International Journal of
Electronic Commerce Studies
Vol.3, No.1 , pp.1-20, 2012

ON-LINE VIDEO SEGMENTATION USING METHODS OF FAULT DETECTION IN MULTIDIMENSIONAL TIME SEQUENCES

Yevgeniy Bodyanskiy
Kharkiv National University of Radio Electronics
pr. Lenina, 14, Kharkiv, Ukraine, 61166
bodya@kture.kharkov.ua

Dmitriy Kinoshenko
Kharkiv National University of Radio Electronics
pr. Lenina, 14, Kharkiv, Ukraine, 61166
kinoshenko@kture.kharkov.ua

Sergii Mashtalir
Kharkiv National University of Radio Electronics
pr. Lenina, 14, Kharkiv, Ukraine, 61166
mashtalir_s@kture.kharkov.ua

Olena Mikhnova
Kharkiv National University of Radio Electronics
pr. Lenina, 14, Kharkiv, Ukraine, 61166
elena_mikhnova@ukr.net

ABSTRACT

Recently, video segmentation with time series has been recognized as an important research topic. Despite great progress in this field, existing approaches have some drawbacks. We first give an overview of existing techniques and approaches, and then we analyze the applicability of the recursive least square method, multidimensional modification of exponentially weighted stochastic approximation algorithm, methods of Kaczmarz, Shown, Brown, Chow, Trigg-Leach, Roberts-Reed, finite and infinite memory algorithms for detection of faults in multidimensional time sequences. At the end we have come to the conclusion that the Trigg-Leach method is preferable for fault detection in video time sequences. Model efficiency has been checked on video containing endoscopic operation.

Keywords: Video Segmentation, Multidimensional Time Sequence, Vector Autoregression Model

1. INTRODUCTION

Nowadays, with the rapid increase of computational capabilities and storage capacities, video plays an important role in lives of every connected human being – it ranges from common desktop tasks (such as watching TV-programs, searching the web for particular video information, saving clips of streaming video while working with Skype, etc.) to complicated enterprise functions (including manufacturing, health care, industrial process control, etc.). Billions of applications require video editing and analysis, search and indexing, segmentation and recognition, summarization and skimming. Not all these tasks have been completely solved yet because of the wide variety of characteristics to be considered. If, for instance, computation speed satisfies, quality of processing leaves much to be desired, and vice versa. These issues are partially covered in Huawei Technologies report [1].

It is also very hard to consider all the video features at the same time, such as: camera motion and activity of captured objects, variety of visual features (color, texture, brightness, intensity, shape, etc.), audio and subtitles, and the semantics of video content. Moreover, it is almost impossible to develop a universal algorithm that would satisfy even a few application domains at once (in case of video recognition, by application domain we mean not only the theme of the video, but also type of camera it was shot with, length and quality, as all of these aspects have a strong impact on processing results).

For effective video processing we usually need to partition the given video into meaningful units (frames, scenes and episodes), and then to segment each unit for monitoring of changes (or faults) in their characteristics[2, 3, 4]. That is why one of the most challenging issues is to find correlation between different parts of video content. An intuitive way lies in modeling video sequence as time series, because video frames follow each other in time. A general definition of time series applied to motion analysis can be found in an unpublished technical report written by D. Bouchard in 2006[5].

Lately, scientists have begun to pay greater attention to video modeling with time series. During the past few years many papers have been written about this issue, though it is considered quite a new area of research. M.

Hoai et al.[6] proposed to classify time series approaches used for video recognition into 2 categories: change point detection and cyclic event periodicity detection. But the present group of researchers prefers to use classification and dynamic programming. As for cyclic event periodicity, in 2000 there emerged a paper concerning exactly this topic. Its authors[7] tried to perform periodic motion analysis in time by estimating an object's self-similarity and by computing recurrence matrices that were then used as a qualitative mechanism for time series analysis. Their work directly concern time-frequency analysis and short-time Fourier transformation, though it deals only with repetitive events (for instance, dogs running can be treated as periodic motion because of their symmetrically moving legs).

Concerning the first approach, mentioned by M. Hoai et al., change point detection was implemented by X. Xuan et. al[8]. They applied the Gaussian graphical model, Bayesian estimation and Fearnhead’s dynamic programming algorithm to multivariate time series. In 2008-2009 Z. Harchaoui et. al. proposed change-point analysis via sliding windows that run across the video sequence to find change-points and they used a test statistic based on the maximum kernel Fisher discriminant ratio as a homogeneity measure between segments[9]. The next year Z. Harchaoui et. al. perfected Least Absolute Shrinkage eStimatOr (LASSO) in least-squares regression introduced for variable selection by Tibshirani. They also provided deep analysis of related research works[10].

