utterances or presented in the function of time. Since parametrical patterns are more explicit than classical spectrograms, the decisions reached by the expert should be more reliable and objective. To test the usefulness of parametrical patterns of voices for visual speaker recognition in forensic applications was the main purpose of the present study.

2. Parametrical patterns of voices

The parameters that are to be applied for speaker recognition should be efficient in representing the individual voice features (i.e. they should vary maximally for different speakers and minimally for a given speaker), should be easy to measure and stable over time, occur frequently in speech, should be resistant to speaking and transmission conditions and not susceptible to mimicry [5]. It is not possible, however, to find parameters that fulfil simultaneously all those conditions, and – in practice – a large variety of speech parameters is utilized. The parameters most widely used include: the speech spectrum, formant frequencies, the fundamental frequency, vocal intensity, LPC parameters, cepstrum coefficients and zero-crossing rates (ZCR). For the purpose of the present study the parameters presented below have been chosen.

For recording the speakers' voices a laboratory recording system consisting of an IBM PC computer, microphone and an acoustic input/output with AD/DA card was used.

The preliminary speech signal coming from a microphone has been low-pass filtered with a 4.5 kHz cut-off frequency, sampled at a rate of 10 kHz and digitized with a 12-bit resolution. A Hamming window of the coefficients $\alpha = 0.54$ and $\beta = 0.46$ was superimposed onto the signal. The window length was $N = 256$ samples with a 128 samples shift. Thus, the parameter vectors were taken from 25.6 ms frames.

2.1. *Spectral parameters*

Description and analysis of the speech in the frequency domain permits to present the utterances in a very useful and often applied form. It results from the fact that the changes in the excitation function and the shape of the vocal tract occurring during speech production are significantly reflected in the spectral parameters' values which are very important for speech perception because the human ear acts like a frequency analyser.

The relation between the frequency domain and time representation of a signal is given by the Fourier transform. FFT algorithms given, by G.E. BERGLAND and M.T. DOLAN [6], were used to calculate the short-term speech spectra. These algorithms satisfy the equation:

$$F(k) = \sum_{n=0}^{N-1} u(n) e^{-j\frac{2\pi}{N}nk}, \tag{1}$$

where $u(n)$ – digital input sequence of the speech signal, $F(k)$ – complex coefficients of FFT, $k = 0, 1, 2, ..., N-1$.

Before FFT was calculated, the signals were standardized to the energy equal to unity.

A commonly applied method of spectral parametrization is the mean spectrum in one-third octave bands measured for stationary speech segments or particular key words. The vectors with components representing the mean amplitude spectra in $P = 16$ one-third octave bands were used as spectral parameters of the utterances under investigation. For each frequency band the components $F(p)$ were calculated according to the formula:

$$F(p) = \frac{1}{K_p - P_p} \sum_{j=P_p}^{K_p} \left\{ [\mathrm{Re}\, F(j)]^2 + [\mathrm{Im}\, F(j)]^2 \right\}, \tag{2}$$

where $P_p = f_p \cdot N/f_s$, $K_p = f_{p+1} \cdot N/f_s$, f_p – boundary frequency between $(p-1)$ and p frequency band, f_s – signal sampling frequency.

2.2. *The density of zero crossings (ZCR-density)*

Directly from the time course of the speech signal the parameters useful for speaker recognition may be extracted. They concern the analysis of the speech signal zero crossings. The results of this analysis may be expressed either as the density of zero crossings (ZCR-density) or as the distribution of time intervals between successive zero crossings (ZCR-distribution).

The mean density of zero crossings ϱ_0 of the speech signal $u(t)$ in the time interval T_p is expressed by the formula:

$$\varrho_0[u(t), T_p] = \frac{1}{T_p} \frac{\int_{-\infty}^{\infty} f^2 P(f)\, df}{\int_{-\infty}^{\infty} P(f)\, df}, \tag{3}$$

where f – frequency, $P(f)$ – power density spectrum.

The measurement of the density $\varrho_0[u(t), T_p] = \varrho_0(p)$ from a speech signal in the digital form $u(n)$ is made according to the relation:

$$\varrho_0(p) = \frac{1}{T_p} \sum_{n=1}^{N} C\left\{u(n), (k-1)T_p + \frac{n}{f_s}\right\}, \tag{4}$$

where

$C(u,t) = 1$, if there exists signal $u(t) = u(n)$ satisfying the conditions (a)–(c), and

$C(u,t) = 0$, if there is no signal $u(t) = u(n)$ satisfying the conditions (a)–(c).

(a) $u(n)u(n-k) < 0$,

(b) $|u(n)| \geq \alpha$ and $|u(n-k)| \geq \alpha$,

(c) $|u(l)| < \alpha$ for $n-k < l < n$,

and α is a threshold level ($\alpha \neq 0$) which prevents from counting additional zero crossings caused by disturbances, p is the index of a signal segment of duration $T_p = N/f_s$.

2.3. Distribution of time intervals between zero crossings (ZCR-distribution)

The moments of successive zero crossings in the speech signal are detected and the lengths of time intervals between these crossings calculated. The intervals are then grouped and ordered with respect to their lengths. They are classified to successive groups according to previously determined time channels. The number of intervals in the individual channels may be determined:

$$x(p) = \{x(t_0, t_1), x(t_1, t_2), ..., x(t_p, t_{p+1}), ..., x(t_{P-1}, t_P)\}, \tag{5}$$

where t_p – threshold values between the channels, P – number of time channels.

An interval with length t_j is classified to the time channel p in accordance with the dependence:

$$x(p) = x(t_{p-1}, t_p) = \begin{cases} x(t_{p-1}, t_p) + 1 & \text{for} \quad t_j \in (t_{p-1}, t_p], \\ x(t_{p-1}, t_p) & \text{for} \quad t_j \notin (t_{p-1}, t_p], \end{cases} \tag{6}$$

where $x(t_{p-1}, t_p)$ – number of time intervals in a given time channel.

The cumulative distribution of time intervals in the P time channels can be presented as:

$$x(p) = \sum_{p=1}^{P} x(t_{p-1}, t_p) \cdot [\mathbf{1}(t - t_{p-1}) - \mathbf{1}(t - t_p)], \tag{7}$$

where

$$\mathbf{1}(t) = \begin{cases} 0 & \text{for} \quad t \leq 0, \\ 1 & \text{for} \quad t > 0. \end{cases}$$

The boundary values of the time channels for the speech signals are $t_{\min} = 0.2\,\text{ms}$ and $t_{\max} = 6.2\,\text{ms}$. If in such a range $P-1$ intermediate threshold values are placed, P time channels will be obtained. In the experiments $P = 16$ time channels with logarithmic distribution of length were applied.

2.4. Linear predictive coding (LPC) parameters

The algorithms given by J.D. Markel and A.H. Gray [6] have been used to calculate the LPC parameters. Linear prediction models the signal spectrum by means of an all-pole filter with the transfer function:

$$H(z) = \frac{G}{1 - \sum_{p=1}^{P} a_p z^{-p}}, \tag{8}$$

where G – gain factor, a_p – prediction coefficient, P – prediction order, z – operator of the **Z** transform.

The Levinson–Durbin [7] recursion was utilized to calculate the prediction and reflection coefficients. It is a recursive-in-model-order solution for the autocorrelation equations applied to the window $n \in (t - N + 1, t)$.

Initialization: $p = 0$

$$E^{(0)(t)} = R(0;t) \tag{9}$$

scaled total energy in the "error" from an "order 0" predictor = average energy in the speech frame $h(n)$ $\{u(n)h(t-n)\}$.

Recursion: For $r = 1, 2, 3, ..., P$,

1. Compute the reflection coefficient,

$$k(r;t) = \frac{R(r;t) - \sum_{p=1}^{r-1} a^{(r-1)}(r;t) \cdot R(r-p;t)}{E^{(r-1)}(t)}. \tag{10}$$

2. Generate the order-r set of LPC parameters,

$$a_{r(r;t)} = k(r;t), \tag{11}$$

$$a_p(p;t) = a^{(r-1)}(p;t) - k(r;t) \cdot a^{(r-1)}(r-p;t), \qquad 1 \le p \le r-1. \tag{12}$$

3. Compute the error energy associated with the order-r solution,

$$E^{(r)}(t) = [1 - k(r;t)]^2 \cdot E^{(r-1)}(t). \tag{13}$$

4. Return to Step 1 with r replaced by $r + 1$ if $r < P$.

In the experiments $P = 12$ prediction coefficients a_p averaged over the utterances were used.

2.5. Amplitude envelope

The amplitude envelope $E(t)$ may be calculated by the following formula:

$$E(t) = \sqrt{u^2(n) + \mathbf{H}^2[u(n)]}, \tag{14}$$

where **H** is the Hilbert transform:

$$\mathbf{H}[u(n)] = \sum_{-N/2}^{N/2} \frac{u(k)}{\Pi(n-k)}. \tag{15}$$

In addition to the parameters presented in points 2.1–2.5, two additional speech signals representations, i.e. the waveforms and the digital spectrograms, were used in the experiments.

3. Experiments under laboratory conditions

The group of speakers consisted of 10 Polish adult males with no speaking defects who uttered two isolated Polish words: "Awek" */avek/* and "logarytm" */logaritm/*. The number of repetitions for each word and each speaker was 20. All the utterances were produced during a single session and recorded under normal laboratory conditions. The recorded speech samples were converted to a digital form and fed into the computer. Each speech sample was labelled by the following code: U0SS01RR, where "U" is the first letter of the utterance, "SS" stands for the speaker number and "RR" for the repetition number.

A specially designed computer program enable to perform and visualize the results of the following speech signal analyses:

- waveforms,
- spectrograms,
- mean spectrum in 16 one-third octave bands,
- distribution of time intervals between zero crossings in 16 time channels,
- linear prediction coefficients of 12-th order,
- amplitude envelope with the printed values of time (first number) and amplitude (second number) of local maxima,
- density of zero crossings with the printed values of local maxima (marking the same as above).

The graphical patterns resulted from the performed analyses were used to evaluate the intra- and interspeaker variations of the speech samples. As an example, the patterns obtained for three repetitions of the word "logarytm" produced by seven speakers are presented in Figs. 1–7. In Fig. 1 the waveforms are presented. As may be seen, the intraspeaker variations are quite small. The interspeaker variations are generally larger, but they depend upon the speakers under comparison – in some cases they are very large (e.g. the waveforms for speaker no 7 do not resemble the waveforms for any of the other six speakers) and in other cases they are comparable to the intraspeaker variations (see e.g. the patterns for the speakers no 3 and 5). Thus, these direct representations of the speech samples provide limited information on intra- and interspeaker variations. In Fig. 2 digital spectrograms are presented. The general conclusion is similar: the intraspeaker variations are rather small, the interspeaker variations are generally larger, but the large number of details cause that the similarities and differences between particular patterns are somewhat obscured. The remaining figures (Fig. 3 – mean spectra, Fig. 4 – distributions of time intervals, Fig. 5 – LPC coefficients, Fig. 6 – amplitude envelopes, Fig. 7 – zero crossing densities) concern parametrical patterns that are much simpler in their graphical form than spectrograms and the comparisons between those patterns may be more easily performed. The mean spectra presented in Fig. 3 are very

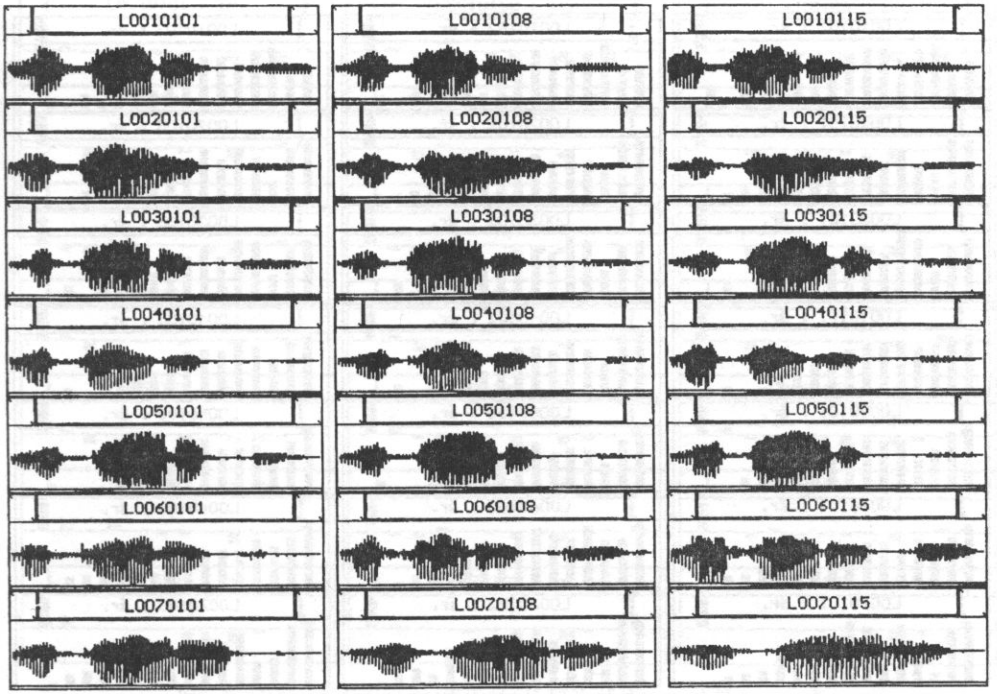


Fig. 1. The waveforms of the word "logarytm" produced three times by seven speakers.

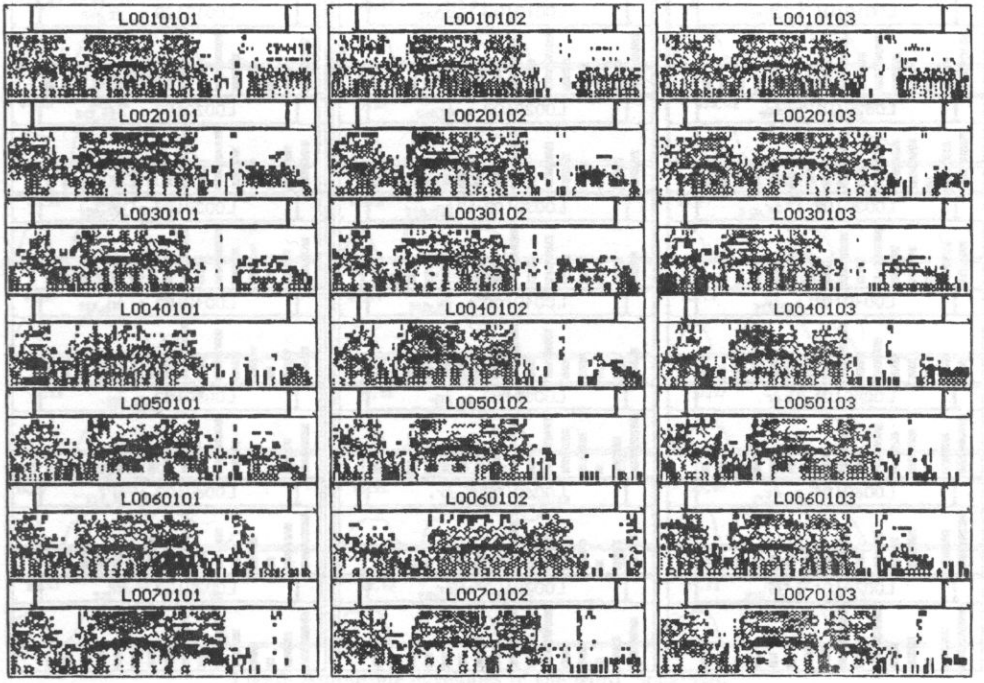


Fig. 2. Digital t-f-a spectrograms of the word "logarytm".

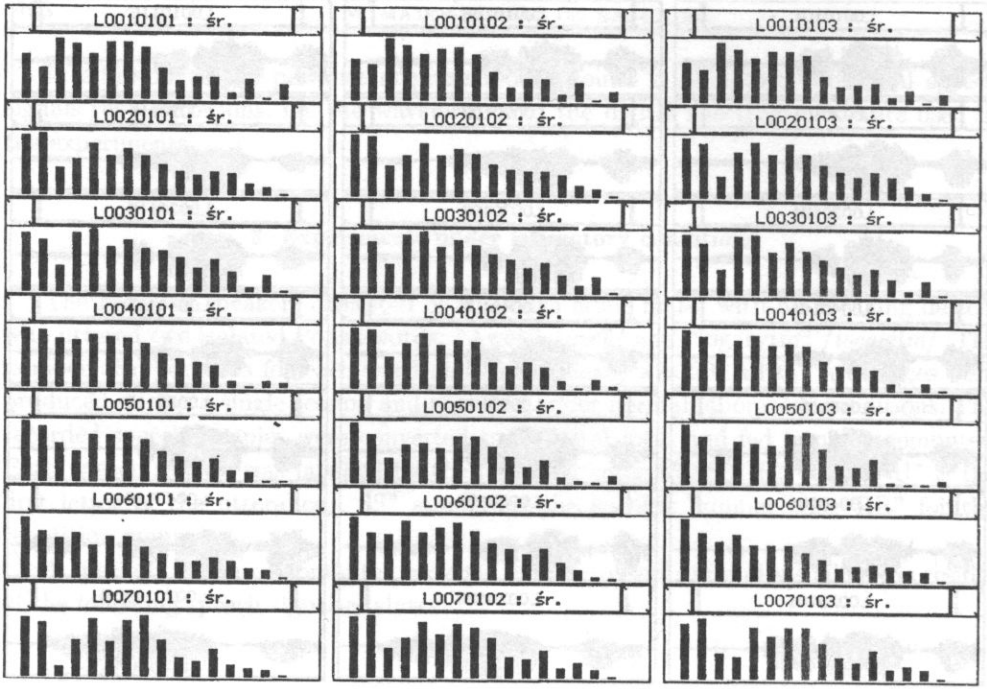


Fig. 3. Mean spectra of the word "logarytm".

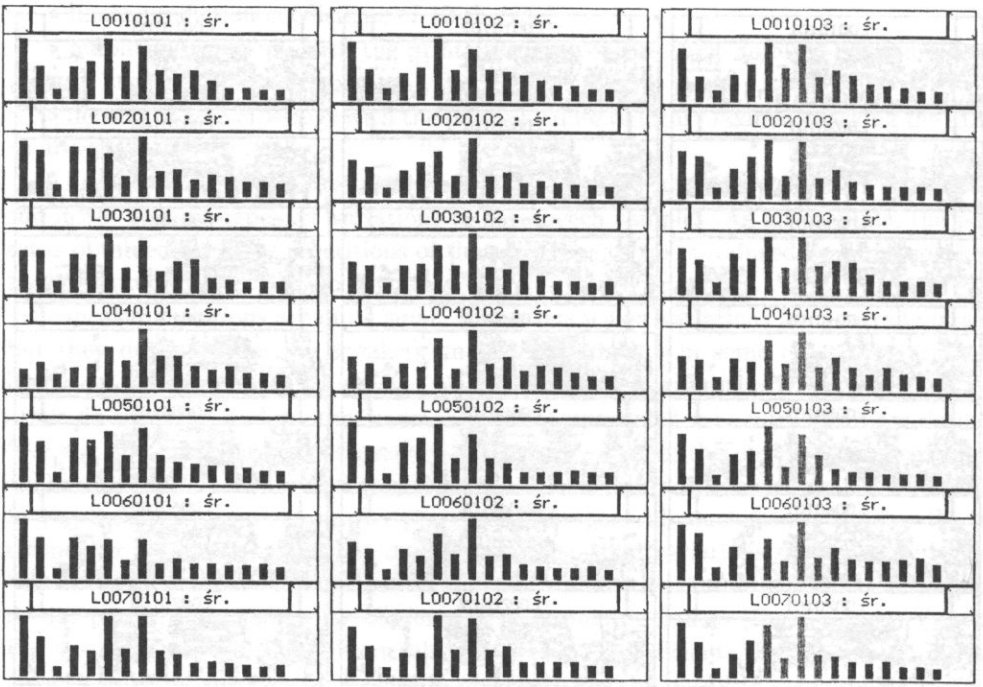


Fig. 4. Distributions of time intervals between zero crossings (ZCR-distributions) of the word "logarytm".

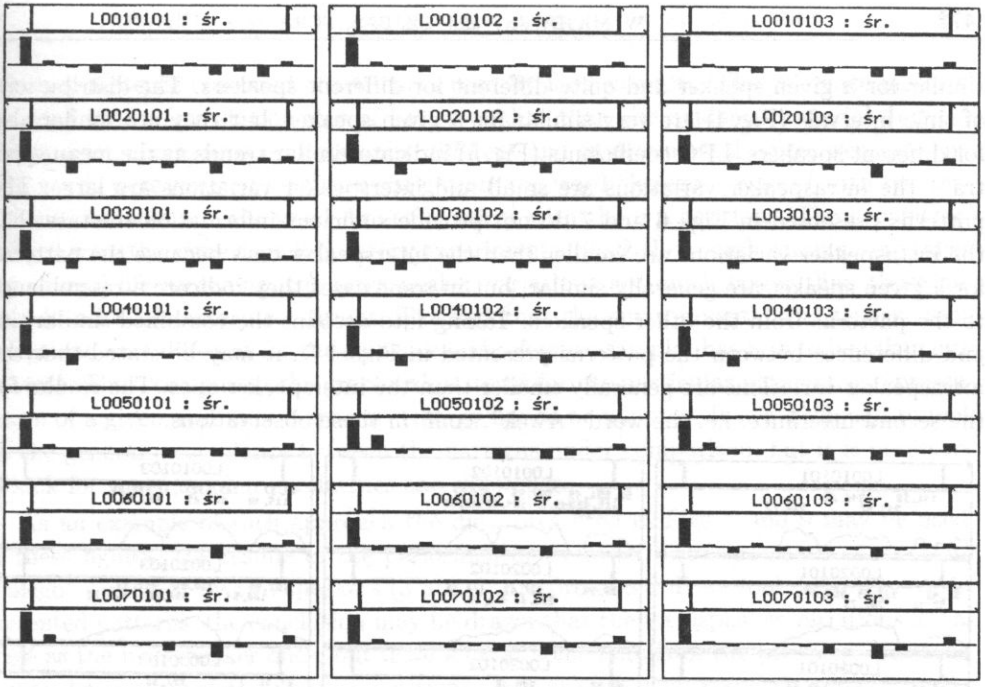


Fig. 5. Mean prediction coefficients (LPC) of the word "logarytm".

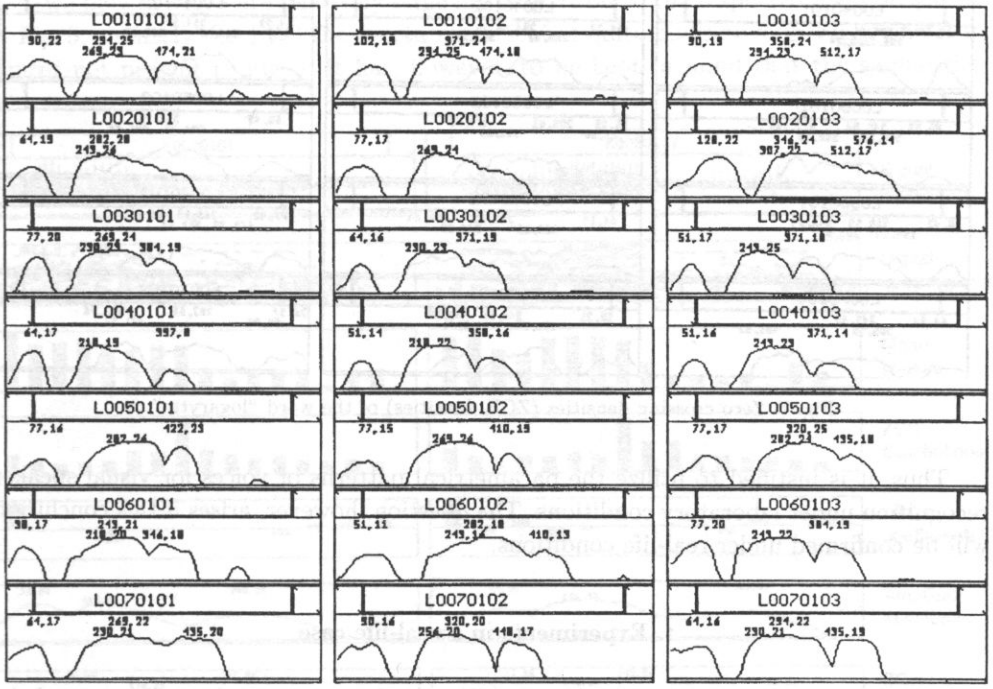


Fig. 6. Amplitude envelopes of the word "logarytm".

similar for a given speaker and quite different for different speakers. The distributions of time intervals (Fig. 4) are very similar for a given speaker, but they are similar also for different speakers. LPC coefficients (Fig. 5) indicate similar trends as the mean spectra – the intraspeaker variations are small and interspeaker variations are large. The patterns presented in Figs. 6 and 7 do not provide sufficient information to state that the intraspeaker variations are smaller than the interspeaker ones because the patterns for a given speaker are generally similar, but in some cases they indicate a resemblance to the patterns from the other speakers. Taking into account the combined similarities and differences between the patterns presented in Figs. 1–7, it may be stated that the intraspeaker variations are generally smaller than the interspeaker ones. The results for the second utterance, i.e. the word "Awek", confirm these observations.

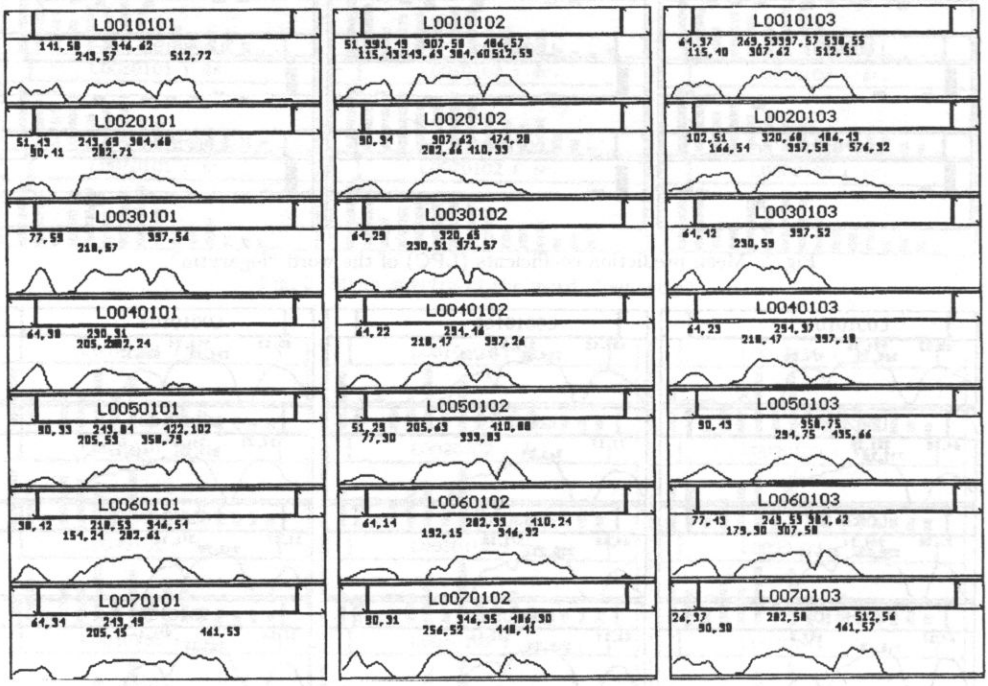


Fig. 7. Zero crossing densities (ZCR-densities) of the word "logarytm".

Thus, it is justified to utilize the parametrical patterns of voices for visual speaker recognition under laboratory conditions. The question, however, arises if this conclusion will be confirmed under real-life conditions.

