GENERALIZED EDGE-PRESERVING SMOOTHING FOR SIGNAL ANALYSIS

Vadim Mottl, Alex Kostin

Tula State University
Lenin Ave. 92, 300600 Tula, Russia
mottl@atm.tsu.tula.ru

Ilya Muchnik

DIMACS, Rutgers University
P.O. Box 1179, Piscataway, NJ 08855, USA
muchnik@dimacs.rutgers.edu

ABSTRACT

Under the generalized smoothing of a signal, a wider class of operations is understood than the plain suppression of noise. We apply this term to all the processing problems which can be interpreted as those of transformation of the original signal, considered as functions of one discrete argument, into a secondary function of another nature on the same carrier, by way of coordinating the local signal-dependent information and a piori smoothness constraints. In this work, a statistical approach to the generalized edge-preserving smoothing is considered on the basis of treating the sought-for result of processing as a realization of a Markov random process, whose Markov continuity is locally broken at the assumed jump-points. The principal idea of the approach consists in finding the break-points one by one and incorporating them into the model of the hidden process as they are found, so that, at each step, the most detectable of not yet legitimated peculiarities is sought for.

1. INTRODUCTION

The problem of edge-preserving smoothing is typical for practice of signal analysis. It constantly arises when there is reason to believe that significant discontinuities of the hidden original signal are disguised by noise in the data array accessible to immediate processing. An edge-preserving smoothing algorithm is inevitably nonlinear, even if the basic smoothing procedure is linear. Many nonlinear smoothing techniques are known to date, among which the leading place occupy, perhaps, median filters [1], and new techniques based on competitive Kalman smoothing [2].

Despite the great number of edge-preserving smoothing techniques, all of them are underlaid by the tacit assumption that any two neighboring edges subject to detection are spaced at a distance not smaller

than the minimal effective width of the window sufficient for achieving the desired smoothing degree. If this assumption is broken, the smoothing algorithm will take two adjacent jumps in the signal for an entire noise pattern and try to suppress them.

However, it is often required to distinguish between closely-spaced peculiarities of the original signal, including salient short pulses, without loss in smoothing degree around them. It is just this problem which roused the authors to seek for alternative ways of preserving discontinuities in signals to be smoothed. The principal idea of the approach being discussed here consists in finding break-points one by one and incorporating them into the model of the hidden process, as they are found, so that at each step the most detectable of not yet legitimated peculiarities is sought for.

In many practical problems of signal analysis, it is expedient to treat the result of a processing of the original signal $Y = (y_t, t \in T)$, $T = \{1, \ldots, N\}$, more widely than the result of smoothing proper, namely, to interpret it as a succession of estimates $\hat{\mathbf{x}}_t$ on the instantaneous vector parameter $\mathbf{x}_t \in \mathsf{R}^n$ of an appropriate local signal model of general kind. In this case, the entire totality of results should be considered as series of mutually coordinated estimates $\hat{\mathbf{X}} = (\hat{\mathbf{x}}_t, \ t \in T)$, smoothed in a general sense with preserving relatively rare abrupt jumps.

Just like in [3], we use here the statistical approach to the generalized smoothing with the outlook of extending it onto the case of images. To be exact, we use the version based on the notion of a Markov model of the hidden signal. It is, in particular, just this approach which underlies the very idea of Kalman filtration and smoothing, therefore, the algorithms described in [2] are immediate progenitors of more general procedures considered in this work.

In Section 2, we start with the presentation of a basic procedure of generalized edge-preserving smoothing. In Sections 3 and 4, we consider two particular versions of the basic procedure designed, respectively, for smoothing proper with preserving eventual breaks

as in the signal itself as well as in its first difference, and for autoregression and time-frequency analysis of signals with abruptly changing spectral properties.

2. GENERALIZED BASIC PROCEDURE

Just as it is widely adopted, we consider the succession to be recovered $\mathbf{X} = (\mathbf{x}_t,\ t \in T)$, $\mathbf{x}_t = (x_t^1, \ldots, x_t^n)^T \in \mathsf{R}^n$, and the original signal $Y = (y_t,\ t \in T)$ as realizations of, respectively, the hidden $\mathcal{X}$ and the observable $\mathcal{Y}$ component of a two-component random process $(\mathcal{X}, \mathcal{Y})$. It is required to estimate the realization of the hidden process from the observed realization $\hat{\mathbf{X}}(Y)$.

