INFORMATICA, 2000, Vol. 11, No. 3, 243–256

Optimal Segmentation of Random Sequences

Antanas LIPEIKA
Institute of Mathematics and Informatics
Akademijos 4, 2600 Vilnius, Lithuania
e-mail: lipeika@ktl.mii.lt

Received: July 2000

Abstract. This paper deals with maximum likelihood and least square segmentation of autoregressive random sequences with abruptly changing parameters. Conditional distribution of the observations has been derived. Objective function was modified to the form suitable to apply dynamic programming method for its optimization. Expressions of Bellman functions for this case were obtained. Performance of presented approach is illustrated with simulation examples and segmentation of speech signals examples.

Key words: optimal segmentation, maximum likelihood, least square, dynamic programming.

1. Introduction

Digital signal processing theory is well developed only for linear time invariant systems (Proakis *et al.,* 1996). Unfortunately, in practice most signals can be regarded as output of a linear but time variant system. Some kinds of signals, e.g., speech signals (Rabiner *et al.,* 1993), to some extent can be interpreted as output of a linear system with abruptly changing parameters.

One problem arising in signals with abruptly changing parameters is change point detection problem (Page, 1958). Various statistical tests were applied for solving this problem (Kander *et al.,* 1965; Bhattacharya *et al.,* 1968) but optimal solutions were obtained only for sequences of independent random variables. Later sequences of dependent random variables were investigated (Telksnys, 1972; 1973), Especially constructive results were obtained for autoregressive (AR) and autoregressive-moving average (ARMA) random sequences (Kligienė, 1974) and (Lipeikienė, 1973). In all these references problem was formulated under assumption that there is one change point in observable random sequence.

Many change point case problem was formulated by Telksnys (1969) but it was solved only for random sequences with some simple covariance matrices. This problem for autoregressive random sequences was formulated by Lipeika (1975) and in Lipeika (1977) explicit expression of the likelihood function as function of many change points was derived. The next problem was that of optimization of a likelihood function depending on many variables. This problem was investigated in Lipeika (1979) and the likelihood function was replaced by another function for which a point (in multivariate space) of the global maximum remains the same and solution reducing amount of computations was given. In Lipeika *et al.* (1987) dynamic programming approach was applied for maximization of likelihood function depending on many variables.

In this paper full constructive solution of maximum likelihood and least square error estimation of change points problem is given and simulation examples are presented. This solution enables reduce amount of computations from order N^M to order $N \times M$, where N is a length of observable random sequence and M – a number of change points in this sequence. Application of this approach to estimation of change points in speech signals is demonstrated.

2. Statement of the Problem

Let us consider random sequence $x = \{x(1), x(2), \dots, x(N)\}$, which is an output of linear discrete time system. Input of the system $v(n)$ is a sequence of Gaussian independent random variables with zero mean and unit variance. The system structure satisfies autoregressive (AR) equation of the form

$$x(n) = -a_1(n)x(n-1) - a_2(n)x(n-2) - \dots - a_p(n)x(n-p) + b(n)v(n), \quad (1)$$

where p is an order of autoregressive system; $A'(n) = [a_1(n), a_2(n), \dots, a_p(n), b(n)]$ is a vector of time-varying parameters of the system at every time instant satisfying system stability conditions. The parameters $A(n)$ of the system are known a priori and are changing according to the rule

$$A(n) = \begin{cases} A_1, & n = \dots, 1, 2, \dots, u_1, \\ A_2, & n = u_1 + 1, \dots, u_2, \\ \dots & \quad \dots \\ A_i, & n = u_{i-1} + 1, \dots, u_i, \\ \dots & \quad \dots \\ A_M, & n = u_{M-1} + 1, \dots, u_M, \\ A_{M+1}, & n = u_M + 1, \dots, N, \dots, \end{cases} \quad (2)$$

where $u = [u_1, u_2, \dots, u_M]$ are random change points of the known system parameters $A(n)$ with a priori distribution $P(u)$ and satisfy the condition $p < u_1 < u_2 \dots < u_M < N$. Input of the system does not depend of change points u.

The problem is to obtain estimates $\widehat{u} = [\widehat{u}_1, \widehat{u}_2, \dots, \widehat{u}_M]$ of change points $u = [u_1, u_2, \dots, u_M]$, when the realization $x = \{x(1), x(2), \dots, x(N)\}$ of the random sequence is available.

