Biomedical Signal Processing and Control 4 (2009) 79–85

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

The use of noise information for detection of temporomandibular disorder

M. Ghodsi [a,*], H. Hassani [b,c], S. Sanei [a], Y. Hicks [a]

[a] Centre of Digital Signal Processing, Cardiff University, UK
[b] School of Mathematics, Cardiff University, UK
[c] Central Bank of the Islamic Republic of Iran, Iran

ARTICLE INFO

Article history:
Received 10 April 2008
Received in revised form 11 September 2008
Accepted 30 October 2008
Available online 4 January 2009

Keywords:
Temporomandibular disorder
Singular spectrum analysis
Separation
Noise
Detection

ABSTRACT

A new method for detection of temporomandibular disorder based on singular spectrum analysis is presented. In this method the motion data of markers placed on the points of special interest on the faces of several subjects is extracted and analysed. The individuals are classified into a group of healthy subjects and a group of those with temporomandibular disorder by extracting the signal components of the original time series and separating the noise using the proposed technique. The results for both simulated and real data verify the effectiveness of the proposed algorithm.

1. Introduction

Temporomandibular disorders (TMDs) occur as a result of problems with jaw, temporomandibular joint (TMJ), and surrounding facial muscles that control chewing and moving the jaw [1,2]. TMJ is the hinge joint that connects the lower jaw (mandible) to the temporal bone of the skull, which is immediately in front of the ear on each side of the head. The joints are flexible, allowing the jaw to move smoothly up and down and side to side. If you place your fingers just in front of your ears and open your mouth, you can feel the joints on each side of your head (illustrated in Fig. 1).

Symptoms of TMD include headaches, tenderness of the chewing muscles, and clicking or locking of the joints [4]. More than 40% of the general population has at least one sign of TMD, yet only one in four of such people is actually aware of, or reports any symptoms [5]. One of the most popular areas of TMD research is developing clear guidelines for diagnosing these disorders. Automatic measurement and classification of TMDs before and during the treatment can assist in early diagnosis, accurate monitoring of treatment, and enhance the efficacy of the treatment.

Current methods for TMD detection involve a physical examination by an expert in the area [4]. A dentist or clinician almost always diagnoses a TMD-based solely on a person's medical history and on a physical examination. Some background work has been reported in the literature for investigating the chewing performance using a multiple regression model of electromyographic and electrognathographic variables [6]. Frontal plane mandibular rotations and corresponding hemimandibular translations were studied in vitro by using direct observations of human cadaver mandible and in vivo by using the indirect observations of rotational electrognathography [7]. In recent research, scientists suggested new methods for detecting TMD on the basis of audio [8] and visual analysis [9,10]. Due to the cyclic nature of jaw movement some algorithmic methods have been developed for analysis of jaw movement based on frequency distribution of the myographic data [11]. Despite various clinical tests and methods, currently there are no scientifically proven tests available to diagnose TMD, and therefore the focus of this research is to provide a new approach verified by clinical trials.

Here we present an alternative method for detection of TMD based on visual analysis of facial movement. For this purpose we attach a number of markers to the points of interest on the individuals' faces and track their positions over a large number of frames in the video sequences. Afterwards, we analyse the motion patterns of the markers and extract their main signal.

In order to evaluate chewing in TMD patients with unilateral and bilateral internal derangement of the TMJ, the envelope of motion and velocity of chewing were recorded and analysed [12]. It was found that patients with bilateral internal derangement demonstrated a significantly restricted range of motion and reduced velocity than patients with unilateral internal derangement or normal subjects. The effect of malocclusion on mandibular movement during speech has also been investigated by using a

* Corresponding author.
E-mail address: ghodsim@cf.ac.uk (M. Ghodsi).

1746-8094/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bspc.2008.10.001

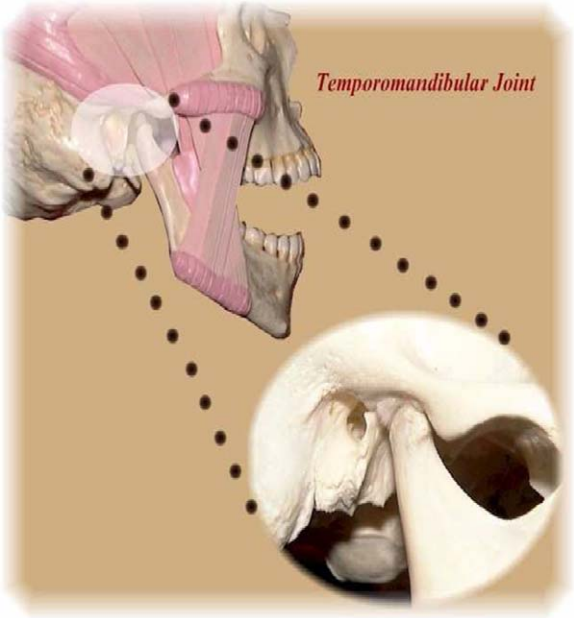

Fig. 1. An illustration of the temporomandibular joint and its location adopted from [3].