4. Experiments in a real-life case

Five speakers were involved in the case. The speaker no 3 was the unknown and the speakers marked with numbers 4–7 were the suspects. The evidence material produced by the unknown consisted of telephone calls recorded on a cassette tape recorder. The

reference material consisted of sentences read by four suspects. The reference material recorded in a police station was of better quality than the evidence material, but the transmission characteristics were quite different and environmental noises were present. Two words were selected for the investigation: the word "kolego" */colego/* with two repetitions available and the word "waler" */valer/* with three repetitions. The parameters utilized were the same as in laboratory experiments. Since the speech sample were taken from a real-life recording, the intraspeaker variations are much larger than those observed under laboratory conditions. The fact that the environmental and transmission conditions were different for the evidence and reference samples has also to be taken into account. Thus, comparing the graphical patterns for the particular parametric representation of a given utterance a special strategy has to be adopted. It is not enough to look for the general resemblance between the patterns under comparison, but it is necessary to look for some peculiar similarities and differences.

As an example of such approach the data presented in Figs. 8 and 9 may be used. In these figures, the results of the performed analyses for two repetitions of the word "kolego" produced by two speakers (3 and 4) are presented. If we look generally at the presented patterns, the conclusion may be drawn that the intraspeaker variations are as large as the interspeaker ones, but if we look at some details of the presented patterns, it may be confirmed that both repetitions were produced by respective speakers. For example: if we compare all four patterns of the mean spectra, it may be found that the last two bars in Fig. 8 are relatively large in comparison to the respective bars presented in Fig. 9; similarly, for LPC the fourth coefficient is quite distinctive in Fig. 8 and it is almost not present in Fig. 9. It has, however, to be kept in mind that the similarities

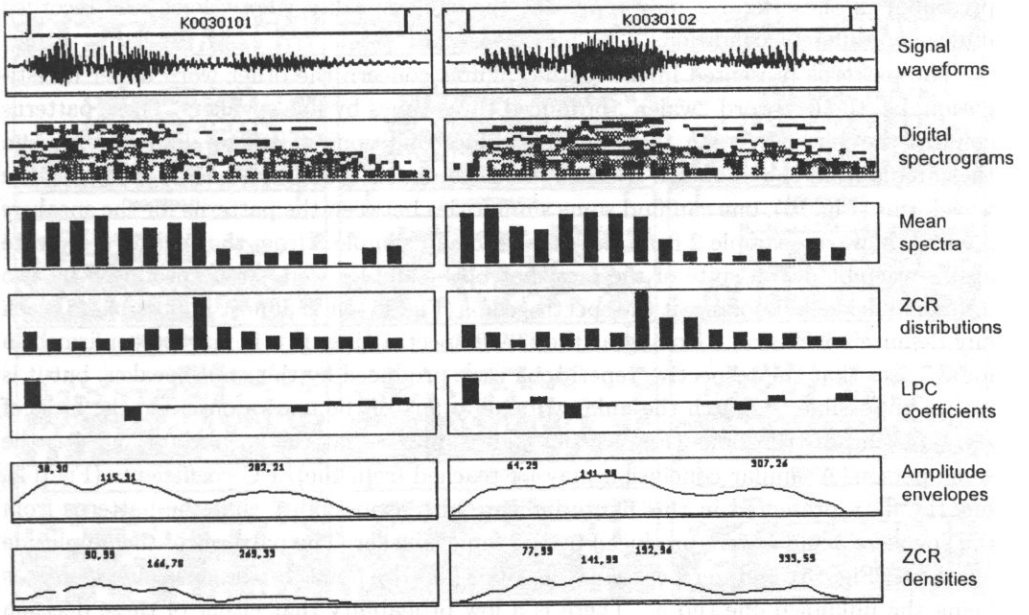


Fig. 8. Seven parametrical representations of the word "kolego" produced two times by speaker no 3.

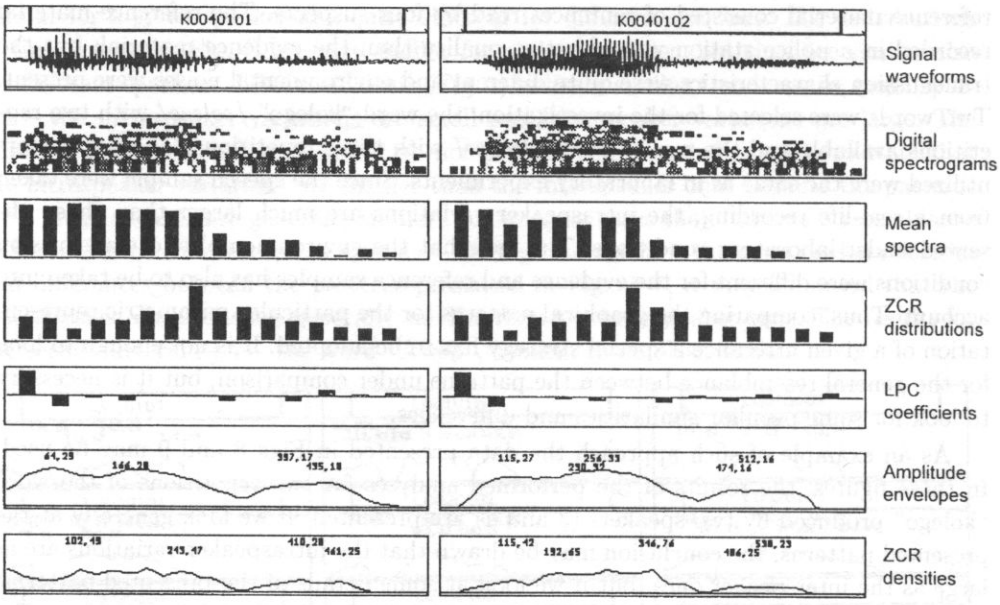


Fig. 9. Seven parametrical representations of the word "kolego" produced two times by speaker no 4.

and differences between the patterns presented in Figs. 8 and 9 may result only from different transmission and recording conditions. Since the speaker no 3 was the unknown and the speaker no 4 was one of the knowns, it is possible that both sets of the patterns presented in these figures may represent the same speaker whose voice was recorded under two different conditions.

The patterns presented in the further figures concern the other word under investigation, i.e. the key word "waler", produced three times by five speakers. These patterns confirm the general notion that under real-life conditions it is very difficult to assign the sample from the unknown speaker (no 3) to one of the known ones. Comparing the waveforms (Fig. 10), one can find some similarities between the patterns for the speakers 3 and 4, however, sample 2 from the speaker 5 and sample 3 from the speaker 6 indicate also a resemblance in spite of the fact that both samples were surely produced by two different speakers. Looking at the spectrograms (Fig. 11), it is almost impossible to reach any definite conclusion. Looking at the mean spectra (Fig. 12), it may be confirmed in many cases that the respective repetitions were produced by the same speaker, but it is rather impossible to match the unknown one to any of the known ones. On the basis of zero crossing distributions (Fig. 13) one may suppose that the speakers 3 and 4 is the same person. A similar conclusion may be reached from the LPC coefficients (Fig. 14), but the data presented in this figure indicate also a possibility that the patterns from the speakers 3 and 5 were produced by the same speaker. The patterns of the amplitude envelope (Fig. 15) and zero crossing density (Fig. 16) point to the speakers 4 or 5 as being the unknown one (no 3). There is a low probability that either of these decision is correct since generally most of the patterns presented in these figures are similar.

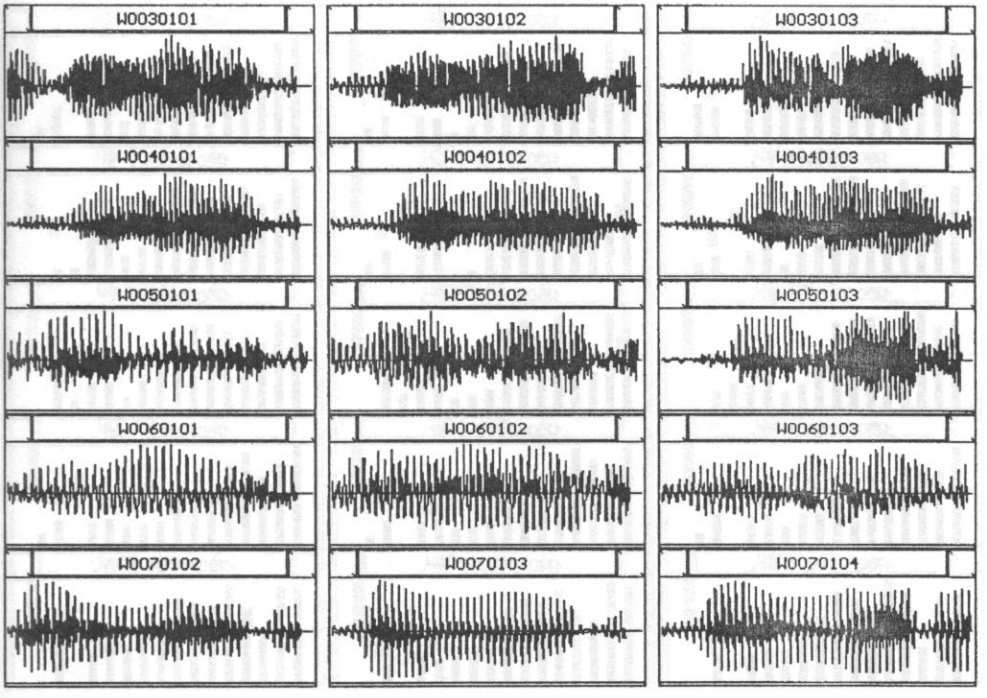


Fig. 10. The waveforms of the word "waler" produced three times by five speakers.

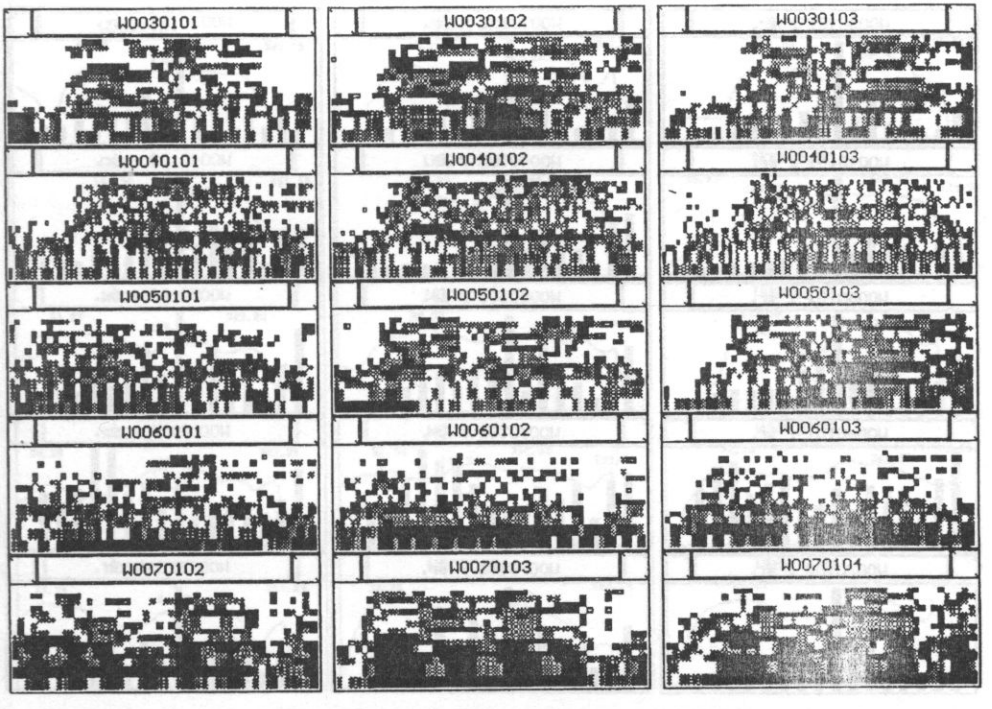


Fig. 11. Digital t-f-a spectrograms of the word "waler".

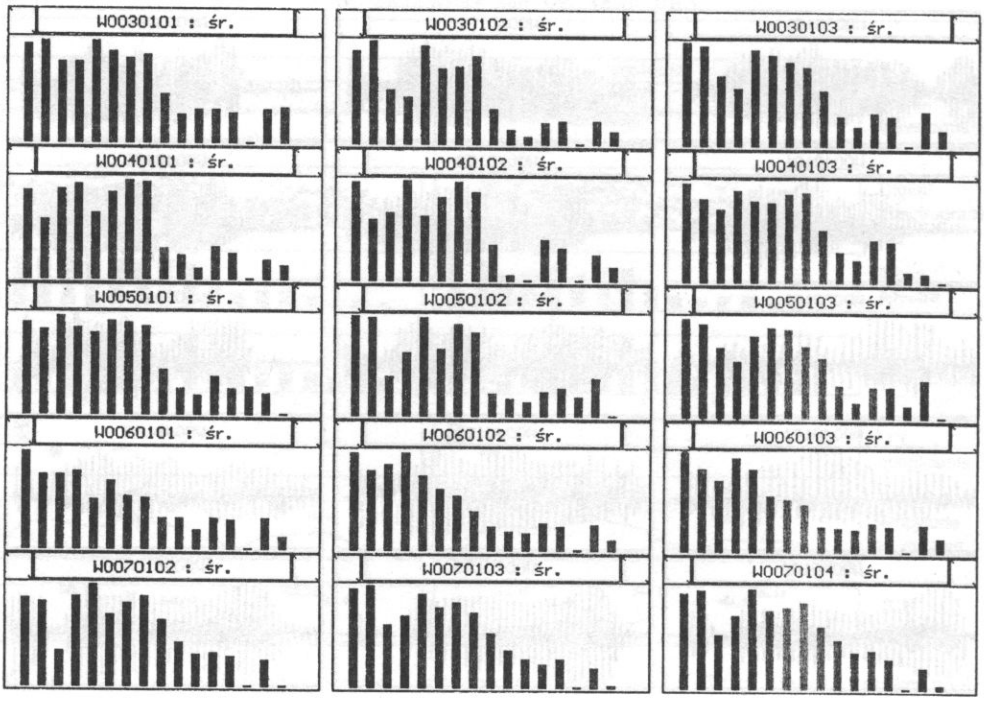


Fig. 12. Mean spectra of the word "waler".

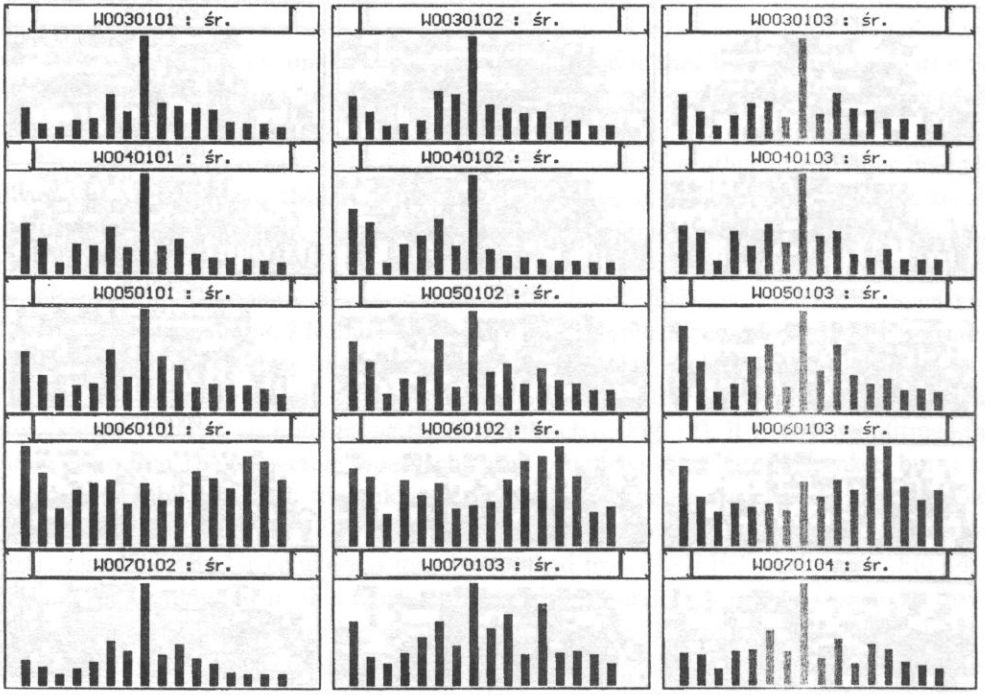


Fig. 13. ZCR-distributions of the word "waler".

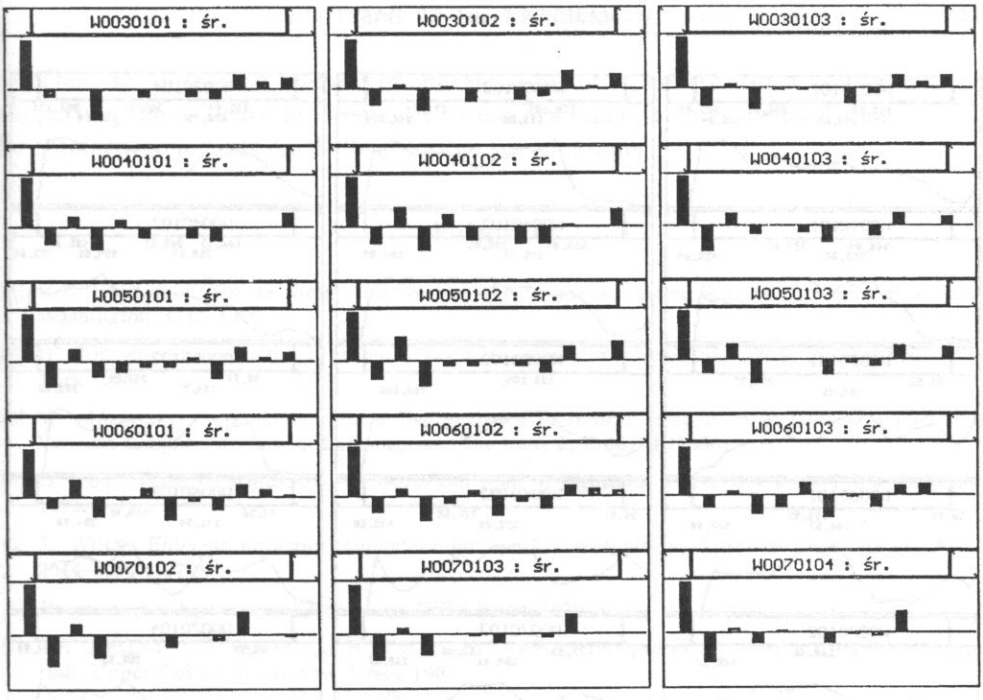


Fig. 14. LPC coefficients of the word "waler".

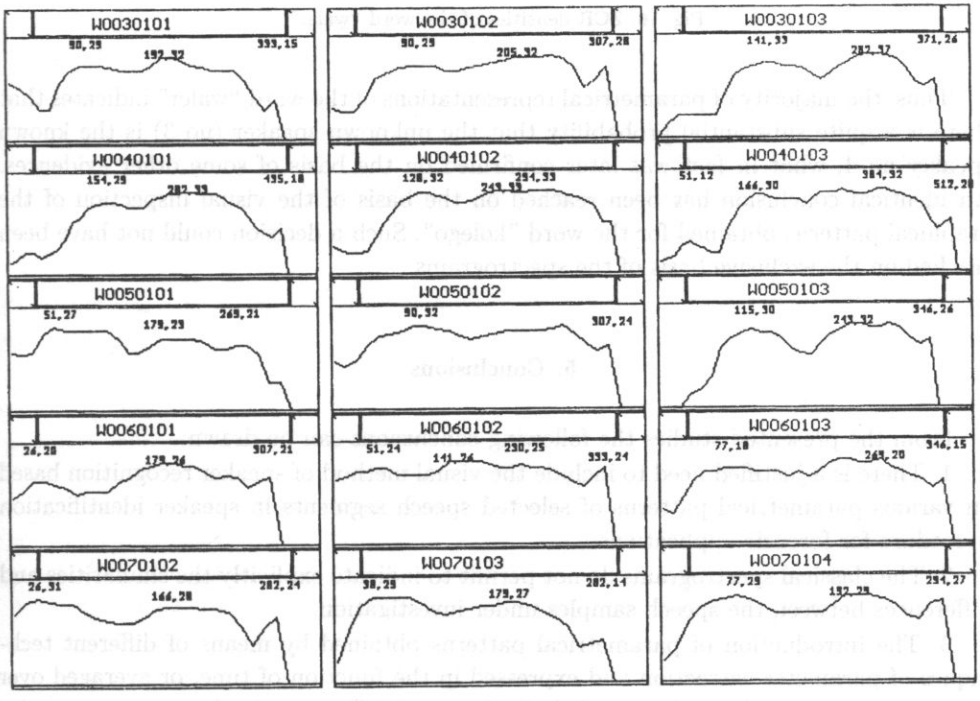


Fig. 15. Amplitude envelopes of the word "waler".

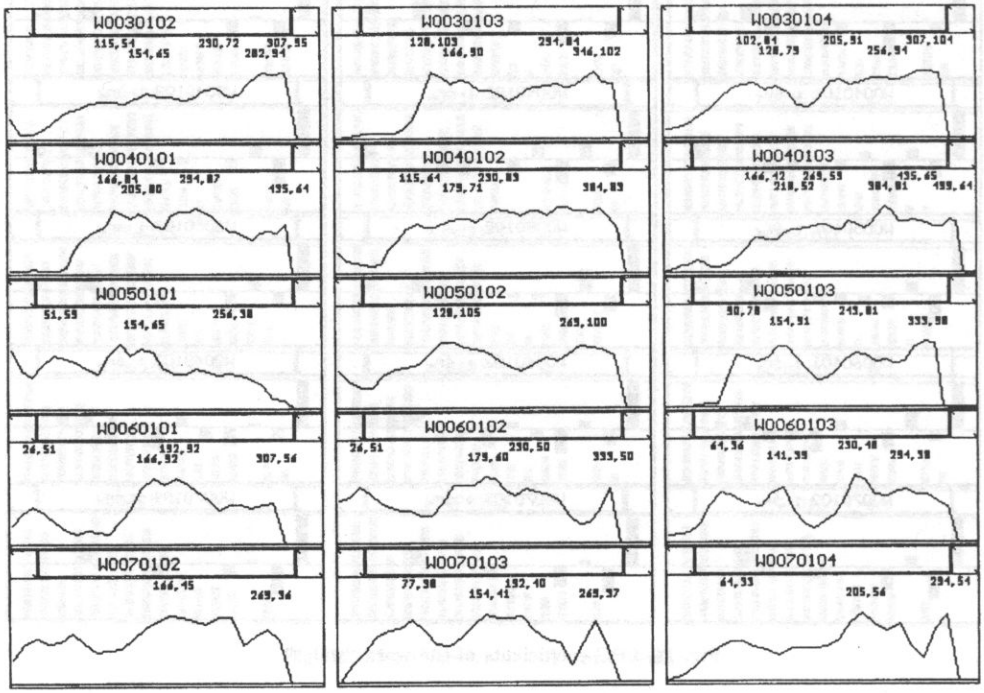


Fig. 16. ZCR-densities of the word "waler".

Thus, the majority of parametrical representations of the word "waler" indicates that there is a quite substantial probability that the unknown speaker (no 3) is the known speaker no 4, what in fact was later confirmed on the basis of some other evidences. An identical conclusion has been reached on the basis of the visual inspection of the graphical patterns obtained for the word "kolego". Such a decision could not have been reached on the exclusive basis of the spectrograms.

5. Conclusions

From the presented studies the following conclusions can be drawn:

1. There is a justified need to include the visual method of speaker recognition based on various parametrical patterns of selected speech segments in speaker identification procedure for forensic applications.

2. The classical spectrograms do not permit to indicate explicitly the similarities and differences between the speech samples under investigation.

3. The introduction of parametrical patterns obtained by means of different techniques of parameter extraction and expressed in the function of time, or averaged over time, permit to visualize substantial similarities and differences in the utterances under investigation that are not visible in the classical spectrograms.

Thus, the presented method of speaker recognition constitutes a very useful tool supporting the evidence proceeding with much a higher confidence level than it could be obtained exclusively on the basis of the spectrograms.

References

[1] R.H. BOLT *et al.*, *On the theory and practice of voice identification*, National Academy of Sciences, Washington, D.C. 1979.

[2] H. HOLLIEN, *The acoustics of crime – The new science of forensic phonetics*, Plenum Press, New York 1990.

[3] W. MAJEWSKI, C. BASZTURA, *Integrated approach to speaker recognition in forensic applications*, Forensic Linguistics: Speech, Language and the Law, **3**, 1, 50–64 (1996).

[4] O. TOSI, *Voice identification – Theory and legal applications*, University Park Press, Baltimore 1979.

[5] J. WOLF, *Efficient acoustic parameters for speaker recognition*, J. Acoust. Soc. Amer., **51**, 6, 2044–2056 (1972).

[6] *Programs for digital signal processing*, IEEE Press, 1979.

[7] J.R. DELLER, J.G. PROAKIS, J.H.L. HANSEN, *Discrete-time processing of speech signals*, Prentice Hall, Upper Saddle River, New Jersey 1993.

ARCHIVES OF ACOUSTICS
23, 4, 481–490 (1998)

VERTICAL DISTRIBUTION OF ROAD TRAFFIC NOISE

R. MAKAREWICZ

Department of Acoustic Design
Kyushu Institute of Design
(9-1, Shiobaru 4-chome, Minami-ku Fukuoka, 815 Japan)
e-mail: makaron@kyushu-id.ac.jp

R. GOŁĘBIEWSKI

Institute of Acoustics
A. Mickiewicz University,
(60-769 Poznań, Matejki 48, Poland)

Road noise generated by freely flowing traffic is assessed by the time-average sound level, L_{AT}. The noise interaction with the ground surface is the only wave phenomenon that modifies geometrical spreading. A simplified model of the ground effect is applied. The model is founded on the Weyl-Van der Pol solution. Prediction of the time-average sound level, at a perpendicular distance less than 100 m and an altitude not exceeding a few tens of meters, requires two simultaneous measurements of L_{AT} at the site of interest.

1. Introduction

There are many models for road traffic noise calculation. Although these models generate precise predictions, they require a tremendous expense in computing time. Thus, there is a need for methods that are practical for everyday use, despite limited accuracy (see e.g. Refs. [2, 9]). In this paper we propose such a method based on two simultaneous measurements at the site of interest.

Above a plane ground surface, without buildings or any other interfering structures, sound propagation depends upon geometrical spreading, ground effect, air absorption, refraction, and scattering by atmospheric turbulence [1]. Close to the road, within the range of 100 m, geometrical spreading is influenced mainly by ground effect. To quantify this influence we use the concept of the time-average sound level, L_{AT}, which is widely used as a measure of road traffic noise,

$$L_{AT} = 10 \log \left\{ \frac{\langle p_A^2 \rangle}{p_0^2} \right\}, \qquad p_0 = 20\,\mu\text{Pa}, \tag{1.1}$$

where the time-average A-weighted squared sound pressure is,

$$\langle p_A^2 \rangle = \frac{1}{T} \int_0^T p_A^2(t)\, dt. \tag{1.2}$$

Because road traffic noise consists of noise events that do not interact with each other, we write,

$$p_A^2(t) = \sum_{i=1}^{N} p_{Ai}^2(t), \tag{1.3}$$

where $p_{Ai}^2(t)$ corresponds to the noise generated by a single vehicle (Fig. 1), and N denotes the number of vehicles passing the receiver during the time period T. Combing Eqs. (1.2), (1.3) yields,

$$\langle p_A^2 \rangle = \frac{1}{T} \sum_{i=1}^{N} E_i\,, \tag{1.4}$$

where

$$E_i = \int_{-\infty}^{\infty} p_{Ai}^2(t)\, dt, \tag{1.5}$$

denotes the sound exposure.