3. Optimization Criterion

The problem of obtaining estimates of the change point's u can be treated as a classification problem. We will solve it using probabilistic approach (Duda *et al.,* 1973), the maximum a posteriori probability clasifier. A posteriori probability $P(u|x)$ can be calculated from the conditional probability distribution density $p(x|u)$ using the Bayes formula as

$$P(u|x) = \frac{p(x|u)P(u)}{p(x)}, \tag{3}$$

where

$$p(x) = \sum_u p(x|u)P(u). \tag{4}$$

Then the maximum probability estimator $\widehat{u}$ of change points u is

$$\widehat{u} = \arg\max_u P(u|x). \tag{5}$$

For the given realization x of the random sequence denominator of (3) is constant. Instead of maximizing (5) we can maximize $p(x|u)P(u)$

$$\widehat{u} = \arg\max_u p(x|u)P(u). \tag{6}$$

When a priori distribution $P(u)$ is unknown (or constant), we can maximize $p(x|u)$

$$\widehat{u} = \arg\max_u p(x|u). \tag{7}$$

This estimator is called maximum likelihood estimator of change points u. We will assume a priori distribution $P(u)$ to be constant and solve the problem of maximum likelihood (ML) estimation of change points u.

4. Derivation of the Conditional Distribution

Since input of the system (1) is a sequence of Gaussian independent random variables (independent of u) with zero mean and unit variance and the system (1) is linear the conditional distribution $p(x|u)$ is also Gaussian.

We can also express $p(x|u)$ in terms of conditional distributions

$$p(x|u) = p(x(1), x(2), \ldots, x(p)) \prod_{n=p+1}^{N} p(x(n)|u, x(n-1), \ldots, x(1)). \tag{8}$$

As output of the system (1) at time instant n depends only of p previous outputs $x(n-1), \ldots, x(n-p)$, we can rewrite (8) in the form

$$p(x|u) = p(x(1), x(2), \ldots, x(p)) \prod_{n=p+1}^{N} p(x(n)|u, x(n-1), \ldots, x(n-p)). \tag{9}$$

Taking into account change points u and the condition $p < u_1 < u_2 \ldots < u_M < N$ we have

$$\begin{aligned} p(x|u) \;=\; & p(x(1), x(2), \ldots, x(p)) \prod_{n=p+1}^{u_1} p(x(n)|u, x(n-1), \ldots, x(n-p)) \\ & \times \prod_{n=u_1+1}^{u_2} p(x(n)|u, x(n-1), \ldots, x(n-p)) \times \ldots \\ & \times \prod_{n=u_M+1}^{N} p(x(n)|u, x(n-1), \ldots, x(n-p)). \end{aligned} \tag{10}$$

Since $p(x|u)$ is Gaussian distribution density, $p(x(n)|u, x(n-1), \ldots, x(n-p))$ is also Gaussian univariate distribution density with conditional mean

$$m(n) = E[x(n)|u, x(n-1), \ldots, x(n-p))] \tag{11}$$

and conditional variance

$$\sigma^2(n) = E[(x(n) - m(n))^2 | u, x(n-1), \ldots, x(n-p))]. \tag{12}$$

Taking into account (1) we can write (11) and (12) as

$$m(n) = -a_1(n)x(n-1) - \ldots - a_p(n)x(n-p), \tag{13}$$

$$\sigma^2(n) = b^2(n), \tag{14}$$

and

$$p(x(n)|u, x(n-1), \ldots, x(n-p)) = \frac{1}{\sqrt{2\pi b^2(n)}} \exp\left\{-\frac{[x(n) - m(n)]^2}{2b^2(n)}\right\}. \tag{15}$$

Substituting (15) into (10) and taking into account (2) we can write (Lipeika, 1977)

$$\begin{aligned} p(x|u) \;=\; & p(x(1), x(2), \ldots, x(p)) \times (2\pi)^{-(N-p)/2} \\ & \times b^{-(u_1-p)}(1) \times b^{-(u_2-u_1)}(2) \times \ldots \times b^{-(N-u_M)}(M+1) \\ & \times \exp\left\{\frac{1}{2b^2(1)} \sum_{n=p+1}^{u_1} \left[\sum_{j=0}^{p} a_j(1)x(n-j)\right]^2 \right. \\ & - \frac{1}{2b^2(2)} \sum_{n=u_1+1}^{u_2} \left[\sum_{j=0}^{p} a_j(1)x(n-j)\right]^2 - \ldots \\ & \left. - \frac{1}{2b^2(M+1)} \sum_{n=u_M+1}^{N} \left[\sum_{j=0}^{p} a_j(1)x(n-j)\right]^2 \right\}, \end{aligned} \tag{16}$$

where we assume $a_0(i) = 1$, $i = 1, 2, \ldots, M+1$.