sirognathograph analyzing system and by looking at the envelope of motion during speech [13]. It was also shown that the signals extracted from the video data as a result of markers' motion during chewing has chaotic behavior [9,10]. The above findings motivated us to use a suitable method which is useful for analysis of nonlinear dynamical system and signal detection.

The study of nonlinear dynamical systems and chaos over the past decade motivated many researchers to reconsider what is meant by noise. In many cases the deterministic signal from a nonlinear system may look like noise when displayed in the time domain. Recently many different noise reduction methods for chaotic time series have been developed [14–18]. Many of them are based on the dynamical and statistical properties which use the embedding methods such as the method of delays [19] and singular value decomposition (SVD) [20].

It is accepted that SVD-based methods are very effective for noise reduction from chaotic signals [21,22] and widely has been used in speech enhancement (see, for example, [23–29]). Such an approach assumes that the signal and the noise are uncorrelated, and their second order statistics are available. In this way, one is able to decompose the noisy signal vector space into two orthogonal subspaces, called the signal subspace and the noise subspace, in order to achieve improved noise reduction.

It should be noted that the signal extracted from TMJ has different types of non-stationarities. Thus, one should use the method such as singular spectrum analysis (SSA) that does not rely on stationarity assumption. The SSA method does not depend on any assumptions based on normality of the residuals, linearity and stationarity of the observed data, and works well for both linear and nonlinear signals [30]. The SSA technique consists of two complementary stages: decomposition and reconstruction. The decomposition step comprises of embedding and SVD operations. The superiority of the SSA technique over traditional digital filtering methods used in biomechanical analysis was shown, with several examples, in the literature [31](for more applications in biomedical data analysis see [32,33]).

The paper is organized as follows. In Section 2, we provide the necessary mathematical background of the SSA technique, then we discuss the experimental data as well as the results of our methods applied to the data in Section 3. Section 4 concludes the paper with a short discussion.

2. Singular spectrum analysis

The SSA technique is a powerful technique for time series analysis incorporating the elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamical systems and signal processing [30]. A thorough description of the theoretical and practical foundations of the SSA technique (with several examples) can be found in [30,34]. In recent years SSA has been developed and applied to many practical problems (see, for example [35–40]).

The SSA technique consists of two complementary stages: decomposition and reconstruction, both of which include two separate steps. At the first stage we decompose the time series and at the second stage we reconstruct the original (noise free) time series. Here we provide a brief discussion of the methodology of the SSA technique.

2.1. Decomposition

The decomposition step comprises of embedding and SVD operations.

2.1.1. Embedding

Consider a noisy signal vector $Y = (y_1, \ldots, y_N)^T$ of N samples. Assume that the noise is white, additive and uncorrelated with the signal:

$$Y = S + N, \tag{1}$$

where S represents the signal and N noise, respectively. Embedding can be regarded as a mapping that transfers a one-dimensional signal $Y = (y_1, \ldots, y_N)$ into a multidimensional signal $X_1, \ldots, X_K$ with vectors $X_i = (y_i, \ldots, y_{i+L-1})^T \in \mathbf{R}^L$, where $K = N - L + 1$. Vectors X_i are called *L-lagged vectors*. The result of this step is the trajectory matrix

$$\mathbf{X} = [X_1, \ldots, X_K] = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} y_1 & y_2 & \cdots & y_k \\ y_2 & y_3 & \cdots & y_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_N \end{pmatrix}. \tag{2}$$

Note that the trajectory matrix $\mathbf{X}$ is a Hankel matrix, which means that all the elements along the diagonal $i + j = C$, where C is a constant, are equal. The only parameter required in the embedding step is the window length L, such that $2 \le L \le N$. We then consider $\mathbf{X}$ as a multivariate data with L characteristics and $K = N - L + 1$ observations. The columns X_j of $\mathbf{X}$, considered as vectors, lie in an L-dimensional space $\mathbf{R}^L$. We also have:

$$\mathbf{X} = \mathbf{S} + \mathbf{N} \tag{3}$$

where $\mathbf{S}$ and $\mathbf{N}$ represent respectively Hankel matrices of the signal S and noise N.