The marginal probabilistic properties of the process $\mathcal{X}$ are assumed to be expressed in Markov terms as conditional normal probability densities

$$q_t^m(\mathbf{x}|\mathbf{x}_{t-1},\mathbf{v}_t) = \mathcal{N}(\mathbf{x}|\mathbf{A}\mathbf{x}_{t-1}, \mathbf{V}_t),\ \ \mathbf{V}_t = \mathrm{Diag}(\mathbf{v}_t),\ \ (1)$$

with the upper index "m" meant as *model*-based. The matrix coefficient $\mathbf{A}$ in the conditional mathematical expectation $\mathbf{A}\mathbf{x}_{t-1}$ and the constant diagonal covariance matrix $\mathbf{V}_t = \mathbf{V} = \mathrm{Diag}(\mathbf{v})$ with sufficiently small variances $\mathbf{v}_t = \mathbf{v} = (v^1, \ldots, v^n)^T = \mathrm{const}$ express the a priori smoothness constraints which are valid beyond the isolated jump-points: the less v^i is, the more smooth the component x_t^i is expected to be. At the points of the assumed jumps, some of the diagonal variances in $\mathbf{V}_t = \mathrm{Diag}(\mathbf{v}_t)$ get infinite for an instant $v_t^i \to \infty$, and cut off, in the respective direction in R^n, the future of the hidden process $(\mathbf{x}_s,\ s \geq t)$ from its past $(\mathbf{x}_s,\ s < t)$.

Let the function $\mathbf{W} = (\mathbf{v}_t,\ t \in T)$, $\mathbf{v}_t = (v_t^1, \ldots, v_t^n)^T$, where at each time moment $v_t^i = v^i \geq 0$ or $v_t^i = \infty$, be called edge function.

As to the conditional probabilistic properties of the observable process $(\mathcal{Y}|\mathcal{X})$, there is no need to preset them in any explicit form. It is sufficient to assume that there is a way, maybe, a heuristic one, to roughly judge about each of the single hidden values $\mathbf{x}_t$ from a fragment Y_t of the original signal around point t, in particular, from its single value y_t. Such local judgments are assumed to be expressed in the form of an instantaneous signal-dependent normal posterior distributions

$$p_t^o(\mathbf{x}|Y_t) = \mathcal{N}(\mathbf{x}|\hat{\mathbf{x}}_t^o, \mathbf{R}_t^o),\qquad (2)$$

where the *observation*-based a posteriori mathematical expectations $\hat{\mathbf{x}}_t^o = \hat{\mathbf{x}}_t^o(Y_t)$ and covariance matrices $\mathbf{R}_t^o = \mathbf{R}_t^o(Y_t)$ are interpreted, respectively, as initial local estimates and their incredibility measures.

Thus, if we knew the edge function $\mathbf{W} = (\mathbf{v}_t,\ t \in T)$, i.e. the locations and directions of jumps, the problem of edge-preserving smoothing would reduce itself to that of finding, at each time moment t, an estimate of the Markov hidden process, for instance, the posterior mathematical expectation $\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_t(Y, \mathbf{W}) = \mathrm{M}(\mathbf{x}_t|Y, \mathbf{W})$ with respect to the entire signal record Y. Under the normality assumptions (1) and (2), the result of such a generalized smoothing $\hat{\mathbf{X}} = (\hat{\mathbf{x}}_t,\ t \in T)$ will be result of a linear operation on the succession of the initial local estimates $\hat{\mathbf{X}}^o = (\hat{\mathbf{x}}_t^o,\ t \in T)$ performed by the usual Kalman filtration-interpolation procedure [4].

