SIGNAL SEGMENTATION AND MODELLING BASED ON EQUIPARTITION PRINCIPLE

Costas Panagiotakis and Georgios Tziritas

Department of Computer Science, University of Crete, Heraklion, Greece, P.O. Box 2208
E-mails:{*cpanag,tziritas*}@*csd.uoc.gr*

ABSTRACT

In this paper, we propose a method for time interval segmentation of signals based on an EquiPartition principle (EP). According to EP, the signal is segmented into segments that give equal errors in reconstruction selecting the most suitable model to describe each segment. Moreover, the segments are equivalent in the content domain, since the signal is segmented into segments that are modelled by the same number of coefficients. The proposed method has been successfully applied on different types of signals like: physiologic, speech, human motion, financial time series. Finally, the proposed methodology is very flexible on changes of error criteria, signal modelling and on signal dimension yielding a robust method for segmentation and modelling of signals.

Index Terms— Signal segmentation, signal modelling, time-frequency analysis, equipartition.

1. INTRODUCTION

Nowadays, there is a tremendous increase of time series data such as audio files, seismic signals, financial time series, medical time series data, etc, being stored, analyzed and transmitted. This explosion of information is combined with a technological evolution in the areas of *terminals, networking and pattern recognition*. All these technological achievements require new services, software methods and tools for searching, understanding, retrieving and browsing signal content. Therefore, curve/signal segmentation and modelling is a challenging problem, particularly for pattern recognition applications. Signal summarization, classification [1] and key frames extraction from a video sequence could lead to a curve segmentation problem in an appropriate feature space. An example of such segmentation approach is the polygonal approximation, which is a well-known and widely studied problem [2]. In [1], sound segmentation and classification into speech and music is done successfully using frequency and energy based features computed on constant time intervals. Earthquake signal segmentation/classification is solved using constant time intervals and the statistical analysis of the estimated features [3].

In many of the above applications it could be interesting to have a uniform representation according to an appropriate
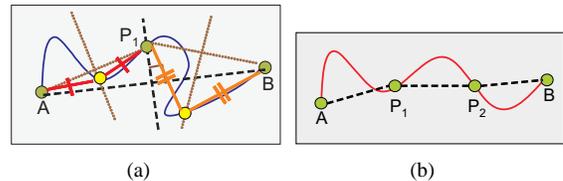

Fig. 1. **(a)** The segment bisector based method has been inductively applied for $N = 4$. The corresponding line segments are not necessary equal. **(b)** A curve equipartition example for $N = 3$, $|AP_1| = |P_1P_2| = |P_2B|$ under Euclidean distance metric.

quality measure. The objective is the partition of the feature sequence into "homogenous" segments with uniform characteristics according to a predefined criterion. We have adopted such an approach for 3D modelling and non articulated motion tracking [4], polygonal approximation [5], key frames detection [6] leading to the curve equipartition problem (EP). More formally, the equipartition problem (EP) is the following [7]. Let $c : [a, b] \to \mathbb{R}^m$ be a continuous, injective curve in m-dimensional space. Let us further suppose that a distance $d(.,.)$ is defined on the m-dimensional space, so that $d(c(t), c(t'))$ is a measure of how dissimilar are the corresponding trajectories. We ask for a sequence of points that are uniformly distributed according to this distance. Let us consider a simple case of EP problem under Euclidean distance metric, where $N = 2$. Then we have to locate a curve point P, so that $|AP| = |PB|$. This point can be given as the intersection of the curve with the AB segment bisector. It holds that when $N > 2$, there is not a trivial method to solve EP (see Fig. 1(a)). Fig. 1(b) illustrates a curve equipartition example for $N = 3$, where $|AP_1| = |P_1P_2| = |P_2B|$.