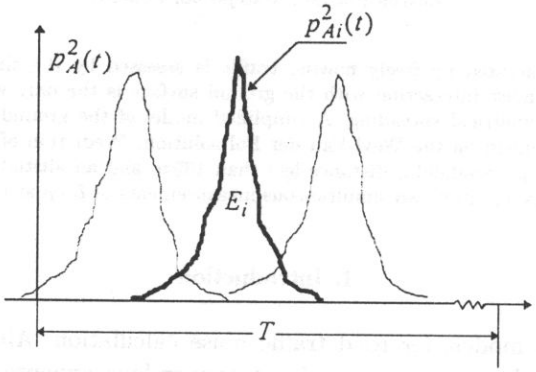


Fig. 1. Road traffic noise consists of noise events, i.e., passage of vehicles.

In Sec. 2 an approximated model of ground effect and the equation for p_A^2 is derived. Random nature of the sound exposure, E_i, is studied in Sec. 3. Then, a two parameter equation for the time-average sound level, L_{AT}, is derived and its application for noise prediction is demonstrated (Sec. 4).

2. Ground effect

A road vehicle can be modeled by a point source [5]. Assuming near grazing propagation (Fig. 2),

$$H \ll d, \qquad H_0 \ll d, \tag{2.1}$$

and neglecting the width of the road so that the ground surface is uniform, i.e., without impedance jumps, the A-weighted squared sound pressure can be written as [7],

$$p_A^2 = \frac{W_A \varrho c}{4\pi d^2} \cdot G(d, H, H_0), \tag{2.2}$$

where: W_A is the A-weighted sound power, ϱc is the characteristic impedance of air, and d is the horizontal distance between the source, S, and receiver, O. There are "exact" equations for the ground factor, G; among others, one based on the Weyl–Van der Pol solution is used [1]. The ground factor meets the principle of reciprocity,

$$G(H, H_0) = G(H_0, H), \tag{2.3}$$

and decreases with the second power of the horizontal distance,

$$\lim_{d \to \infty} G \propto d^{-2}. \tag{2.4}$$

To find a simple model of ground effect, we make use of the result obtained by Li *et al.* [6]: the ground factor, G, is a function of the grazing angle (Fig. 2),

$$\Psi = \cot^{-1}\left(\frac{d}{H + H_0}\right). \tag{2.5}$$

Taking into account Eqs. (2.3)–(2.5) we obtain the following approximation [8],

$$\widetilde{G} = \beta \cdot \left[1 + \gamma \cdot \frac{d^2}{(H + H_0)^2}\right]^{-1}. \tag{2.6}$$

Far away from the source, $d \to \infty$, the condition defined by Eq. (2.4) is met. Close to the source, $d \to 0$, the above formula yields, $\widetilde{G} \to \beta$, so parameter β may be interpreted as a measure of noise reflection from the road surface beneath the source (e.g., for $\beta = 2$ the sound level increases by 3 dB).

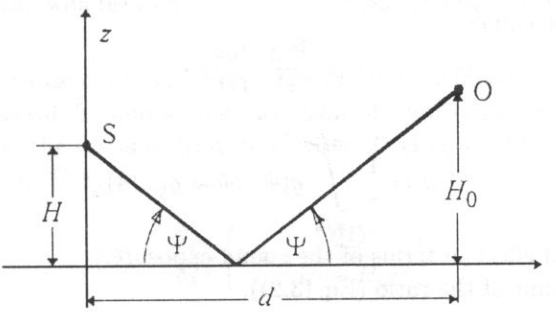


Fig. 2. Source, S, receiver, O, geometry is determined by the heights, H and H_0, horizontal distance, d, and the grazing angle, Ψ.

Introducing the perpendicular distance, D, and the angle Φ (Fig. 3), we get (Eqs. (2.2), (2.6)),

$$p_A^2 = \frac{\beta W_A \varrho c \cdot \cos^2 \Phi}{4\pi D^2} \cdot g(\Phi), \tag{2.7}$$

where

$$g(\Phi) = \left[1 + \gamma \cdot \frac{D^2}{(H + H_0)^2 \cos^2 \Phi}\right]^{-1}. \tag{2.8}$$

The *ground coefficient*, γ, characterizes the noise variations due to ground effect within the vertical plane ($y = D$, $z = H_0$).

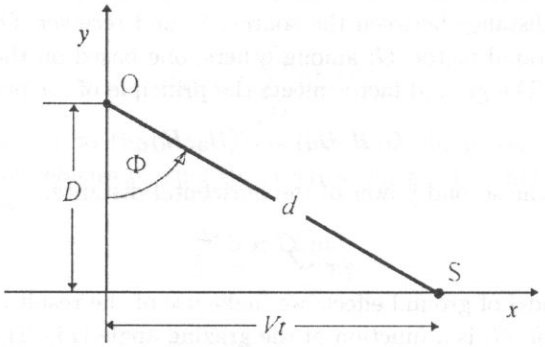


Fig. 3. Source, S, and receiver, O, in the horizontal plane, $x - y$, with the perpendicular distance, D, and the angle, Φ.

3. Sound exposure

If a source is moving with a steady speed, V, along the x axis, at the perpendicular distance, D, from the receiver, then the integral Eq. (1.5) takes the form,

$$E = \frac{D}{V} \int_{-\pi/2}^{\pi/2} \frac{p_A^2(\Phi)}{\cos^2 \Phi} \, d\Phi, \tag{3.1}$$

and sound exposure equals,

$$E = \frac{W_A}{V} \frac{\beta \varrho c}{4D} \cdot J, \tag{3.2}$$

where (Eq. (2.8))

$$J = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} g(\Phi) \, d\Phi \approx g(\pi/4) \tag{3.3}$$

accounts for ground effect in terms of the sound exposure.

Note, that the unit of the ratio (Eq. (3.2)),

$$S = W_A/V, \tag{3.4}$$

is *Joule* per *meter*. Thus, S is called the *linear density of energy*. This energy is spend during the passage of a vehicle. Making use of the reference power, $W_0 = 10^{-12}$ Watts, we introduce the *reference linear density of energy*,

$$S_0 = W_0/V_0, \tag{3.5}$$

where $V_0 = 1\,\mathrm{m/s}$.

Passing vehicles yield random values of E. The reasons for the random values are the variations of the sound power, W_A, velocity, V, height, H, distance to the receiver, D, and ground parameters, β and γ. Therefore, for the i-th vehicle we can write (Eqs. (2.8), (3.2)–(3.4)),

$$E_i = \frac{S_i \beta_i \cdot \varrho c}{4D_i} \left[1 + \gamma_i \frac{2D_i^2}{(H_i + H_0)^2} \right]^{-1} . \tag{3.6}$$

Consequently, the time-average A-weighted squared sound pressure is (Eq. (1.4)),

$$\langle p_A^2 \rangle = \frac{N}{T} \langle E \rangle , \tag{3.7}$$

where

$$\langle E \rangle = \frac{1}{N} \sum_{i=1}^{N} E_i , \tag{3.8}$$

denotes the average sound exposure.

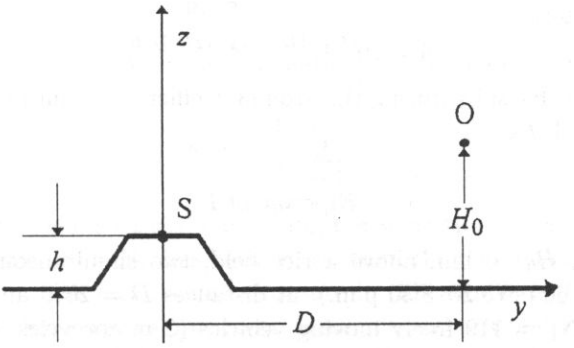


Fig. 4. Vertical plane with the receiver, $O(y = D, z = H_0)$, and the source, $S(y = 0, z = h)$.

Now we make two assumptions. First: the ground coefficient, γ, remains constant during the time interval, T, and second: location of each vehicle's track is determined by the height, h, and the distance from the center, D (Fig. 4). Thus, the average sound exposure is (Eqs. (3.6), (3.8)),

$$\langle E \rangle = \frac{\widetilde{S} \varrho c}{4D} \left[1 + \gamma \frac{2D^2}{(h + H_0)^2} \right]^{-1} , \tag{3.9}$$

where

$$\widetilde{S} = \frac{1}{N} \sum_{i=1}^{N} S_i \beta_i , \tag{3.10}$$

expresses the average of the linear density of energy S, which is modified by the ground reflection, β. Because of variations of mix of vehicles and weather conditions, the values of $\widetilde{S}$ and γ might change during the day and night (see the next section).

4. Time-average sound level

To obtain the time-average sound level we apply Eqs. (1.1), (3.5), (3.7) and (3.9),

$$L_{AT} = L_S + 10\log\left\{\frac{Nt_0}{T}\,\frac{l_0}{4D}\right\} - 10\log\left\{1+\gamma\frac{2D^2}{(h+H_0)^2}\right\}, \tag{4.1}$$

where $l_0 = 1\,\mathrm{m}$ and $t_0 = 1\,\mathrm{s}$. The *energy density level* (Eqs. (3.5), (3.10)),

$$L_S = 10\log\left\{\widetilde{S}/S_0\right\}, \tag{4.2}$$

and the ground coefficient, γ, are unknown. To determine them, two simultaneous measurements of L_{AT}, at the perpendicular distances, D and $2D$, at the same height, H_0, are needed. Considering Eq. (4.1) as a theoretical prediction with two adjustable parameters, we obtain,

$$\gamma = \frac{1}{2}\,\frac{(h+H_0)^2}{D^2}\cdot\frac{A-2}{8-A}, \tag{4.3}$$

where the parameter,

$$A = 10^{[L_{AT}(D)-L_{AT}(2D)]/10}, \tag{4.4}$$

is fully determined. By substituting the ground coefficient, γ, into Eq. (4.1) we find the energy density level, L_S.

Experiment I.

At the height, $H_0 = 1\,\mathrm{m}$, above a rice field, two simultaneous measurements of L_{AT} during $T_1 = 600\,\mathrm{s}$ (5:20–5:30 p.m.), at distances $D = 25\,\mathrm{m}$ and $2D = 50\,\mathrm{m}$, have been performed. $N_1 = 119$ freely moving vehicles (3 motorcycles, 88 automobiles, 28 trucks) passed the microphones. There was cloudy and windless weather. Substituting the results of measurements, $L_{AT}(D) = 61.4\,\mathrm{dB}$ and $L_{AT}(2D) = 56.7\,\mathrm{dB}$ into Eq. (4.4) gives $A = 2.95$, and then, for $h = 0$ (horizontal road) and $H_0 = 1\,\mathrm{m}$, we obtain: $\gamma_1 = 0.00015$ and $L_{S1} = 89.2\,\mathrm{dB}$.

When the numerical values of N_1/T_1, L_{S1}, and γ_1 are available, one can predict the time-average sound level at any distance from the road, $y = D$, and the height above the ground, $z = H_0$ (Eq. (4.1)),

$$L_{AT}(y,z) = L_{S1} + 10\log\left\{\frac{N_1 t_0}{T_1}\cdot\frac{l_0}{4y}\right\} - 10\log\left\{1+\gamma_1\frac{y^2}{(h+z)^2}\right\}. \tag{4.5}$$

The graphs of $L_{AT} = \mathrm{const}$ are given by,

$$z = y\cdot\left[\frac{\gamma_1 y}{y_0 - y}\right]^{1/2} - h, \tag{4.6}$$

where

$$y_0 = l_0\frac{N_1 t_0}{4T_1}10^{(L_{S1}-L_{AT})/10}. \tag{4.7}$$

Experiment II.

The same measurements used in Experiment I were repeated 1 hour later during $T_2 = 600\,\mathrm{s}$ (6:20–6:30 p.m.). The weather conditions were the same. This time noise was produced by $N_2 = 88$ vehicles (1 motorcycle, 74 automobiles, 13 trucks). Making use of the results of the measurements, $L_{AT}(D) = 57.5\,\mathrm{dB}$ and $L_{AT}(2D) = 53.6\,\mathrm{dB}$, we get (Eqs. (4.1), (4.3)–(4.4)): $\gamma_2 = 0.000066$ and $L_{S2} = 86.2\,\mathrm{dB}$.

(L_{S1}, γ_1) and (L_{S2}, γ_2) characterize traffic noise during the "short" time intervals, T_1 (5:20–5:30 p.m.) and T_2 (6:20–6:30 p.m.), respectively. What are the noise characteristics for longer time interval, T (e.g. 5:00–7:00 p.m.)? In other words, how to predict the time-average sound level for a long time interval, L_{AT}, when the values of (L_{S1}, γ_1) and (L_{S2}, γ_2) are available?

To answer this question, suppose n is the number of vehicles passing the receiver during the long time interval, T, so that the time-average A-weighted sound pressure is (Eqs. (3.7), (3.8)),

$$\langle p_A^2 \rangle_T = \frac{n}{T} \cdot \langle E \rangle_T \,, \tag{4.8}$$

where the mean sound exposure in the population of n noise events is defined by,

$$\langle E \rangle_T = \frac{1}{n} \cdot \sum_{i=1}^{n} E_i \,. \tag{4.9}$$

By using two measurements carried out during the short time intervals, T_1 and T_2, the sample of noise events, $N_1 + N_2$, has been drawn from the population of n noise events, where $N_1 + N_2 \ll n$. The mean sound exposure in this sample is defined by Eq. (3.8) with $N = N_1 + N_2$. Since the sample mean is an unbiased estimator of the population mean [3], one can write,

$$\langle E \rangle_T = \frac{1}{N_1 + N_2} \sum_{i=1}^{N_1+N_2} E_i \,. \tag{4.10}$$

The above average can be rewritten as (Eq. (4.8)),

$$\langle E \rangle_T = a_1 \cdot \langle E \rangle_1 + a_2 \cdot \langle E \rangle_2 \,, \tag{4.11}$$

where

$$\langle E \rangle_1 = \frac{\widetilde{S}_1 \varrho c}{4D} \left[1 + \gamma_1 \frac{2D^2}{(h + H_0)^2} \right]^{-1} , \tag{4.12}$$

$$\langle E \rangle_2 = \frac{\widetilde{S}_2 \varrho c}{4D} \left[1 + \gamma_2 \frac{2D^2}{(h + H_0)^2} \right]^{-1} , \tag{4.13}$$

and the relative numbers of vehicles are,

$$a_1 = N_1/(N_1 + N_2), \qquad a_2 = N_2/(N_1 + N_2). \tag{4.14}$$

Finally, for $y = D$ and $z = H_0$, we arrive at the time-average sound level (Eqs. (1.1), (3.4), (3.5), (4.8), (4.11)),

$$L_{AT} = 10\log\left\{\frac{nt_0}{T}\cdot\frac{l_0}{4y}\right\} + 10\log\left\{\frac{a_1\cdot 10^{L_{S1}/10}}{1+\gamma_1\dfrac{2y^2}{(h+z)^2}} + \frac{a_2\cdot 10^{L_{S2}/10}}{1+\gamma_2\dfrac{2y^2}{(h+z)^2}}\right\}, \tag{4.15}$$

where n denotes the number of vehicles passing the receiver during a long time interval, T. The contours $L_{AT} = \text{const}$ can be found from,

$$z = y\cdot[f(y, L_{AT})]^{1/2} - h, \tag{4.16}$$

where

$$f = \frac{\gamma_1(y-y_2)+\gamma_2(y-y_1)}{y_1+y_2-y} + \frac{\sqrt{[\gamma_1(y-y_2)+\gamma_2(y-y_1)]^2 + 4\gamma_1\gamma_2 y(y_1+y_2-y)}}{y_1+y_2-y}, \tag{4.17}$$

with

$$y_1 = l_0\frac{a_1 nt_0}{4T}10^{(L_{S1}-L_{AT})/10} \quad \text{and} \quad y_2 = l_0\frac{a_2 nt_0}{4T}10^{(L_{S2}-L_{AT})/10}. \tag{4.18}$$

Figure 5 shows the contours of $L_{AT} = 50$, 55, and 60 dB calculated for the output data of the above experiments: $a_1 = 0.57$, $\gamma_1 = 0.00015$, $L_{S1} = 89.2$ dB and $a_2 = 0.43$, $\gamma_2 = 0.000066$, $L_{S2} = 82.2$ dB, with the traffic flow during a long time interval, $n/T = (N_1+N_2)/(T_1+T_2) = 0.17$ [veh./s] and the road's height, $h = 0$.

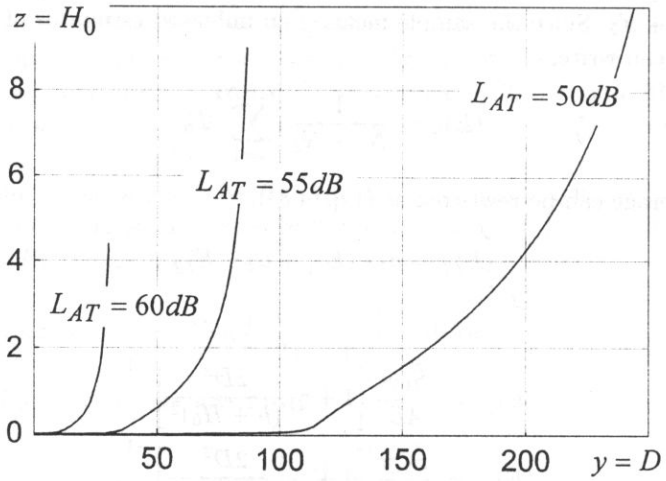


Fig. 5. Contours of $L_{AT} = \text{const}$ calculated from Eq. (4.16).

If m measurements of L_{AT} were made during the short time intervals $T_1, T_2, \ldots, T_i, \ldots, T_m$ (Fig. 6), with the relative numbers of vehicles, $a_1, a_2, \ldots, a_i, \ldots, a_m$ (Eq. (4.14)),

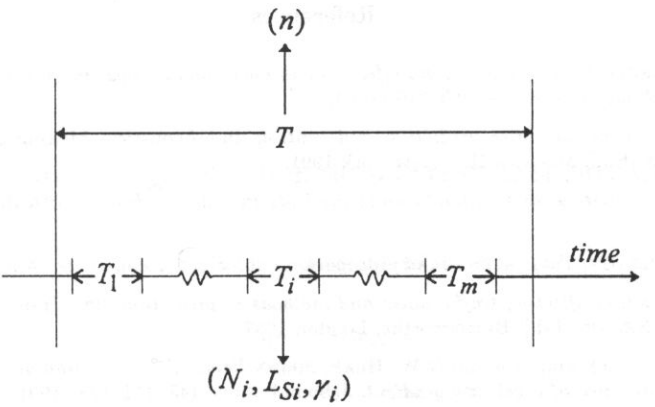


Fig. 6. During a short time interval, T_i, passages of N_i vehicles generate noise which is characterized by L_{Si} and γ_i.

respectively, then Eqs. (4.1), (4.3), and (4.4) yield the noise characteristics, (L_{S1}, γ_1), (L_{S2}, γ_2) ... (L_{Si}, γ_i) ... (L_{Sm}, γ_m), and the time-average sound level for a long time interval, $T > T_1 + T_2 + \ldots T_i + \ldots T_m$, can be calculated from,

$$L_{AT} = 10 \log \left\{ \frac{nt_0}{T} \cdot \frac{l_0}{4y} \right\} + 10 \log \left\{ \sum_{i=1}^{m} a_i \cdot 10^{L_{Si}/10} \cdot \left[1 + \gamma_i \frac{2y^2}{(h+z)^2} \right]^{-1} \right\}. \tag{4.19}$$

5. Conclusions

The model of road traffic noise has been developed under the following conditions:

- vehicles are moving with steady speeds along a straight road,
- the road width and receiver height ($z = H_0$) are significantly less than the perpendicular distance between the road center and the receiver ($y = D$),
- geometrical spreading is modified only by ground effect.

Two adjustable parameters of the model, i.e., energy density level, L_S, and the ground coefficient, γ, can be estimated from two measurements *in situ* of the time-average sound level, $L_{AT}(D)$ and $L_{AT}(2D)$. Making use of Eqs. (4.6), (4.16), and (4.19), one can calculate, $L_{AT}(y, z)$ at the distance, $y = D$, less than 100 m and at the height, $z = H_0$ not exceeding a few meters (Eq. (2.1)).

Acknowledgment

The authors are grateful to Mr. S. MAEDA for performing the measurements.

References

[1] K. ATTENBOROUGH, *Review of ground effects on outdoor sound propagation from continuous broadband noise*, Appl. Acoust., **24**, 289–319 (1988).

[2] W. BOWLBY, *Highway noise prediction and control*, [in:] Acoustical Measurements and Control, C.M. HARRIS [Ed.], McGraw-Hill, New York 1991.

[3] S. BRANDT, *Statistical and computational methods in data analysis*, North-Holland, Amsterdam 1976.

[4] T.F.W. EMBLETON, *Tutorial on sound propagation outdoors*, J. Acoust. Soc. Am., **100**, 31–48 (1996).

[5] B. FAVRE, *Factors affecting traffic noise and methods of prediction*, [in:] Transportation Reference Book, P.M. NELSON [Ed.], Butterworths, London 1987.

[6] K.M. LI, K. ATTENBOROUGH and N.W. HEAP, *Source height determination by ground effect inversion in the presence of a velocity gradient*, J. Sound Vibr., **145**, 111–128 (1991).

[7] R. MAKAREWICZ, *Near grazing propagation above a soft ground*, J. Acoust. Soc. Am., **82**, 1706–1710 (1987) and **88**, 1172–1175 (1990).

[8] R. MAKAREWICZ and P. KOKOWSKI, *Simplified model of ground effect*, J. Acoust. Soc. Am., **101**, 372–376 (1997).

[9] K. TAKAGI and K. YAMAMOTO, *Calculation methods for road traffic noise propagation proposed by ASJ*, Inter-Noise 94, 289–294 (1994).

ARCHIVES OF ACOUSTICS
23, 4, 491–498 (1998)

TUNING IN THE AMPLITUDE MODULATION RATE DOMAIN

A.P. SĘK and E.B. SKRODZKA

Institute of Acoustics,
A.Mickiewicz University,
(60-769 Poznań, Matejki 48/49, Poland)

This paper is concerned with a certain form of masking that seems to exist in the modulation rate domain. It was shown that clearly audible changes in the amplitude of a 4 kHz sinusoidal carrier signal produced by a sinusoidal "probe" modulator were inaudible (masked) in the presence of amplitude changes in the same carrier produced by a "masking" modulator which was a 10 Hz wide band of white noise. This type of masking was most effective when the centre rate of the masking modulator was close or equal to the modulation rate of the probe modulator. The pattern of results showed a tuning effect in the modulation rate domain. These findings are generally consistent with the concept of a second stage of the filtering that may take place in the auditory system.

1. Introduction

Temporal resolution of the auditory system has been extensively discussed in the psychoacoustics literature, [7, 8, 14]. In the most general case, models of temporal resolution of an auditory system consist of four stages [7]. The first one is a bandpass filtering reflecting the action of the auditory filters. The output of those filters is fed to the second stage of the model, namely a non-linear (compressive) device describing the transduction process from excitation to the neural activity [9]. The next one is a temporal integration which sums energy (information) of acoustic stimuli over a certain time interval. The final stage is a decision device in that decisions are made.

VIEMEISTER [14] suggested that the temporal integration may be approximated by a lowpass filtering. He suggested that the characteristic of this lowpass filtering could be based on a Temporal Modulation Transfer Function, TMTF. The TMTF was determined using a threshold for detecting the sinusoidal amplitude modulation of a broadband noise. The other way of modelling of the temporal integration takes advantage of a temporal integrator [8, 9]. The temporal integrator is a time window usually described by a roex(t) function, sliding in the time domain. Energy falling into the window is weighed and summed up.

More recently a new hypothesis concerning the modelling of this stage of the signal transformation in the auditory system has been proposed. This hypothesis assumes that the temporal integration may be considered as a result of a bandpass filtering that is observed in a set of overlapping bandpass filters tuned to different modulation rates,

i.e. in the so-called "modulation filter bank". This idea comes from experiments concerning across channel processes, i.e. processes of hearing of sounds exciting different auditory filters well separated in the frequency domain. In many experiments it was shown that outputs from the auditory filters tuned to well spaced frequencies had an effect on each other resulting in an increase in the modulation detection (or discrimination) threshold, as, for example, in the modulation detection/discrimination interference (MDI) phenomenon. MDI can be characterised by "tuning" in the rate of amplitude (or frequency) change domain and reaches its maximum when the modulation rate of a "target" signal is close to that of an "interfering" one. The frequency selectivity in the modulation rate domain found by HOUTGAST [4] for broadband noise and the masking in the modulation domain, found by BACON and GRANTHAM [1] for broadband stimuli, seem to confirm the hypothesis of the modulation filter bank. DAU [2, 3] showed that taking into account a second stage of filtering enabled successful theoretical interpretation of the amplitude modulation detection thresholds of noise bands and the shape of the temporal modulation transfer function.

A strong argument supporting the concept of the modulation filter bank comes from neurophysiological experiments. Each auditory neurone has its own characteristic frequency. An analysis of the signal envelope takes place probably on higher stages of the auditory system which contains neurones "tuned" to modulation rates [5]. The arrays of neurones tuned to different modulation rates are found in different places of the auditory system [6, 11, 13], especially in inferior colliculus. These arrays are suggested to be a basis for the modulation filter bank.

If the auditory system is considered as containing a second stage of filtering, i.e. the modulation filter bank, then we should be able to observe some effects in the modulation rate domain similar to those found in the first stage of filtering and performed by the auditory filters. Thus, it should be possible to observe in the modulation rate domain phenomena such as excitation pattern, tuning, masking, ringing, time integration and so on, not only for broadband stimuli but for any carrier signal.

The main purpose of this paper was to examine the concept of the modulation filter bank. Especially, we have intended to show a masking effect in the modulation rate domain and to determine modulation tuning curves for a sinusoidal carrier. These curves were determined in experiments similar to those of VOGTEN [15] who measured the so-called psychophysical tuning curves (PTC). In Vogten's experiment a fixed low-level "probe" signal was masked by a masker. The level of the masker, that just masked the probe, as a function of its centre frequency determines PTC. However, our experiments were carried out in the modulation rate domain instead of the audible frequency domain. Fixed amplitude changes of a sinusoidal carrier evoked by one modulator were masked by amplitude changes produced by a second one. Both modulators were applied to the **same** carrier signal.

2. Method

A 4 kHz, 70 dB (SPL) pure tone carrier was amplitude modulated by a signal consisting of two modulators whose schematic spectral structure is presented in Fig. 1. One

of them was a sinusoidal (probe) modulator of a frequency $f_p = 30$ Hz or 100 Hz. Clearly audible amplitude changes produced by this modulator were twice as large as the threshold changes (i.e. $2m_{th}$) and they were fixed in the experiments. A second modulator, a masking one, was a 10 Hz wide band of white noise whose centre frequency, f_c, was varied over the frequency region up to only 100 Hz since in this frequency range AM thresholds of random amplitude changes evoked by a noise-like modulator are roughly constant [10]. The maximum amplitude in the 1 s samples of the noise band used as modulating signals could be changed up to $1 - 2m_{th}$ (if m is expressed in terms of its peak value). This enabled the masking of the amplitude changes produced by the probe modulator by a limited dynamic range of the masker and avoiding an overmodulation effect ($m > 1$).

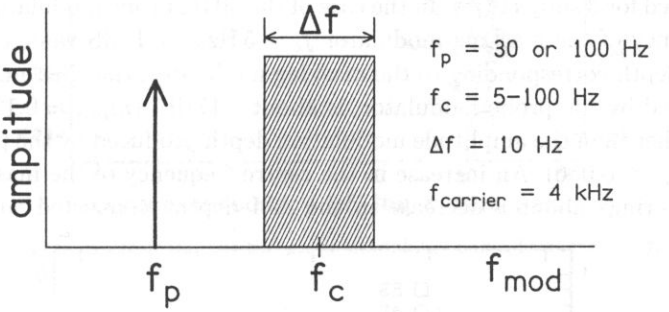


Fig. 1. A schematic illustration of the spectral structure of the modulator.