The final result of interpolation at each point t of the discrete time interval $T = \{1, \ldots, N\}$ can be shown to be representable as linear combination

$$\begin{aligned}
\hat{\mathbf{x}}_t &= \mathbf{R}_t\Big[(\mathbf{A}\mathbf{R}_{t-1}^l\mathbf{A}^T + \mathbf{V}_t)^{-1}\mathbf{A}\hat{\mathbf{x}}_{t-1}^l + (\mathbf{R}_t^r)^{-1}\hat{\mathbf{x}}_t^r\Big],\\
\mathbf{R}_t &= \Big[(\mathbf{A}\mathbf{R}_{t-1}^l\mathbf{A}^T + \mathbf{V}_t)^{-1} + (\mathbf{R}_t^r)^{-1}\Big]^{-1},
\end{aligned} \qquad (3)$$

where

$$\begin{aligned}
\hat{\mathbf{x}}_{t-1}^l &= \mathrm{M}(\mathbf{x}_{t-1}|y_1, \ldots, y_{t-1}, \mathbf{W}),\\
\mathbf{R}_{t-1}^l &= \mathrm{Cov}(\mathbf{x}_{t-1}|y_1, \ldots, y_{t-1}, \mathbf{W}),
\end{aligned} \qquad (4)$$

and

$$\begin{aligned}
\hat{\mathbf{x}}_t^r &= \mathrm{M}(\mathbf{x}_t|y_t, \ldots, y_N, \mathbf{W}),\\
\mathbf{R}_t^l &= \mathrm{Cov}(\mathbf{x}_t|y_t, \ldots, y_N, \mathbf{W})
\end{aligned} \qquad (5)$$

are posterior mathematical expectations and covariance matrices of the hidden process values found at two neighbouring points $t - 1$ and t by two independent Kalman filters running, respectively, from the left $t = 1, \ldots, N$ and from the right $t = N, \ldots, 1$. At the moments of jumps already legitimated in $\mathbf{W}$, the infinite diagonal elements of $\mathbf{V}_t$ prevent, at least, partially, the averaging of the respective estimates in (3).

It should be borne in mind that the hidden model (1) must be inverted in time before using it in the backward Kalman filter (5): $q_t^m(\mathbf{x}|\mathbf{x}_{t+1},\mathbf{v}_{t+1}) = \mathcal{N}(\mathbf{x}|\mathbf{A}^{-1}\mathbf{x}_{t+1}, \mathbf{V}_{t+1})$.

But the positions and directions of jumps are unknown in the general case and subject to an estimation along with the realization of the hidden process. In this work, a multistage procedure of estimating the edge function is proposed on the basis of the idea to seek, at each step of signal processing, for only one jump in the hidden signal. Let the edge function $\mathbf{W} = (\mathbf{v}_t,\ t \in T)$ be known up to the only last jump which still remains to be found. Let us, sequentially for all the points of the discrete time interval $T = \{1, \ldots, N\}$, test the hypothesis that the current point t is just the sough-for remaining jump-point, i.e. that the difference $\hat{\mathbf{u}}_t = \hat{\mathbf{x}}_t^r - \mathbf{A}\hat{\mathbf{x}}_{t-1}^l$ between the left-side $\mathbf{A}\hat{\mathbf{x}}_{t-1}^l$ and right-side $\hat{\mathbf{x}}_t^r$ estimates on the same hidden random value $\mathbf{x}_t$ is too large to be referred to the legitimate stochasticity of the hidden model in accordance with $\mathbf{W}$.

If we assume that no a priori information on the random difference $\mathbf{u}_t = \mathbf{x}_t - \mathbf{A}\mathbf{x}_{t-1}$ is available, its a

posteriori probability distribution will be normal with mathematical expectation $\hat{\mathbf{u}}_t = \mathbf{x}_t^r - \mathbf{A}\mathbf{x}_{t-1}^l$ and covariance matrix $\mathbf{Q}_t = \mathbf{A}\mathbf{R}_{t-1}^l\mathbf{A}^T + \mathbf{V}_t + \mathbf{R}_t^r$ defined by (4) and (5). On the other hand, such an a priori information exists and is carried jointly by the hidden model (1) and the jump function $\mathbf{W}$, which prescribe the difference $\mathbf{u}_t$ to be a random vector with zero mathematical expectation and covariance matrix $\mathbf{V}_t$.