Piecewise continuous signal segmentation and reconstruction can be solved by first detecting the segments and then applying the modelling [8]. In [9], discontinuities separating consecutive time intervals of the original signal are initially detected by measuring the curvature and arc length of the smoothed signal. Otherwise, signal segmentation and reconstruction can be solved simultaneously based on energy optimization methods [10]. These methods compute the time segments (t_i) and the approximations (g_i) per segment by mini-

mizing an energy function $E(g)$.

$$E(g) = \sum_{i=1}^{N} \int_{t_i}^{t_{i+1}} (g_i(t) - f(t))^2 + \mu(\dot{g}_i(t))^2 dt \quad (1)$$

The optimization problem is solved using level sets algorithms applying successfully in piecewise continuous signals. The parameter μ controls the smoothness of g_i. Another class of segmentation methods is based on Bayesian inference [11, 12]. These algorithms consist of defining appropriate prior distributions for the unknown signal parameters (including the change points between the different segments) and estimating these unknown parameters from their posterior distributions. Finally, these methods provide a possibility of change for each frame of given signal. The method proposed in [11] was applied in speech data. In [12], the segmentation procedure allows joint segmentation of signals recorded by different sensors using astronomical time series data.

Most of the above mentioned approaches address the signal segmentation problem focusing either on a restricted signal content (e.g. piecewise continues signals) under predefined and specific modelling or minimizing specific energy criteria. On the contrary, in this paper, signal segmentation and reconstruction-modelling is performed simultaneously on continuous or piecewise continuous signals by the use of an innovative computational geometry algorithm [7], which equally partitions the signal resulting segments with equal errors in reconstruction. The estimated segments, that are modelled by the same number of coefficients, are *equivalent* in the content domain. The proposed method can be used under different energy criteria.

In [13], it is shown that many signals with a finite rate of innovation can be sampled and perfectly reconstructed using physically realizable kernels of compact support and a local reconstruction algorithm. The rate of innovation corresponds to the number of degrees of freedom per unit of time. These signals may be not bandlimited, but still have a finite number of degrees of freedom per unit of time, such as, for example, nonuniform splines or piecewise polynomials. In this work we segment the signal into segments with the same number of degrees of freedom.

The rest of the paper is organized as follows: Section 2 gives the problem formulation describing the proposed modelling, energy criteria. Section 3 presents the proposed signal segmentation and reconstruction method. The experimental results are given in Section 4. Finally, conclusions and discussion are provided in Section 5.

2. PROBLEM FORMULATION

2.1. Signal modelling and segmentation

Let us assume a digital signal $f(t)$, $t \in \{0, 1, \cdots, T-1\}$, where T denotes the number of samples. The goal of this work is to automatically segment and to reconstruct each segment of the given signal using the most appropriate model for each time segment. However, the performance of reconstruction depends on the selection of basis. For example, a reconstruction of piecewise polynomial signal yields more accurate and compact results when a polynomial basis is used. So, an interesting problem is the selection of basis where the sparsity of signal is high. It is possible that a signal consists of piecewise continuous segments that admit a sparse representation in different bases. The proposed method selects for each detected segment a compact representation so that the signal is segmented into segments with equal reconstruction errors and number of coefficients per segment solving simultaneously the signal segmentation and modelling.

Let $g(t)$ be a modelling of $f(t)$. Let N be the number of the time segments $[0, t_1] \cup [t_1, t_2] \cup \cdots \cup [t_{N-1}, T-1]$, where $t_i \leq t_{i+1} \in \{0, \cdots, T-1\}, i \in \{1, \cdots, N-1\}$ be the time that define the end time of $i-1$ segment and the start time of i segment ($t_0 = 0, t_N = T-1$). The error between the segment $[t_i, t_{i+1}]$ of f and the corresponding segment of g is given by $d(t_i, t_{i+1})$. Then the global approximation error of $f(t)$, $g(t)$ is defined as the maximum error between the segments of f and their corresponding segments of g,

$$E(f, g) = \max_{i \in \{0, 1, 2, \cdots, N-1\}} d(t_i, t_{i+1}) \quad (2)$$

This definition of error has been also used on polygonal approximation problem [5]. Therefore, the goal of the method is to select simultaneously the modelling and segments so that the global approximation error $E(f, g)$ is minimized. A near optimal solution of the segmentation problem is achieved, when the approximation errors per segment are equal, as the error is shared between all the segments,

$$\epsilon = d(0, t_1) = d(t_1, t_2) = \cdots = d(t_{N-1}, T-1) \quad (3)$$

The approximation error ϵ will be minimum or close to minimum. This is the equal errors (EE) criterion. A detailed analysis of EE criterion is given in [5]. The straightforward implementation of the EP method provides directly N. According to the proposed methodology, the segmentation and the signal modelling are given at the same time by the reduction to the EP problem using as input the function $d(.,.)$. In the next sections, $d(.,.)$ is defined and the basis selection is described, under polynomial and Fourier bases.