A two-alternative forced-choice (2AFC) procedure was used in the investigations. Pairs of signals, 1 s each (including rise/fall times of 50 ms), were presented in random order. One of these signals was modulated by a signal consisting of both the probe and the masking modulator. The other signal of the pair was modulated only by the masking modulator. The RMS value of the amplitude modulation depth connected with the masking modulator was kept the same in both the intervals, although the masking modulator was not a frozen noise. The task of the three normal hearing subjects was to indicate the interval containing the probe modulator, i.e. to detect specific fluctuations in the amplitude of the carrier which were clearly audible (without masking modulator) in the first 5 pairs presented at the beginning of each measuring run. Two successive correct answers caused an increase in the masking modulator amplitude, while one incorrect answer caused its decrease. This procedure tracks the threshold corresponding to 71% of correct. 12 turnpoints were determined in a single run and the threshold, as a geometric mean, was calculated based on the 8 last turnpoints. The results presented in this paper were calculated on the base of at least 4 separate measurements.

The reason for choosing a high carrier frequency was the fact that for this frequency constant values of the AM threshold are observed for a wide range of the modulation rate. For a sinusoidal modulator this range extends up to 200 Hz [2, 3, 10]. However the auditory system seems to be much more sensitive for random changes in the amplitude at rates higher than 100 Hz [10]. Thus the range of the modulation rates was limited to 100 Hz.

3. Results and discussion

The data gathered in the experiment are presented in Figs. 2 and 3. In general, it can be stated that the pattern of results obtained for three normal hearing subjects are similar. Figures 2 and 3 show the results obtained for 30 Hz and 100 Hz probe modulators respectively, and for a masking modulator which was a 10 Hz bandpass noise. In each figure the RMS values of the amplitude modulation depth connected with the masking modulator (i.e. m_{rms}), that just masked amplitude fluctuations evoked by the periodic probe modulator, are plotted as a function of the centre frequency of the masking modulator. The right vertical axis, on the other hand, shows $20\log(m_{\mathrm{rms}})$. Asterisks in the Figs. 2 and 3 indicate the frequency and the modulation depth evoked by the probe modulator averaged for 3 subjects ([1]). In the case of the 30 Hz probe modulator, for the lowest centre frequency of the masking modulator $f_c = 5$ Hz, the RMS value of the amplitude modulation depth corresponding to the masking modulator, that just masked amplitude changes evoked by the probe modulator, is about -17 dB, ($m_{\mathrm{rms}} \approx 0.12$). This value is markedly higher than the amplitude modulation depth produced by the probe modulator (-25 dB, $m_{\mathrm{rms}} \approx 0.056$). An increase in the centre frequency of the masking modulator up to 30 Hz brings about a decrease in the AM depth connected with the masking

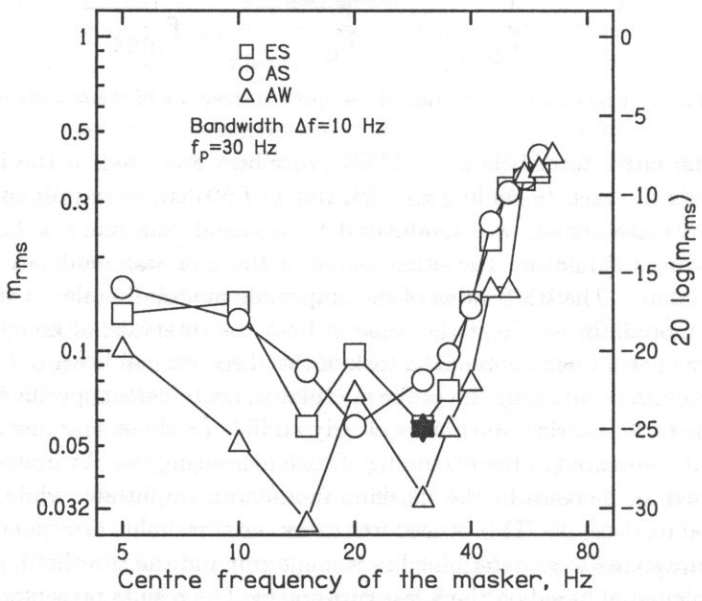


Fig. 2. Amplitude modulation depths, expressed as RMS values (m_{rms}), connected with the masking modulator that just masked amplitude fluctuations evoked by the periodic probe modulator, as a function of the centre frequency of the masking modulator. The asterisk indicates the RMS value of the AM depth related to the probe modulator. The masking modulator is a 10 Hz wide noise band. The probe modulator had a frequency of $f_p = 30$ Hz.

([1]) Thresholds for detecting the sinusoidal amplitude modulation at 30 and 100 Hz were measured for each subject in a preliminary experiment. Then, the individual values of these thresholds were used in the main experiment.

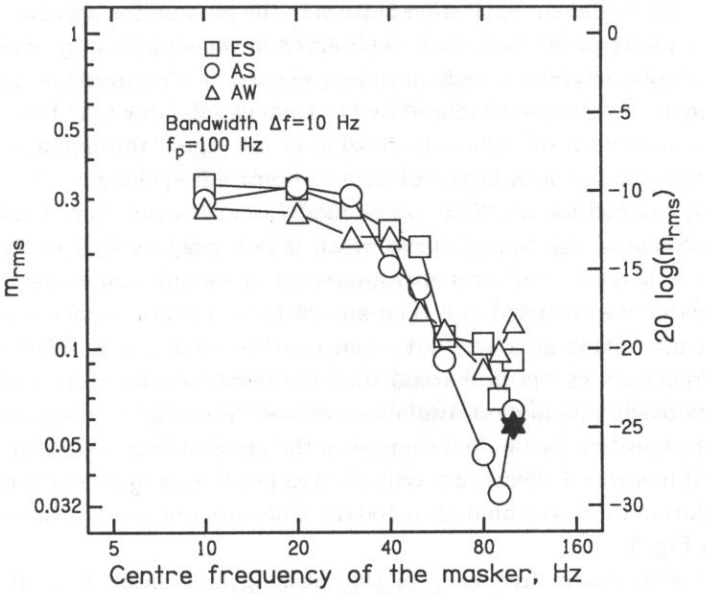


Fig. 3. The same as in Fig. 2 Hz, but the probe modulator had a frequency of $f_p = 100$ Hz.

modulator sufficient to mask amplitude changes evoked by the probe modulator. For centre frequency of the masker $f_c = 30$ Hz, the value of the AM depth connected with the masking modulator reaches its minimum which is close to the amplitude modulation depth connected with the probe modulator. When the centre frequency of the masking modulator increases above 30 Hz, a much higher amplitude of the masking modulator is required to mask periodic changes corresponding to the probe modulator. The highest centre frequency at which the masking effect is still measurable and overmodulation does not occur is about 65 Hz. The RMS values of the AM depth obtained for this centre frequency are about -7 dB, ($m_{\text{rms}} \approx 0.4$). It was impossible to obtain a masking effect for higher centre frequencies of the masking modulator without an overmodulation. For high differences between the frequencies of the probe and the masker the sensations of amplitude changes caused by these two modulators were heard separately. This may be a form of frequency selectivity in the modulation rate domain.

The results of the experiment presented are broadly consistent with the concept of a second stage of filtering in the auditory system, i.e. with the modulation filter bank concept, where signals corresponding to the envelopes of amplitude changes are filtered in some way. These data plotted as a function of the modulation rate could be called a "modulation tuning curve". The low-frequency side of this curve has a smaller slope than the high-frequency one. This means that the periodic changes in amplitude occurring at a rate of 30 Hz can be quite easily masked by amplitude changes at a lower rate, although the effectiveness of this masking tends to be the highest when the difference between the modulator frequencies is close to zero. The shape of those curves can be interpreted assuming that the excitation pattern in the modulation frequency domain is

unsymmetrical, i.e. that the excitation pattern in the modulation domain has a shallower slope at the high-frequency side, as it is observed in the audible frequency domain case. Thus, if the probe modulator falls into area excited by the masking modulator, then the small amplitude changes produced by the probe modulator may be entirely masked. Such excitation could result from the passing of the signal through a system of broad bandpass filters whose width increases with the centre frequency.

The data obtained for the 30 Hz probe modulator illustrate only a short part of the low-frequency side of the tuning curve which is not steep enough to show the tuning effect clearly. Therefore, the next measurements of modulation masking for a 100 Hz probe modulator were carried out. The aim of these measurements was to show that the masking modulator at a very low centre frequency is not an effective masker for the higher frequency of the probe and that the low-frequency side of the modulation tuning curves reaches its highest available level too. Since the auditory system seems to be much more sensitive for random changes in the amplitude at rates higher than 100 Hz [10] (where also spectral effects, not only the temporal ones, may pay a role), the range of the modulation rates was limited to 100 Hz. The data obtained in this experiment are presented in Fig. 3.

For the lowest centre frequency of the masking modulator $f_c = 10$ Hz, the RMS value of the amplitude modulation connected with the masking modulator, that just masked amplitude changes corresponding to the periodic probe modulator, was about -10 dB ($m_{\mathrm{rms}} \approx 0.3$). On the other hand, for centre frequencies ranging from 50 Hz to 100 Hz the values of the AM depth connected with the masking modulator are very close to those connected with probe modulator. Thus, an increase in the centre frequency of the masking modulator brings about that smaller values of the masking modulator amplitude are required to mask effectively the periodic changes in amplitude of the carrier evoked by the probe modulator. These data are also consistent with the concept of the modulation filter bank. However, in spite of the previous result, these data illustrate the low-frequency side of the modulation tuning curve in the whole available dynamic range of the masking modulator. They are also consistent with the idea that a bandwidth of the modulation filter, related to its centre frequency, is rather broad in comparison to that of the auditory filter.

4. Conclusions

The most important experimental results of this paper are as follows:

1. A masking effect in the modulation rate domain can be observed in the auditory system: amplitude changes (of a sinusoidal carrier) at a given rate (less than or equal to 100 Hz) may be masked by amplitude changes at different rates within the range of $5 \div 100$ Hz.

2. The effectiveness of this form of masking is the highest when modulation frequencies of the probe and the masking modulators are close to each other, and decreases when the difference between these frequencies increases. A similar masking pattern is observed in the audible frequency domain: a tone at a given frequency is masked most

by a masker whose centre frequency is close to the frequency of that tone as indicated by the psychophysical tuning curves (PTCs). The masking effect observed in the experiments presented, by analogy to the psychophysical tuning curves, may be considered to be a tuning in the modulation rate domain. This effect reflects also the frequency (rate) selectivity in the modulation frequency domain.

3. For large differences between the frequencies of the probe and the masking modulator, amplitude fluctuations corresponding to each of these modulators produce well separated sensations. This suggests that the components of the modulator could be clearly resolved, for certain frequency differences between them.

The concept of auditory filters is broadly used for description of the auditory system in the auditory frequency domain. Thus, it seems that the concept of a second stage of filtering based on modulation filters could be applied to describe the auditory system in the modulation rate domain. In the first stage of the two-stage filtering model, an auditory input is primarily filtered in the auditory filters. Then, in the second stage of filtering, the modulation filter bank is applied to the envelope of the amplitude changes of the signal.

Acknowledgements

We thank an anonymous reviewer for helpful comments on an earlier version of this paper.

References

[1] S.P. Bacon and D.W. Grantham, *Modulation masking: effects of modulation frequency, depth and phase*, J. Acoust. Soc. Am., **85**, 2575–2580 (1989).

[2] T. Dau, B. Kollmeier and A. Kohlrausch, *Modelling modulation perception: modulation low-pass filter or modulation filter bank?*, [in:] Psychoacoustics, Speech and Hearing Aids, B. Kollmeier [Ed.], World Scientific, Singapore 1996.

[3] T. Dau and D. Püschel, *A quantitative model of the effective signal processing in the auditory system*, [in:] Contributions to Psychological Acoustics, A. Schick [Ed.], Bibliotheks- und Informationssystem der Universität Oldenburg, Oldenburg 1993.

[4] T. Houtgast, *Frequency selectivity in amplitude-modulation detection*, J. Acoust. Soc. Am., **85**, 1676–1680 (1989).

[5] R.H. Kay, *Hearing of modulation in sounds*, Physiol. Rev., **62**, 894–975 (1982).

[6] C. Lorenzi, C. Micheyl and F. Berthommier, *Neuronal correlates of perceptual amplitude-modulation detection*, Hear. Res., **90**, 219–227 (1995).

[7] B.C.J. Moore, *An introduction to the psychology of hearing*, 4th Ed., Academic Press, London 1997.

[8] B.C.J. Moore, B.R. Glasberg, C.J. Plack and A.K. Biswas, *The shape of the ear's temporal window*, J. Acoust. Soc. Am., **83**, 1102–1116 (1988).

[9] A.J. Oxenham and B.C.J. Moore, *Modelling the additivity of nonsimultaneous masking*, Hear. Res., **80**, 105–118 (1994).

[10] E. Ozimek, J. Konieczny, Y. Suzuki and T. Sone, *Random changes in envelope of AM tones and their detection*, J. Acoust. Soc. Jpn. (E), **19** (1998).

[11] A.R. PALMER, *Neural signal processing*, [in:] Hearing, B.C.J. MOORE [Ed.], Academic Press, San Diego 1995.

[12] R.D. PATTERSON and B.C.J. MOORE, *Auditory filters and excitation patterns as representations of frequency resolution*, [in:] Frequency Selectivity in Hearing, B.C.J. MOORE [Ed.], Academic, London 1986.

[13] A. REES and A.R. MOLLER, *Responses of neurones in the inferior colliculus of the rat to AM and FM tones*, Hearing. Res., **10**, 301–310 (1983).

[14] N.F. VIEMEISTER, *Temporal modulation transfer functions based on modulation thresholds*, J. Acoust. Soc. Am., **66**, 1364–1380 (1979).

[15] L.L.M. VOGTEN, *Pure-tone masking: A new result from a new method*, [in:] Facts and Models in Hearing, E. ZWICKER and E. TERHARDT [Eds.], Springer-Verlag, Berlin 1974.

ARCHIVES OF ACOUSTICS
23, 4, 499–511 (1998)

APPLICATION OF THE FRACTAL ANALYSIS IN THE SEA BOTTOM RECOGNITION

Z. ŁUBNIEWSKI and A. STEPNOWSKI

Technical University of Gdańsk
Acoustics Department
(80-952 Gdańsk, Narutowicza 11/12)
e-mail: lubniew@eti.pg.gda.pl, astep@pg.gda.pl

The paper proposes a newly developed simple method of the sea bottom typing using elements of fractal analysis. The fractal dimension was calculated as a box dimension for sampled envelopes of echo signals from four types of sea bottom recorded during mobile acoustic surveys carried out in Lake Washington. The histograms of obtained box dimension values for particular bottom types as well as scatter diagrams in box dimension-echo energy parameter space were constructed and analysed. The obtained results show that this simple method can be used both alone and combined with other methods for on board sea bed recognition in real time with accuracy similar to that of other methods.

1. Introduction

The problem of a proper sea bottom identification is important in many fields, for instance in hydrography, marine engineering, environmental sciences, fisheries and other domains. Acoustic methods of sea bottom typing, which use the information retrieved from the bottom echo, have advantages over the other methods (e.g. geological cores or remotely operated vehicles with TV cameras), as being faster, non-invasive and more cost effective.

In general, the following approaches are used in the acoustic methods of bottom typing:

- measurement of energy ratio of the first and second bottom echo ("Roxann" method) [3] or division of the first echo signal [1],
- comparison of the actual cumulative echo envelopes with theoretical patterns [7],
- analysis of a set of values of acoustic and statistical parameters of the echo envelope using cluster analysis [12, 13] or artificial neural networks [4, 8],
- using parametric sources and wideband "chirp" signals [2, 4].

All acoustic methods of bottom identification are based on the assumption, that the received echo pulse contains the significant and possible to retrieve information about the bottom characteristics. The authors have developed the simple method in which fractal dimension of an echo envelope is used as a signature of the bottom type. It is known that the surface of sea bottom is one of examples of a fractal structured object in

nature [9]. Taking this into consideration we made an implicit assumption that fractal structure of the bottom is transferred onto its image observed in the reduced form as an envelope of sonar echo. The fractal dimension of the bottom surface can be treated as a measure of its complexity and roughness, which is correlated with hardness of the sediment, as the harder bottom, the more irregular, corrugated shape it has. That is why we predicted that the envelopes of echoes from harder bottom would have greater fractal dimension values than those of echoes from softer bottom.

2. The outline of the fractal analysis

It was found that the shapes and structures of many objects in nature usually show no regularities that are characteristic of simple figures, which can be easily described in terms of the Euclidean geometry. That is why this kind of geometry is not the most adequate tool for describing such objects. On many occasions, however, nature has proved to accommodate various types of elements with the fractal structure, e.g. the structure of plants' leaves, corrugated sea surface or bottom surface [9], which suggests that fractal analysis methods are the proper methods for studying and describing such elements.

The Koch snowflake (Fig. 1a) could be an example of a fractal figure. It is defined as a limit of the sequence of simple iterative steps: Starting with the equilateral triangle, each consecutive stage is constructed by replacing each line segment with copy of the broken line of the following shape: $\_/\!\!\backslash\_$. In the random version of the Koch snowflake (Fig. 1b), in every stage the replacement by $\_/\!\!\backslash\_$ or $\backslash\!\!/$ may be made, with probability of 0.5 for each case. The Cantor set [9] can be another example of the fractal structure.

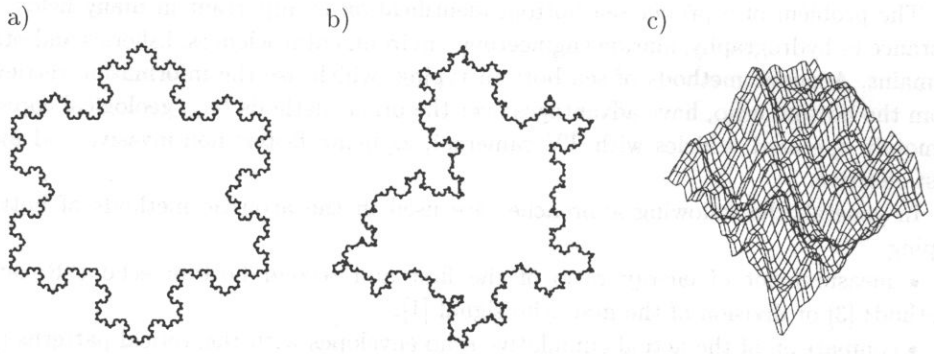


Fig. 1. The examples of fractal 2D and 3D figures; a) the regular Koch snowflake, b) the randomized Koch snowflake, c) fractal surface.

Fractal sets are defined as scale-invariant (self-similar) geometric objects. A geometric object is called scale-invariant, if it can be written as a union of rescaled copies of itself. Regular fractals, such as the Koch snowflake, Cantor set or Sierpiński triangle, demonstrate exact self-similarity [9, 11]. Random fractals show a weaker, statistical version of self-similarity.

The interesting feature of fractal figures is its similarity to some objects occurring in nature. For instance, the random version of the Koch snowflake from Fig. 1b could remind a mountain lake, and the computer generated fractal surface (Fig. 1c) has the shape similar to that of corrugated sea surface or sea-bed.

To be able to investigate and describe real objects in nature with use of the fractal analysis, one should have a method of measuring the magnitude of their dimensions and comparing them. Standard methods, that consist of measuring of length or area of 1D and 2D figures, are not appropriate here. The Koch snowflake, for instance, has an area equal to zero, but its length tends to infinity when the size of the measuring step tends to 0.

One of the used fractal dimensions, so-called Hausdorff dimension [9, 11] may be the solution to this problem, as it can be used as a measure of many very general sets, including fractals. The Hausdorff dimension of a subset X of Euclidean space is defined as a limit

$$D = \lim_{r \to 0} \frac{-\log N(r)}{\log r}, \tag{1}$$

where $N(r)$ denotes the smallest number of open balls of radius r needed to cover subset X; an open ball $B(p,r) = \{x : \text{dist}\,(x,p) < r\}$, where $\text{dist}\,(x,p)$ is the distance between points x and p. For example, it could be shown [9], that the Koch snowflake has the Hausdorff dimension equal to $\log 4/\log 3 \approx 1.262$.

It is easy to see that the dimension defined by formula (1) measures of the complexity of a given figure. In the case of a sea bottom echo envelope, it should be an indicator of the complexity or variability of this waveform, which may imply its use as a signature of the type of investigated sea-bed.

3. Materials and methods

The bottom echoes data we used to calculate the fractal dimension were recorded during acoustic surveys on Lake Washington with the use of digital DT4000 BioSonics echosounder with two operating frequencies: 38 and 120 kHz. Simultaneously the current position of the research vessel has been controlled and recorded using the GPS system. Data acquisition was performed both while the vessel was moving along the selected transects, and while the ship was anchored. In each case the type of sea bottom in the given water region was known from ground samples taken by divers. This enabled a validation of the developed and applied method of bottom typing. The length of the sounding pulse of the echosounder was 0.4 ms and the frequency of sampling of the echo envelope at the echosounder output was equal to 41.66 kHz.

It is not easy to calculate the fractal dimension of a figure following the definition of the Hausdorff dimension (1). Therefore we decided to use the box dimension [9], that can replace the Hausdorff dimension for many sets, including shapes of our echo pulses. The box dimension of a plane figure of investigated echo waveforms is defined as follows. Let $N(\Delta s)$ denote the number of boxes in a grid of the linear scale Δs which meet the

set X on a plane. Then X has a box dimension

$$D = \lim_{\Delta s \to 0} \frac{-\log N(\Delta s)}{\log \Delta s}. \tag{2}$$

The method of evaluating the box dimension of a bottom echo envelope is explained in Fig. 2. The grid of square boxes of side Δs is superimposed on the graph of echo envelopes and the number of boxes which consist of the fragments of the envelope graph is counted and denoted as $N(\Delta s)$. It must be pointed out, that here we cannot use the accurate definition of box dimension with limit given by formula (2), because a digitized echo pulse consists of a finite set of straight sections and it is not a real fractal. That is why one would always obtain the box dimension value equal to 1. However, it is possible to approximately evaluate this dimension for such a figure calculating it not as a limit, but assuming Δs to have a finite fixed value:

$$\widetilde{D}_{\text{box } \Delta s} = \frac{-\log N(\Delta s)}{\log \Delta s} \tag{3}$$

or substituting Δs by $1/N_s$, where N_s is both horizontal and vertical number of boxes in the grid:

$$\widetilde{D}_{\text{box } N_s} = \frac{\log N(N_s)}{\log N_s}. \tag{4}$$

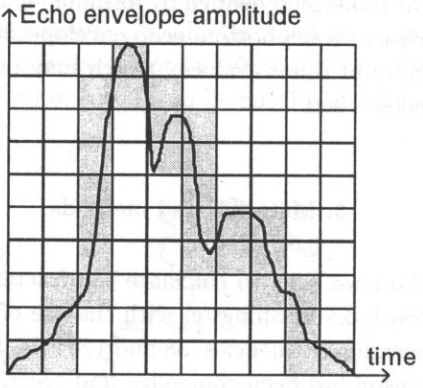


Fig. 2. Illustration of the box dimension evaluation. In the considered case, $\Delta s = 0.1$, $N_s = 10$, $N(\Delta s) = 30$.

We evaluated the box dimension of the echo pulses envelopes acquired during the surveys using the above concept with $\Delta s = 1/36$, after normalizing each echo pulse to the standard length and height. Additionally, we tested a modified version of the box dimension evaluation procedure, which has two following differences from the one discribed above.

Firstly, we used the grid of rectangular, instead of square boxes, what meant that the numbers of boxes in the grid along the horizontal (N_{sx}) and vertical (N_{sy}) axis were not equal, as we wanted to investigate the irregularity, or complexity, of an echo

envelope using different horizontal and vertical scale. In such a case, the formula (3) or (4) used for evaluating the box dimension had to be modified, to take into account the difference between N_{sx} and N_{sy}. Following the principle, that the least irregular, straight horizontal line – the graph of a constant function – has the box dimension equal to 1 and such the line would cause N_{sx} rectangular boxes of the grid to be marked as dark, and the second principle, that the most composed line, which would cause all of boxes $(N_{sx} \cdot N_{sy})$ to be darked, has the box dimension approximately equal to 2, we derived the more general formula for the box dimension:

$$\widetilde{D}_{\text{box}\,N_{sx},N_{sy}} = \frac{\log\left[N(N_{sx},N_{sy}) \cdot \dfrac{N_{sy}}{N_{sx}}\right]}{\log N_{sy}}. \tag{5}$$

The second modification in procedure of calculating the box dimension of echo envelopes was made to avoid normalizing, what means enlarging or squeezing all echo pulses to one standard length, what could affect the final results. Because of differences in length of particular echo pulses, we applied the concept of moving window of fixed length W samples, not greater than length of the shortest echo pulse in the data set. For each pulse we evaluated the box dimension as an average of the box dimensions calculated for parts of signal with length W, sequentially being chosen from the whole pulse:

$$\widetilde{D}_{\text{box}} = \frac{1}{N-W+1} \sum_{i=1}^{N-W+1} \widetilde{D}_{\text{box}}(i, i+W-1), \tag{6}$$

where N is the whole echo pulse length in samples and $\widetilde{D}_{\text{box}}(i,j)$ denotes the box dimension calculated for a part of the pulse starting at i-th sample and ending at j-th sample. We used the window of $W = 48$ samples in length.

Three sets of data for testing the proposed sea bottom typing method were used:

1) echoes recorded on the anchored ship, at echosounder frequency 120 kHz,

2) echoes recorded on the anchored ship, at echosounder frequency 38 kHz,

3) echoes recorded during moving along transects, at echosounder frequency 120 kHz.

Each of these sets contained digitized envelopes of echo pulses from four different types of sea bottom: soft mud, soft sand, hard sand and rock. There were more than 600 echoes from each type of bottom in each data set. Before calculations, only 10% of pings of the highest amplitude were selected from each data set in order to take into account only the echoes from pulses with the most likely normal incidence to the bottom (and drop the others), reducing the effect of a ship's pitching and rolling in this way.

The amplitude threshold used for the analysed signals was −70 dB. The box dimension was calculated for each echo pulse separately, and histograms of its values for each type of bottom were constructed and analysed. Additionally, other echo pulse parameters, connected with the echo energy, were calculated in some cases and scatter diagrams with points in the 2-dimensional parameter space were constructed.