The ratios $z_t^i = (\hat{u}_t^i)^2/q_t^{ii}$ of the squared elements of $\hat{\mathbf{u}}_t$ to their variances in $\mathbf{Q}_t$ appear to be appropriate measures of the local "strain" in the respective components $\hat{x}_t^i$ of the estimated hidden signal when the model is attempting to adjust them without jump at the current point. The most "stressed" point, if the maximal ratio exceeds a preset threshold,

$$t^* = \arg\max_{1 \le t \le N} z_t, \ z_t = \max_{i=1,\ldots,n} z_t^i, \ z_{t^*} > h, \quad (6)$$

is to be pronounced the new jump-point. Mark, that this may happen only at a time moment t where no jump had been found at the previous stages of processing, and, so, the covariance matrix $\mathbf{V}_t$ in $\mathbf{W}$ is still matrix $\mathbf{V} = \text{Diag}(\mathbf{v})$ with small diagonal elements $\mathbf{v} = (v^1, \ldots, v^n)^T$ which preset the normal smoothness degree of the hidden process. This matrix is to be replaced by a matrix $\mathbf{V}_{t^*}$ with one or several infinite diagonal elements to mark the components of $\mathbf{x}_t$ which must be broken. The choice of such components depends on the particular hidden model.

The procedure starts with the initial edge function $\mathbf{W} = \left(\mathbf{V}_t = \mathbf{V} = \text{Diag}(\mathbf{v}), \ t \in T\right)$ which presupposes no jumps in the signal, and is to be repeated again and again, each time with the renewed $\mathbf{W}$, while the most stressed point of T satisfies the condition (6).

In contrast to [2], where at each point only one of two independent estimates (4) and (5) is taken as the result of smoothing, the approach described here should be called rather collaborative than competitive, because at the most part of points, where no jumps are detected, both of the filtration results participate in forming the final decision on the hidden signal value. As a result, the presented approach provides more pronounced edge-preserving effect, but the necessity to detect jumps in an explicit form is the inevitable payment for such an improvement. Since the plain comparison of two estimates at each point provides for an accurate finding of only one eventual jump in the hidden process, we are forced to use a multistage procedure.

Fortunately, there is an effective way to significantly accelerate the procedure without any appreciable loss in the accuracy of estimating the locations of jumps. This way consists in finding, at each stage of processing, several sufficiently distant local maxima of the estimated strain z_t at once, instead of only one point

t^* of the absolutely maximum strain (6). As the new legitimate jumps, all the points t^* are pronounced at which z_t achieves its absolute maximum simultaneously in two adjacent intervals of a preset length Δ at the left and at the right of the current point:

$$t^* = \arg\max_{(t^*-\Delta)\le t\le(t^*+\Delta)} z_t, \ z_{t^*} > h. \quad (7)$$

As it will be illustrated in the next Section, such a procedure is able to find jump-points located closer to each other than Δ, but it detects them sequentially, at different stages of signal processing.

3. EDGE-PRESERVING SUPPRESSION OF NOISE

In this Section, we consider the original signal y_t as sum $y_t = x_t + \eta_t$ of a hidden signal x_t to be recovered and additive white noise η_t with zero mean and a supposed variance w. The aim of processing is suppression of the noise with respect to the assumed smoothness of the hidden signal almost everywhere except some isolated points where jump-like discontinuities may occur in the signal itself x_t and (or) in its differences $\nabla^j x_t$ up to the preset highest order $j = 1, \ldots, k$.

We keep here to the approach depicted in [2], where the supposed smoothness of the hidden signal is expressed by the assumption that all the differences of the hidden signal equal zero $\nabla^j x_t = 0$, $j = 1, \ldots, k-1$, except that of the highest order k which is dealt with as white noise $\nabla^k x_t = \xi_t$ with variance v taken essentially smaller than that of the noise in observation $v \ll w$.

In accordance with such an approach, it is natural to take the hidden process in the generalized model (1) as vector $\mathbf{x}_t = (x_t^1, \ldots, x_t^n)^T \in \mathsf{R}^n$, $n = k + 1$, whose elements are the hidden signal itself $x_t^1 = x_t$ and its differences $x_t^i = \nabla^{i-1}x_t$, $i = 2, \ldots, n$, so that $x_t^1 = x_{t-1}^1 + x_{t-1}^2, \ldots, x_t^{n-1} = x_{t-1}^{n-1} + x_{t-1}^n$, $x_t^n = x_{t-1}^n + \xi_t$.