2.1.2. Singular value decomposition

The second step, the SVD step, calculates the singular value decomposition of the trajectory matrix and represents it as a sum of rank-one bi-orthogonal elementary matrices. SVD of $\mathbf{X}$ provides us with the collections of L eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_L \ge 0$ and the corresponding eigenvectors $\mathbf{U} = (U_1, U_2, \ldots, U_L)$ where U_i is the normalized eigenvector corresponding to the eigenvalue $\lambda_i (i = 1, \ldots, L)$.

The SVD of the trajectory matrix $\mathbf{X}$ can be written as

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^T \tag{4}$$

where $\mathbf{U} \in \mathbf{R}^{L\times K}$, $\mathbf{V} \in \mathbf{R}^{K\times K}$, and $\Sigma = \mathrm{diag}(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L)$. We can also partition the SVD of $\mathbf{X}$ as follows:

$$\mathbf{X} = [\mathbf{U}_1 \quad \mathbf{U}_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} \tag{5}$$

where $\mathbf{U}_1 \in \mathbf{R}^{L\times r}$, $\Sigma_1 \in \mathbf{R}^{r\times r}$ and $\mathbf{V}_1 \in \mathbf{R}^{K\times r}$. We can also represent SVD of the Hankel matrix of the signal S as

$$\mathbf{S} = [\mathbf{U}_{1s} \quad \mathbf{U}_{2s}] \begin{bmatrix} \Sigma_{1s} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_{1s}^T \\ \mathbf{V}_{2s}^T \end{bmatrix} \tag{6}$$

It is clear that the Hankel matrix $\mathbf{S}$ cannot be accurately reconstructed if the matrix is perturbed by noise. The least squared (LS) estimate of $\mathbf{S}$ is obtained when we approximate $\mathbf{S}$ by a matrix of rank r [41]:

$$\hat{\mathbf{S}}_{\mathrm{LS}} = \arg\min \|\mathbf{X} - \mathbf{S}_{\mathrm{LS}}\|_F^2 \tag{7}$$

where $\|\cdot\|_F$ denotes *Frobenius* norms. The solution of (7) can be obtained directly from SVD of $\mathbf{X}$ as

$$\mathbf{S}_{\mathrm{LS}} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^T \tag{8}$$

In fact the trajectory matrix $\mathbf{X}$ is decomposed into a sum of rank-one bi-orthogonal elementary matrices, $\mathbf{E}_i (i = 1, \ldots, d \leq L)$,[1] by applying the singular value decomposition on $\mathbf{X}$. Thus, the SVD of the trajectory matrix $\mathbf{X}$ can be written in the following format:

$$\mathbf{X} = \mathbf{E}_1 + \cdots + \mathbf{E}_d, \tag{9}$$

SVD of (9) is optimal in the sense that among all the matrices, the matrix $\mathbf{X}^{(r)} = \sum_{i=1}^{r} \mathbf{E}_i$ of rank $r < d$, provides the best approximation to the trajectory matrix $\mathbf{X}$, so that $\|\mathbf{X} - \mathbf{X}^{(r)}\|_F$ is minimum.

2.2. Reconstruction

The reconstruction part includes grouping and diagonal averaging operations as follows.

2.2.1. Grouping

The grouping step corresponds to splitting the elementary matrices into several groups and summing the matrices within each group. Let $I = \{i_1, \ldots, i_p\}$ be a group of indices $i_1, \ldots, i_p$. Then the matrix $\mathbf{X}_I$ corresponding to the group I is defined as $\mathbf{X}_I = \mathbf{E}_{i_1} + \cdots + \mathbf{E}_{i_p}$. The spilt of the set of indices $J = \{1, \ldots, d\}$ into disjoint subsets $I_1, \ldots, I_m$ corresponds to the representation

$$\mathbf{X} = \mathbf{X}_{I_1} + \cdots + \mathbf{X}_{I_m}. \tag{10}$$

The procedure of choosing the sets $I_1, \ldots, I_m$ is called the eigentriple grouping.