In 2011 there emerged many works connected with time series for video segmentation[11, 12]. The main procedures are often implemented with the help of dynamic programming. In the next section we also present an approach concerning video recognition with time series. In our case we use a *fault* term by which we mean changes of some means, variance parameters (like variance), i.e. internal structure of multidimensional time series. The proposed technique copes with the problem of unknown sample size, which is very important in video recognition. Moreover, medical data, that we are targeted for processing, require high computational quality we are trying to achieve due to combining and analyzing several segmentation and modeling techniques.

This article is organized as follows: Section 2 describes current time sequence models and different approaches to parameter estimation, Section 3 depicts our proposed adaptive model used for fault detection in multidimensional time sequences, in Section 4 the application of our model to medical video is presented, and the conclusions are given in Section 5.

2. TIME SEQUENCE MODELS

Time sequence models provide mathematical description of observed processes that change over time. In this section non-stationary mathematical models of time sequences will be given, adapted to the situation where the number of observations is unknown. We assume that these models help to find significant changes of video content with the passage of time by analyzing salient video properties.

Video sequence can be presented in structured and reduced forms via econometric models[13, 14]. The structured form can be written as follows

$$\sum_{l=0}^{p} B_l x(k-l) + Dz(k) = \eta(k) \tag{1}$$

(where B_l – matrix coefficients for endogenous variables, B_0 – non-singular matrix for endogenous variables of running time, D – matrix of coefficients for exogenous variables, $z(k)$ – vector of exogenous variables with their delays, $\eta(k)$ – vector of disturbing signal with zero mathematical expectation and bounded second moments). The reduced form of video sequence can be rewritten as follows

$$x(k) = -B_0^{-1}(\sum_{l=1}^{p} B_l x(k-l) + Dz(k) - \eta(k)) \tag{2}$$

or

$$x(k) = CZ(k) + \xi(k) \tag{3}$$

where $Z(k)$ is a vector which represents exogenous variables and delays of endogenous variables

$$\xi(k) = B_0^{-1}\eta(k).$$

Parameters in (1-3) can be found by indirect, two-step or three-step least squares methods[13]. Sequential algorithms discussed in cannot efficiently operate in real time14. Time sequence $x(k)$, introduced in, can be defined with an infinite set of the following equations[15].

$$x(k) = \sum_{l=1}^{p} B_l x(k-l) + \sum_{p=1}^{q} D_p z(k-p) + F\psi(k-1) + \xi(k) \tag{4}$$

$$B(z^{-1})x(k) = D(z^{-1})z(k-1) + F\psi(k-1) + G(z^{-1})\eta(k) \tag{5}$$

where unknown video sequence coefficients are included in matrices B_l, D_p, F or matrix polynomials $B(z^{-1})$, $D(z^{-1})$, $G(z^{-1})$ of backward shift operator z^{-1}, $\psi(k)$ – function that describes $x(k)$ signal trend.

To estimate parameters in these equations 3 different approaches can be used: The Bayesian approach, the method of maximum likelihood, and the limited information method. The first two techniques are realized in batch, and the third technique is a form of recursive least squares method which enables sequential processing. Since the common recursive least squares method possesses infinite memory, it cannot be used for non-stationary video sequences that change over time. Instead, vector autoregression models are implemented to observe changes in video. The autoregression model binds signal observations $x(k)$ as shown below[16, 17]

$$x(k) = B_0 + \sum_{l=1}^{p} B_l x(k-l) + \xi(k) \tag{6}$$

where $B_0 = \{b_{0i}\} - (n \times 1)$ represents vector of means, $B_l = \{b_{l_{ij}}\}$ – $(n \times n)$ matrices of parameters, p – the order of the autoregression model.

State space description of autoregression model (which is shown below) enables one to use Kalman filter to process video signal.

$$\begin{cases} x(k) = \Pi x(k-1) + \Pi_0 + E(k), \\ \qquad y(k) = Cx(k) \end{cases} \tag{7}$$

where

$$x(k) = \begin{pmatrix} x(k) \\ x(k-1) \\ \vdots \\ x(k-p+1) \end{pmatrix}, \Pi_0 = \begin{pmatrix} B_0 \\ \vec{0} \\ \vdots \\ \vec{0} \end{pmatrix}, \Pi = \begin{pmatrix} B_1 & \cdots & B_{p-1} & B_p \\ I_n & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & I_n & 0 \end{pmatrix},$$

$$E(k) = \begin{pmatrix} \xi(k) \\ \vec{0} \\ \vdots \\ \vec{0} \end{pmatrix}, \quad C = (I_n, 0, \ldots, 0), \quad x(k) - (np \times 1)$$

– vector of states, $\Pi - (np \times np)$

– transfer matrix, $\vec{0}$ and 0
– $(n\times 1)$ zero-vector and $(n\times n)$ zero-matrix respectively.