4. The results of the box dimension evaluation

Figures 3a, 3b, 4 and 5 show the histograms of box dimension values obtained for echo pulses for four types of bottom. Figures 3a and 3b show the results obtained for the first

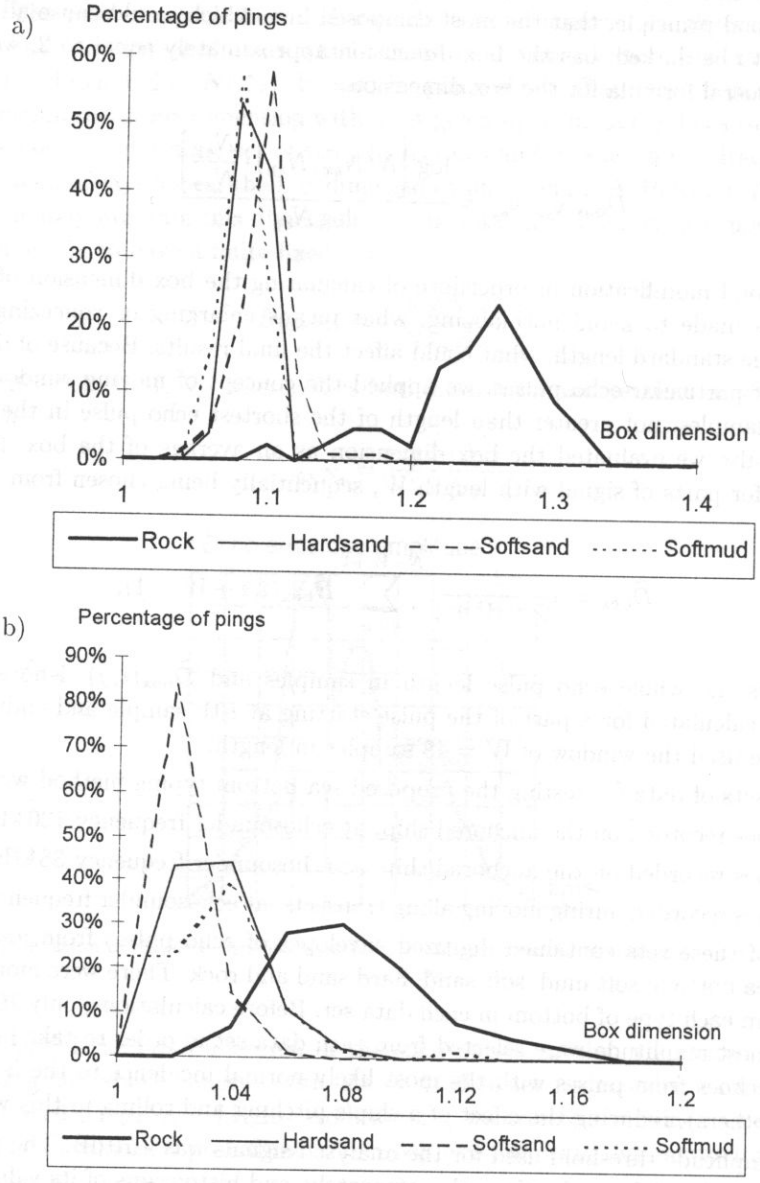


Fig. 3. a. The histogram of the box dimension evaluated from bottom echo pulses recorded on an anchored ship at frequency 120 kHz (first method – square boxes). b. The histogram of the box dimension evaluated from bottom echo pulses recorded on an anchored ship at frequency 120 kHz (second method – rectangular boxes and moving window).

data set, Fig. 4 the results for the second data set, and Fig. 5 for the third. Figures 3a, 4 and 5 present histograms of the box dimension obtained by applying the first method, using the grid of square boxes and without applying the moving window. Figure 3b presents the results obtained by applying the second, modified method of box dimension evaluation, using rectangular boxes ($N_{sx} = 48$, $N_{sy} = 16$) and the moving window. As one can see by comparing Fig. 3a and Fig. 3b, there are no significant differences between the results of both methods, as far as the shapes of histograms and mutual relations between distributions of box dimension of echoes from particular types of bottom are concerned. We found this also true for two other sets of data, so we decided to stay with the first method results for further analysis. However, the visible differences between absolute values of box dimension obtained in the first and second method (Fig. 3a and Fig. 3b) need the explanation. This difference is the result of applying grids with boxes of different sizes in each of these two cases, what certainly cannot affect the results of box dimension calculation for mathematical, "real" fractal figures, but can cause differences between obtained values in the case of an echo envelope which is in fact a broken line and not a fully featured fractal.

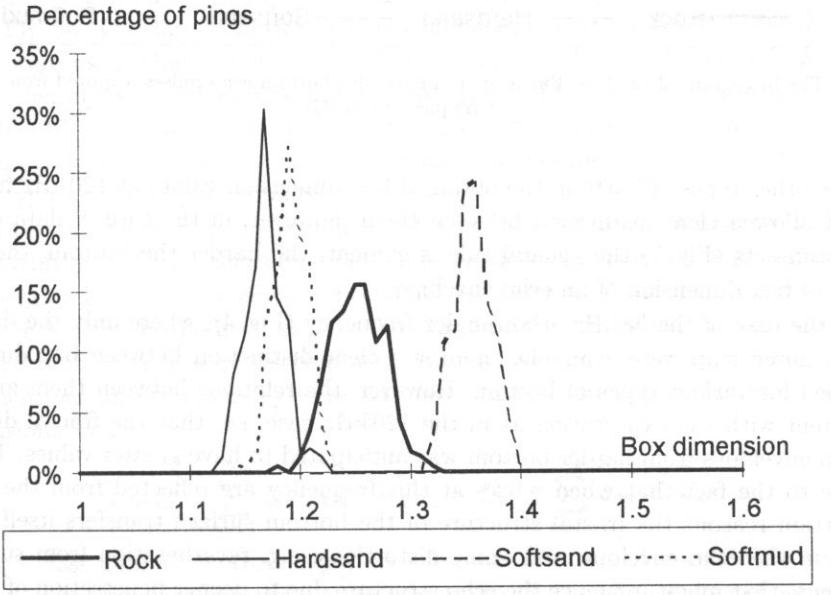


Fig. 4. The histogram of the box dimension evaluated from bottom echo pulses recorded on an anchored ship at frequency 38 kHz.

The presented histograms of box dimension values show that in the case of 120 kHz frequency of the echosounder (Fig. 3) there is a clear difference between the values of the box dimension for a rocky bottom and for the other types of seabed, especially for the data collected on the anchored ship. In the case of rock the box dimension values are significantly greater ($1.2 \div 1.3$ vs. ca 1.1 for the other types of bottom), what is consistent with expectations, because the surface of rocky bottom is more corrugated and irregular.

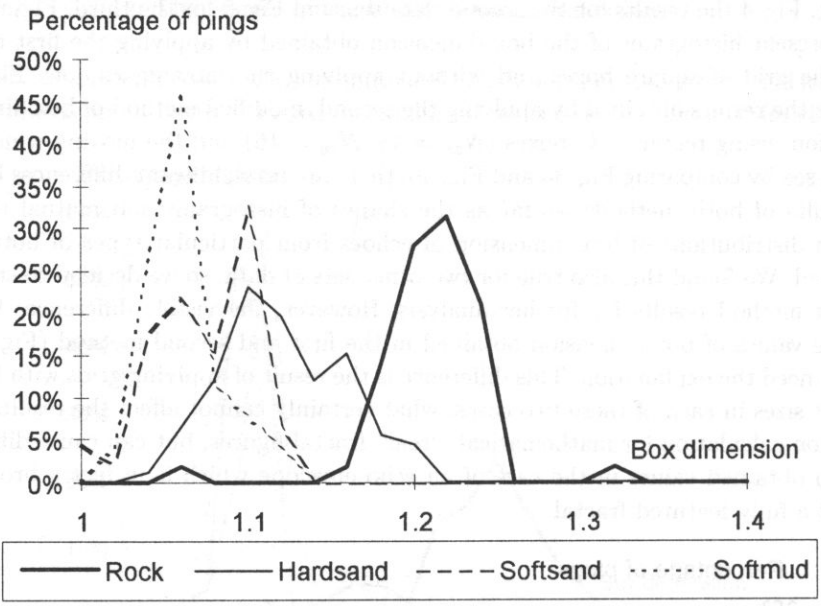


Fig. 5. The histogram of the box dimension evaluated for bottom echo pulses acquired from transects at frequency 120 kHz.

For the other types of bottom, the obtained box dimension values at 120 kHz frequency do not allow a clear distinction between them, however, in the case of data obtained from transects (Fig. 5) the general rule is evident: the harder the bottom, the greater values of box dimension of an echo envelope.

In the case of the 38 kHz echosounder frequency (Fig. 4), where only the data from the anchored ship were available, there is a clear distinction between box dimensions obtained for various types of bottom. However, the relations between them are not as consistent with the expectation as in the 120 kHz case, i.e. that the fractal dimension of echo envelopes from harder bottom was anticipated to have greater values. This may be due to the fact that when pulses at this frequency are reflected from the bottom, for certain reasons the fractal structure of the bottom surface transfers itself onto its image in the echo envelope with some distortions, e.g. reverberation from subbottom sediments that might influence the echo structure due to deeper penetration of pulses of lower frequency.

5. Combining the box dimension parameter with the echo energy. Application of the minimum distance classifier

The authors predicted that performance of the described fractal dimension method of sea bottom typing might be improved by combining it with other methods, using the different parameters of received bottom echo. To test this prediction, besides the

evaluation of box dimension of an echo envelope, the parameters containing information about the echo energy were calculated. Subsequently, the scatter diagrams of data points in the two parameter space were constructed, in similar way as in Roxann or the first echo division method [1, 3]. Three parameters connected with the echo energy were calculated for each pulse, viz.:

- the energy of the attack phase of an echo E_1', which may be treated as an indicator of the hardness of bottom,
- the energy of the decay and release phases of an echo E_1 – the indicator of the roughness of bottom,
- the total energy of the first echo $E = E_1' + E_1$ [1]:

$$E_1' = \int_a^{\tau} V^2(t)\,dt, \qquad E_1 = \int_{\tau}^{b} V^2(t)\,dt, \qquad E = \int_a^{b} V^2(t)\,dt, \tag{7}$$

where V – echo voltage, a – start point of an echo pulse, b – end point of an echo pulse, τ – position of the maximum echo amplitude.

The scatter diagrams of the echoes in the 2-dimensional parameter space (E, D_{box}) are shown in Fig. 6: the space of parameter E_1' vs. D_{box} in Fig. 6a and the space of E vs. D_{box} in Fig. 6b. The data from the anchored ship acquired at the echosounder frequency 120 kHz were used. One can see on the both diagrams, that while the box dimension allows to distinguish between rock and the other types of bottom, the value of the second parameter, E_1' or E, could be in this case used for discrimination between hard sand and mud. These diagrams show that using more than one parameter of an echo improves the bottom typing method performance.

In addition, the authors constructed and tested an automatic procedure of bottom type recognition – the 2-dimensional minimum distance classifier [10] based on values of box dimension and the total energy of echo pulses, which used the data from the scatter diagram in Fig. 6b. In the first stage of the classification procedure, 70% of echo pulses from each bottom type were selected as a training set and the average values of their box dimensions and total energy were calculated as co-ordinates of cluster centers in a 2-dimensional space:

$$\overline{D}_{\text{box}\,k} = \sum_{m=1}^{M_k} D_{\text{box}\,km} \Big/ M_k\,, \qquad \overline{E}_k = \sum_{m=1}^{M_k} E_{km} \Big/ M_k\,, \tag{8}$$

where k – class number (bottom type), m – number of echo pulse within a given class, $D_{\text{box}\,km}$ – box dimension value of a given echo pulse, E_{km} – total energy of a given echo pulse, M_k – number of echoes in a given class in entire training set (70% of all echoes in a class), $\overline{D}_{\text{box}\,k}$ – average value of $D_{\text{box}\,km}$ for a given class in the training set, $\overline{E}_k$ – average value of E_{km} for a given class in the training set.

Subsequently, the procedure of classification was carried out, following the rule of minimum distance from a cluster center i.e. each echo pulse was assigned to that class, for which the distance between the echo point in the 2-dimensional (E, D_{box}) space and the cluster center of that class was minimal, what can be expressed as:

$$C_{km} = l\colon \text{dist}\left(X_{km}, \overline{X}_l\right) = \min_{\kappa=1\ldots K} \left(\text{dist}\left(X_{km}, \overline{X}_\kappa\right)\right), \tag{9}$$

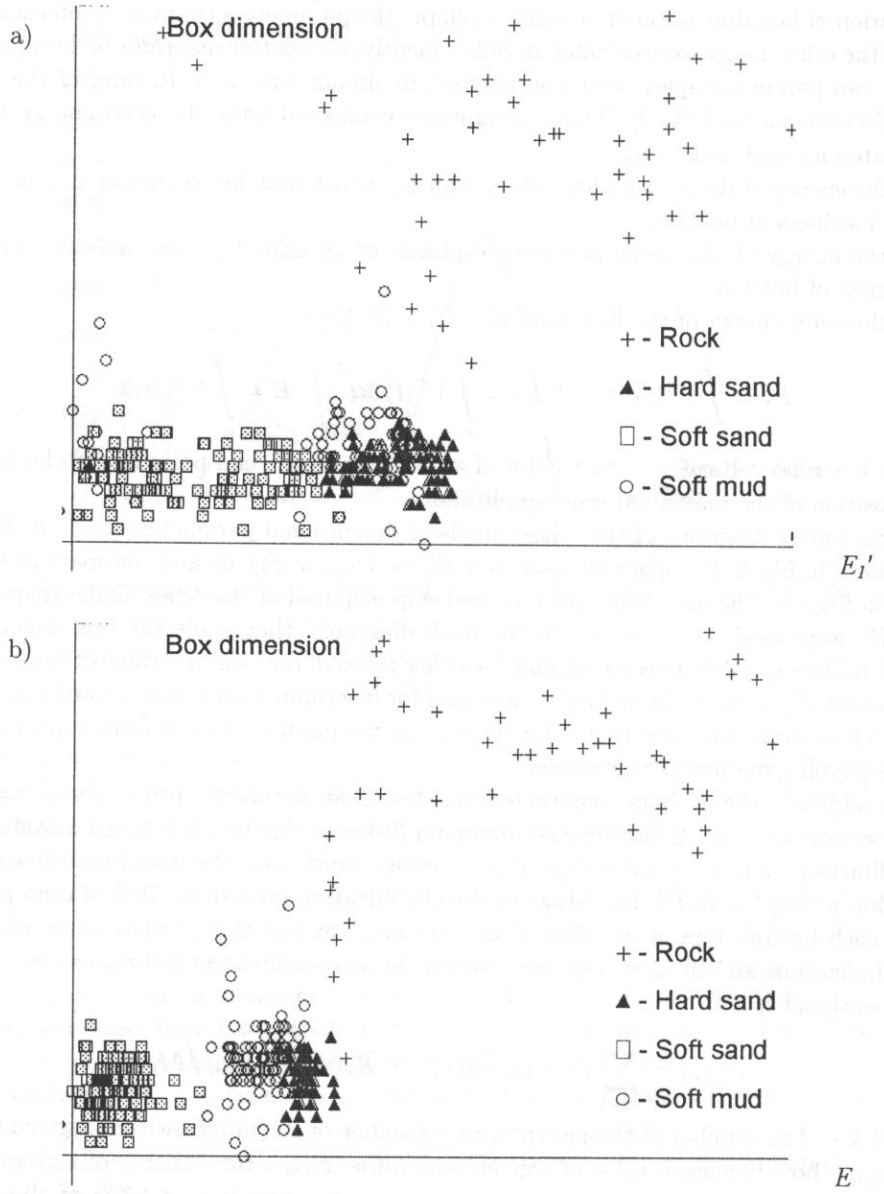


Fig. 6. a. Scatter diagrams of bottom echoes from anchored ship for frequency 120 kHz in the 2-dimensional space of parameters E_1' and D_{box}. b. Scatter diagrams of bottom echoes from anchored ship for frequency 120 kHz in the 2-dimensional space of parameters E and D_{box}.

where K – number of classes, C_{km} – class number assigned to the pulse, X – point in the 2-parameter space corresponding to an echo or to a cluster center: $X(E, D_{\text{box}})$.

The standard formula for calculating the distance between two points in the parameter space (E, D_{box}) was slightly modified and each square of the difference of points'

co-ordinates was additionally divided by the variance of a given parameter in the training set:

$$\text{dist}\,(X_{km}, \overline{X}_l) = \sqrt{\left(E_{km} - \overline{E}_l\right)^2 / \sigma^2_{E_l} + \left(D_{\text{box}\,km} - \overline{D}_{\text{box}\,l}\right)^2 / \sigma^2_{D_{\text{box}\,l}}}, \tag{10}$$

where $\sigma^2_{E_l}$ and $\sigma_{D_{\text{box}\,l}}$ denote the variance of energy and box dimension values respectively within l-th class in the training set.

Additionally, in the preliminary stage of the classification procedure, all values of both parameters E and D_{box} were linearly mapped to [0, 1] interval, to avoid the influence of great differences in ranges of E and D_{box}.

The classification results are shown in Table 1 in a form of the percentage values in confusion matrix.

Table 1. Confusion matrix for the tested minimum distance classifier based on values of box dimension and total energy of echo pulses.

True class	Assigned class (%)			
	Soft mud	Soft sand	Hard sand	Rock
Soft mud	**100**	0	0	0
Soft sand	0	**92.59**	7.41	0
Hard sand	0	27.59	**72.41**	0
Rock	0	7.69	0	**92.31**
Total percentage of correct assignments: 89.32%				

The classification results are satisfactory as the particular classes are quite well separable, what is easily to seen in Fig. 6b. In general, only echoes from soft and hard sand were confused, as the result of the E and D_{box} parameters distributions overlap for these types of bottom.

6. Inverse filtering

An extension of the described method could be an application of the deconvolution of the spatio-temporal scattering impulse response of seabed from bottom echo, before the fractal dimension calculation. In general, the spatio-temporal acoustic field $p(t, \mathbf{r})$ measured after passing the system characterised by scattering function $h(t, \mathbf{r})$, is related to transmitted acoustic field $g(t, \mathbf{r})$ via the integral equation [5]:

$$p(t, \mathbf{r}) = \int h(t - \tau, \mathbf{r}) \cdot g(\tau, \mathbf{r})\, d\tau. \tag{11}$$

If one is given $g(t, \mathbf{r})$ and $p(t, \mathbf{r})$ functions, the scattering function $h(t, \mathbf{r})$ can be obtained via direct or iterative inverse filtering techniques. However, if only time description $p(t)$ of a received acoustic pressure field $p(t, \mathbf{r})$ is available (in a form of the echo envelope), a special model of sound scattering on a rough bottom surface will be helpful for obtaining the bottom scattering function $h(t, \mathbf{r})$.

7. Conclusion

The results of the presented investigation are promising, as they show that evaluation of the fractal dimension of acoustic echo signal scattered at the seabed may be a useful simple (single- or two- parameter) method of bottom typing. The results are not worse than those obtained concurrently using other methods. The fractal box dimension method when combined with other parameters of bottom echo envelope can add useful information and improve the reliability of sea bottom typing. However, it must be noticed that the data used in paper were acquired only from relatively small water region, so the method should be verified for a larger area and for different bottom types.

The developed method was implemented (among other methods) in Visual Bottom Typer [8] – the computerized system for the seabed identification and imaging, developed in Acoustics Department, Technical University of Gdask, in co-operation with BioSonics Inc. In this system fractal dimension method is combined with the technique based on the first echo energy division [1]. The data points are plotted in the (E_1, D_{box}) space in run time and the bottom type is assigned referring to regions in that space defined by user as shown in Fig. 7, similarly to Roxann method [3].

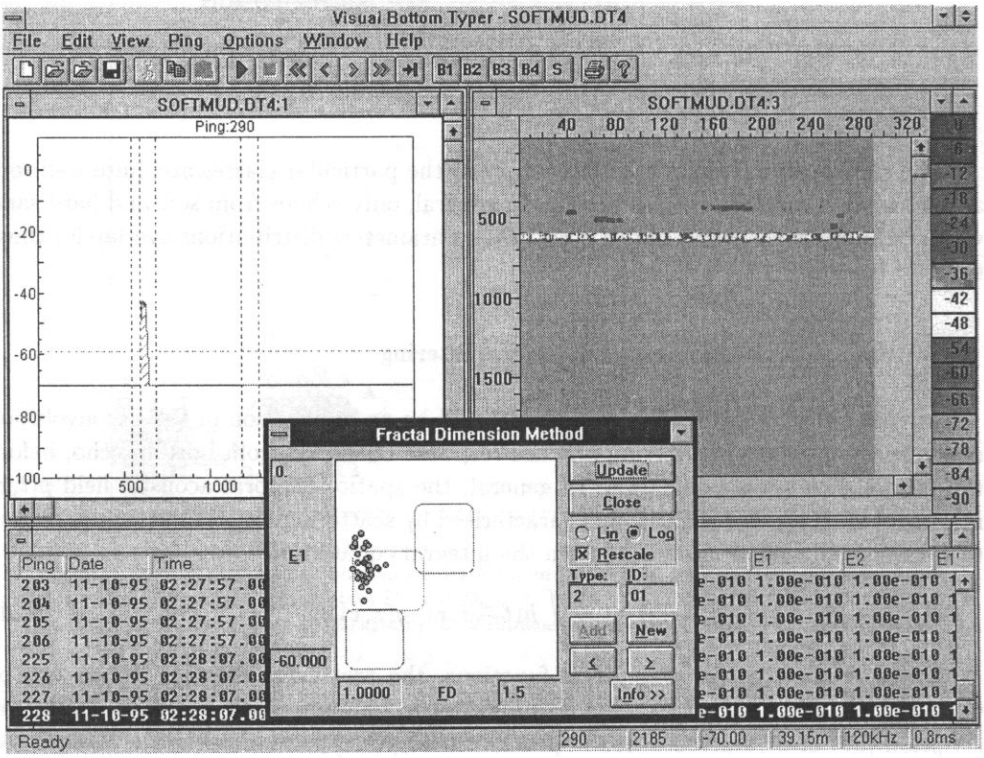


Fig. 7. Sample screen of VBT (Visual Bottom Typer) system for sea bed identification and imaging, with Fractal Dimension Method window.

References

[1] D. Bakiera, A. Stepnowski, *Method of the sea bottom classification with a division of the first echo signal*, Proceedings of the XIIIth Symposium on Hydroacoustics, Gdynia-Jurata, 55–60 (1996).

[2] L.R. LeBlanc, L. Mayer, M. Rufino, S.G. Schock, J. King, *Marine sediment classification using the chirp sonar*, J. Acoust. Soc. Am., **91**, 1, 107–115 (1992).

[3] R.C. Chivers, *Acoustical sea-bed characterisation*, XIth Symposium on Hydroacoustics, Jurata (1994) [invited paper].

[4] M. Gensane, H. Tarayre, *Tests of sea-bottom discrimination with a parametric array*, Acoustic Letters, **16**, 5, 110–115 (1992).

[5] W.C. Knight, R.G. Pridham, S.M. Kay, *Digital signal processing for sonar*, Proc. IEEE, **69**, 11, 1451–1506 (1981).

[6] Z. Łubniewski, A. Stepnowski, *Sea bottom typing using fractal dimension*, Proceedings of the International EAA/FASE Symposium on Hydroacoustics and Ultrasonics, Gdańsk-Jurata, 131–134 (1997).

[7] E. Pouliquen, X. Lurton, *Sea-bed identification using echosounder signal*, European Conference on Underwater Acoustics, Elsevier Applied Science, London and New York, 535 (1992).

[8] A. Stepnowski, M. Moszyński, R. Komendarczyk, J. Burczyński, *Visual real-time Bottom Typing System (VBTS) and neural networks experiment for sea bed classification*, Proceedings of the 3rd European Conference on Underwater Acoustics, Heraklion, Crete, 685–690 (1996).

[9] H. M. Hastings, G. Sugihara, *Fractals. A user's guide for the natural sciences*, Oxford University Press, Oxford, New York, Tokyo (1994), pp. 7–77.

[10] K. Jajuga, *Statistical theory of pattern recognition* [in Polish], PWN, Warszawa 1977, pp. 119–124.

[11] B.B. Mandelbrot, *The fractal geometry of nature*, Freeman, San Francisco 1982.

[12] *anon.*, Quester Tangent (1997), Quester Tangent Marine Instrumentation, 13 February 1997, http://www.questercorp.com/

[13] J. Tęgowski, *Characteristic features of backscattering of the ultrasonic signals from the sea bottom at the Southern Baltic* [in Polish], Ph.D. Thesis, Institute of Oceanology of Polish Academy of Sciences, Sopot 1994.

ARCHIVES OF ACOUSTICS
23, 4, 513–520 (1998)

A MULTILAYER METHOD FOR LINEARITY DETERMINATION OF THE PVDF HYDROPHONE FOR PRESSURES UP TO 2.3 MPa

L. FILIPCZYŃSKI, J. ETIENNE, T. KUJAWSKA,
R. TYMKIEWICZ and J. WÓJCIK

Department of Ultrasound
Polish Academy of Sciences,
Institute of Fundamental Technological Research
(00-049 Warszawa, Świętokrzyska 21, Poland)

A new method of linearity measurements of a PVDF membrane hydrophone was elaborated. High pressure pulses up to 6 MPa were obtained in the focus of a concave PZT transducer with the frequency of 3 MHz. The method is based on the pressure decrease of these pulses transmitted through metal layers with known acoustic impedances, immersed in water. In this way one obtained pressure changes at constant shapes (spectrum) of the pulses which were measured by means of the hydropbone under investigation. The measurement results confirmed the linearity of the hydrophone up to pressures equal to 2.3 MPa. Correlation coefficients of 5 measured relations were in average equal to $r = 0.994$.

1. Introduction

The PVDF (polyvinylidene fluoride) hydrophones became nowadays very useful for acoustic measurements of ultrasonic diagnostic devices. Ultrasonic pressures, obtained in such cases, are sometimes higher than 1 MPa (about 10 atm), being 7 orders of magnitude higher than acoustic pressures corresponding to the laudness level of 80 dB. Parameters of these hydrophons are given by producers except their linearity properties. In the literature there was published a paper [3] showing a good linearity of such a hydrophone up to 30 MPa, and a little worse for pressures up to 70 MPa. This problem is for us especially important since we are carring out measurements of nonlinear effects caused by nonlinear wave propagation. In our PVDF hydrophone, which was fabricated in laboratory [1], after some years of use we have noticed a partial damage of gold electrodes caused probably by cavitation. The thin plastic foil could be also partially damaged although there are no visible signs of destruction. So the authors decided to check the hydrophone's linearity to ensure the correctness of future measurements of nonlinear wave propagation. The question of the possible decrease of its sensitivity will be discussed in an another paper. To solve the linearity problem a multilayer method was elaborated, where metal layers with various acoustical properties [2], immersed in water, are used to decrease the measured pressure pulse without changing its shape (spectrum).

2. Decrease of the pressure pulse penetrating metal layers

When mesuring by means of our PVDF hydrophone pressure pulses with the same spectrum, however with different pressures, different sections of the linear or nonlinear hydrophone's characteristics are used.

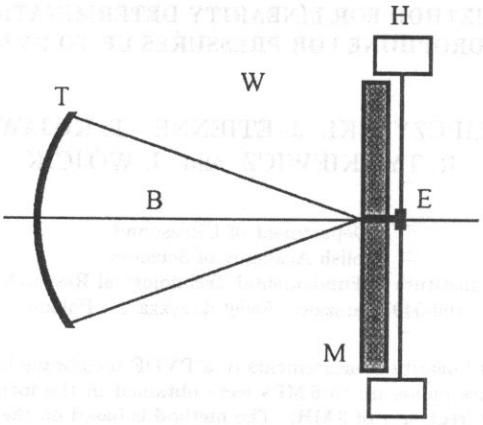


Fig. 1. The system used in measurements. T – transmitting PZT transducer, W – water, P – metal layer, B – ultrasonic focused beam, H – membrane PVDF hydrophone, E – its sensitive electrode.

The measurements were carried out in the system shown in Fig. 1. The PZT concave, circular transducer 2 cm in diameter, with 3 MHz frequency and the focal length of 65 mm was used as the source of ultrasonic pulses. The electrical voltage applied to the transducer was changed from $47V_{pp}$ to $320V_{pp}$ (Fig. 2). Pressure pulses were measured by means of the PVDF hydrophone under investigation. The maximum pressure obtained in the focus was equal to several MPa. Its axial distribution showed the constant pressure in the distance from 65 to 70 mm from the PZT transducer, while its perpendicular focal width was 2 mm (−6 dB). The measurements were carried out with layers of metal 5 mm in thickness, made of steeel, aluminium alloy and electron (alloy of magnesium with addition of aluminium, zink and manganese). Acoustic impedances of the materials listed above are gradually decreasing as shown in Table 1.

Table 1. Measured parameters of metal plates.