2.2.2. Diagonal averaging

The purpose of diagonal averaging is to transform a matrix to the form of a Hankel matrix which can be subsequently converted to a time series. If z_{ij} stands for an element of a matrix $\mathbf{Z}$, then the k-th term of the resulting series is obtained by averaging z_{ij} over all i, j such that $i + j = k + 1$. This procedure is called *diagonal averaging*, or Hankelization of matrix $\mathbf{Z}$. The result of Hankelization of a matrix $\mathbf{Z}$ is the Hankel matrix $\mathcal{H}\mathbf{Z}$, which is the trajectory matrix corresponding to the series obtained as a result of the diagonal averaging.

[1] d is the number of nonzero eigenvalues.

By applying the Hankelization procedure to all matrix components of (10), we obtain another expansion: $\tilde{\mathbf{X}} = \tilde{\mathbf{X}}_{I_1} + \cdots + \tilde{\mathbf{X}}_{I_m}$, where $\tilde{\mathbf{X}}_{I_i} = \mathcal{H}\mathbf{X}_{I_i} (i = 1, \ldots, m)$ and $\mathcal{H}$ is the Hankel operator. This is equivalent to the decomposition of the initial series $Y = (y_1, \ldots, y_N)$ into a sum of m series:

$$y_n = \sum_{k=1}^{m} \tilde{y}_n^{(k)} \tag{11}$$

where $\tilde{Y}^{(k)} = (\tilde{y}_1^{(k)}, \ldots, \tilde{y}_N^{(k)})$ corresponds to the matrix $\mathbf{X}_{I_k}$. The procedure of computing the time series $\tilde{Y}^{(k)}$ (that is, building up the group I_k plus diagonal averaging of the matrix $\mathbf{X}_{I_k}$) is called *reconstruction* of a series $Y^{(k)}$ by the eigentriples with indices in I_k.

3. Application

To examine the algorithm, a set of simulated data is generated and tested by the algorithm. Next, the performance of the method is evaluated on the TMD data.

3.1. Simulated data

The SSA technique can be applied to various time series. As our simulated example, let us consider the application of SSA for analyzing the following series in detail:

$$y(t) = 0.7\sin(2\pi t/17) + 0.5\sin(2\pi t/7) + 0.2\sin(2\pi t/3) + u_t I_{u_t} + 0.5\,\epsilon(t) \tag{12}$$

where $\epsilon(t)$ is a white noise process, the random variable u_t is distributed uniformly between 0.5 and 1, and I_{u_t} is defined as

$$I_{u_t} = \begin{cases} 1 & \text{if } t = 40k \ \ (k = 1, \ldots, 4) \\ 0 & o.w \end{cases} \tag{13}$$

In fact the random variable $u_t I_{u_t}$ adds some peaks to the time series systematically. Obviously, the behavior of these peaks are dynamic and completely different from random variable $\epsilon(t)$. However, this part of the signal can be simply considered as a noise component whilst contains useful information. Fig. 2 shows the series. Visual inspection of Fig. 2 indicates that the depicted series has a complicated structure and looks like a noise signal. Next, we extract the added peaks from the signal $y(t)$ using the SSA technique and show that these peaks appear in the noise series.

3.1.1. Decomposition: seasonality and noise

A general descriptive model of the series that is considered in SSA is an additive model where the components of the series are oscillatory signals and noise. In addition, the oscillatory components are subdivided into periodic and quasi-periodic components, while noise components are, as a rule, *aperiodic* series. The sum of all additive components, except for the noise, is called signal. We thus decompose the series into the signal (oscillations) and noise. Note that SSA does not require any priori parametric model for the oscillatory signals.

As was mentioned earlier, the window length L is the only parameter to be used in the decomposition stage. Theoretical results advise us to choose L large enough but not greater than $N/2$. Using these recommendations, we take $L = 100$ (in our case $N = 200$). Thus, based on this window length and on the SVD of the trajectory matrix, we have 100 eigentriples, ordered by their contributions (shares) into the decomposition stage. The leading eigentriple describes the general tendency of the series. Since in most cases the eigentriples with small shares are related to the noise component of the series, we need to identify the set of leading eigentriples. Let us consider the result of the SVD step.