5. Modification of the Likelihood Function

Instead of using $p(x|u)$ (expressions (7) and (16)) in implementation it is more convenient to use $\log p(x|u)$. It helps us to avoid exceeding dynamic range of the computer and reduces computation score. Taking $\log p(x|u)$ we can write (7) and (16) as

$$\widehat{u} = \arg\max_u p(x|u) = \arg\max_u \log p(x|u), \tag{17}$$

and

$$\begin{aligned}\log p(x|u) \;=\; & \log p(x(1), x(2), \ldots, x(p)) - (N-p)/2 \log(2\pi) \\ & - (u_1 - p)\log b(1) - (u_2 - u_1)\log b(2) - \ldots \\ & - (N - u_M)\log b(M+1) - \frac{1}{2b^2(1)} \sum_{n=p+1}^{u_1} \left[\sum_{j=0}^{p} a_j(1)x(n-j)\right]^2 \\ & - \frac{1}{2b^2(2)} \sum_{n=u_1+1}^{u_2} \left[\sum_{j=0}^{p} a_j(2)x(n-j)\right]^2 \\ & - \ldots - \frac{1}{2b^2(M+1)} \sum_{n=u_M+1}^{N} \left[\sum_{j=0}^{p} a_j(M+1)x(n-j)\right]^2.\end{aligned} \tag{18}$$

For the given realization of the random sequence instead of maximizing (18) we can maximize objective function $\Theta(u|x)$ (Lipeika, 1979), which differs from (18) by an additive constant not depending on u, i.e.,

$$\widehat{u} = \arg\max_u \log p(x|u) = \arg\max_u \Theta(u|x), \tag{19}$$

where

$$\Theta(u|x) = L_1(u_1|x) + L_2(u_2|x) + \ldots + L_M(u_M|x). \tag{20}$$

Each of functions $L_i(u_i|x)$, $i = 1, 2, \ldots, M$ depends only on one unknown change point u_i and can be expressed as

$$\begin{aligned}L_i(k|x) = & -\; (k-p)\log b(i) - (N-k)\log b(i+1) \\ & -\; \frac{1}{2b^2(i)} \sum_{n=p+1}^{k} \left[\sum_{j=0}^{p} a_j(i)x(n-j)\right]^2 \\ & -\; \frac{1}{2b^2(i+1)} \sum_{n=k+1}^{N} \left[\sum_{j=0}^{p} a_j(i+1)x(n-j)\right]^2, \\ & i = 1, 2, \ldots, M; \quad k = p+1, 2, \ldots, N,\end{aligned} \tag{21}$$

or can be calculated recursively

$$
\begin{aligned}
L_i(k|x) \;=\; & L_i(k-1|x) - \log b(i) + \log b(i+1) \\
& - \frac{1}{2b^2(i)}\Bigg[\sum_{j=0}^{p} a_j(i)x(n-j)\Bigg]^2 \\
& + \frac{1}{2b^2(i+1)}\Bigg[\sum_{j=0}^{p} a_j(i+1)x(n-j)\Bigg]^2 \\
& i = 1,2,\ldots,M; \quad k = p+1,2,\ldots,N,
\end{aligned} \tag{22}
$$

with the initial conditions

$$L_i(p|x) = 0, i = 1,2,\ldots,M.$$

6. Maximization of the Objective Function Using Dynamic Programming Method

Since the function $\Theta(u|x)$ consists of the sum of partial functions $L_i(u_i|x)$, $i = 1,2,\ldots,M$ and each of these partial functions depends only on one variable, we can use the dynamic programming method to determine place of the global maximum (Lipeika e*t al.,* 1987) of this function.

According to the dynamic programming method let us define the Bellman functions (Cooper e*t al.,* 1981). Procedure fully depends on restrictions in the range of the maximization variables. In our case the restrictions are

$$p < u_1 < u_2 < \ldots < u_M < N. \tag{23}$$

Let us begin from maximization of $L_1(u_1|x),\ p < u_1 < u_2$. We can see that maximum of $L_1(u_1|x)$ depends only on u_2. So we can define the function $g_1(u_2|x)$ such that

$$g_1(u_2|x) = \max_{\substack{u_1 \\ p<u_1<u_2}} L_1(u_1|x), \quad u_2 = p+2,\ldots,N. \tag{24}$$

We can see that the global maximum of the of the function $g_1(u_2|x)$ (and also of the function $L_1(u_1|x)$) depends only on u_2 (one variable) for all possible values of $u_3, u_4, \ldots, u_M$, satisfying the condition $u_2 < u_3 < u_4 < \ldots < u_M < N$.

Now let us maximize $L_1(u_1|x) + L_2(u_2|x),\ p < u_1 < u_2 < u_3$. We can see that maximum of $L_1(u_1|x) + L_2(u_2|x)$ depends only on u_3. So we can define the function $g_2(u_3|x)$ such that

$$g_2(u_3|x) = \max_{\substack{u_2 \\ p+1<u_2<u_3}} \big[L_2(u_2|x) + g_1(u_2|x)\big], \quad u_3 = p+3,\ldots,N. \tag{25}$$

The global maximum of the of the function $g_2(u_3|x)$ (and also of the function $L_1(u_1|x) + L_2(u_2|x)$) depends only on u_3 (one variable) for all possible values of $u_4, u_5, \ldots, u_M$, satisfying the condition $u_3 < u_4 < u_5 < \ldots < u_M < N$.