	Density ϱ [g/cm^3]	Velocity c [m/s]	Acoust. impedance ϱc [MRayl]	Attenuation α [dB/cm]
Stainless steel	8.59	4890	42	
Carbon steel	7.87	5926	46.6	$\alpha_{3\,\mathrm{MHz}} = 0.24$ $\alpha_{10\,\mathrm{MHz}} = 0.66$
Dural 1	2.83	5897	16.7	
Dural 2	2.81	6350	17.8	
Elektron	1.86	6366	11.9	

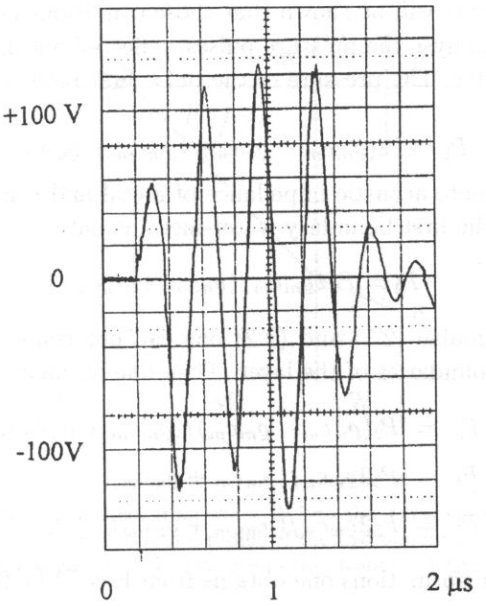


Fig. 2. Example of an electrical pulse with the voltage of $265V_{pp}$ generated by the transmitter. Vertical scale – 20 V. Horizontal scale – 0.2 μs.

Figure 3 shows in coordinates z, t the reflections of the pressure pulse P_0 incident on the metal layer. The pulse width is short enough to avoid standing waves in the layer. So formulae for propagating plane waves penetrating the metal layer will be assumed.

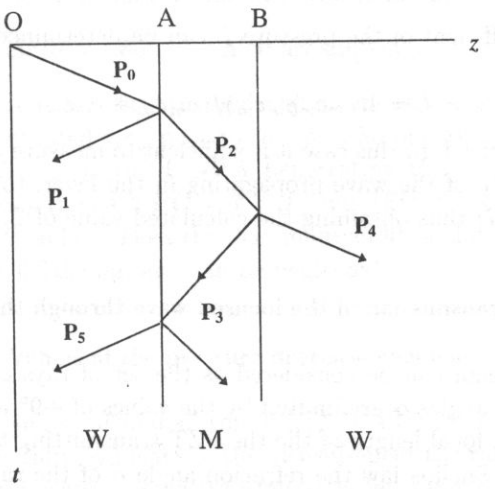


Fig. 3. Reflections of the plane wave pressure pulse in a metal layer M immersed in water W. z, t – distance and time coordinates, AB – metal layer thickness, for $A < z > 0$ and $z > B$ there is water, P_0 – incident pulse, P_1, P_5 – pulses reflected from the front surface of the layer, P_2, P_3 – pulses penetrating into the layer, P_4 – pulse transmitted through the layer.

In the following section it will be shown that those conditions are justified.

It is necessary to analyse the pressure pulses, reflected and propagated in the metal layer, immersed in water. The pressure of the pulse first reflected from the metal layer equals

$$P_1 = P_0(\varrho_m c_m - \varrho_w c_w)/(\varrho_m c_m + \varrho_w c_w), \tag{2.1}$$

where $\varrho_m c_m$, $\varrho_w c_w$ denote acoustic impedance of metal and water. The pressure of the pulse P_2 penetrating the first boundary water-metal equals

$$P_2 = P_0 2\varrho_m c_m/(\varrho_m c_m + \varrho_w c_w). \tag{2.2}$$

Using respectively formulae (2.1) and (2.2) one can determine all the waves reflected and penetrating the boundaries of the layer. Then one obtains

$$P_3 = P_2(\varrho_w c_w - \varrho_m c_m)/(\varrho_m c_m + \varrho_w c_w), \tag{2.3}$$

$$P_4 = P_2 2\varrho_w c_w/(\varrho_m c_m + \varrho_w c_w), \tag{2.4}$$

$$P_5 = P_3 2\varrho_w c_w/(\varrho_m c_m + \varrho_w c_w). \tag{2.5}$$

After some simple transformations one obtains from Eqs. (2.1)–(2.5) the relation

$$P_4/P_0 = -P_5/P_1\,. \tag{2.6}$$

It means that the decrease coefficient of the pressure pulse $L = P_4/P_0$, transmitted through the metal layer, equals $-P_5/P_1$; the negative ratio of the second pulse reflected from the front surface of the layer to the first pulse reflected from the back surface of the layer. The relation (2.6) makes it possible to determine the value of L by means of one simple direct mesurement using for this purpose only one plane transmit-receive transducer.

The decrease coefficient of the pressure L can be determined also directly from the equation

$$L = 4(\varrho_w c_w \varrho_m c_m)/(\varrho_m c_m + \varrho_w c_w) \tag{2.7}$$

that can be easily derived. In this case it is sufficient to measure the metal density ϱ, the longitudinal velocity c of the wave propagating in the layer, to introduce these values into the equation (2.7) thus obtaining the calculated value of L.

3. Transmission of the focused wave through the layer

The ultrasonic beam can be considered as the set of rays crossing themself in the focus. Their incident angles α are limited by the values of $+9°$ and $-9°$, resulting from the diameter and the focal length of the the PZT transmitting transdcuer.

According to the Snelius law the refracion angle β of the ray penetrating the layer (Fig. 4) equals

$$\beta = \arcsin(4\sin\alpha), \tag{3.1}$$

since the ratio of wave velocities in metals and in water equals $\cong 4$ (see Tabl. 1).

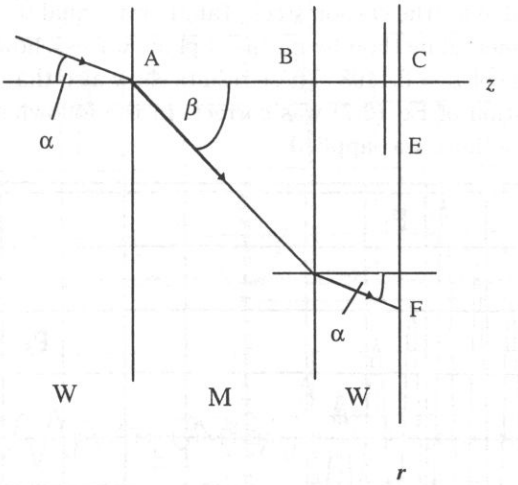


Fig. 4. The path of the ray incident from water W into the metal layer M. $AB = d = 5\,\text{mm}$, $BC = d' = 1.5\,\text{mm}$. $CE = 0.25\,\text{mm}$ – radius of the hydrophone's sensitive electrode, $CF = \Delta r$ – displacement of the ray, α and β – incident and refraction angles.

So one can determine the refraction angles β for all the incident angles α to find the ray displacement Δr due to the pulse transmission through the layer $AB = d$ (Fig. 4)

$$\Delta r = d \operatorname{tg} \beta. \tag{3.2}$$

The radius of the sensitive electrode of the PVDF hydrophone equals 0.25 mm, so the equations (3.1) and (3.2) make it possible to determine the following limiting angles for rays, which arrive the electrode and in this way will be measured by the hydrophone

$$\beta' = \operatorname{arc\,tg} \Delta r/d = 2.86°, \qquad \alpha' = \operatorname{arc\,sin}[(\sin \beta)/4] = 0.72° \tag{3.3}$$

Therefore, one can conclude that the hydrophone measures only the rays for $\alpha < \alpha'$; and $\beta < \beta'$ which are situated very near to the axis. So our assumption of plane waves in derivation of Eq. (2.7) is justified. After penetrating the layer, the pressure pulse propagates in water along the path $BC = d' = 1.5\,\text{mm}$ (Fig. 4) causing only a very small changes of the angles β' and α'. Then the maximum additional displacement for limitng rays equals only $\Delta r' = 0.019\,\text{mm}$, and can be neglected.

4. Values of the pressure decrease coefficient L

The decrease of L can be calculated from Eq. (2.7), after having measured the values of ϱ and c, or experimentally by means of the measurement the ratio P_5/P_1 according to the relation (2.6). In this case a plane transmitt-receive transducer 10 mm in diameter with the resonanse frequency of 3.7 MHz was used. The elecrical voltage applied to the transducer equaled $140 V_{pp}$. The measurements were performed in the near field of the beam.

The calculated value for the carbon steel (Tab. 1) was equal to $L = 0.119$. The results of the second, experimental method by means of plane waves, shown in Fig. 5, have given almost the same value of $L = 0.1198$. These results show also that our assumption of the plane wave in derivation of Eq. (2.7) was correct. In the following work the first, much simpler calculation method, was applied.

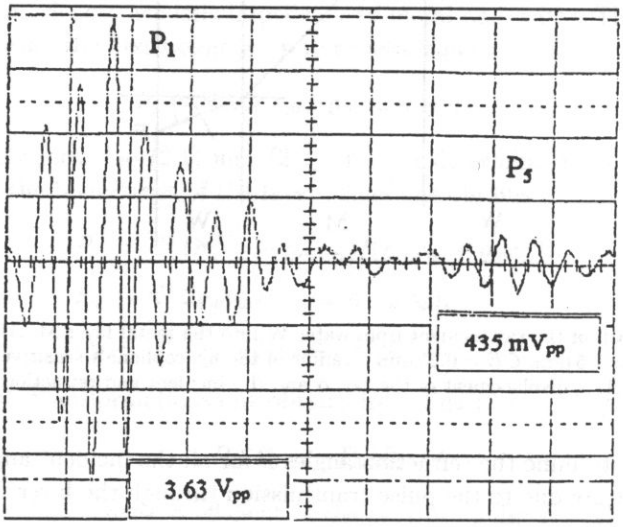


Fig. 5. The first P_1 and the second P_5 pressure pulses reflected from the layer of carbon steel immersed in water (see Fig. 3). Vertical scale – 50 mV, horizontal – 0.5 μs.

Our expectation regarding the shape and the pressure value of pulses, penetrating the various metal layers, were fulfiled as can be shown in Fig. 6. The shapes of pulses presented in Fig. 6 are identically distorted due to nonlinear propagation in water, especially in the focus, where the pressures are maximum.

Using the described technique the diagram in Fig. 7 was drawn. The horizontal axis represents the value L (the pressure pulse decrease coefficient – see Table 2). The vertical axis shows the maximum values of the pressure signals [in mV_{pp}] obtained from our investigated PVDF hydrophone The measurements were carried out for 320, 275, 230, 140 and $47V_{pp}$ applied to the transmitting transducer PZT. Figure 7 shows the straight lines that best fit the measurement points. At every line the corresponding correlation coefficient is given. They are equal from $r = 0.999$ to 0.987, in average 0.994, demonstrating the linearity of the hydrophone. In such a case one can try to extrapolate this linearity for higher pressures. The highest pulse pressure measured in the focus of the system of Fig. 1 (without the metal layers) was 6 MPa. So taking into account the maximum value of the pressure decrease coefficient $L = 0.39$ (see Fig. 7) one obtains the maximum pulse pressure equal to 2.3 MPa, for which the linearity of the hydrophone was really demonstrated.

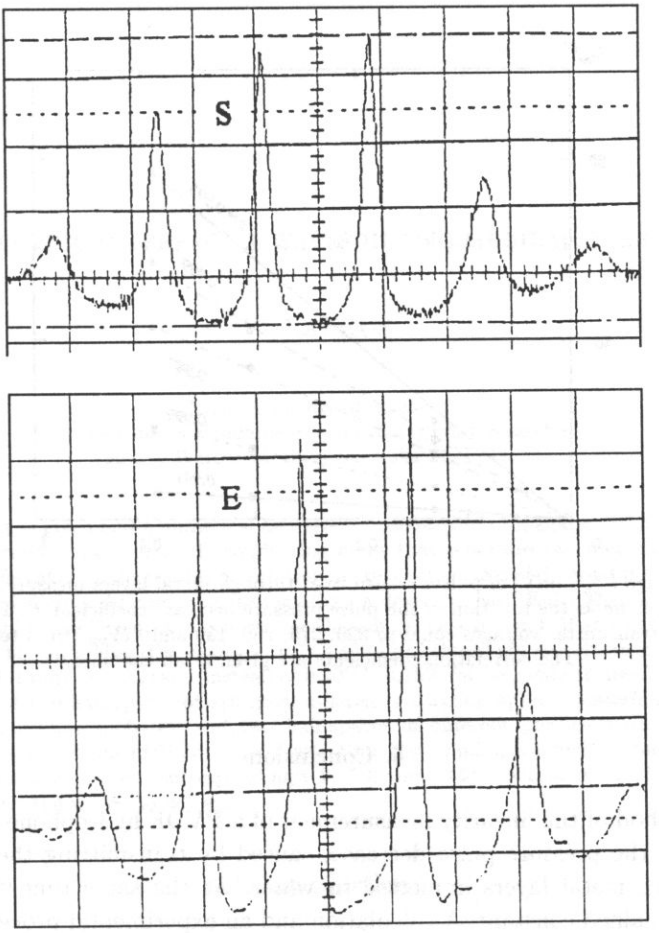


Fig. 6. Pressure pulses P_4 transmitted through the layer of stainless steel (S) (vertical scale 5 mV) and of electron (E) (vertical scale 10 mV) for the transmitter voltage of 320 V_{pp}. Horizontal scale in both cases – 0.2 μs.

Table 2. Values of the pressure decrease coefficient L.

	ϱc [MRayl]	L
Stainless steel	42	0.13
Carbon steel	46.6	0.12
Dural 1	16.7	0.3
Elektron	11.9	0.39

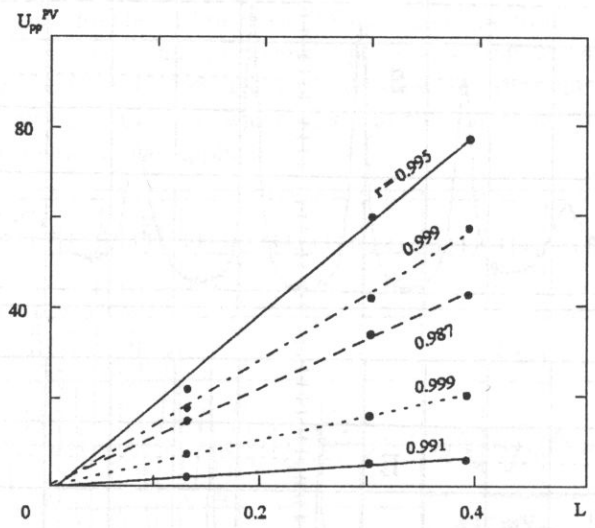


Fig. 7. The voltage U_{pp}^{PV} [mV] of pulses transmitted through metal layers measured by means of the PVDF hydrophone as the function of the pulse pressure decrease coefficient L. The results were obtained at transmitter voltages equal to 320, 275, 230, 130 and $47V_{pp}$ (from top to bottom). The correlation coefficients are given at all the lines.

5. Conclusions

A new method of the linearity measurement of a PVDF hydrophone was elaborated. It is based on the pressure pulse decrease, caused by transmitting the pressure pulse through various metal layers immersed in water. At the same time the pulse shape (spectrum) remained constant. A calculation and an experimental procedures for determination of the pressure pulse decrease – essential for this method – were given. Using these procedures, it was possible to determine the linearity of the hydrophone up to pressures equal to 2.3 MPa. The correlation coefficients of 5 measured relations were found to be in average equal to $r = 0.994$.

Acknowledgments

The authors thank prof. dr hab. J. DEPUTAT for measurements in his laboratory wave velocities and attenuation of metal plates used in this study.

References

[1] L. FILIPCZYŃSKI, J. ETIENNE, G. ŁYPACEWICZ, T. WASZCZUK, *Measurement technique of shock wave pulses at extremely high pressures*, Archives of Acoustics, **21**, 37–51, 1996.

[2] J.H. KRAUTKRAMMER, *Ultrasonic material testing*, Springer Verlag, Berlin 1990.

[3] S. MEEKS, R. TING, *The evaluation of* PVF_2 *for underwater shock-wave sensor application*, J. Acoust. Soc. Am., **75**, 1010–1012, 1984.

ARCHIVES OF ACOUSTICS
23, 4, 521–531 (1998)

PIEZOELECTRIC INTERFACIAL WAVES IN TWO-MEDIUM STRUCTURES

W. LAPRUS

Polish Academy of Sciences,
Institute of Fundamental Technological Research
(00-049 Warszawa, Świętokrzyska 21, Poland)

The theory is given and a numerical analysis is presented of propagation of piezoelectric interfacial waves along a perfectly conducting plane that separates two piezoelectric half-spaces of different crystallographic orientations or two half-spaces of different piezoelectrics. In the case of slightly different orientations (half a degree in each Euler angle), it is found that the coupling coefficient is in general less than in the case of two half-spaces of the same orientation, but still large. For example, it is 3.40% for lithium niobate (instead of 3.50%), 0.85% for dilithium tetraborate (instead of 0.90%), 0.15% for langasite (instead of 0.20%). In the case of two half-spaces of arbitrary different crystallographic orientations (lithium niobate and lithium niobate), and two half-spaces of different piezoelectrics (dilithium tetraborate and lithium niobate), it is found that the coupling coefficient may be quite large for some crystal cuts (correspondingly 2.25% and 1.45%), although for the majority of crystal cuts the waves do not exist.

1. Introduction

The piezoelectric interfacial wave (PIW) is guided by a perfectly conducting plane embedded in a piezoelectric medium [3, 4]. The guiding structure can be produced by cutting a piezoelectric crystal into two pieces, covering the faces of the cut with metal, and bonding the two pieces together. In this process, some misalignment of the crystallographic axes of the two pieces is inevitable (up to about half a degree in each Euler angle). In view of applications, it is important to know if such an error does not affect adversely the existence and properties of PIW.

In the paper, a numerical analysis of misalignment is presented for several piezoelectrics. The electro-mechanical field equations are transformed to a system of first-order ordinary differential equations [5, 11]. The associated eigenvalue problem is solved numerically for many orientations of the conducting plane with respect to the crystallographic axes and for different types of misalignment. It is found that piezoelectric interfacial waves with a high coupling coefficient are relatively insensitive to misalignment. In general, however, the domains of existence in the space of Euler angles are smaller than in the case without misalignment.

Next, two-medium guiding structures of two kinds are investigated. The first consists of a perfectly conducting interface and two piezoelectric half-spaces of arbitrary different

crystallographic orientations (this is a generalization of the basic structure with misalignment). The second consists of a perfectly conducting interface and two half-spaces of different piezoelectrics. Calculations are performed for variable crystallographic orientation of one half-space (lithium niobate or dilithium tetraborate) and a fixed orientation of the other (lithium niobate). The domains of existence of PIW are determined, and wave propagation parameters are calculated for selected orientations. It is found that small isolated domains may exist in the angle space, where PIW piezoelectric coupling is high.

2. Theoretical background

We assume that the electro-mechanical field depends on time through the factor $\exp(j\omega t)$. In the system of coordinates (x, y, z) the field satisfies the equations given in Ref. [11]. We shall adopt the notations and conventions of Ref. [11].

The medium is described by the following constants: elastic tensor c_{ijkl}, piezoelectric tensor e_{kij}, dielectric tensor ε_{ki}, and mass density ϱ. The field variables are: particle displacement u_i, electric potential ϕ, stress tensor T_{ij}, and electric displacement D_k.

Let us consider two homogeneous piezoelectric half-spaces (for $z > 0$ and $z < 0$), and the perfectly conducting plane $z = 0$, as shown in Figure 1. We assume that the field is independent of y, and that the dependence on x and z is given by the factor $\exp(-j\omega r x - j\omega s z)$.

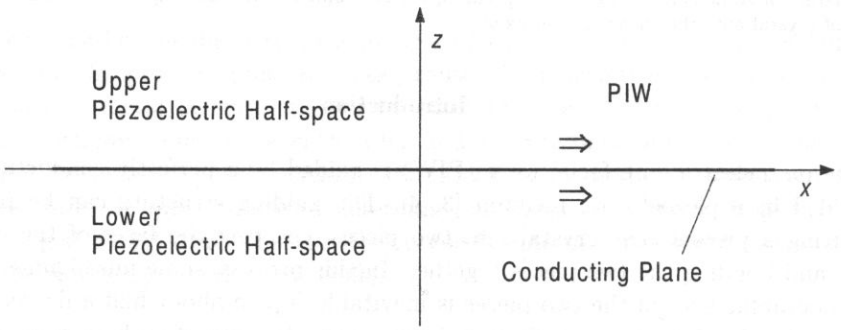


Fig. 1. The system of coordinates.

In each half-space the field equations can be reduced to a system of eight linear algebraic equations [11]. The following eight field variables will be used: particle displacement u_i, electric potential ϕ, traction force $T_i = T_{3i}$ at the interface, and normal (to the interface) component D_3 of the electric displacement D_i. We have

$$H_{KL}(r)F_L = qF_K\,, \tag{1}$$

where $K, L = 1, \ldots, 8$, $(F_K) = (j\omega r u_i, j\omega r\phi, T_i, D_3)$, and $q = s/r$. To each half-space corresponds a matrix H_{KL} which is real and nonsymmetric (r is assumed to be real). It depends on material constants [11].

After solving the eigenvalue problem defined by Eq. (1) we get eight eigenvectors $\tilde{F}_K^{(J)}(r)$ corresponding to eight eigenvalues $q^{(J)}(r)$ for $J = 1, \ldots, 8$. The Jth eigenwave has the form

$$F_K^{(J)} = \tilde{F}_K^{(J)} \exp(j\omega t - j\omega r(x + q^{(J)} z)). \tag{2}$$

The solution of the field equations in each half-space is a linear combination of the eigenwaves corresponding to that half-space.

The eigenvector $\tilde{F}_K^{(J)}$ will be called upper (lower) if either $\operatorname{Im} q^{(J)} < 0$ ($\operatorname{Im} q^{(J)} > 0$) or $\operatorname{Im} q^{(J)} = 0$ and the z component of the real part of the Poynting vector is positive (negative). Note that $\operatorname{Im} q^{(J)} \neq 0$ for $J = 1, \ldots, 8$ if $r > r_c$ where r_c is the cutoff slowness of bulk waves. In each half-space the set of eigenvectors consists of four upper and four lower eigenvectors.

The surface wave field vanishes at infinity, and there is no energy flux to the boundary (no sources in the space). Therefore, the solution F_K^+ of the field equations in the upper half-space consists of upper eigenwaves for that half-space, and the solution F_K^- in the lower half-space consists of lower eigenwaves for that half-space. At the plane $z = 0$, the complex amplitudes of the two solutions are

$$\tilde{F}_K^{\pm} = \pm \sum_J^{\pm} C_J \tilde{F}_K^{(J)} \tag{3}$$

where the set of the eight eigenvectors $\tilde{F}_K^{(J)}$ for $J = 1, \ldots, 8$ consists of four upper eigenvectors for the upper medium and four lower eigenvectors for the lower medium. The plus (minus) superscript of the sum symbol means that the summation is performed over J such that $\tilde{F}_K^{(J)}$ is an upper (lower) eigenvector for the upper (lower) half-space. The minus sign before the sum symbol in Eq. (3) is used to simplify some formulas considered later. The coefficients C_J are to be determined from boundary conditions.

The field equations can be solved in each half-space separately, provided appropriate boundary conditions are imposed at the plane $z = 0$. In general, the boundary conditions may determine either the jumps of amplitudes $\Delta\tilde{F}_K = \tilde{F}_K^+ - \tilde{F}_K^-$, or the amplitudes $\tilde{F}_K^+$ (or $\tilde{F}_K^-$). Below, it will be shown that the two sets of values are related to each other.

Let $R_{KJ} = \tilde{F}_K^{(J)}$, and denote by L_{JK} the inverse of the matrix R_{KJ}. From Eq. (3) it follows that

$$\Delta\tilde{F}_K = R_{KJ} C_J \,, \tag{4}$$

and hence

$$C_J = L_{JK}\, \Delta\tilde{F}_K \,. \tag{5}$$

Using Eq. (5) we write Eq. (3) as

$$\tilde{F}_K^{\pm} = Z_{KL}^{\pm}(r)\, \Delta\tilde{F}_L \,, \tag{6}$$

where

$$Z_{KL}^{\pm}(r) = \pm \sum_J^{\pm} R_{KJ} L_{JL} \,. \tag{7}$$

The plus or minus superscript of the sum symbol has the same meaning as in Eq. (3) (columns of the matrix R_{KJ} are eigenvectors of the two matrices H_{KL}). R_{KJ} and L_{JL},

and thereby $Z^{\pm}_{KL}$, depend on r. The matrices Z^{+}_{KL} and Z^{-}_{KL} are related to each other. Indeed, we have

$$\Delta\tilde{F}_K = \tilde{F}^{+}_K - \tilde{F}^{-}_K = (Z^{+}_{KL} - Z^{-}_{KL})\,\Delta\tilde{F}_L\,, \tag{8}$$

which means that

$$Z^{+}_{KL} - Z^{-}_{KL} = I_{KL}\,, \tag{9}$$

where I_{KL} is the identity matrix. Thus, off-diagonal elements of the two matrices are equal.

Equation (6) is a fundamental relation in the problem of PIW propagation. It should be satisfied by the field variables at the boundary of each piezoelectric half-space, irrespective of what boundary conditions are imposed there. The matrices $Z^{+}_{KL}(r)$ and $Z^{-}_{KL}(r)$ are determined completely by the set of eigenvectors for a given r. Therefore, they are characteristic of the both media rather than of any particular solution of the field equations.

We assume that all the field variables are continuous across the plane $z = 0$ except D_3. The amplitude $\tilde{D}_3$ suffers a jump $\Delta\tilde{D}_3$ (assumed to be real). For $K = 4$, Eq. (6) gives

$$j\omega r\tilde{\phi} = Z(r)\,\Delta\tilde{D}_3\,, \tag{10}$$

where $\tilde{\phi} = \tilde{\phi}^{+} = \tilde{\phi}^{-}$ is the amplitude of the electric potential at the interface, and $Z(r) = Z^{+}_{48}(r) = Z^{-}_{48}(r)$ is a complex-valued function. 1/Z has the physical dimension of the dielectric permittivity. This is more clearly seen if we rewrite Eq. (10) in the form

$$\tilde{E}_1 = Z(r)\,\Delta\tilde{D}_3\,. \tag{11}$$

$E_1 = -\partial\phi/\partial x$ is the tangential (to the interface) component of the electric field. This component should vanish at the perfectly conducting interface which implies that $\phi = 0$ for $z = 0$.

The function $Y(r) = 1/Z(r)$ is a counterpart of the effective permittivity function in the theory of surface acoustic waves [1, 2].

It can be shown that $\operatorname{Re} Z(r)$ is equal to zero for $r > r_c$, and that $\operatorname{Im} Z(r)$ tends to a positive number as $r \to \infty$. Here and below, by r_c we mean the greater of the two values of r_c corresponding to the two half-spaces. PIW exists if the dispersive equation $Z(r) = 0$ is satisfied for a particular value of r, say r_p. Then $\tilde{\phi} = 0$, i.e. ϕ satisfies the boundary condition for $z = 0$, and r_p is the slowness of PIW.

3. Numerical analysis

In each half-space the orientation of crystallographic axes is determined by a triplet of Euler angles. Triplets of angles may be represented as points in a three-dimensional space. The solution of the dispersive equation $Z(r) = 0$ depends on the crystallographic orientation of the both piezoelectric half-spaces, i.e. it depends on two triplets of Euler angles.

The effect of misalignment is calculated as follows. First, assuming that there is no misalignment, we scan the three-dimensional space of Euler angles at discrete points

so that the angles change in steps of 4°. Then, for each scanned point, we rotate the upper and/or lower half-space in a sequence of six separate rotations which correspond to differences of 0.5° in Euler angles between the two half-spaces: differences in one Euler angle (three rotations), and differences in two Euler angles (three rotations). These rotated configurations are intended to simulate random misalignment.