2.2. Basis description

In this work, we have used two types of bases, polynomial and Fourier, to model the given signal. The proposed method can include more bases without changes in methodology. The model coefficients in each segment can be used for signal summarization. Let u, v, $0 \leq u \leq v \leq T-1$ be the start and the end time of a time segment of the given signal $f(t)$. $f(t)$, $t \in [u, v]$, has been modelled by the S most important

Fourier coefficients $g_1(t)$, using a polynomial of $2 \cdot S - 1$ degree $g_2(t)$, getting equal rate innovations $\frac{2 \cdot S}{v - u + 1}$ for the two representations [1].

$$f(t) \approx g_1(t) = \frac{1}{N} \sum_{k=1}^{S} w_k \cdot e^{2\pi i \cdot f_k \cdot t} \qquad (4)$$

$$f(t) \approx g_2(t) = \sum_{k=0}^{2 \cdot S - 1} a_k \left(\frac{t}{T}\right)^k \qquad (5)$$

w_k and f_k denote the weight and the frequency of k Fourier coefficient, respectively. a_k denotes the weight of $k - degree$ monomial t^k. We have used the S highest in energy coefficients of Fast Fourier Transform (FFT).

These coefficients correspond to a robust reconstruction of the signal, reducing noise and providing at the same time meaningful information of the signal. S is a parameter of the proposed method that can be defined by the user (e.g. $S = 5$). In polynomial approximation, we keep the first $2 \cdot S - 1$-degree monomials in order to model low degree polynomials like step functions, polygonal functions, splines, etc., where the Fourier basis fails. The polynomial and Fourier series modelling have been used as an example showing that we can use more than one models/bases.

2.3. Error criterion

Let $d_j(u, v), j \in \{1, 2\}$ be the approximation error between $g_j(t)$ and $f(t)$. Then, $d(u, v)$ corresponds to minimum error, since we select the basis that gives the approximation of minimum error per segment.

$$d(u, v) = min_j d_j(u, v) \qquad (6)$$

According to the EP problem requirements [5], $d(u, v)$ should satisfy the following properties:

1. $d(u, v) = 0 \Leftrightarrow u = v$ (isolation).

2. $d(u, v) = d(v, u)$ (symmetry).

These properties are satisfied by the difference in mean energy of signal $f(t)$ and $g_j(t)$.

$$d_j(u, v) = \frac{1}{v - u + 1} \left| \sum_{t=u}^{v} f^2(t) - \sum_{t=u}^{v} g_j^2(t) \right|, j \in \{1, 2\} \qquad (7)$$

In addition, $d(u, v)$ should be continuous that means that small changes in u, v correspond to small changes in $d(u, v)$ (discrete case). By the equation 6, it holds that $d(u, v)$ is continuous, if and only if, $d_j(u, v)$ are continuous functions. According to the definition of $g_1(t)$, we select the S highest

in energy coefficients, so small changes in u, v correspond to small changes in energy of the $g_1(t)$. Therefore, $d_1(u, v)$ is continuous. Similarly, $d_2(u, v)$ is also continuous function.

$d(u, v)$ is a good metric to test if the proposed model fits well to the data. In the case of Fourier basis, it holds that $\sum_{t=u}^{v} g_1^2(t) = \frac{\sum_{k=1}^{S} |w_k|^2}{v - u + 1} \leq \sum_{t=u}^{v} f^2(t)$. Therefore, $d_1(u, v) = 0$, if and only if, $f = g$. If $d_1(u, v)$ is getting high, it means that the signal cannot be described well by Fourier basis. Concerning the polynomial approximation, it minimizes the square error $\sum_{t=u}^{v} (f(t) - g(t))^2$, fitting to $f(t)$. If $\sum_{t=u}^{v} (f(t) - g(t))^2$ is getting low, then $d_2(u, v)$ will also be low, which means that the signal is well described by polynomial basis.