In this work, we restrict our consideration, for the sake of simplicity only, to the first-order model of the hidden signal. It means that the hidden process is assumed to take values from the two-dimensional space formed by the signal itself and its first difference $\mathbf{x}_t = (x_t, \nabla x_t)^T \in \mathsf{R}^2$. In this case, matrices $\mathbf{A}$ and $\mathbf{V}_t$ in (1) will have the form

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \mathbf{V}_t = \begin{pmatrix} v_t^1 & 0 \\ 0 & v_t^2 \end{pmatrix} = \text{Diag}(\mathbf{v}_t), \quad (8)$$

where parameter values $v_t^1 = 0$ and $v_t^2 = v$ express the smoothness assumption on the hidden signal $x_t = x_t^1$ beyond the jumps. As to the jump-points, the first-order model allows for simulating as jumps in the signal itself (ruptures) by instantaneous assignment $v_t^1 = \infty$
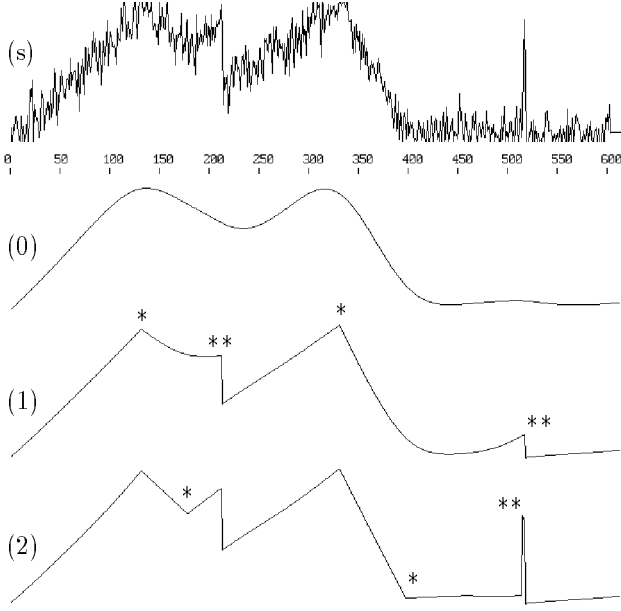
Figure 1: Original simulated signal (s), result of its continuous smoothing (0), and two steps (1) and (2) of finding essential discontinuities: fractures * and ruptures **.

and $v_t^2 = \infty$, as well as those in the signal's first difference (fractures) by assignment $v_t^1 = 0$ and $v_t^2 = \infty$.

There is no need to take the signal-dependent posterior mathematical expectations in the observation model (2) in a more sophisticated form than

$$\hat{\mathbf{x}}_t^o = \left[y_t, \frac{1}{2}(y_{t-1} - y_{t+1}) \right]^T, \qquad (9)$$

because the current values of the observed signal and its first difference are natural unbiased local estimates on the respective constituents of the hidden process.

As to the posterior covariance matrices, there is no reason, at least, in the case of elementary smoothing, to place different reliance on the local estimates $\hat{\mathbf{x}}_t^o$ at different time moments t, so, all the $\mathbf{R}_t^o$ should be taken equal to each other. In addition, in the particular case under consideration, it is sufficient to take them diagonal with identical posterior variances $\mathbf{R}_t^o = \text{Diag}(w, w)$, because values of $y_t = x_t + \eta_t$ at different time moments and, hence, different elements of $\hat{\mathbf{x}}_t^o$ (9) are formed by independent identically distributed samples of the white noise. Finally, the resulting degree of smoothing is determined by ratio v/w, and, so, we can put $w = 1$ without any loss in generality. Thus, the unitary matrix will be a good choice for all the posterior covariance matrices in the observation model $\mathbf{R}_t^o = \mathbf{I}$.

Fig. 1 shows an example of stepwise finding jumps in the hidden process in accordance with the rule (7) where $\Delta = 100$. Only two steps turned out to be required for finding all the discernible discontinuities in the original signal.

In such a procedure, the smoothed curve manifests new and new discontinuities, originally disguised by noise, like an overloaded elastic rod of limited strength gets fractured or ruptured, by stages, every time at the point of maximal local stress, as precedent breaks relieve former stress concentrations and create thereby new ones.