Faults in time sequences can be seen from observation of $x_i(k), i=1,2,...,n$. This observation may lead to the next cases:

a) change of component mean values (while $l \le p$)

$$b_{0i}(k)=\begin{cases} b_{0i}, if\ k<k_a, \\ b_{0i}^a, if\ k\ge k_a \end{cases} \tag{8}$$

b) change of characteristics (like dispersion) of disturbances $\xi_i(\sigma_i^2)$ (while $l \le p$)

$$x_i(k)=\begin{cases} b_{0i}+\sum_{l=1}^{p}\sum_{j=1}^{n} b_{lij}x_j(k-l)+\xi_i(k), if\ k<k_a, \\ b_{0i}+\sum_{l=1}^{p}\sum_{j=1}^{n} b_{lij}x_j(k-l)+\xi_i^a(k), if\ k\ge k_a \end{cases} \tag{9}$$

c) change of b_{lij} that results in autocorrelation property change in video

$$x_i(k)=\begin{cases} b_{0i}+\sum_{l=1}^{p}\sum_{j=1}^{n} b_{lij}x_j(k-l)+\xi_i(k), if\ k<k_a, \\ b_{0i}+\sum_{l=1}^{p}\sum_{j=1}^{n} b_{lij}^0 x_j(k-l)+\xi_i(k), if\ k\ge k_a \end{cases} \tag{10}$$

where k_a – time of change in properties.

In the next section we introduce an adjustable model that allows the existing limitation of predefined interval of changes, as changes cannot be predicted a priori in video sequence.

3. ADJUSTABLE MODEL FOR OBSERVING FAULTS IN MULTIDIMENSIONAL TIME SEQUENCES

Faults in video occur in real time, that is why it is supposed that they are adaptively identified by the autoregression model (6) presented via time sequence $x(k)$. For simplification purposes we assume that $B=(B_0 \vdots B_1 \vdots ... \vdots B_p)$ is a composite matrix and

$X(k) = (1, x^T(k-1), ..., x^T(k-p))^T$ – a composite vector with dimension $((pn+1)\times 1)$. These assumptions allow a rewrite of (6) in the following way

$$x(k) = BX(k) + \xi(k) \tag{11}$$

where B is a matrix of video properties. We fit the following adjustable model to the video signal autoregression model

$$\hat{x}(k) = B(k-1)X(k), \tag{12}$$

assuming that matrix of parameters $B(k)$ is refined at each time interval k by minimizing the identification criterion. The criterion is presented as the difference between the estimated $\hat{x}(k)$ and experimental time sequence $x(k)$. The proposed adjustable model (12) also aids prediction. Predictability violation indicates appearance of recursive procedures which can be written as follows

$$\begin{cases} B(k) = B(k-1) + \gamma(k)e(k)X^T(k), \\ e(k) = x(k) - \hat{x}(k) = x(k) - B(k-1)X(k) \end{cases} \tag{13}$$

where $\gamma(k)$ – learning role factor of the algorithm that depends on the applied criterion, $e(k)$ – vector identification error.

In practice, algorithms with sum of squares minimum identification error criterion are used, as shown below

$$I(k) = \sum_{u=1}^{k} \beta(u)\|e(u)\|^2 = \sum_{u=1}^{k}\sum_{i=1}^{n} \beta(u)e_i(u)^2 \tag{14}$$

where $\beta(u)$ – system of weights.

Least squares method is the most popular one with equal significance of all weights, i.e.

$$I(k) = \sum_{u=1}^{k} \|e(u)\|^2. \tag{15}$$

The following algorithm corresponds to (13), (15)

$$\begin{cases} B(k) = B(k-1) + \dfrac{e(k)X^T(k)P(k-1)}{1+X^T(k)P(k-1)X(k)}, \\ P(k) = P(k-1) - \dfrac{P(k-1)X(k)X^T(k)P(k-1)}{1+X^T(k)P(k-1)X(k)}. \end{cases} \tag{16}$$

Unfortunately, the recursive least square method is not suitable for fault detection in multidimensional time sequences. An alternative for (15), (16) is given below

$$B(k) = B(k-1) + \frac{e(k)X^T(k)}{X^T(k)X(k)}. \quad (17)$$

This procedure is generated by the criterion

$$I(k) = \|e(k)\|^2.$$

It is a generalization of the Kaczmarz algorithm for the vector-matrix model (12)[18]. Despite good performance, the procedure (17) cannot distinguish signal faults or the influence of the stochastic component $\xi(k)$. That is why it is preferable to use finite memory algorithms with smoothing and tracking properties. When the predictability of model (12) diminishes and the need for algorithm memory rearrangement emerges, it indicates fault occurrence.