The symmetry of piezoelectric crystals implies that it is sufficient to scan the following ranges: 0°–30°, 0°–180°, 0°–180° for trigonal crystals (and orthorhombic Rochelle salt), and 0°–45°, 0°–180°, 0°–180° for cubic crystals.

For each chosen set of six Euler angles, the two eigenvalue problems are solved numerically (using EISPACK routines [6]) for several values of r in search of the two r_c (simple zero-finding routine). Next, the eigenvalue problems are solved for several values of r in a small neighborhood of the greater of the two r_c, just above it, in search of r_p (zero-finding routine applied to the function $Z(r)$). If $Z(r)$ has no zero for a rotated configuration, then calculations are stopped and it is considered that PIW does not exist at that particular scanned point.

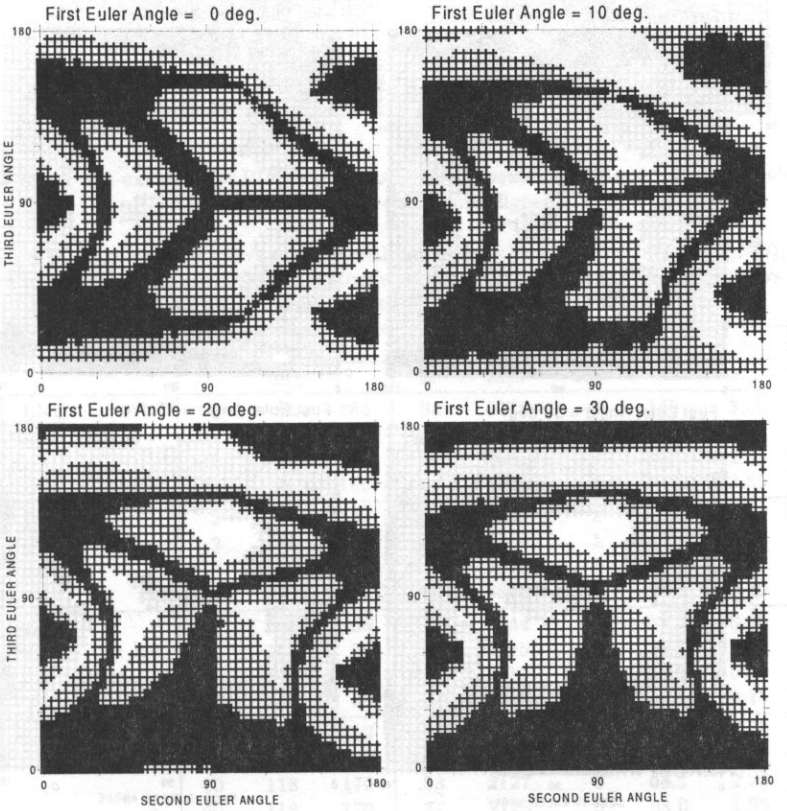


Fig. 2. Domains of non-existence of PIW for lithium niobate (*black squares*: half-spaces aligned, *black squares* and *crosses*: half-spaces misaligned).

If PIW exists then its parameters are calculated: $\delta = (r_p - r_c)/r_c$, phase velocity $v_p = 1/r_p$, beam steering angle ψ, effective permittivity ε_{eff}, piezoelectric coupling coefficient κ (definitions of the last three parameters are given in Ref. [11]). The calculations are performed for $\omega = 10^6\,s^{-1}$. The material constants for the piezoelectrics are taken from Refs. [7] and [8].

The effect of misalignment for lithium niobate, dilithium tetraborate, langasite, and Rochelle salt is illustrated in Figures 2 to 5. Some asymmetry of the patterns (especially well seen for the first Euler angle equal to 0° and to 30° or 45°) is an effect of finite precision of calculation. Domains of non-existence of PIW without misalignment are shown more precisely (due to scanning in steps of 2°) in figures which can be found in Refs. [9]-[11]. The figures are detailed symbolic maps of PIW properties (without misalignment) for lithium niobate, dilithium tetraborate, langasite, and other piezoelectrics. Color versions of these maps are published on the Internet [12].

Table 1 compares PIW parameters without misalignment to those with misalignment. For each medium, parameters at two points of the angle space are given (representative

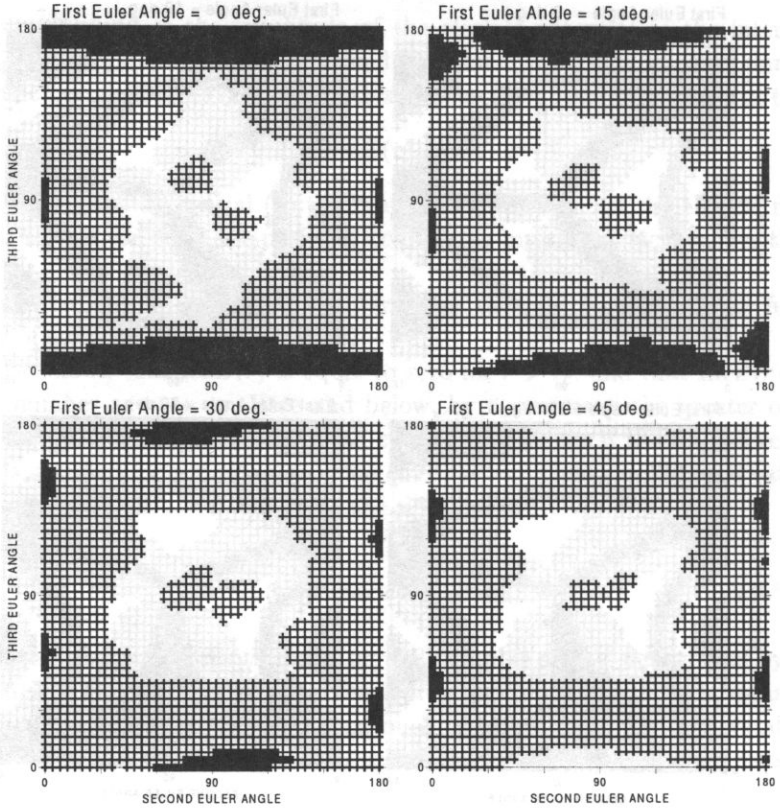


Fig. 3. Domains of non-existence of PIW for dilithium tetraborate (*black squares*: half-spaces aligned, *black squares* and *crosses*: half-spaces misaligned).

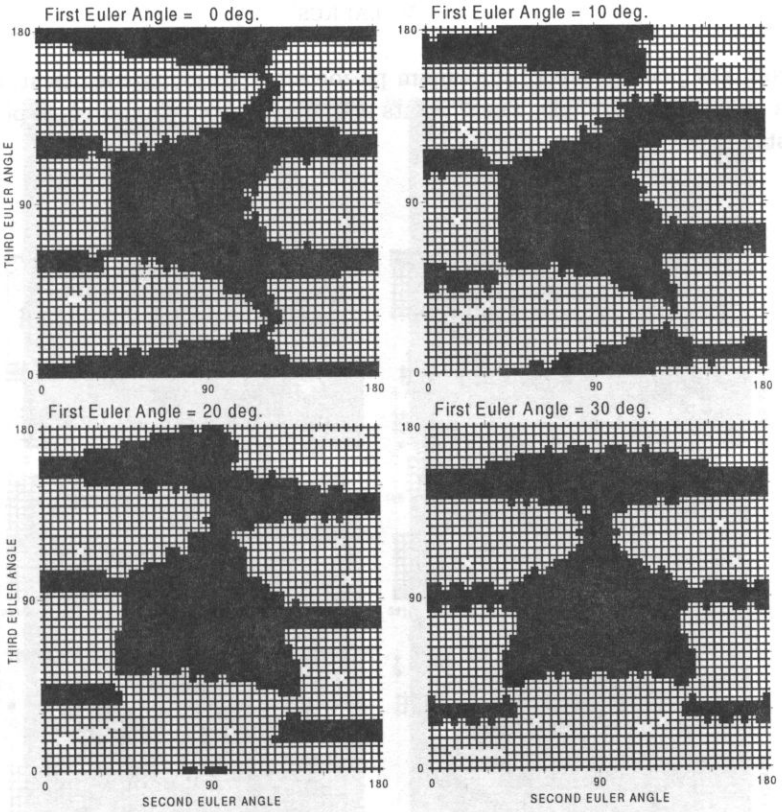


Fig. 4. Domains of non-existence of PIW for langasite (*black squares*: half-spaces aligned, *black squares* and *crosses*: half-spaces misaligned).

Table 1. Effect of misalignment (rows marked with an asterisk) for lithium niobate (LNO), dilithium tetraborate (LBO), langasite (LGS), and Rochelle salt (LRS). For reference, parameters of PIW without misalignement are also given (rows without asterisk).

medium	Euler angles			δ	v_p	ψ	ε_{eff}	κ
	deg			%	m/s	deg	–	%
LNO	30	88	164	.38	4075	−3	149.	3.50
*	30	88	164	.39	4074	−3	149.	3.40
	28	96	166	.04	4081	0	150.	2.30
*	28	96	166	.05	4081	0	149.	2.65
LBO	2	88	124	.48	3299	2	18.7	0.90
*	2	88	124	.33	3298	1	18.7	0.85
	10	90	90	.39	3873	0	18.8	0.75
*	10	90	90	.06	3868	4	18.8	0.50
LGS	8	144	36	.13	2944	−1	58.0	0.20
*	8	144	36	.01	2945	−9	58.0	0.10
	10	148	38	.11	2917	0	59.1	0.20
*	10	148	38	.03	2919	−4	59.1	0.15
LRS	60	66	12	.47	2120	0	70.0	2.00
*	60	66	12	.50	2120	1	70.1	2.40
	60	118	170	.53	2127	0	68.1	2.20
*	60	118	170	.37	2126	0	68.0	1.95

of two groups of points): one of maximum piezoelectric coupling coefficient, and one of minimum beam steering angle. These points are selected from thousands of points where PIW exists.

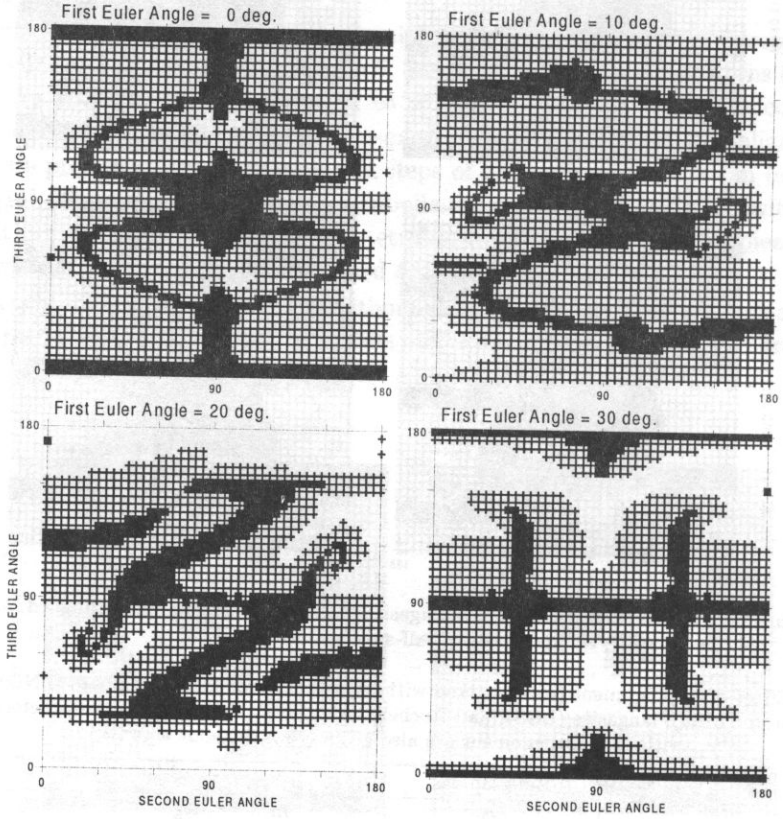


Fig. 5. Domains of non-existence of PIW for Rochelle salt (*black squares*: half-spaces aligned, *black squares* and *crosses*: half-spaces misaligned).

For two half-spaces of arbitrary different crystallographic orientations and two half-spaces of different piezoelectrics, PIW propagation parameters are calculated using the procedures that are described above. In the both cases, the lower piezoelectric is lithium niobate of a fixed orientation (Euler angles: 30°, 88°, 164°), and the upper piezoelectric has a variable orientation (the scanning of the angle space in steps of 4° is performed). The fixed orientation gives high piezoelectric coupling of PIW in lithium niobate without misalignment, i.e. when the both half-spaces have that orientation.

The maps of PIW properties for each case are shown in Figures 6 and 7. PIW parameters are given in Table 2 for each case at two points of the angle space: one of maximum piezoelectric coupling coefficient, and one of minimum beam steering angle.

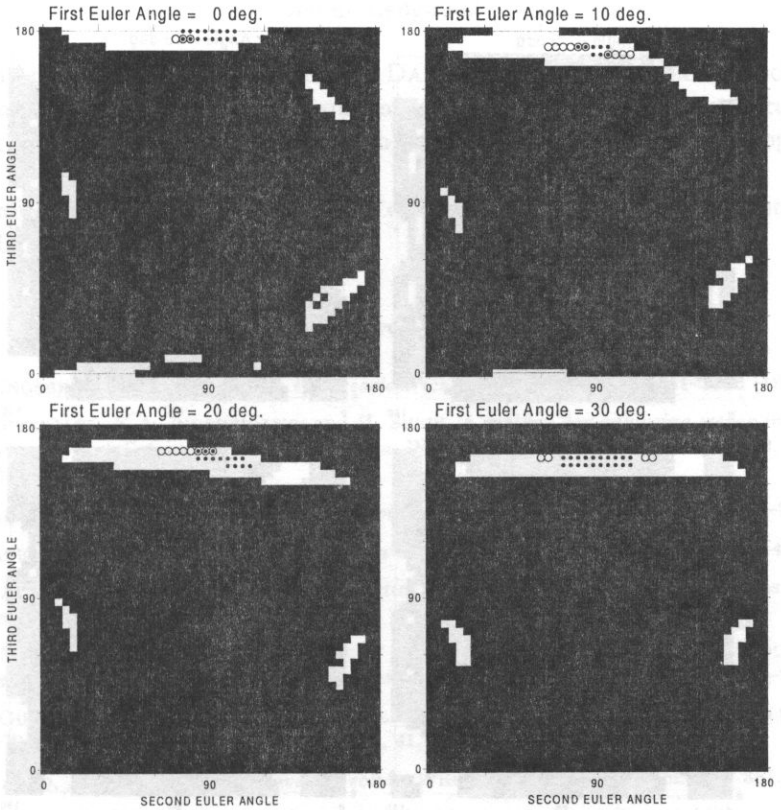


Fig. 6. Maps of PIW properties for lithium niobate of fixed orientation in the lower half-space (Euler angles: 30°, 88°, 164°) and lithium niobate of variable orientation in the upper half-space (*black squares*: no PIW, *dots*: PIW coupling coefficient > 2%, *circles*: PIW coupling coefficient > 1% and PIW beam steering angle < 0.5°).

Table 2. PIW parameters for two piezoelectric half-spaces of different orientations (lithium niobate and lithium niobate) and of different media (dilithium tetraborate and lithium niobate). In the both cases lithium niobate in the lower half-space has the same orientation (Euler angles 30°, 88°, 164°) and the orientation of the upper half-space is given in the table.

medium	Euler angles			δ	v_p	ψ	ε_{eff}	κ
	deg			%	m/s	deg	–	%
LNO/LNO	30	88	164	0.38	4075	−3	149.	3.45
	20	88	168	0.12	4074	0	151.	2.25
LBO/LNO	0	84	152	1.05	4048	−31	76.6	2.60
	15	20	152	0.56	4067	0	74.9	1.45

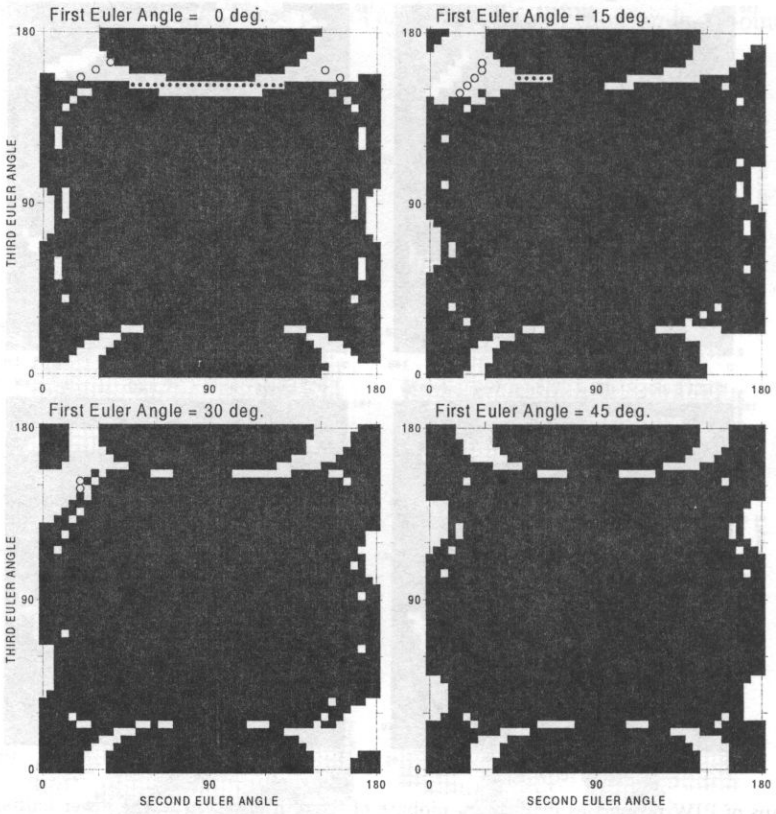


Fig. 7. Maps of PIW properties for lithium niobate of fixed orientation in the lower half-space (Euler angles: 30°, 88°, 164°) and dilithium tetraborate of variable orientation in the upper half-space (*black squares*: no PIW, *dots*: PIW coupling coefficient > 2%, *circles*: PIW coupling coefficient > 1% and PIW beam steering angle < 0.5°).

4. Conclusion

From Figures 2 to 5 it is seen that misalignment affects mainly PIW propagation for those points in the angle space (i.e. those crystal cuts) for which piezoelectric coupling is low. Domains of existence of PIW are smaller, sometimes much smaller, than in the case without misalignement. Nevertheless, as Table 1 proves, the piezoelectric coupling coefficient remains large if it is large without misalignment. Therefore, in applications, when crystall cuts of high coupling are selected, misalignment is not very important.

In contrast to one-medium structures, two-medium structures are characteristic of extremely small domains of existence of PIW. In fact, for other pairs of piezoelectrics, which have also been analyzed, PIW seems not to exist at all (perhaps a scanning in smaller steps would reveal PIW existence). But still there are small domains of high piezoelectric coupling, as seen in Figures 6 and 7. The piezoelectric coupling coefficients shown in Table 2 are comparable with those for one-medium structures.

Acknowledgment

The author would like to thank Prof. E. DANICKI for suggesting the research of PIW propagation in the presence of misalignment, and for discussing technical problems of producing crystal cuts of a given orientation, which is important in view of applications of PIW.

This work was supported by Committee of Scientific Research (Poland) under Grant No. 2P03B 171 12.

References

[1] K.A. INGEBRIGTSEN, *Surface waves in piezoelectrics*, J. Appl. Phys., **40**, 2681–2686 (1969).

[2] K. BLØTEKJAER, K.A. INGEBRIGTSEN and H. SKEIE, *A method for analyzing waves in structures consisting of metal strips on dispersive media*, IEEE Trans. Electron Devices, **ED-20**, 1133–1138 (1973).

[3] E. DANICKI, *New interfacial shear waves in piezoelectrics*, Appl. Phys. Lett., **64**, 969–970 (1994).

[4] E. DANICKI, *Piezoelectric interfacial waves in* $LiNbO_3$", Appl. Phys. Lett., **66**, 3439–40 (1995).

[5] E.L. ADLER, *SAW and pseudo-SAW properties using matrix methods*, IEEE Trans. Ultrason. Ferroel. Freq. Control, **41**, 699–705 (1994).

[6] B.T. SMITH, J.M. BOYLE, B. S. GARBOW, Y. IKEBE, V. C. KLEMA, AND C. B. MOLER, *Matrix eigensystem routines – EISPACK Guide*, Lecture Notes in Computer Science, **6**, 149–258 (1974).

[7] J.G. GUALTIERI, J.K. KOSINSKI and A. BALLATO, *Piezoelectric materials for SAW applications*, IEEE Trans. Ultrason. Ferroel. Freq. Control, **41**, 53–59 (1994).

[8] B.A. AULD, *Acoustic fields and waves in solids*, Wiley, New York 1973, Vol. 1, p. 357.

[9] E. DANICKI and W. LAPRUS, *Piezoelectric interfacial waves in langasite and dilithium tetraborate*, [in:] 1995 IEEE Ultrasonics Symposium Proceedings, M. LEVY, S.C. SCHNEIDER and B.R. MCAVOY [Eds.], IEEE, New York 1995, pp. 1011–1014.

[10] E. DANICKI and W. LAPRUS, *Piezoelectric interfacial waves in langasite and dilithium tetraborate*, Archives of Acoustics, **21**, 99–107 (1996).

[11] W. LAPRUS and E. DANICKI, *Piezoelectric interfacial waves in lithium niobate and other crystals*, J. Appl. Phys., **81**, 855–861 (1997).

[12] www.ippt.gov.pl/~wlaprus.

ARCHIVES OF ACOUSTICS
23, 4, 533–539 (1998)

AN ACOUSTIC METHOD OF DETERMINATION OF LEAKAGE COORDINATES IN GAS PIPELINES

J. BUTKUS, L. JAKEVIČIUS and O. TUMŠYS

Ultrasound Research Center
Kaunas University of Technology
(3031 Kaunas, Studentu 50, Lithuania)

The present paper describes an important leak-proofness of pipe-lines transporting gas, oil and other chemical liquids. It is revealed that one of the more perspective methods of detection of leakage coordinates of gas-lines is the acoustic correlation method. The employment of whistling acoustic noise propagating in the gas media filling the pipeline is proposed for the detection of leakage coordinates of gas pipelines. In accordance with the length of the pipe section under control and the desired precision of the determination of the leakage coordinates in the pipes, a method of choicing an optimal bandwidth of the filter and a sampling rate of received signals was developed. The results of experimental investigations are submitted. The method can be effectively used in checking the state of the pipe network as well as for solving the problems of management of the power economy objects and environmental protection.

1. Introduction

The control of leak-proofness of pipelines transporting natural and liquefied gases, oil and other chemical products is of great importance in modern technology. Leak-proofness is very significant in performing commerce and technological operations and satisfying the requirements of environmental protection [1].

At present the leak test is mostly performed by the methods of chemical analysis. But the accuracy of the chemical methods, when used for the determination of the coordinates of the leakage in gas pipelines, is low. Thus the mentioned above methods did not satisfy the requirements of modern technology. Methods of ultrasonic non-destructive testing and acoustic emission methods of searching of gas leakage in pipelines were applied only within the last decade. The methods of ultrasonic non-destructive testing, in spite of the good results, are of low capacity and are not acceptable due to the great length of the modern gas pipelines. Good results of leak tests are obtained by acoustic emission methods [1, 2]. They are used for periodical control of the leak-proofness of pipelines transporting oil products and for pipelines of head-supply systems. For this purpose there are often used portable ultrasonic units. The principle of operation of those units is based on the analysis of acoustic noises generated in the sites of leakage of the pipeline.

2. Methodology and instrumentation

There are two groups of ultrasonic devices for the determination of leakage of pipelines [2]:

1. Devices in which the intensity of acoustic noises is measured.
2. Devices in which the correlation measurement method is utilized.

Measuring units of the first group have many limitations. The main of them is the low sensitivity and noise immunity. The level of acoustic noises generated in unhermetic places must be much higher than that of received noises and the random noises of the environment. By using this method, it is difficult to determine the coordinates of the damaged place of the pipeline. These shortcomings are not significant for devices in which the correlation measurement method is employed [3, 4]. Piezoelectric accelerometers are used, as a rule for the detection of vibrations propagating in the walls of the pipeline. However, the intensity of the noises generated in the walls of the pipeline filled by gases under low pressure (0, 8 ... 4 kPa) is small. In practice the intensity of acoustic noises propagating in the walls of a gas pipeline a distance of few tens of meters from the damaged place is commensurable with the receiver noises. Besides, the gas pipelines of low pressure usually lays in inhabited places of towns and their structure is very complicated. This leads to multiple reflections of the acoustic signals propagating in the walls of the gas pipeline and the level of random noises is increased. For this reason, a correlative method based on the analysis of acoustic noises propagating in the walls of a gas pipeline is not acceptable for the control of low pressure gas pipelines. Furthermore, gas pipelines in the towns have usually an underground structure which lays under side-walks and other structures. This restricts the direct acoustic contact of the electroacoustical transducers with the walls of the pipeline. In order to solve this problem, a correlation measurement method was developed in which acoustic signals propagating in the gas media filling the low pressure pipeline are used. The employment of a whistling acoustic noise propagating in the gas media filling the pipeline is proposed for the detection of the leakage of low pressure gas pipelines.

Experimental investigations showed that the maximum of the energy of the acoustic noise generated at the damaged sites of the gas pipeline is concentrated in the band of 30 ... 50 kHz. It is almost independent of the gas pressure inside the pipeline and of the shape and linear dimensions of the damaged places. It is necessary to note that the spectrum of the acoustic noise is changed when the signal propagates in the pipeline. The components of high frequency are suppressed and maximum of the energy spectrum of the noises moves toward lower frequency. For example, at a distance of one hundred meters from the site of a leakage, the maximum of the signal energy lays usually in the band of 2 ... 3 kHz. For the registration of signals of such a low frequency piezoelectric transducers of flexural vibrations were constructed. The piezoelectric transducers are connected with the pipe under control by special tube-like insertions which are screwed to the inspection pits of the gas pipeline. This ensures a direct contact of the piezoelectric transducer to the gas media filling the pipeline under control. A block diagram of the correlation system for the determination of the leakage coordinates in the pipelines is shown in Fig. 1.

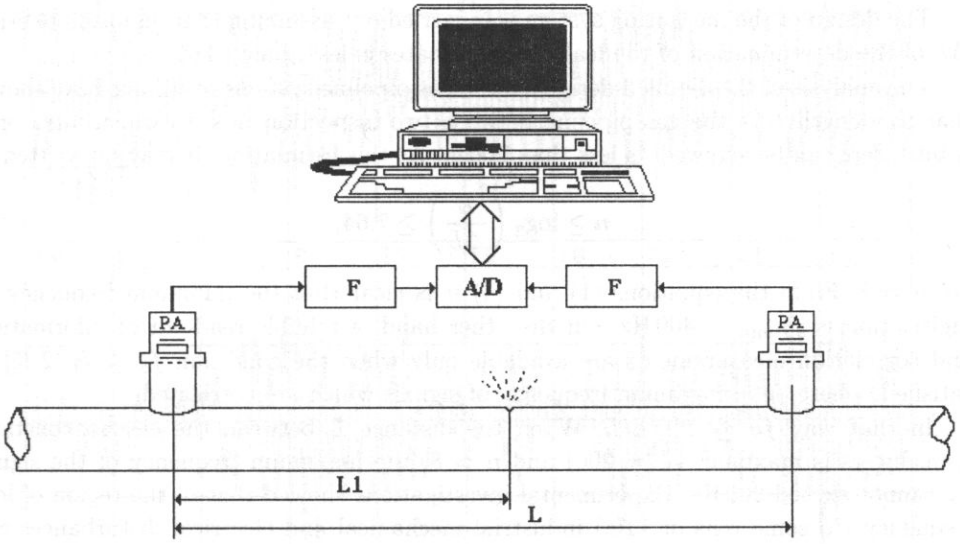


Fig. 1. Block diagram of the correlation system for the determination of a leakage in pipelines.