4. EDGE-PRESERVING TIME-FREQUENCY ANALYSIS

Within the bounds of the correlation theory, any stationary and ergodic random process y_t on the infinite discrete time axis is completely representable by its spectral density $S(f) \geq 0$, which is the distribution of the full variance $D = \mathsf{M}\left\{[y_t - \mathsf{M}(y_t)]^2\right\}$ over the Nyquist frequency interval $0 \leq f < 1/2$, so that

$$\int_{f_1}^{f_2} S(f) df = D(f_1, f_2) \qquad (10)$$

is the variance of the result of ideal band-pass filtration of y_t in the frequency band (f_1, f_2) and $\text{D}=\text{D}(0,1/2)$.

If a stationary random process y_t is represented by its parametric autoregression model $y_t = \mathbf{a}^T \mathbf{y}_{t-1} + \xi_t$, where $\mathbf{y}_{t-1} = (y_{t-1}, \ldots, y_{t-n})^T$, $\mathbf{a} \in \mathsf{R}^n$, and ξ_t is white noise with a constant variance, the spectral density is completely determined by autoregression coefficient vector $\mathbf{a}$ and the variance of the process D, i.e. $S(f) = S(f; \mathbf{a}, D)$. The relevant formulas are to be found in [5].

To bring such a parametric approach in more close accordance with the needs of practical signal analysis, it appears natural to consider the model parameters as variables and, in doing so, extend the autoregression model onto the case of nonstationary random processes $y_t = \mathbf{a}_t^T \mathbf{y}_{t-1} + \xi_t$, where the variance of the white noise is no longer assumed to remain constant in time, either. Such a model presupposes the existence of a hidden vector process formed by the instantaneous parameter values $\mathbf{x}_t = (\mathbf{a}_t, D_t) \in \mathsf{R}^{n+1}$ whose assumed smoothness can be naturally expressed in terms of a hidden Markov random process (1) with eventual abrupt discontinuities described by an appropriate jump function.

From such a standpoint, the processing of a signal $Y = (y_t, t \in T)$ boils down to estimation of the realization of the hidden process $\mathbf{X} = (\mathbf{x}_t, t \in T)$ along with the jump function $\mathbf{W}$ in full accordance with the generalized technique described in Section 2, and to

recalculation of the obtained estimates into the succession of the respective instantaneous variance spectra $\hat{S}_t(f) = S(f; \hat{\mathbf{x}}_t)$, $\hat{\mathbf{x}}_t = (\hat{\mathbf{a}}_t, \hat{D}_t)$. It remains only to specify the matrix parameters $\mathbf{A}$ and $\mathbf{V}_t$ of the hidden model (1) and to choose the observation model (2).

There is no reason to presuppose any a priori interdependence between the elements of the hidden vector process $\mathbf{x}_t = (x_t^1, \ldots, x_t^{n+1})^T$, therefore, the matrix in the conditional mathematical expectation in (1) should be taken unitary $\mathbf{A} = \mathbf{I} \, [(n+1) \times (n+1)]$.

Likewise, there is no reason to assume different smoothness properties of the elements of the autoregression parameter vector $\mathbf{a}_t = (a_t^1, \ldots, x_t^n)^T$, whereas the smoothness degree of the process variance D_t should be set individually. So, it is sufficient to determine the diagonal covariance matrix $\mathbf{V} = \mathrm{Diag}(v^1, \ldots, v^{n+1})$ of smoothness constraints, valid within the continuity intervals, by two parameters $v^i = v^a$, $i = 1, \ldots, n$, and $v^{n+1} = v^D$, responsible for the smoothness of, respectively, $\mathbf{a}_t$ and D_t. There are no grounds, either, to consider more than one kind of breaks, so, it is sufficient to mark each break-point by infinite values of all the elements of the covariance matrix $\mathbf{V}_t = \mathrm{Diag}(\infty, \ldots, \infty)$.