The exponentially weighted least squares method with the following criterion gained its popularity in the above group of algorithms

$$I(k) = \sum_{u=1}^{k} \beta^{k-u} \|e(u)\|^2,$$

its recursive setup procedure is shown below

$$\begin{cases} B(k) = B(k-1) + \dfrac{e(k)X^T(k)P(k-1)}{\beta + X^T(k)P(k-1)X(k)}, \\ P(k) = \dfrac{1}{\beta}\left(P(k-1) - \dfrac{P(k-1)X(k)X^T(k)P(k-1)}{\beta + X^T(k)P(k-1)X(k)}\right) \end{cases} \quad (18)$$

where $0 < \beta \le 1$ is a forgetting factor.

In the general case the identifier with exponential forgetting is unstable, which leads to an increase in covariance matrix parameter values. That is the case when time sequence $x(k)$ reaches high dimensions. Thus, the standard exponentially weighted recursive least squares method is inapplicable because of ill-conditioning of the information matrix

$$\sum_{u=1}^{k} \beta^{k-u} X(u)X^T(u)$$

with high correlation between $x_i(k)$.

This problem can be solved using a pseudo-inverse operation with the Greville theorem instead of using the weighted information matrix reversal operation based on the Sherman-Morrison formula, but it is computationally demanding (especially for high n). According to the above, we offer a multidimensional modification of an exponentially weighted stochastic approximation algorithm

$$\begin{cases} B(k) = B(k-1) + \dfrac{e(k)X^T(k)}{\beta r(k-1) + \|X(k)\|^2}, \\ r(k) = \beta r(k-1) + \|X(k)\|^2. \end{cases} \tag{19}$$

The modification is a compromise between (17) and (18). It possesses smoothing and tracking properties. Dynamic properties of (18) and (19) depend on memory which influences the refinement of matrix of current estimates $B(k)$. As mentioned above, algorithms with infinite memory (for instance, recursive least squares method) possess the best filtering properties. At the same time, these procedures have poor tracking properties, as the characteristics of a video signal sequentially change. In this case, low-memory and noise irresistible algorithms (like Kaczmarz algorithm) are considered preferable.

Thus, algorithm memory capacity should be chosen based on the compromise between its smoothing and tracking properties. Unfortunately, time sequence characteristics are unknown a priori, and they constantly change. Variable memory algorithms are preferred in this case above those with fixed memory. Taking the above into consideration, we propose a smoothing parameter β adjustment method which controls statistics characterizing regular signal prediction error. It is assumed that model parameter setup is made via the exponentially weighted Kalman-Mayne algorithm which can be written as follows for i-th component $x_i(k)$

$$\begin{cases} b_i(k) = b_i(k-1) + \dfrac{(e(k)X^T(k)P_i(k-1))_i}{\beta\sigma_i^2 + X^T(k)P_i(k-1)X(k)}, \\ P_i(k) = \dfrac{1}{\beta}(P_i(k-1) - \dfrac{P_i(k-1)X(k)X^T(k)P_i(k-1)}{\beta\sigma_i^2 + X^T(k)P_i(k-1)X(k)}) \end{cases} \tag{20}$$

where $b_i(k)$ – i-th row of matrix $B(k)$, $(\circ)_i$ – i-th row of corresponding matrix product.

When variances of some components $\xi_i(k)$ of the disturbance vector are unknown, the following estimate can be used for (20):

$$\begin{cases} \sigma_i^2(k) = \sigma_i^2(k-1) + p_i(k-1)(\sigma_i^2(k-1) - l_i^2(k)), \\ p_i(k) = \dfrac{1}{\beta}(p_i(k-1) - \dfrac{p_i^2(k-1)}{\beta + p_i(k-1)}). \end{cases}$$

The model provides an acceptably accurate description of a controlled signal on time intervals of S observations, and parameters can change discontinuously at any time k_a. Varying value $\beta(k)$ is adjusted via the following statistics

$$T_i(k) = \sum_{u=k-S}^{k} \frac{l_i^2(u)}{\beta(k-1)\sigma_i^2 + X^T(u)p_i(u-1)X(u)}$$

with χ^2 distribution and S degrees of freedom, provided that $\beta(0) = 1$. Adjustment of $\beta(k)$ is made in discrete moments of time tS using the following rule:

$$\beta(k) = \begin{cases} 1\ when\ k < S,\ k = tS\ and\ T_i(k) \le \chi_j^{\ 2}, \\ \beta(k-1) - \Delta\beta\ when\ \ k = tS\ and\ T_i(k) > \chi_j^{\ 2}, \\ \beta(k-1)\ when\ tS < k < (t-1)S,\ t = 1, 2, ... \end{cases} \tag{21}$$

where $\chi_j^{\ 2}$ – quantile of χ^2 law with respect to significance level j, $\Delta\beta$ – step of adjustment.