Similarly we can maximize the function

$$\begin{aligned} &L_1(u_1|x) + L_2(u_2|x) + \ldots + L_i(u_i|x), \\ &p < u_1 < u_2 < u_3 < \ldots < u_i < u_{i+1}, \end{aligned} \tag{26}$$

as the maximum of $L_1(u_1|x) + L_2(u_2|x) + \ldots + L_i(u_i|x)$ depends only on u_{i+1}. So we can define the function $g_i(u_{i+1}|x)$ such that

$$\begin{aligned} &g_i(u_{i+1}|x) = \max_{\substack{u_i \\ p+i-1<u_i<u_{i+1}}} \big[L_i(u_i|x) + g_{i-1}(u_i|x)\big], \\ &u_{i+1} = p+i+1, \ldots, N. \end{aligned} \tag{27}$$

The global maximum of the of the function $g_i(u_{i+1}|x)$ (and also of the function $L_1(u_1|x) + L_2(u_2|x) + \ldots + L_i(u_i|x)$) depends only on u_{i+1} (one variable) for all possible values of $u_{i+2}, u_{i+3}, \ldots, u_M$, satisfying the condition $u_{i+1} < u_{i+2} < u_{i+3} < \ldots < u_M < N$.

And for maximization of

$$\Theta(u|x) = L_1(u_1|x) + L_2(u_2|x) + \ldots + L_M(u_M|x)$$

we have

$$\begin{aligned} &g_M(u_{M+1}|x) = \max_{\substack{u_M \\ p+M-1<u_M<u_{M+1}}} \big[L_M(u_M|x) + g_{M-1}(u_M|x)\big], \\ &u_{M+1} = p+M+1, \ldots, N, \end{aligned} \tag{28}$$

where u_{M+1} is additional variable, which means the possible length of the available realization of the random sequence. Since the length of the realization is N, the value $g_M(N|x)$ is the global maximum of the $\Theta(u|x)$. The functions $g_i(n|x)$ are not decreasing functions with respect to $n = p+1, \ldots, N$.

Now the maximum likelihood estimator $\widehat{u} = [\widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_M]$ of the change points u is obtained in the following way

$$\widehat{u}_k = \min\Big[\arg \max_{\substack{n \\ p+k\leqslant n\leqslant \hat{u}_{k+1}}} g_k(n|x)\Big], \quad k = M, M-1, \ldots, 2, 1, \tag{29}$$

where, for convenience, we made a notation $\widehat{u}_{M+1} = N$.

For further reduction of computation amount, we can compute the functions $g_i(u_{i+1}|x)$, $i = 1, \ldots, M$ recursively (Lipeika *et al.,* 1992)

$$g_1(u_2|x) = \max\big[g_1(u_2-1|x), L_1(u_2-1|x)\big], \quad u_2 = p+3, \ldots, N, \tag{30}$$

with the initial condition $g_1(p+2|x) = L_1(p+1|x)$ and for $i = 2, \dots, M$

$$g_i(u_{i+1}|x) = \max\Big\{g_i(u_{i+1}-1|x), \big[g_{i-1}(u_{i+1}-1|x) + L_i(u_{i+1}-1|x)\big]\Big\},$$
$$u_{i+1} = p+i+2, \dots, N, \tag{31}$$

with the initial conditions

$$g_i(p+i+1|x) = L_i(p+i|x) + g_{i-1}(p+i|x), \quad i = 2, \dots, M. \tag{32}$$

Finally we obtained maximum likelihood estimator of change points $\widehat{u} = [\widehat{u}_1, \widehat{u}_2, \dots, \widehat{u}_M]$. By the way, it is interesting to note, that by assuming $b(i) = 1,\ i = 1, 2, \dots, M+1$ and using the same optimization method we obtain least square (LS) estimates of change points $u = [u_1, u_2, \dots, u_M]$. According to needs of implementation there are several possible modifications of the optimization algorithm (this problem is not discussed in this paper).

This solution enables to reduce amount of computations from the order N^M (full search) to the order $N \times M$ (dynamic programming), where N is a length of an observable random sequence and M is the number of change points in this sequence. E.g., if $N = 10000$, $M = 5$, we have $N^M = 10^{15}$ for full search and $N \times M = 50000$ for dynamic programming approach. This result is very important for "real world" applications, e.g., speech signal segmentation into phonemes, EEG (electroencephalogram) analysis, etc.

7. Simulation Examples

We solved the following simulation problem. We generate a realization of the random tenth order autoregressive sequence with parameters corresponding to the Lithuanian word "namas" as listed in Table 1.