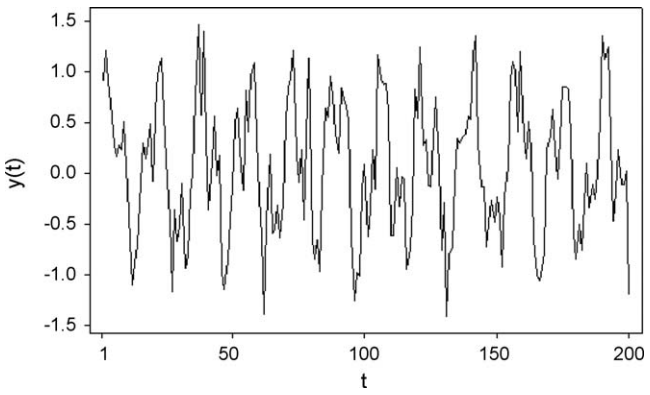

Fig. 2. Simulated series.

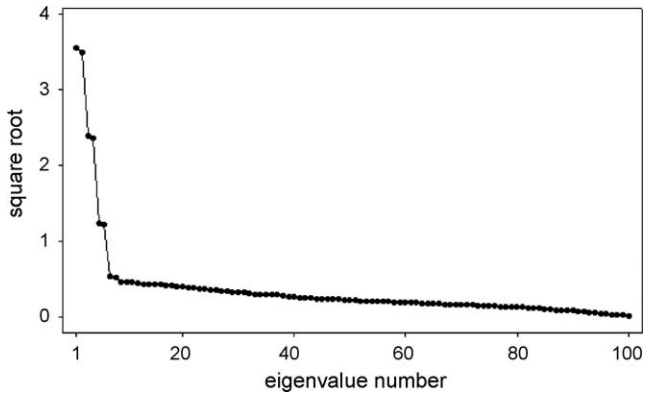

Fig. 3. Square roots of the 100 eigenvalues.

A useful insight is provided by checking breaks in the eigenvalue spectra. As a rule, pure noise produces a slowly decreasing sequence of singular values. If such noise is added to a signal, described by a few eigentriples with large singular values, then a break in the eigenvalue spectrum can distinguish the signal eigentriples from the noise. Fig. 3 depicts the plot of the square roots of the 100 eigenvalues for the simulated series.

Six evident pairs with almost equal leading singular values, correspond to five (almost) harmonic components of the series: eigentriple pairs 1–2, 3–4 and 5–6 are related to harmonics with specific periods.

The main concept in studying SSA properties is 'separability', which characterizes how well different components can be separated from each other. SSA decomposition of the series Y_N can only be successful if the resulting additive components of the series are separable from each other. The following quantity (called the weighted correlation or *w-correlation*) [30] is a natural measure of dependence between two series $Y^{(1)}$ and $Y^{(2)}$:

$$\rho_{12}^{(w)} = \frac{\langle Y^{(1)}, Y^{(2)} \rangle_w}{\|Y^{(1)}\|_w \|Y^{(2)}\|_w}$$

where $\|Y^{(i)}\|_w = \sqrt{\langle Y^{(i)}, Y^{(i)} \rangle_w}$, $\langle Y^{(i)}, Y^{(j)} \rangle_w = \sum_{k=1}^{N} w_k y_k^{(i)} y_k^{(j)} (i, j = 1, 2)$ and $w_k = \min\{k, L, N-k\}$ (here we assume $L \leq N/2$).

If two reconstructed components have zero *w*-correlation it means that these two components are separable. Large values of *w*-correlations between the reconstructed components indicate that the components should possibly be gathered into one group and correspond to the same component in the SSA decomposition.

Fig. 4 shows the *w*-correlations for the 100 reconstructed components in a 20-grade grey scale from white to black corresponding to the absolute values of correlations from 0 to 1.

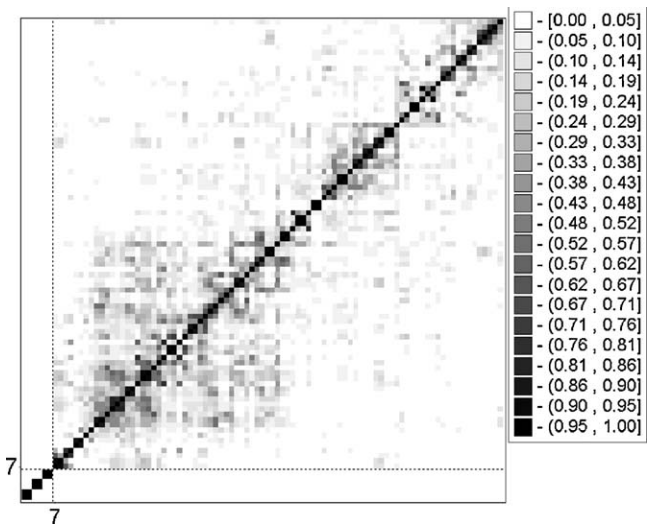

Fig. 4. Matrix of *w*-correlations for the 100 reconstructed components.