The rule (21) assumes the change of $\beta(k)$ when values of k are multiples of S. When values of k are intermediate, the value of $\beta(k)$ remains unchangeable, and a fault is registered when the second ratio from (21) occurs. Such a lengthy procedure calls for the search of a more effective fault detection technique.

In (19) a smoothing parameter β adjustment method is proposed based on the Mann-Whitney U-test with the following controllable characteristic:

$$\sum_{u=k-S+1}^{k} sign(x_i(u) - \hat{x}_i(u)) \ge \gamma \tag{22}$$

where γ is a threshold value, S is a value of a moving control window.

$$sign(x_i(u)-\hat{x}_i(u)) = \begin{cases} 0 \; when \; x_i(u)=\hat{x}_i(u), \\ +1 \; when \; x_i(u) > x_i(u), \\ -1 \; when \; x_i(u) \le \hat{x}_i(u). \end{cases}$$

The control window moves from $\beta(1)=0$ that equals maximum algorithm performance (19). The exponentially weighted recursive least squares method is invalid for such a purpose. During this process different situations may occur, as described below.

$$\sum_{u=k-S+1}^{k} sign(x_i(u)-\hat{x}_i(u)) \ge \delta.$$

This means that signal x_i stochastic component ξ_i prevails over the "drift" component. In this case, the necessity to improve smoothing properties of the algorithm arises, i.e. we need to increase the algorithm by the rule given next:

$$\beta(k)=\beta(k-1)+\Delta\beta.$$

In (22) the signal drift component prevails, and the algorithm has no ability to monitor faults on-line. In such case we need to reduce memory

$$\beta(k)=\beta(k-1)-\Delta\beta$$

in order to register fault occurrence.

To control multidimensional time sequence faults, the authors propose the use of the following modification of (22):

$$\max_i\left(\sum_{u=k-S+1}^{k} sign(x_i(u)-\hat{x}_i(u))\right) \ge \gamma,$$

This modification has the ability to simultaneously control all of the components and register fault occurrence, if the value of at least one component $x_i(k)$ changes significantly.

In practice heuristic procedures are used instead of stochastic. The methods of Shown, Brown, Chow, Trigg-Leach, Roberts-Reed, etc. belong to the heuristic procedures[19, 20, 21]. The main principle of heuristic methods lies in specification of a set of values for smoothing parameter β (for instance, $0; 0.05; 0.1; ...0.95; 1$) with a set of characteristics that determine identification quality. Very often the sets are the following:

$$e_i(k,\beta) = x_i(k) - \hat{x}_i(k,\beta)$$

– current error of estimation for i -th component,

$$S_i(k,\beta) = e_i(k,\beta) + S_i(k-1,\beta)$$

– cumulative sum of errors,

$$d_i(k,\beta) = (1-\beta)\left|e_i(k,\beta)\right| + \beta d_i(k-1,\beta)$$

– mean absolute error value,

$$\overline{e}_i(k,\beta) = (1-\beta)e_i(k,\beta) + \beta\overline{e}_i(k-1,\beta)$$

– mean error,

$$\tilde{e}_i(k,\beta) = (1-\beta)\frac{e_i(k,\beta)}{x_i(k)} + \beta\tilde{e}_i(k-1,\beta)$$

– mean relative error,

$$\overline{e}_i^2(k,\beta) = (1-\beta)e_i^2(k,\beta) + \beta\overline{e}_i^2(k-1,\beta)$$

– mean square error.

If the controlled value exceeds some threshold, γ, then the decision is made about fault occurrence, and the necessity to refine smoothing parameter β arises.

Stationary stochastic signal usually collides with β in the interval $0.7 \le \beta \le 0.99$ [22]. The simplest way is to decrease smoothing parameter by the following rule when the value $\tilde{e}_i(k,\beta)$ with some β exceeds the 0.05 threshold:

$$\beta(k) = \beta(k-1) - \Delta\beta.$$

And then, the process of identification continues with a new $\beta(k)$. If β exceeds the threshold 0.7 $(\beta(k) \le 0.7)$, then the decision is made about fault occurrence.