Consecutively, if the error is low, it means that the segmentation is good and the content description of the signal by the proposed descriptors is valid. Moreover, it means that the segment is homogenous in content, since it can be described by a small number of coefficients that are related to the signal content.

Fig. 2 illustrates the approximation of synthetic signals using Fourier basis and polynomial basis, for $S = 5$. In the first example, the reconstruction error was 96.97% lower using Fourier basis, while in the second example, the reconstruction error was 81.48% lower using polynomial basis. In each example, the most robust approximation is given by the model that yields the minimum error in reconstruction.

3. SEGMENTATION AND RECONSTRUCTION

3.1. Iso-Level Algorithm (ILA)

The signal segmentation and reconstruction is done simultaneously using the Iso-Level Algorithm (ILA). The input of the proposed method is the number of segments N. In addition, it needs the matrix $d(t_k, t_l), k, l \in \{0, 1, \cdots, T - 1\}$ of distortions. The detailed description of the algorithm used can be found in [5]. It is a recursive method. Thus, when it is executed for N segments, it uses the precomputed results for $N - 1$ segments. According to our analysis [7], the equipartition problem (EP) always admits at least one solution.

N can be given by the user or can be estimated automatically by terminating the EP algorithm, when the estimated "distortion" exceeds a predefined error, similar with the problem of minimum number of segments $(min - \#)^2$ in polygonal approximation [5]. Both cases are solved in $O(N \cdot T^2)$ steps thanks to the property of the method that solves the problem for values less than N without additional cost [5].

3.2. Selection of the Number of Segments

The selection of N could be done by the user to fit specific preferences and information needs. A 'semiautomatic' com-

---

[1] The number of stored parameters for S Fourier coefficients is $2 \cdot S$, since they are complex numbers having real and imaginary part. The number of stored parameters for $2 \cdot S - 1$ degree polynomial is $2 \cdot S$

[2] $min - \#$ is used for the problem of finding the minimum number of segments that gives error lower than the given error.
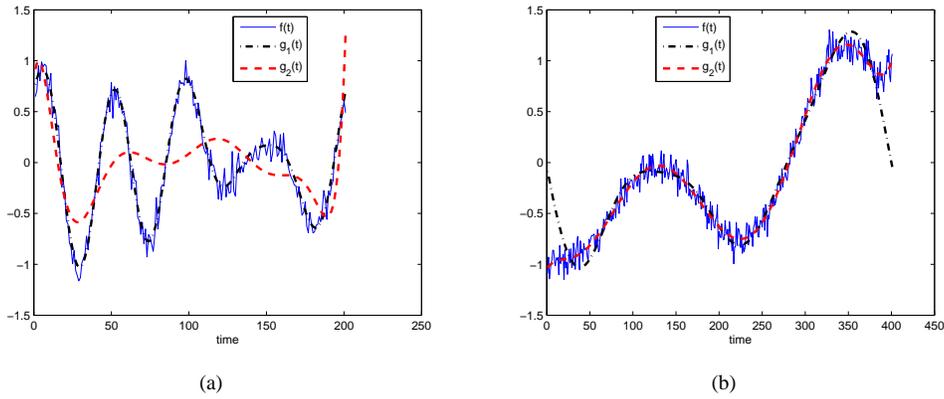
Fig. 2. The given synthetic signals (blue line) are approximated by $g_1(t)$ and $g_2(t)$ with $S = 5$. **(a)** $d_1(0, 200) = 0.0063$ and $d_2(0, 200) = 0.2082$. **(b) (a)** $d_1(0, 400) = 0.0086$ and $d_2(0, 400) = 0.0465$

putation of N can be done by terminating the EP algorithm, when the estimated "distortion" (error) exceeds a predefined error. However, it is crucial to develop a mechanism able to automatically estimate the most appropriate number of segments N. In [14] the number of segments are automatically determined by permutation tests. According to the tests, a segmentation algorithm is executed for increasing number of segments and the error is measured. When we exceed the true number of segments, the error keeps decreasing as well, but the reduction of error is due to noise in the data (overfitting).