It is clearly seen that splitting of all the eigentriples into two groups, from the first to the sixth and the rest, gives rise to a decomposition of trajectory matrix into two almost orthogonal blocks, with the first block corresponding to the signal and the second block corresponding to noise.

3.1.2. Reconstruction: grouping and diagonal averaging

Reconstruction is the second stage of the SSA technique. As mentioned above, this stage includes two separate steps: grouping (identifying signal component and noise) and diagonal averaging (using grouped eigentriples to reconstruct the new series without noise).

The problem of finding a refined structure of a series by SSA is equivalent to identification of the eigentriples of the SVD of the trajectory matrix of this series, which corresponds to various oscillatory components and noise. From the practical point of view, a natural way of noise extraction is by grouping of the eigentriples, which do not seemingly contain elements of oscillations. Based on the above information, we thus classify eigentriple 7–100 as a part of noise. Fig. 5 shows the extracted noise obtained by considering the eigentriples 7–100.

As it appears from Fig. 5, the noise series, $n(t)$, consists of the added peaks. This confirms that the SSA technique can be used as a powerful method for detection of regular or dynamical peaks. Note that there is evidence that the signal extracted from the TMJ of an individual with TMD consists of such regular peaks during opening

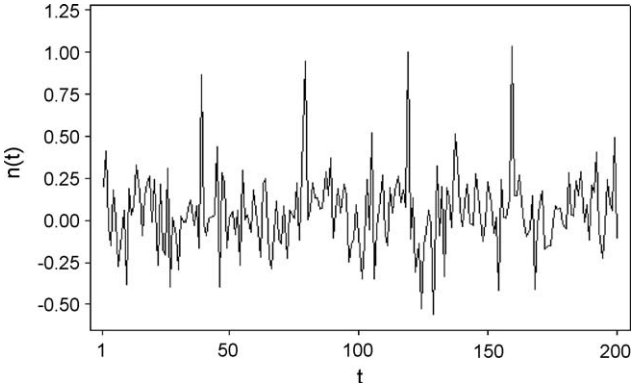

Fig. 5. Noise series (eigentriples 7–100).

and closing of the mouth. Therefore, beside the advantages of SSA mentioned in the Introduction, SSA can be considered as a suitable candidate to detect these peaks.

3.2. TMJ signal

We apply the SSA technique to TMJ signal to illustrate the capability of the technique to detect TMD. In this study, we used 14 subjects. We have four individuals with TMD in the left side of their faces and the rest are healthy subjects. We only represent the results for an individuals with TMD and a healthy subject. The patients used in our experiments were examined by our clinical expert collaborator.[2]

Different types of imaging systems may be utilized for diagnosis of TMD. Arthrography and magnetic resonance imaging (MRI) are the most popular ones. As the advantages of MRI over arthrography it is known that MRI is non-invasive, readily obtains multiplanar images in an infinite array of anatomic sections, and allows direct visualization of soft tissue components. However, MRI is costly and it cannot show the perforations of the posterior attachment or of the disc, and using these images it is difficult to assess accurate jaw position for the initiation or adjustment of splint therapy. Other imaging modalities such as computerised tomography are costly, hazardous and cause discomfort for the patients. Here we captured the video of subjects' faces from the left and right sides in frontal-lateral direction by two cameras (illustrated in Fig. 6).

Each subject was captured performing four cycles of chewing motion using two high resolution (640×480 pixels) color video cameras at 30 fps. On average, 400 video frames were obtained per subject. We placed four blue round markers at the locations of interest on each subject's face. The size of each marker is 6 mm. We used blue color because it was easy to detect in the images due to the contrast between it and the color of subjects' faces.

We attached two markers on the TMJ at the left and right side of the face. We also attached two additional markers of the same color on nose and chin level (illustrated in Fig. 7).

We used image processing methods to extract the position of the markers in video frames. The important locations with significant changes during mouth movement are around the TMJ. Thus we analysed the motion pattern of TMJs marker for each subject. Therefore, for each marker we obtained a time sequence representing its movement, in the horizontal direction, in the video sequences.

Fig. 8 shows the time sequence of a TMJ's marker for individual with TMD (left side) and healthy one (right side) in the original scale. As it appears from Fig. 8, there are different types of non-stationarity involved and it is not possible to detect which series is related to individuals with TMD through visual inspection of the signals.