The Chow method is more effective, but more complicated as well. Three models with smoothing parameters β, $\beta + \Delta\beta$ and $\beta - \Delta\beta$ are used in his method. If the best result is obtained with smoothing parameter $\beta(k) = \beta + \Delta\beta$ at moment k, then three new parameters β, $\beta + \Delta\beta$ and

$\beta + 2\Delta\beta$ are used the next time. If the best model is obtained with smoothing parameter $\beta(k) = \beta - \Delta\beta$, then triple β, $\beta - \Delta\beta$ and $\beta - 2\Delta\beta$ is formed. If the best result is achieved with smoothing parameter $\beta(k) = \beta$, then the set β, $\beta + \Delta\beta$ and $\beta - \Delta\beta$ continues to be used.

Control tracking signal methods are simpler[22]. When the tracking signal exceeds a predefined threshold, it indicates fault occurrence. R. Brown proposed the following definition for the tracking signal[20]:

$$T_i^B(k) = \frac{\sum_{u=1}^{k} e_i(u)}{\sqrt{\sigma_i^2(k)}}.$$

In the above equation the sum of errors varies around 0 (because of error randomness), if the adjusted model matches the controlled signal. The sum of errors does not exceed some threshold which is set a priori for a particular level of probability with definite dispersion of prediction error sum, and the value to which it tends is given by the following equation:

$$\lim_{k\to\infty} \sigma_i^2(k) = \frac{1}{1-(1-\beta)^{2(pn+1)}} \sigma_i^2.$$

R. Brown proposed using the mean absolute deviation instead of mean square errors σ_i for practical calculations

$$d = \int_{-\infty}^{\infty} \left| e - M\{e\} \right| p(e) de \approx \sum_{u=0}^{n} \left| e(u) - M\{e\} \right| P_i.$$

The deviation is proportional to the mean square value of error, because

$$\Delta = (e - M\{e\}) / \sigma_i,$$
$$d_i = \sigma_i \int_{-\infty}^{\infty} |\Delta| \, p(\Delta) d\Delta,$$

as proportionality coefficient changes insignificantly for a wide variety of distributions (for normal distribution $\frac{d_i}{\sigma_i} = \sqrt{2\pi} = 0,7979$).

If d_i is calculated the following way

$$d_i(k) = (1-\beta)\left| e_i(k) \right| + \beta d_i(k-1)$$

with tracking signal

$$T_i^{\beta}(k) = \frac{\sum_{u=1}^{k} e_i(u)}{d_i(k)},$$

then $(1-P_i)\%$ threshold can be presented as shown below:

$$\pm\tau_{i,1-P_i} = \pm\frac{\sigma_i}{2}\sqrt{\frac{\pi(1-\beta)}{1-(1-\beta)^{2(pn+1)}}}.$$

D. Trigg and A. Leach suggested the following definition of a tracking signal[21]:

$$T_i^{TL}(k) = \frac{T_i'(k)}{d_i(k)}$$

where $T_i'(k) = (1-\beta)e_i(k) + \beta' T_i'(k-1)$ is not a total sum of deviations, but a smoothed error, provided that the following inequality is observed

$$\beta' \le \beta.$$

When $\beta' = \beta$, the tracking signal varies between –1 and +1. In order to incorporate automatic feedback, D. Trigg and A. Leach proposed to calculate the smoothing parameter according to the ratio:

$$\beta(k) = 1 - \left|T^{TL}(k)\right|,$$

and to register signal faults when $\beta(k)$ changes significantly. The modification shown here is also well known:

$$\beta(k) = 1 - \left|T^{TL}(k-1)\right|.$$

To reduce the gap between the model and controlled sequence (that is seen from tracking signal increase), smoothing parameter value must be decreased.

It is better to use the Brown method for processes with smooth drifts. Strong jumps again are better distinguished via the Trigg-Leach tracking signal. That is why this form of tracking signal can be helpful for fault detection in time sequences given by (8)-(10).

4. APPLICATION IN MEDICAL VIDEO

To check the effectiveness of the proposed approach, an experiment has been performed on a medical video which contains 550 frames. The video sequence was divided into separate frames, as shown in figure 1. These sequences of frames have been modeled as a two-dimensional time sequence.