Change points were selected as $u = [100, 200, 300, 400, 500]$, length of the realization $N = 600$. The maximum likelihood (ML) estimate of the change points was obtained $\widehat{u}_{ML} = [101, 200, 300, 400, 499]$, the least square (LS) estimate was $\widehat{u}_{LS} =$

Table 1

Parameters of the autoregressive model and their meaning

Parameters of the AR model	Meaning
A_1	Background noise
A_2	Sound 'n'
A_3	Sound 'a'
A_4	Sound 'm'
A_5	Sound 'a'
A_6	Sound 's'

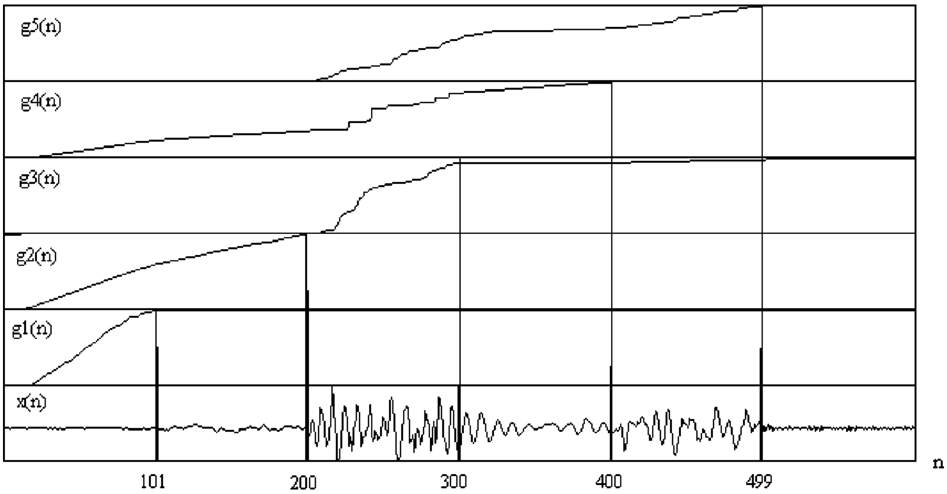

Fig. 1. Generated realization and the Bellman functions. Estimated change points are $\widehat{u}_{ML} = [101, 200, 300, 400, 499]$, the maximum likelihood estimate. The true change points are $u = [100, 200, 300, 400, 500]$.

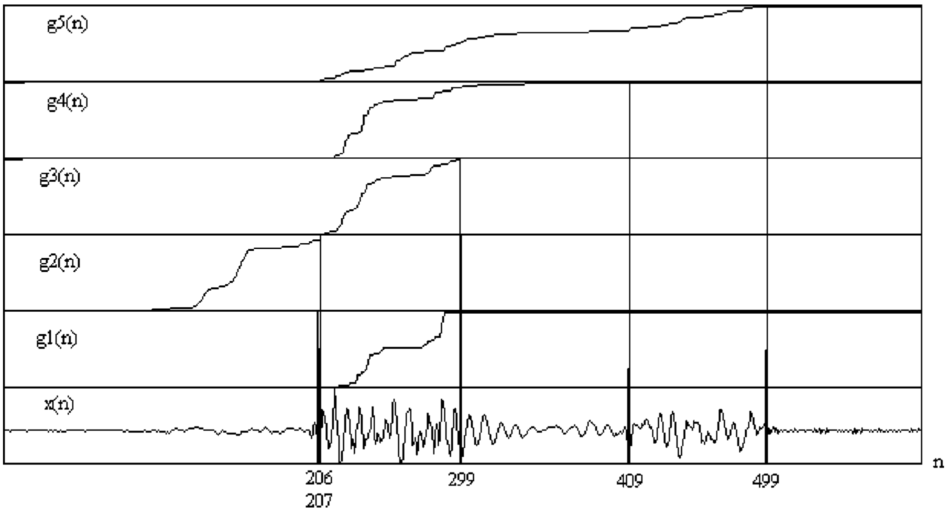

Fig. 2. Generated realization and the Bellman functions. Estimated change points are $\widehat{u}_{LS} = [206, 207, 299, 409, 499]$, the least square estimate. The true change points are $u = [100, 200, 300, 400, 500]$.

$[206, 207, 299, 409, 499]$. In Fig. 1 the realization of generated random sequence and the Bellman functions are shown for the maximum likelihood estimate. Vertical lines indicate estimated change points.

Results of the least square estimation for the same realization are displayed in Fig. 2. From these figure we can see that the algorithm was not able to catch transition from background noise parameters to the sound 'n' parameters.

The same experiment was performed with longer realization, $N = 6000$. Change points were selected $u = [1000, 2000, 3000, 4000, 5000]$. The maximum likelihood estimate of the change points was $\widehat{u}_{ML} = [1002, 2001, 3000, 4001, 4997]$ and the least square estimate was $\widehat{u}_{LS} = [1002, 2009, 3000, 3996, 4997]$. In this case change from background noise to the sound 'n' was caught immediately even by the least square algorithm ($u_1 = 1000, \widehat{u}_{1LS} = 1002$). In Fig. 3 a realization of generated random sequence and the Bellman functions are shown for the maximum likelihood estimate.