3.2.1. Decomposition: window length and SVD

As mentioned before, for the simulated series L should be large enough but not greater than $N/2$. We take $L = 40$. Based on this window length and the SVD of the trajectory matrix (40×40), we have 40 eigentriples, ordered by their contribution (share) in the decomposition.

Fig. 9 shows the w-correlations for 40 reconstructed components in a 20-grade grey scale from white to black corresponding to the absolute values of correlations from 0 to 1. It is clearly seen that the splitting of all eigentriples into two groups, from the first to the fifth and the rest, gives rise to decomposition of the trajectory matrix into two almost orthogonal blocks, with the first block corresponding to the refined signal and the second block corresponding to noise.

[2] We are grateful to Prof. S. Dunne of the Kings College London, Dental Institute.

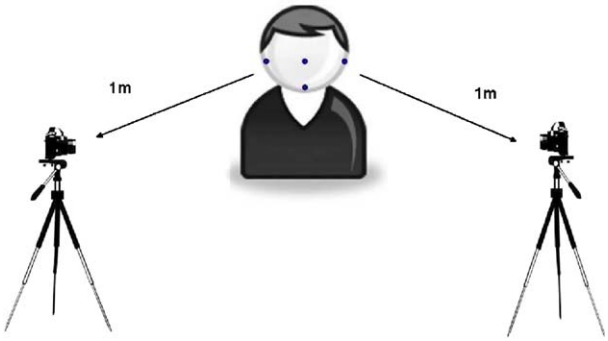

Fig. 6. The diagram of cameras and subject position.

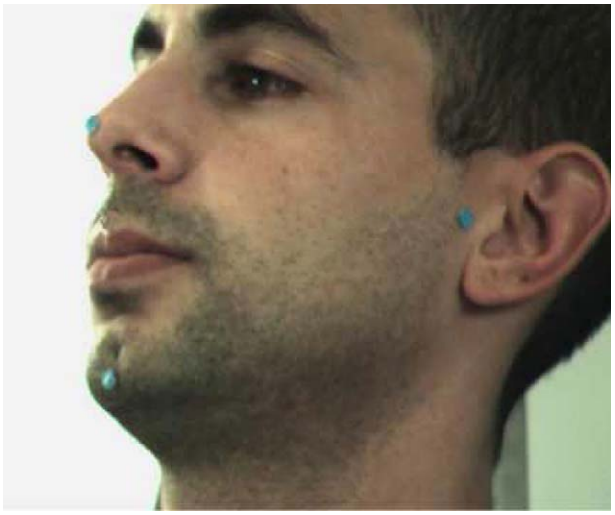

Fig. 7. The original image.

3.2.2. Reconstruction: grouping and diagonal averaging

Based on information given for the simulated data we consider the eigentriples 1–5 as signal component and eigentriples 6–40 as noise. Thus, we have two groups ($m = 2$).

Fig. 10 shows the noise series for an individual with TMD (left side) and a healthy subject (right side) after extracting the signal from the original series. Large peaks are clearly visible in Fig. 10 for the individual with TMD.

The analysis of chewing velocity pattern suggested that opening and closing patterns with obvious peak velocity was significantly more frequent in patients with TMD than healthy subjects. Here the behavior of the peaks is dynamic, confirming that the series with this structure in the noise series relates to the individuals with TMD. Note that the peaks appear in Fig. 10 are the points in the TMD signal when the individual has a problem during closing of the mouth. Therefore, we conclude that the noise with such a structure is related to individual with TMD. In fact, our detection method is based on the separated noise, not the signal, and the peaks are important observations.

3.3. Statistical test

To acquire a better understanding of the accuracy of the results, we designed the following statistical test. In many applications, it is important to detect the outliers, i.e., unusual abnormal values. Here outliers can be considered as the peaks in the noise series. In medicine, unusual values may indicate the diseases (see, e.g., [42]). One approach to outlier detection is to start with N observations