A similar idea is used on the proposed method so that N can be provided without any user interaction estimating, if the segments at iteration l of the algorithm suffice to approximate the given signal. Let Q_l is a measurement of the distortion between the given signal and its approximation using the l segments. As Q_l we use the error of the approximation (ϵ_l). Q_l is usually decreasing as l increasing, $Q_l \geq 0$. Q_l has characteristics of a convex function, that is, if we smooth it we will get a convex function. Therefore, we have to introduce a new criterion instead of minimum of this function. Thus, we propose to select the appropriate level l so that the numerical approximation of the second derivative of Q_l, $\ddot{Q}_l$, is maximized.

$$\ddot{Q}_l = Q_{l+1} + Q_{l-1} - 2 \cdot Q_l, \ \ l \in \{2, 3, \cdots, T - 2\} \quad (8)$$

This is due to the fact that the second derivative expresses a measure of the curvature of the content curve.

4. EXPERIMENTAL RESULTS

4.1. Dataset content description

In this section, the experimental results of the proposed algorithm are presented. We have tested the proposed algorithm on a data set consisting of different types of signals, like physiologic signals [15, 16], speech signals, human motion signals,

financial time series and synthetic signals, in order to show that our method can be used under any type of signal without any constraint. The goal on physiologic signals is to segment them into homogenous in content time segments, detecting in electrocardiogram (ECG) irregular states like arrhythmia, tachycardia, bradycardia, etc. In speech signals, the proposed method can be applied for phoneme recognition [17], since during a phoneme the speech signal is almost stationary. Human motion tracking data have been used for human action recognition [18] using data from athletics meeting. An important task for human action recognition is to segment the human motion signal into homogenous in content segments that correspond to sequential human actions. For example, in a high jump sequence we have to determine the running, jumping, falling and standing actions. The segmentation of financial time series [19] is another challenging problem. The goal is to segment the signal into intervals, where the market characteristics are homogeneous.

Fig. 3 illustrates signals that have been used in the article with blue curves. Fig. 3(a) shows a bradycardia electrocardiogram (ECG) from a defibrillator [15]. Fig. 3(b) shows an 100 msec speech signal [15] (sample rate = 6 Khz). The waveform changes during phonemes. Fig. 3(c) shows the human major axis angle ϕ [18, 20] during a high jump sequence. In this case, the signal corresponds to several sequential phases (actions) like running ($\phi \simeq \frac{\pi}{2}$), jumping ($0 \leq \phi \leq \frac{\pi}{2}$), falling ($-\frac{\pi}{2} \leq \phi \leq 0$) and standing ($\frac{\pi}{2} \leq \phi \leq \frac{\pi}{2}$) actions. Fig. 3(d) shows a 450-day EURO-USD exchange rate (from 04/16/2004 to 07/10/2005). Fig. 3(e) shows a synthetic signal. The signal of Fig. 3(f) is produced by added Gaussian white noise (12 dB Signal-to-noise ratio (SNR)) in the signal of Fig. 3(e). The synthetic signal consists of 4 segments, a sinus series, a polynomial and two sequential sinus series.
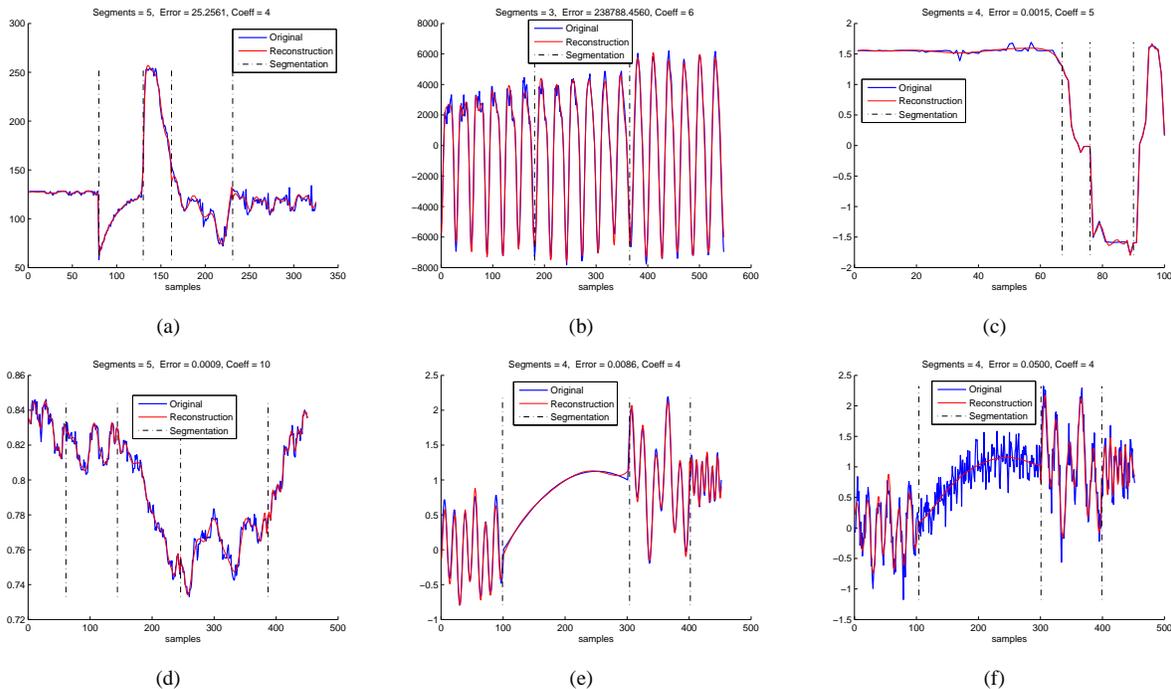
Fig. 3. Original signal (blue curves) and segmentation (black dotted lines) - modelling (red curves) results of the proposed method on: **(a)** electrocardiogram (ECG), **(b)** speech signal, **(c)** human major axis angle signal, **(d)** EURO-USD exchange rate, **(e)** synthetic signal, **(f)** synthetic signal with noise.