To get insight about a behavior of partial likelihood functions $L_i(n|x)$ we provided the following experiment. We generated random sequence with parameters

$$A(n) = \begin{cases} A_1 - noise, & n = 1, 2, \ldots, 1000, \\ A_2 - 'n', & n = 1001, \ldots, 2000, \\ A_5 - 'a', & n = 2001, \ldots, 3000, \\ A_6 - 's', & n = 3001, \ldots, 4000, \end{cases}$$

(used in previous examples) but gain parameter $b(n)$ was made equal to 1, i.e., $b(n) = 1$. In this case maximum likelihood estimate coincides with least square estimate $\widehat{u}_{ML} = \widehat{u}_{LS} = [1002, 1993, 2997]$. Partial likelihood functions $L_i(n|x)$ for this experiment are displayed in Fig. 4. From this figure we can see that ensemble mean of partial likelihood function, corresponding to particular change point, increases as n approaches to change point and decreases as n moves away from the change point (Lipeika, 1973). When n is more than second change point and is increasing, ensemble mean of partial likelihood function can increase or decrease, it depends on particular system parameters corresponding to this time segment.

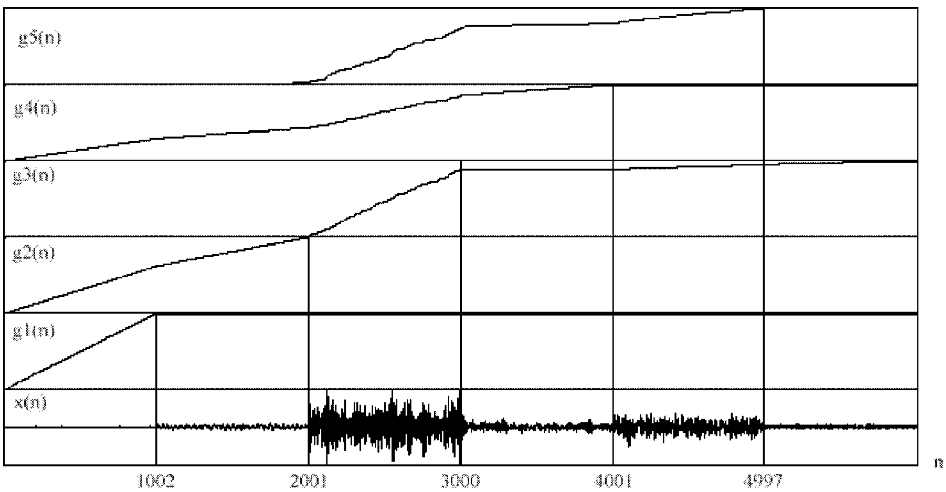

Fig. 3. Generated realization and the Bellman functions. Estimated change points are $\widehat{u}_{ML} = [1002, 2001, 3000, 4001, 4997]$, the maximum likelihood estimate. True change points are $u = [1000, 2000, 3000, 4000, 5000]$.

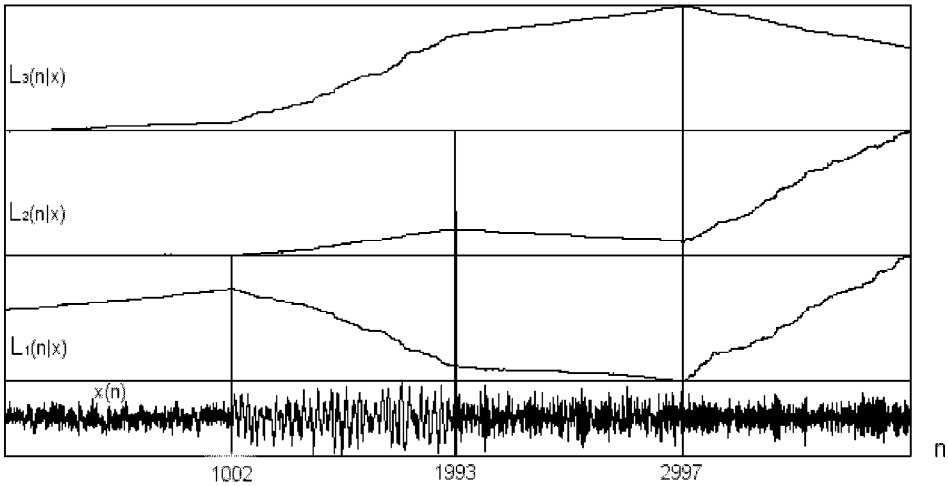

Fig. 4. Generated realization and the partial likelihood functions $L_i(n|x)$. Estimated change points are $\widehat{u}_{ML} = \widehat{u}_{LS} = [1002, 1993, 2997]$. True change points are $u = [1000, 2000, 3000]$.