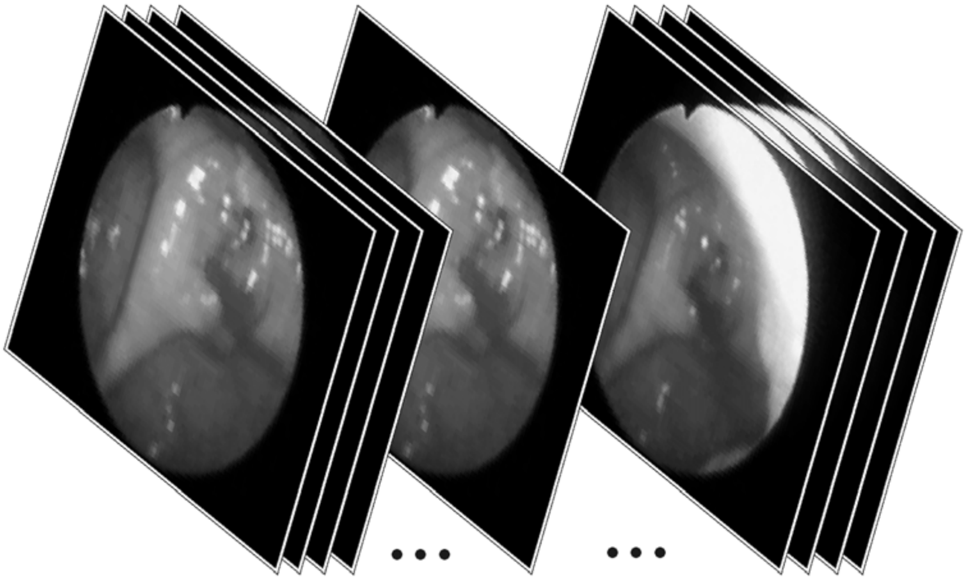

Figure 1. Video partitioning into multidimensional time sequence

For more effective control of changes in the video we have to perform a segmentation, as shown in figure 2. A texture algorithm (JSeg) was used to implement the segmentation. Segmentation enabled the researchers to determine initial video characteristics for fault detection. We used segmentation based on geometrical properties (area, perimeter, angle of segment slope, etc.) of stable images.

Figure 2. Example of frame segmentation

The analysis of segments from video frames has led us to the results shown in figure 3. This figure shows that strong changes in frames 40-55 and 270-311 influence strong decrease of parameter values on the plot. Consequently, salient changes in a scene have been accurately detected with the proposed approach. It was also noted that with instant appearance/disappearance of noise, parameter values react quickly as well.

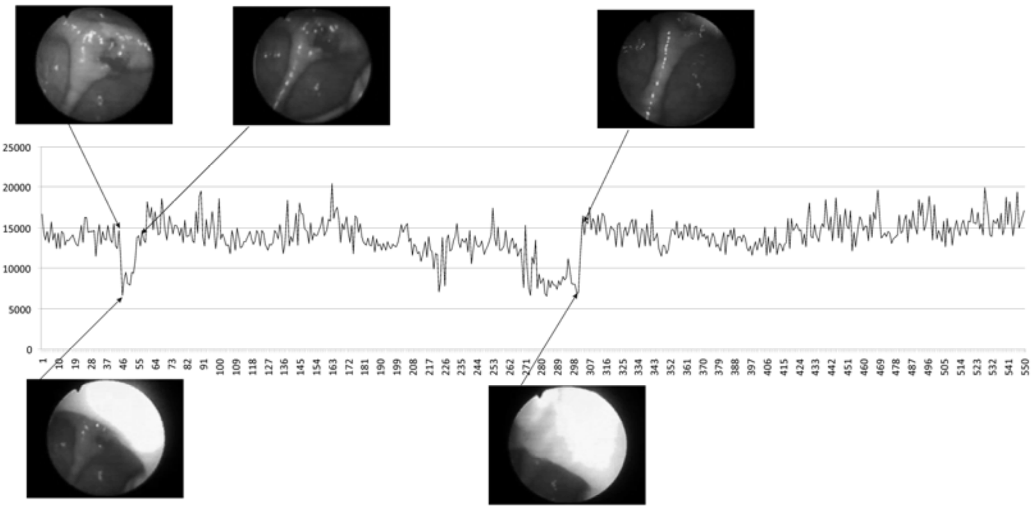

Figure 3. Video analysis results

The proposed approach allows the detection of salient changes in video. It can be used for effective shot boundary detection, information search in video databases, video interpretation, summarization and skimming.

5. CONCLUSION

Time series for video segmentation has been widely used during recent years. Despite clear progress, existing approaches are either computationally heavy or produce poor results in terms of quality. Moreover, it is very hard to develop a unique approach that would satisfy the needs of at least several applications at the same time. We first analyzed existing techniques and approaches, and then we introduce an adjustable model for detection of faults in multidimensional time sequences.

The adjustable model enables tracking video changes. It deals with both – smooth changes and strong jumps. One of the main disadvantages of the proposed model is that it has a large number of adjustable parameters, which potentially may lead to serious processing problems for large amounts of data (like HD-content processing). Despite that this model was shown to be fully applicable for medical diagnostics, as proved by an endoscopic operation analysis that has been carried out.