4.2. Segmentation and Modelling

In this section, we discuss the robustness of segmentation results of the proposed algorithm. Fig. 3 illustrates the original signals (blue curves) and segmentation (black dotted lines) - modelling (red curves) results of the proposed method. In each case, N was automatically estimated by the method of Section 3.2. According to the proposed method, it holds that the number of estimated parameters (modelling) of each segment are equal, as well as the errors of reconstruction for each segment. Therefore, the given signals have been modelled by the S most important Fourier coefficients, or using a polynomial of $2 \cdot S - 1$ degree. We have used different values for S, like $S = 4$ (Figs. 3(a), 3(e), 3(f)), $S = 5$ (Fig. 3(c)), $S = 6$ (Fig. 3(b)) and $S = 10$ (Figs. 3(d)).

In each example, the segmentation gives homogenous in content segments, due to equipartition principle. For example, the method divides the signal of Fig. 3(c) into 4 segments, that corresponds to 4 different actions of high jump (first segment: running, second segment: jumping, third segment: falling, fourth segment: standing). The method divides the signal of Fig. 3(d) into 5 segments, that corresponds to 5 different phases of the exchange rate (first segment: stable-decreasing, second segment: stable, third segment: fast decreasing, fourth segment: unstable, fifth segment: fast increasing). Moreover, the segmentation and modelling results

of Figs. 3(e) and 3(f) are almost the same yielding robustly the equation of original synthetic signal, showing that the method is robust to noise effects. In addition, we get accurate modelling with low number of coefficients, due to the selection of most appropriate basis to model each segment. For example, the last segment of Fig. 3(a) as well as the segments of Fig. 3(b), that correspond to periodic signals, have been efficiently modelled by Fourier basis, while the other segments of Fig. 3(a) have been modelled by polynomials.

5. CONCLUSIONS

Consecutively, an EP based method for time interval segmentation and modelling of signals is described. According to the proposed method the signal is segmented into segments that give equal errors in reconstruction, selecting the most suitable model to describe each segment. Moreover, the segments are *equivalent* in the content domain yielding segments of the same number of degrees of freedom. Therefore, each segment has the same significance under the used principle. Concerning, the number segments, two automatic approaches are proposed each exploiting supervised or unsupervised content based criteria. The proposed methodology is very flexible on changes of error criteria, signal modelling and on signal dimension yielding a robust method for segmentation and modelling of signals.

Experimental results on a large data set of different types of signals like physiologic signals, speech signals, financial time series and synthetic signals have been obtained to demonstrate the efficiency and the robustness of the proposed schema. In all cases, the appropriate number of segments N, the segmentation and modelling as obtained by the proposed automatic processes is close to the human's perception with respect to the content fluctuation.