8. Estimation of Change Points in Speech Signals

Now will demonstrate a few examples with change points estimation in speech signals. In Fig. 5 utterance of the Lithuanian word "namas" and the Bellman functions are shown. Similarly in Fig. 6 utterance of the Lithuanian word "mama" and the Bellman functions are shown.

In Fig. 7 utterance of the Lithuanian word "saulė" pronounced by female speaker and the Bellman functions are shown. In this case the maximum likelihood estimate was not able to "catch" transition from sound 'u' to 'l' but the least square estimate solved the

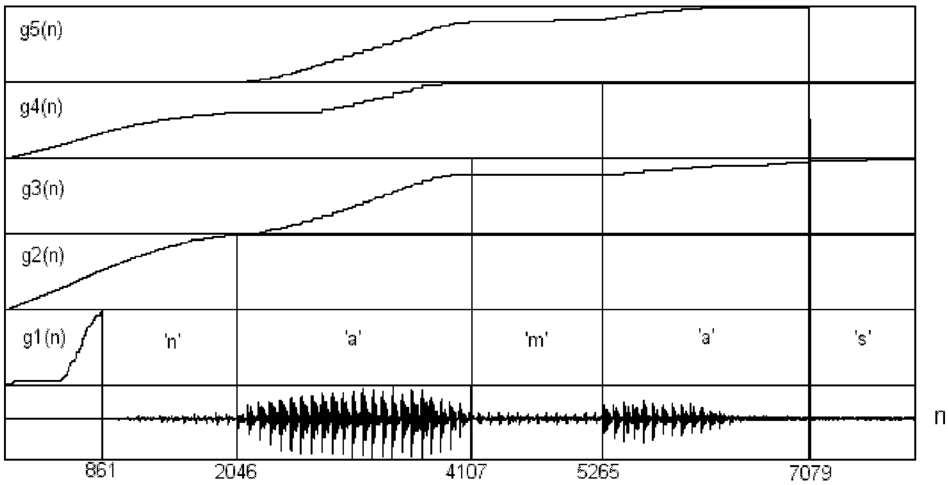

Fig. 5. Segmentation of the Lithuanian word "namas", the maximum likelihood estimate.

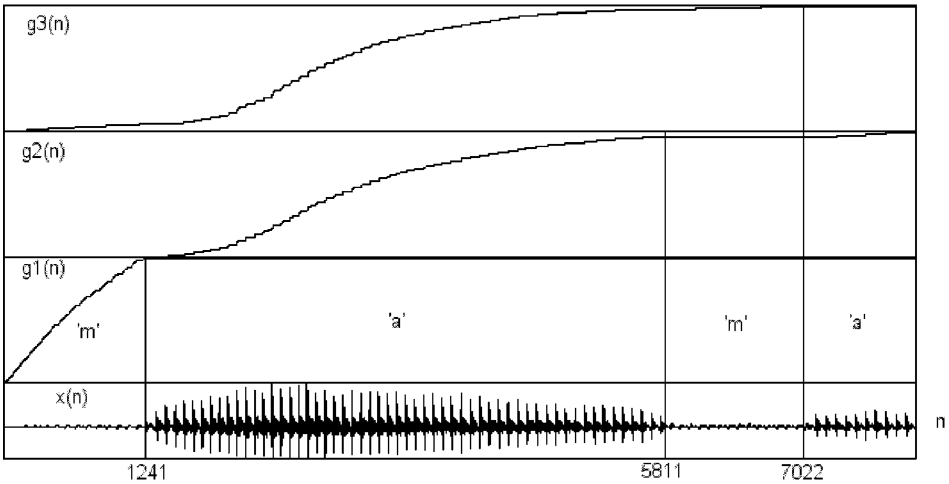

Fig. 6. Segmentation of the Lithuanian word "mama", the maximum likelihood estimate.

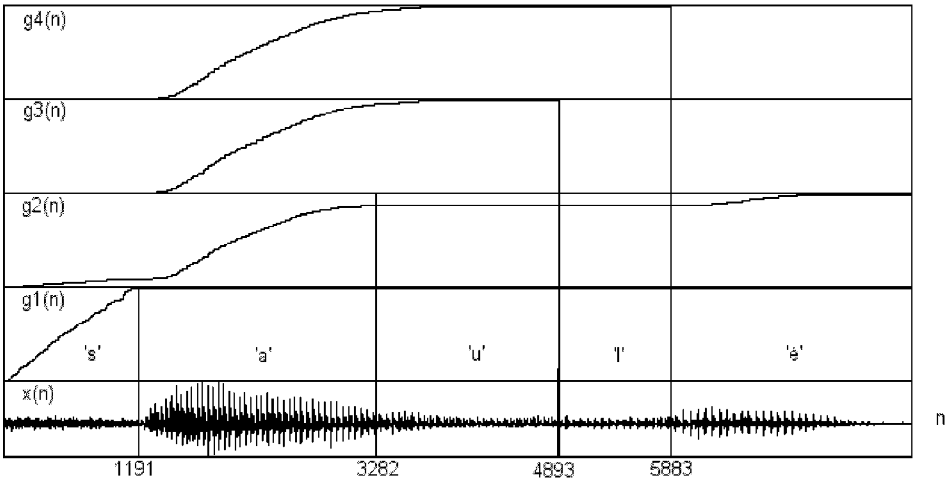

Fig. 7. Segmentation of the Lithuanian word "saulė", the least square estimate.

problem. The main reason of mistakes in change points determination in speech signals is parameter variability in different utterances of the same word even pronounced by the same speaker and time varying nature of the phonemes itself. These problems we leave for further investigation.