The acoustic noise signals received by the piezotransducers, after amplification in a primary amplifiers (PA), are sent by coaxial cables to the filters of the signal processing unit. This unit consists of filters (F), an analogue-digital (A/D) converter and a personal IBM computer. After introducing the data of the velocity of sound c in the gas under control and of the distance L between the electroacoustical transducers, the computer determines the frequency f_T of the digitisation of the analogue-digital converter. The digitisation frequency f_T is chosen in accordance with the time needed by the acoustic signal to travel the distance L in the controllable section of the pipeline:

$$f_T = \frac{2^n c}{L}; \tag{1}$$

where 2^n is the quantity of countings of the received signal; $n = 1, 2, 3, ...$; c is the speed of sound in the controllable gas filling the pipeline.

As a result of the correlative analysis of the received signals, the difference in the input time Δt of two similar signals is determined. The distance L_1 from one of the transducers to the damaged site of the pipeline is given by

$$L_1 = \frac{L - \Delta t \cdot c}{2}. \tag{2}$$

When processing, the function of correlation is obtained by the algorithm

$$K(j) = \sum_{i=1}^{M} (X(i) + Y(i+j)), \tag{3}$$

where $j = 1, 2, 3, ..., N$; $i = 1, 2, 3, ..., M$; X and Y are the digitised values of the received signals; $N = 2^n$ is the number of correlative points of the signal; M is the number of digitised points of the signal.

The design of the measuring system was carried out assuming that an absolute error ΔL of the determination of the leakage coordinates is less than ± 1 m.

The analysis of the detailed designs of the gas pipeline systems of Vilnius have shown that the length L of the gas pipeline between two inspection pits, to which ultrasonic transducers can be screwed, is less than 200 m. At this estimation, it may be written

$$n \geq \log_2 \left(\frac{L}{\Delta L} \right) \geq 7.64, \tag{4}$$

i.e. $n = 8$. From the equations (1) and (4) it is clear that the minimum frequency of digitisation is $f_{T\,\min} = 400$ Hz. On the other hand, a reliable read-out of information and correlation measurements are available only when the condition $f_S \leq f_T/2$ [5] is satisfied, where f_S – maximum frequency of signals which are correlated.

In that way $f_S \leq 2^{n-1}c/L$. When the distance L between the electroacoustical transducers is maximal, ($L = 200$) and $n = 8$, the maximum frequency of the signal f_S cannot exceed 200 Hz. Experimental investigations showed that in the region of low frequency (to some tens of 1 Hz) industrial mechanical and electrical disturbances are very considerable. Therefore this region of frequencies cannot be used for correlation measurements; they need much higher frequencies. Thus the number n must be increased, but this implies a considerably increased measuring time.

Experimental investigations showed that $n = 9$ is the optimum number. The band of passing frequencies required for the correlation measurements is supported by a band-pass filter. The lower boundary of the transmission frequency depends on the intensity of industrial disturbances and lays in the region of 20...80 Hz. The higher boundary of the passing frequencies is chosen twice less than the frequency of digitisation and depends on the length of the pipeline under control. When the length of a section of the pipeline is $L = 200$ m, the frequency of digitisation is $f_T = 440$ Hz. If the length of the pipe section under control is less, the frequency of digitisation increases to a few kHz.

3. Results and discussion

The software of the system developed permits the operator to watch the noise signals on the screen of the computer (Fig. 2) and to adjust their levels before the digitisation. Besides, according to the level of the received signals, the operator can preliminary decide on the condition of the gas pipeline under control.

In order to provide, reliability of the correlation measurements, the signals before their correlation ought to have similar frequency spectra. The spectrum analysis of the signals conveyed to the analogue-digital converter enables one to solve this problem. The frequency band of the signals passing the filter is chosen in accordance with the similarity of the spectra of the received signals in both the channels (Fig. 3).

After that the correlation curve and the numerical information of the location of the leakage of the gas pipeline appears on the screen of the computer (Fig. 4). In the case of a high level of environment noises and when the correlation curve has not a single clearly expressed maximum, a summing of the results of separate correlation measurements

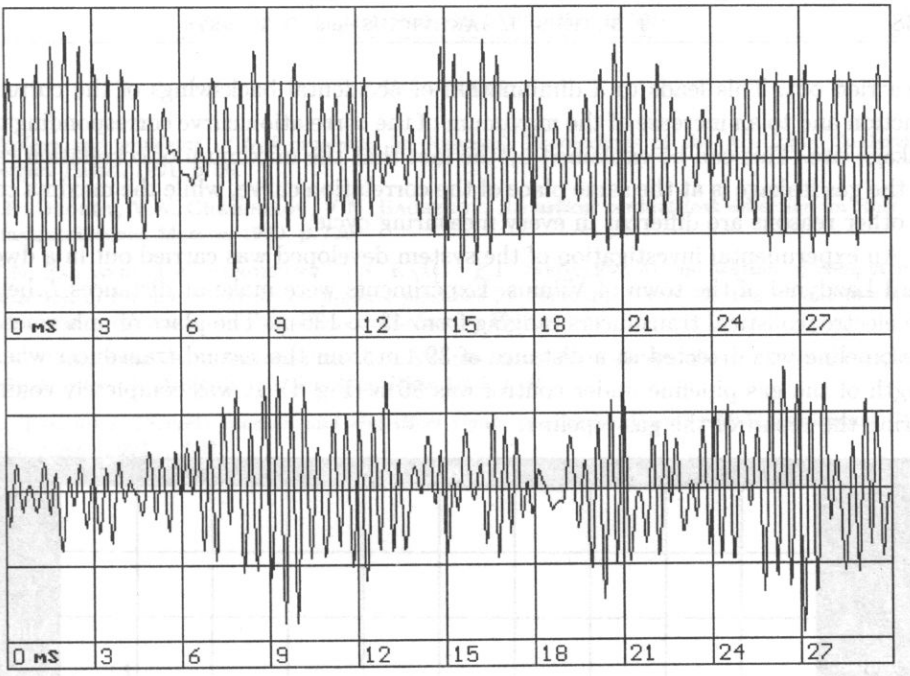


Fig. 2. Noise signal on the screen of the computer.

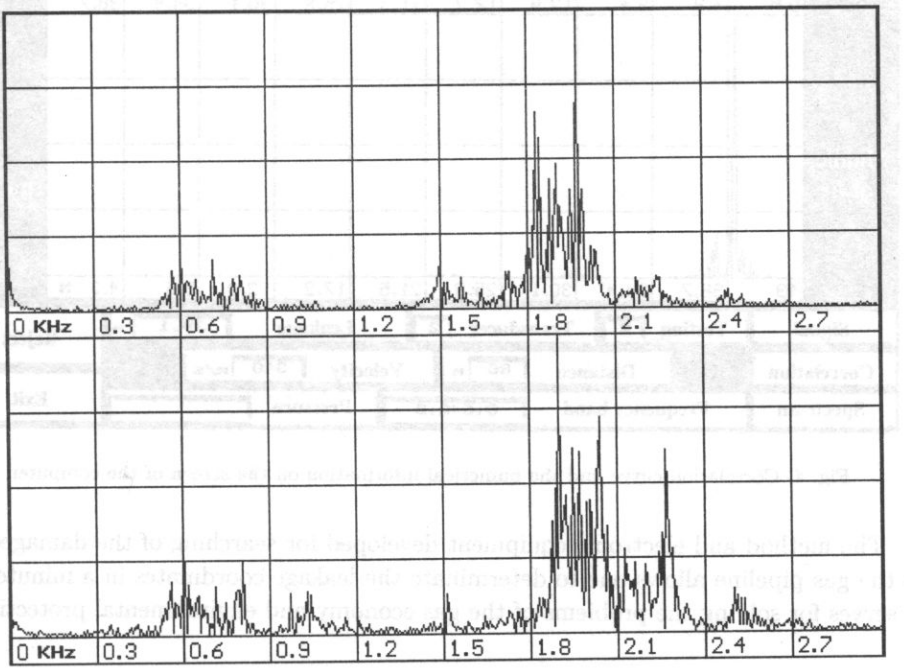


Fig. 3. Spectra of the received acoustic signals in both channels.

is carried out. This leads to a diminnation of accidental backswings of the correlation function and to an increase of the maximum of the correlation curve corresponding to the leakage site. This can be explained by the fact that the maximum of correlation caused by the gas leakage is at the same place of the correlation curve, while the maxima caused by other reasons are different in every measuring cycle.

An experimental investigation of the system developed was carried out in a dwelling ward Lazdynai of the town of Vilnius. Experiments were make at distances L between the electroacoustical transducers ranging from 43 to 136 m. The place of unhermeticy of the pipeline was detected at a distance of 39.1 m from the second transducer when the length of the gas pipeline under control was 86 m (Fig. 4). It was completely confirmed during the repair of the gas pipeline.

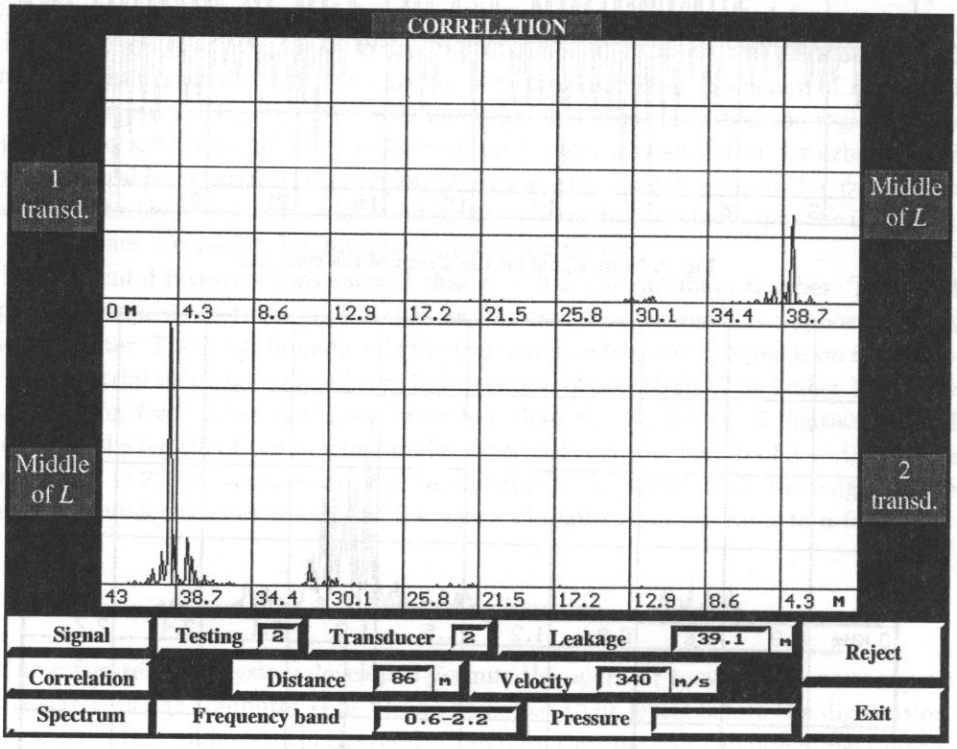


Fig. 4. Correlation curve and the numerical information on the screen of the computer.

The method and electronic equipment developed for searching of the damaged sites of the gas pipeline allows one to determinate the leakage coordinates in a minute time. It serves for solving the problems of the gas economy and environmental protection.

Acknowledgment

This project is sponsored by the Lithuanian Science and Studies Foundation.

References

[1] *Ecology and problems of leak-proofness* [in Russian], Proceedings of the All-Union Conference, Sergiev Pasad, 1991, p. 86.

[2] J.B. DROBOT, V.A. GRESHNIKOV, V.N. BACHEGOV, *Acoustical contact leak detection* [in Russian], Mashinostroenie, Moscow 1989, p. 120.

[3] V.N. ZAGOSKIN, J.A. VENGERCEV, A.S. KAZAK, E.I. JAKOVLEV, *An application of computer for detection of leakage of petroleum product pipelines* [in Russian], Series: Transport and storage of petroleum products and hydrocarbon raw, Moscow 1989, p. 60.

[4] V.S. GODLEWSKIJ, A.A. VLADIMIRSKIJ, I.A. VLADIMIRSKIJ, *Methodology of application of correlative leak detector* "KORSHUN" [in Russian], Kiew 1992, p. 46.

[5] C. THERRIEN, *Discrete random and statistical signal processing*, Prentice-Hall, Englewood Cliffs, New Jersey 1992, p. 392.

ARCHIVES OF ACOUSTICS
23, 4, 541–548 (1998)

C H R O N I C L E

104th AES Convention,

Amsterdam 16–19 May, 1998

Again, the 104th Convention held in Anmsterdam in May 1998, gave an opportunity, like the 103rd one in New York, to celebrate the golden anniversary of the first 50 years of the Audio Engineering Society. The site for this European golden gala Covention was very well chosen. Amsterdam, the venerable city, with ancient traditions and precious monuments, created, like it did four years ago, a delighful atmosphere of the truly international friendship and cooperation.

Opening ceremonies, as well the special celebration meeting, held on Sunday Evening, May 17, were devoted in particular to memories concerning the AES history and the most spectacular achievements attained by AES officers and contributors. The most significant recent accomplishments were honoured by AES citation awards, presented to deserving contemporaries, during this Sunday meeting. The award presentation was of special interest to the Polish AES Section, as well as, probably, to all Polish acousticians, due to the fact that among eight persons awarded this time was Marianna Sankiewicz, the cofounder of the Polish AES Section and its first chairperson.

Mrs Sankiewicz received a Citation Award for her outstanding contribution to European Regional development. It may be added here, that during her Vice-Presidency for AES Central Europe Region, she originated and helped organizing several new AES sections (in Russia, Lithuania, Belarus, Ukraina, Poland), thus enlarging the area of AES influence and augmenting the number of AES members by several hundreds. The moment of award presentation is shown on the enclosed photograph. The AES President Subir Pramanik has just handed the Citation Address to Mrs Sankiewicz, both on the celebration stage, being warmly applauded by the numerous audience.

The results of the 104th Convention were as successful as those of the past ones. 126 papers, grouped into sixteen paper-sessions, revealed actual scientific progress in the most important domains of audio-engineering. 14 workshop-presentations gave opportunities to participants for more close and detailed insight into technical subjects connected to particular enginering problems. Many other special-type sessions were organized with accompanying presentations, in particular those prepared for students.

Several meetings, social events, technical tours were also arranged in a traditional way as during previous Conventions.

118 paper-preprints were edited prior to debates, and distributed among Convention participants. An very rich exhibition with attendance of 296 leading companies from all over the world, completed the achievements of Organizing Committee of the 104th AES Convention.

The truly international character of the Convention activities finds its reflection in the provenience of paper authors. They came from 25 countries, mostly from Europe (18). The highest number of papers was presented by authors from United Kingdom (32), next to them were German authors (25), Dutch ones (16) and those from USA(11). Helas Polish authors participation decreased in comparison to previous European Convention (102nd in Munich) – only four papers were presented in Amsterdam, all of them by authors from Wrocław, who, thereby, are worthy of being laudatorily mentioned.

B.W. KULESZA, W. RYSZKOWSKI, *The sound reinforcement system of large halls*, Sound Reinforcement, P12-3, Preprint 4665.

R. BOLEJKO, P. PRUCHNICKI, *Sound diffuser based on number theory with random variations of surface acoustic impedance*, Architectural Acoustics, P8-3, Preprint 4711.

M.J. KIN, J. RENOWSKI, *Pitch perception of frequency-modulated tones*, Psychoacoustics, P 14-3, Preprint 4670.

K. BAŚCIUK, *Measurements of logatom intelligibility in digital chains with delta modulation coding*, Quality evaluation of sound systems, P16-5, Preprint 4707.

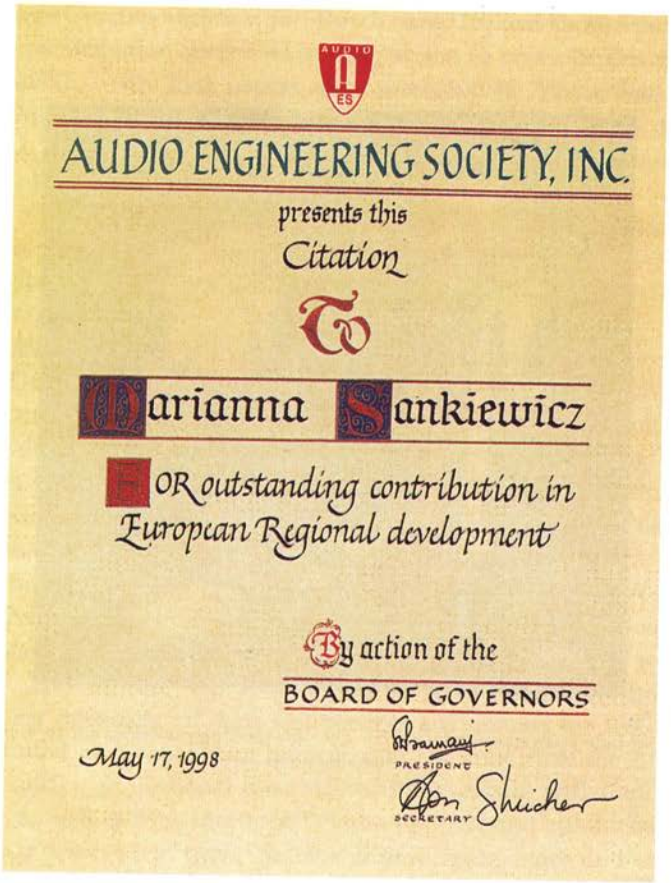
The next Convention have been announced to be held in San Francisco – the 105th Fall'98 Convention, and in Munich – the 106th Spring '99 Convention. Let us hope the participation of Polish authors will argument, especially in Munich due to the relative easiness of access from Poland, and with reference to the successful Polish participation there, during the 102nd Convention.

There will be, besides, two more International AES Conferences soon: – the 15th in Copenhagen, Denmark (October 31–November 2), entitled "Audio, Acoustics and Small Spaces", and the 16th in Rovaniemi, Finland (April 10–12), devoted to "Spatial Sound Reproduction". May be they will attire attention of our colleagues working on similar problems. Further international AES Conferences are planned for 1999 and 2000.

It may be stated, as a concluding remark, that the activity of the Audio Engineering Society in organizing professional meetings of its members with representants of the audio industry, as well as with ther broad circles of interested public, has recently grown in a substantialy way, while them Society is developing more and more into a fully international organization. The Polish AES Section is supporting this development with all possible means.

Jan A. Adamczyk

Marianna Sankiewicz is awarded citation for her outstanding contribution to European Regional development.


AUDIO ENGINEERING SOCIETY, INC.
presents this
Citation
To
Marianna Sankiewicz
FOR outstanding contribution in
European Regional development
By action of the
BOARD OF GOVERNORS
May 17, 1998
PRESIDENT
SECRETARY

Seminar on
Ultrasound in Biomeasurements, Diagnostics and Therapy

2–5 September 1998, Warsaw

The International Seminar on Ultrasound in Biomeasurements, Diagnostics and Therapy was the third seminar organised in the period between 1992 and 1998 by the International Center of Biocybernetics (ICB) in Warsaw.

The presented papers covered the following topics: sonography, tissue characterization, signal processing, transducers, Doppler, physics of diagnostic ultrasound, and physics of high intensity ultrasound.

In the opening ceremony Prof. M. NAŁĘCZ, the Director of the International Center of Biocybernetics in Warsaw, welcomed the participants of the seminar describing shortly the organization of the ICB and the seminars on several subjects of biomedical engineering organized regularly.

Next Prof. A. NOWICKI, the Chairman of the Seminar welcomed the participants and opened the Seminar.

The following papers were presented during the Symposium:

W. JAKUBOWSKI (Warsaw, Poland), *The activity of Polish Ultrasound Society.*

M. THIJSSEN, A.P. BERKHOFF (Nijmegen, The Netherlands), *Ultrasonic imaging in aberrating media.*

K. BOM (Rotterdam, The Netherlands), *Aspects of intravascular imaging.*

K. BOM (Rotterdam, The Netherlands), *3D imaging: principles, limitations and results.*

K. CHIHARA (Nara, Japan), *Three-dimensional ultrasound imaging system.*

W. BICZ, Z. GUMIENNY, D. KOSZ, S. GOMULIŃSKI (Wrocław, Poland), *Ultrasonic equipment for fingerprint recognition applied to tomographical testing of materials and biological objects.*

E. SZYMAŃSKA, A. NOWICKI, K. MLOSEK, J. LITNIEWSKI, M. LEWANDOWSKI, W. SECOMSKI (Warszawa, Poland), *Skin imaging with high frequency ultrasound.*

J.M. THIJSSEN, H.J. HUISMAN (Nijmegen, The Netherlands), *In vivo ultrasonic tissue characterization of liver metastases.*

P. LAUGIER, G. BERGER (Paris, France), *Quantitative ultrasonic imaging of bone for the diagnosis of osteoporosis: fundamental and clinical aspects.*

J. LITNIEWSKI, A. NOWICKI (Warszawa, Poland), *Bones characterization with high and low frequency ultrasounds.*

M. SCHUBERT, O. LENKEIT, W. GRILL (Leipzig, Germany), *Volume imaging by three dimensional scanning acoustic microscopy.*

T. JĘDRZEJEWICZ (Mountain View, USA), *Ultrasound speckle.*

R. OLSZEWSKI, A. NOWICKI, J. ETIENNE, P. KARŁOWICZ, W. MARCINIAK, J. ADAMUS (Warsaw, Poland), *Echogen and harmonic imaging in assessment of left ventricular function.*

P.A. LEWIN (Philadelphia, USA), *Pulse shaping and deconvolution processing in ultrasound image optimization.*

M. MOSZYŃSKI (Gdańsk, Poland), *Wavelet techniques in fish target strength estimation.*

Z. ŁUBNIEWSKI, A. STEPNOWSKI (Gdańsk, Poland), *Fractal analysis in sea bottom identification and monitoring.*

W. LIS, R. SALAMON J. ZIENKIEWICZ (Gdańsk, Poland), *Application of processed acoustic emission data for investigation of the temporomandibular joint pathology.*

T. GUDRA, K. OPIELIŃSKI (Wrocław, Poland), *Some problems of ultrasonic transmission tomography.*

K.K. SHUNG, T. RITTER, P. LOPATH (University Park, USA), *Recent advances in ultrasonic transducer engineering.*

T. JĘDRZEJEWICZ (Mountain View, USA), *Broadband phased array transducer design with frequency- controlled two-dimensional capability.*

K. KYCIA, A. NOWICKI, G. ŁYPACEWICZ (Warszawa, Poland), *FEM modelling of composite transducers.*

G. GUIDI, C. LICCIARDELLO (Firenze, Italy), *Acoustic field mapping by means of Doppler detection.*

P. TORTOLI, F. CURRADI, E. MAIONE (Firenze, Italy), *Behaviour of contrast agents in high-intensity beams.*

T. POWAŁOWSKI, Z TRAWIŃSKI, L. BOROWSKA (Warszawa, Poland), *Non-invasive ultrasonic examination of the forward and reflected pressure wave in human arteries.*

P. TORTOLI, F. GUIDI, G. GUIDI, C. MENICUCCI, C. ATZENI (Firenze, Italy), *Real-time detection of velocity profiles in human arteries.*

J. IHNATOWICZ, Z. KALINA (Katowice, Poland), *The computer aided quantitative analysis of the colour flow mapped (CFM) ultrasound Doppler images.*

K. KAŁUŻYŃSKI, D. LIEPSCH, B. LEŚNIAK (Warszawa, Poland), *Laser and ultrasonic Doppler flow studies in carotid bifurcation models with irregular wall surface.*

L. HILGERTNER, T. POWAŁOWSKI, Z. TRAWIŃSKI (Warszawa, Poland), *Common carotid artery wall elasticity in patients with unilateral carotid occlusive disease.*

Z. TRAWIŃSKI, T. POWAŁOWSKI, L. HILGERTNER (Warszawa, Poland), *Influence of atherosclerosis risk factors on the common carotid artery wall elasticity.*

R. OLSZEWSKI, A. NOWICKI, J. ETIENNE, P. KARŁOWICZ, W. MARCINIAK, J. ADAMUS (Warszawa, Poland), *The value of TDI in the estimation of myocardial velocity gradients and early detection of thrombi.*

J. WÓJCIK (Warszawa, Poland), *A new theory of absorption for finite amplitude disturbances.*

L. FILIPCZYŃSKI, J. ETIENNE, T. KUJAWSKA, R. TYMKIEWICZ, J. WÓJCIK (Warszawa, Poland), *Measurements of linearity of a PVDF hydrophone for ultrasonic high pressures*

L. FILIPCZYŃSKI, J. ETIENNE, T. KUJAWSKA, R. TYMKIEWICZ, J. WÓJCIK (Warszawa, Poland), *Nonlinear effects of ultrasonic focused beams of high intensity measured by means of PVDF and electromagnetic hydrophones.*

J. WÓJCIK (Warszawa, Poland), *Numerical results of temperature elevations computed on the basis of a new absorption theory.*

A. KOCHAŃSKI, J. MEJNAROWICZ, A. LATOS-BIELEŃSKA, J. ETIENNE, L. FILIPCZYŃSKI (Poznań, Warszawa, Poland), *The damage of genomic DNA by shock waves generated by the lithotripter.*
G. GRELOWSKA, E. KOZACZKA (Gdynia, Poland), *Investigation of nonlinear focused acoustic fields.*

All speakers were accommodated in the at the ICB Hotel next to the Seminar place. This gave the opportunity for further afternoon and evening discussions and comments on scientific papers which number equalled to 37.

Apart from the technical programme several social; events like ice-breaking party, a concert and excursions to interesting Warsaw historical places were organized.

The chairmen of the Symposium Prof. Prof. L. FILIPCZYŃSKI, A. NOWICKI and P. TORTOLI hope to welcome all the researches, interested in the subjects presented during this meeting, in the next Seminar, which will be organized probably in the year 2001.

2nd EAA International Symposium on Hydroacoustics

Gdańsk – Jurata 1999, May 24-27

First Announcement and Call for Papers

The 2nd EAA International Symposium on Hydroacoustics will be held in Jurata near Gdańsk, Poland from 24 to 27 May 1999.

The Symposium is jointly organised by the Technical University of Gdańsk and the Naval Academy in Gdynia under the auspices of the European Acoustical Association, and sponsored by the Polish National Committee for Scientific Research.

The Symposium will include invited and contributed papers on the following topics:

- Ambient and ship noise,
- Reverberation and scattering,
- Sound propagation modelling,
- Sonar systems and data processing,
- Non-linear acoustics,
- Acoustics in fisheries,
- Transducers and instrumentation.

The working language of the Symposium will be English (for both the manuscripts and the oral presentations).

We are asking interested participants to submit an abstract for review by the organising committee before **November 30, 1999**.

All authors will be notified by **January 15, 1999** on acceptance on their abstracts. Full text of papers must be submitted by **March 15, 1999**.

The abstracts and the pre-registration form can be sent by regular mail or via our Web Site at:

WWW: www.hydro.eti.pg.gda.pl

The more detailed facts about the conference, guidelines for abstract and papers editing and the deadlines of their submission can be also found on our WWW pages.

The conference proceedings will be published and distributed among the participants during the conference.

Organising Committee:

Prof. A. Stepnowski	President
Prof. R. Salamon	Vice President
Prof. J. Blauert	EAA Representative
Prof. H. Lasota	
Dr. G. Grelowska	
Dr. J. Dobrzeniecki	
Dr. L. Kilian	
Mr A. Partyka	

Address:
Secretariat of the 2nd EAA International Symposium on Hydroacoustics,
Technical University of Gdańsk,
Department of Acoustics,
ul. G. Narutowicza 11/12
80-952 Gdańsk, POLAND
tel.: (+48 58) 3472320, (+48 58) 3472511
fax: (+48 58) 3471535, (+48 58) 3416132
e-mail: hydro@eti.pg.gda.pl