A possible extension of the proposed methodology may include more models (e.g. wavelets) and other minimization criteria. The proposed methodology can also be applied to multidimensional signals (e.g. produced by many sensors). Finally, we plan to apply the proposed method on signal denoising and compression problems.

Acknowledgments

6. REFERENCES

[1] C. Panagiotakis and G. Tziritas, "A speech/music discriminator based on RMS and zero-crossings," *IEEE Trans. on Multimedia*, vol. 7, no. 1, pp. 155–166, Feb. 2005.

[2] H. Imai and M. Iri, "Polygonal approximations of a curve (formulations and algorithms)," *Computational Morphology*, pp. 71–86, 1988.

[3] C. Panagiotakis, E. Kokinou, and F. Vallianatos, "Automatic p phase picking based on local maxima distribution," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2280–2287, 2008.

[4] C. Panagiotakis and G. Tziritas, "Snake terrestrial locomotion synthesis in 3d virtual environments," *Visual Computer*, vol. 22, no. 8, pp. 562–576, 2006.

[5] C. Panagiotakis and G. Tziritas, "Any dimension polygonal approximation based on equal errors principle," *Pattern Recogn. Lett.*, vol. 28, no. 5, pp. 582–591, 2007.

[6] C. Panagiotakis, A. Doulamis, and G. Tziritas, "Equivalent key frames selection based on iso-content principles," *IEEE Trans. Circuits Syst. Video Techn. (accepted), 2008*.

[7] C. Panagiotakis, G. Georgakopoulos, and G. Tziritas, "On the curve equipartition problem: a brief exposition of basic issues," in *European Workshop on Computational Geometry*, 2006.

[8] A. Gelb and E. Tadmor, "Spectral reconstruction of piecewise smooth functions from their discrete," in *Math. Model. Numer. Anal*, 2002, pp. 155–175.

[9] S. Mahmoodi and B. S. Sharif, "Signal segmentation and denoising algorithm based on energy optimisation," *Signal Processing*, vol. 85, no. 9, pp. 1845–1851, 2005.

[10] S. Mahmoodi and B. Sharif, "Noise reduction, smoothing and time interval segmentation of noisy signals using an energy optimisation method," *IEE Proceedings Vision, Image and Signal Processing*, vol. 153, no. 2, pp. 101–108, April 2006.

[11] P. Fearnhead, "Exact bayesian curve fitting and signal segmentation," *IEEE Transactions on Signal Processing*, vol. 53, no. 6, pp. 2160–2166, 2005.

[12] N. Dobigeonand, J.-Y. Tourneret, and J. D. Scargle, "Joint segmentation of multivariate astronomical time series: Bayesian sampling with a hierarchical model," *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 414–423, 2007.

[13] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling Moments and Reconstructing Signals of Finite Rate of Innovation: Shannon Meets Strang-Fix," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1741–1757, 2007.

[14] K. T. Vasko and H. T. T. Toivonen, "Estimating the number of segments in time series data using permutation tests," in *IEEE International Conference on Data Mining*, 2002, pp. 466–473.

[15] ," Available at http://ecow.engr.wisc.edu/cgi-bin/get/bme/463/tompkins.

[16] ," Available at http://www.physionet.org/.

[17] P. McCourt, S. Vaseght, and N. Harte, "Multi-resolution cepstral features for phoneme recognition across speech subbands," *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, vol. 1, pp. 557–560, 1998.

[18] C. Panagiotakis, E. Ramasso, G. Tziritas, M. Rombaut, and D. Pellerin, "Shape-motion based athlete tracking for multilevel action recognition," in *Proc. of Int. Workshop on Articulated Motion and Deformable Objects*. 2006, vol. 4069 of *Lecture Notes in Computer Science*, pp. 385–394, Springer.

[19] E. Otranto, "Clustering heteroskedastic time series by model-based procedures," Tech. Rep., 2006.

[20] C. Panagiotakis, I. Grinias, and G. Tziritas, "Automatic human motion analysis and action recognition in athletics videos," in *14th European Signal Processing Conf.*, 2006.