9. Concluding Remarks

The problem of maximum likelihood and least square segmentation of autoregressive random sequences with abruptly changing parameters was investigated. Dynamic programming method was applied for maximization of objective functions. Simulation examples and speech signal segmentation examples illustrate performance of the proposed approach.

Future research should be concentrated on application of this approach to "real world" signal segmentation, adaptation to intra-speaker variability in speech and text-dependent speaker recognition.

References

Bhattacharya G.K., R.A. Johnson (1968). Non-parametric test for shift at an unknown time point, *Ann. Math. Stat.* , **39,** 1731–1743.

Cooper L., M. Cooper (1981). *Introduction to Dynamic Programming*, Pergamon Press.

Duda R.O., P.E. Hart (1973). *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York.

Kander Z., S. Zacks (1965). Test procedures for possible changes in parameters of statistical distributions occurring at unknown time points. *Ann. Math. Stat.*, **37,** 999–1018.

Kligienė N. (1974). On the estimation of the change point in the autoregressive sequence. In *Proc. of the 7th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, Prague.

Lipeika A. (1975). Estimation of the accuracy of determination of the time moment of change in the properties of random series. *Statistical Problems of Control,* **7**, 76–87 (in Russian).

Lipeika A. (1975). Determination of changes in the properties of autoregressive sequences. A two dimensional case. *Statistical Problems of Control,* **12**, 27–41 (in Russian).

Lipeika A. (1977). Determination of changes in the properties of an autoregressive sequence. *Statistical Problems of Control,* **24**, 27–71 (in Russian).

Lipeika A. (1979). On the determination of changes in the properties of an autoregressive sequence. *Statistical Problems of Control,* **39**, 9–24 (in Russian).

Lipeika A., J. Lipeikienė (1987). Application of dynamic programming method for change detection in the properties of random autoregressive sequences, *Statistical Problems of Control,* **80**, 9–23 (in Russian).

Lipeika A., J. Lipeikienė (1992). Estimation of many change-points in long autoregressive sequences. *Informatica,* **3**(1), 37–46.

Lipeikienė J. (1973). Determination of the change in the properties of the random sequence described by the autoregressive-moving average model. *Statistical Problems of Control,* **7**, 65–75 (in Russian).

Page E.S. (1958). A Test for a change in parameter occurring at an unknown point. *Biometrika*, **44,** 248–252.

Proakis J.G., D.G. Manolakis (1996). *Digital Signal Processing: Principles, Algorithms, and Applications.* Prentice Hall.

Rabiner L., B.-H. Juang (1993). *Fundamentals of Speech Recognition.* Prentice Hall.

Telksnys L. (1969). About application of optimal Bayes training algorithm to determination change points in the properties of random signals, *Avtomatika i Telemechanika*, **6** (in Russian).

Telksnys L. (1972). Determination of changes in the properties and recognition of random processes with a complicated structure. In *Proc. of the 5th IFAC World Congress*, Paris 12–17 June.

Telksnys L. (1973). Determination of changes in the properties of random processes by the Bayes method. In *Proc. of the 3rd IFAC Symp.,* Netherlands.

A. Lipeika is a Doctor of Technical Sciences, a senior researcher of the Recognition Processes Department at the Institute of Mathematics and Informatics and an Associate Professor at the Radioelectronics Department of Vilnius Technical University. Scientific interests include processing and recognition of random processes, detection of changes in the properties of random processes, signal processing, speech processing and speaker recognition.

Atsitiktinių sekų optimali segmentacija

Antanas LIPEIKA

Darbe yra nagrinėjama autoregresinių atsitiktinių sekų su šuoliais besikeičiančiais parametrais segmentacija naudojant maksimalaus tikėtinumo ir vidutinės kvadratinės klaidos kriterijus. Gauta stebėjimų sąlyginio pasiskirstymo išraiška. Tikslo funkcija yra modifikuota taip, kad optimizacijai galima būtų panaudoti dinaminio programavimo metodą. Šiai tikslo funkcijai gautos patogios taikymui Belmano funkcijų išraiškos. Metodo darbingumas yra iliustruojamas modeliavimo ir realių kalbos signalų segmentavimo pavyzdžiais.