However, it should be noted that real time calculations are difficult to implement at the present time, since video sequence segmentation takes considerable time. Consequently, it makes sense to consider possibilities for speeding up the procedure. We can move toward such a possibility by "partial" data extraction, i.e. taking of separate shots at certain intervals from a video sequence. Another variant is based on the changing of segmentation parameters (however it is necessary to consider the influence of these changes on proposed methods). It is also possible to replace the segmentation method by a quicker one, but in this case care must be taken not to decrease segmentation quality.

In addition, an optimal vector of segment characteristics should be chosen. Implementation of all these characteristics at the same time should provide maximum help for detection of initial data changes. Thus, it is necessary to observe characteristics that have the greatest impact on detection of changes (these are characteristics that provide incomplete or contradictory information). It is also necessary to incorporate more weight coefficients for the most meaningful characteristics.

6. REFERENCES

[1] Y. Huang, *Adaptive nonlinear video editing: Retargeting, replaying, repainting and reusing (R4)*, Bridgewater: Huawei Technologies, 40, 2009.

[2] D.R. Goldman, *Framework for video annotation, visualization, interaction: Doctoral Thesis*, Washington, 140, 2007.

[3] T. Tsoneva, *Automated summarization of movies and TV series on a semantic level*: *Doctoral Thesis*, Eindhoven, 126, 2007.

[4] K. Ngan, and H. Li, *Video segmentation and its applications*. NY: Springer, 2011.

[5] D. Bouchard, *Automated time series segmentation for human motion analysis*, Center for Human Modeling and Simulation, University of Pennsylvania, 2006.

[6] M. Hoai, Z. Lan, and F. Torre, Joint segmentation and classification of human actions in video. *Presented at IEEE Conference on Computer Vision and Pattern Recognition*, Colorado Springs, June 20- 25, 2011.

[7] R. Cutler, and L. Davis, Robust real-time periodic motion detection, analysis and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), p781-796, 2000.

[8] X. Xuan, and K. Murphy, Modeling changing dependency structure in multivariate time series. *Presented at 24-th International Conference on Machine Learning*, Corvallis, June 20-24, 2007.

[9] Z. Harchaoui, F. Bach, and E. Moulines, Kernel change-point analysis. *Presented at 23-rd Annual Conference on Neural Information Processing Systems,* Vancouver, December 7-10, 2009.

[10] Z. Harchaoui, and C. Lévy-Leduc, Multiple change-point estimation with a total variation penalty. *Journal of the American Statistical Association*, 105(492), p1480-1493, 2010.

[11] D. Angelosante, and G. Giannakis, Sparse graphical modeling of piecewise-stationary time series. *Presented at 36-th International Conference on Acoustics, Speech and Signal Processing*, Prague, May 22-27, 2011.

[12] I. Budvytis, V. Badrinarayanan, and R. Cipolla, Semi-supervised video segmentation using tree structured graphical models. *Presented at IEEE Conference on Computer Vision and Pattern Recognition*, Colorado Springs, June 20- 25, 2011.

[13] R. Isermann, Process fault detection based modeling and estimating methods – A survey. *Automatica,* 20(4), p387-404, 1984.

[14] I.V. Nikiforov, Sequential detection of changes in stochastic systems. *Lecture Notes in Control and Information Sciences*, 77, p216-258, 1986.

[15] P.C. Badavas, *Real-time statistical process control*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[16] A.D. Pouliezos, and G.S. Stavrakakis, *Real time fault monitoring of industrial processes*. Dordrecht: Kluver Academic Publishers, 1994.

[17] K. Juselis, *The cointegrated var-model*. N.Y.: Oxford University Press, 1994.

[18] S. Kaczmarz, Approximate solution of system of linear equations. *International Journal of Control*, 57(5), p1269-1271, 1993.

[19] E.Y. Chow, and A.S. Willsky, Issues in the development of a general design algorithm for reliable failure detection. *Presented at 19-th IEEE Conference of Decision and Control,* Albuquerque, December 10-12, 1980.

[20] R.G. Brown, *Smoothing, forecasting, and prediction of discrete time series*. Englewood Cliffs, NJ: Prentice Hall, 1963.
[21] D.W. Trigg, and A.G. Leach, Exponential smoothing with an adaptive response rate. *Operational Research Quarterly*, 18(1), p53-59, 1967.
[22] D.C. Montgomery, I.A. Johnson, and J.S. Gardiner, *Forecasting and time series analysis*. NY: Mc Graw-Hill, 1990.