If we assume the current signal fragment $Y_t = (y_t, \mathbf{y}_{t-1})$ to be the only source of information on the instantaneous hidden value $\mathbf{a}_t$, the posterior mathematical expectation and inverse posterior covariance matrix of the autoregression coefficient vector can be shown, respectively, to equal $\hat{\mathbf{a}}_t^o = (y_t/\mathbf{y}_{t-1}^T \mathbf{y}_{t-1})\mathbf{y}_{t-1}$ and to be proportional to $\mathbf{B}_t^o = \mathbf{y}_{t-1}\mathbf{y}_{t-1}^T$. Similarly, if the current value y_t is considered as the only source of information on D_t, the squared value y_t^2 will be just the posterior mathematical expectation of the hidden random variance of the observed process at this time moment, and as to the posterior variance, it can be treated as "proportional to unity". We needn't take care of the proportionality coefficients, because they affect only the smoothness degree and always can be compensated by an appropriate choice of smoothness parameters v^a and v^D. We needn't pay attention, either, to the degeneracy of the covariance matrices for all t, because the Kalman filter in the basic procedure refers only to their inversions when evaluating (4) and (5). So are the reasons for taking the signal-dependent posterior mathematical expectation $\hat{\mathbf{x}}_t^o \in \mathsf{R}^{n+1}$ and covariance matrix $\mathbf{R}_t^o \, [(n+1) \times (n+1)]$ in (2) in the form, respectively, $\hat{\mathbf{x}}_t^o = (\hat{\mathbf{a}}_t^o, y_t^2)^T$ and $(\mathbf{R}_t^o)^{-1} = \begin{pmatrix} \mathbf{B}_t^o & 0 \\ 0 & 1 \end{pmatrix}$.

In Fig. 2, the results of time-frequency analysis of a simulated signal with changing spectral properties are shown in an abridged form as diagrams of partial signal variances estimated in three adjacent frequency bands (low, medium and high) $\hat{D}_t(0, 1/6)$, $\hat{D}_t(1/6, 1/3)$ and
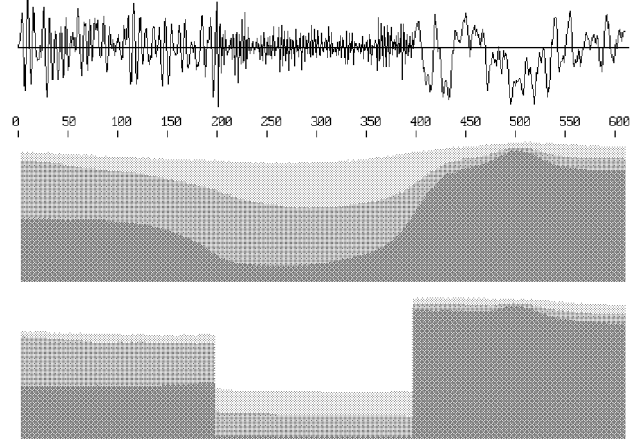

Figure 2: Results of continuous and edge-preserving time-frequency analysis of a simulated signal: diagrams of current estimates on the instantaneous partial variances in the low (dark-gray), medium (gray) and high (light-gray) frequency bands as the summands of an estimate on the full current variance of the signal.

$\hat{D}_t(1/3, 1/2)$ (10) making the Nyquist interval. The spectral densities $\hat{S}_t(f) = S(f; \hat{\mathbf{a}}_t, \hat{D}_t)$ were calculated through the autoregression of the second order $n = 2$.

Both jumps in the spectral properties were found on the first run by the rule (7) with $\Delta = 100$.

5. REFERENCES

[1] E.J. Coyle, J.H. Lin, and M. Gabbouj. Optimal stack filtering and the estimation and structural approaches to image processing. *IEEE Trans. on Acoust., Speech, Signal Processing*, Vol. 37, 1989, pp. 2037-2066.

[2] M. Niedzwiecki, and W.A. Sethares. Smoothing of discontinuous signals: The competitive approach. *IEEE Trans. on Signal Processing*, Vol. 43, No. 1, January 1995, pp. 1-13.

[3] V.V. Mottl, I.B. Muchnik, A.B. Blinov, and A.V. Kopylov. Hidden tree-like quasi-Markov model and generalized technique for a class of image processing problems. *13th International Conference on Pattern Recognition*. Vienna, Austria, August 25-29, 1996. Track B, pp. 715-719.

[4] A.P. Sage, and J.T. Melse. *Estimation Theory and Application to Communication and Control*. McGraw Hill, N.Y., 1972.

[5] G.E.P. Box, and G.M. Jenkins. *Time Series Analysis. Forecasting and Control*. Holden-Day, 1970.