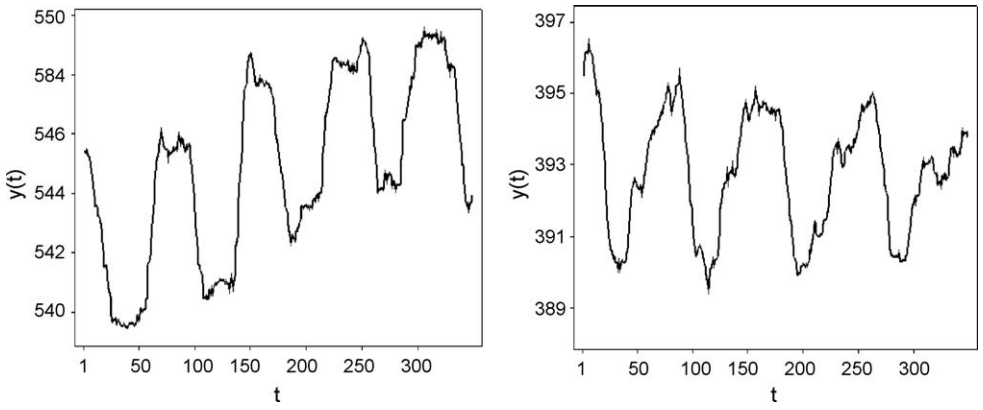

Fig. 8. Original time series of TMJ marker from an individual with TMD (in the left) and that of a healthy individual (in the right).

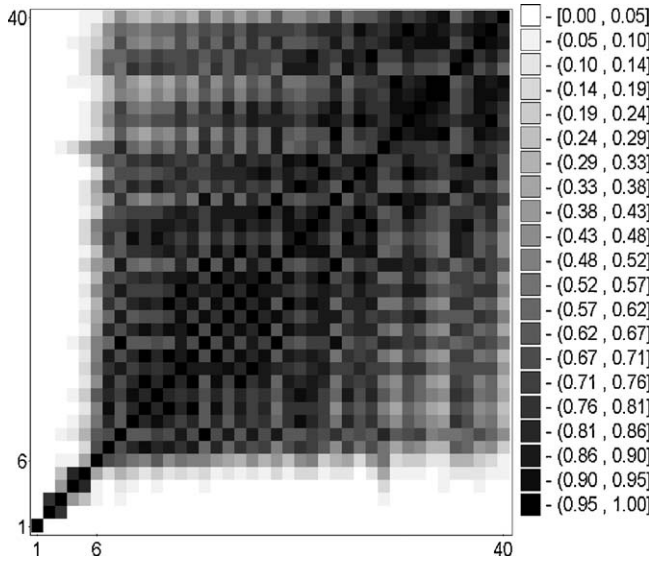

Fig. 9. Matrix of w-correlations for 40 reconstructed components.

$n_1, \ldots, n_N$, compute the sample average $\bar{n}$, the sample standard deviation σ, and then mark a value n_i as an outlier if n_i is outside the interval $(\bar{n} - a\sigma, \bar{n} + a\sigma)$ (for some preselected number a). We can therefore identify the outliers as those values that are outside the $a\sigma$ intervals (for an application of this method in engineering, see, e.g., [43]). Here, we selected $a = 3$. Unlike the mean, the median is not influenced by outliers at the extremes of the series. For this reason, the median often is used when there are a few extreme values. Here we also used median in place of mean as there are some outliers/extreme values in our observations. Let m_i for $i = 1, \ldots, N$ take values 1 if $|n_i| \geq 3\sigma$ in each period of opening and closing the mouth, and 0 otherwise. Therefore, variable m_i follows a Bernoulli distribution with parameter p. Now a Bernoulli experiment is performed n times. Here n is four as we consider only four periods, and also there are, in an ideal situation, two peaks in each period. The periods are mutually independent, and also p is a constant in each period. Therefore, $B = \sum_{i=1}^{n} m_i$ has a Binomial distribution with parameters n and p. The parameter p is estimated by

$$\hat{p} = \frac{1}{2n} \sum_{i=1}^{n} m_i \tag{14}$$

Now we are able to test whether $\hat{p}$ obtained from a noise series of an individual is statistically different from expected proportion p_0 or not. Therefore, we consider the following hypothesis

$$\begin{aligned} H_0 &: \quad \hat{p} = p_0 \\ H_1 &: \quad \hat{p} \neq p_0 \end{aligned} \tag{15}$$

In an ideal situation p_0 is equal to 1 for an individual with TMD as we are concerned with regular peaks; two peaks in each period. Thus, if the method recognizes the peaks properly then one should have $\hat{p} = p_0 = 1$. Otherwise, one needs to check whether the discrepancy is statistically significant or not. Here we assume that $p_0 = 0.95$. Table 1 shows the results of the Binomial test for 14 subjects. The first four rows are related to the individual with TMD, labeled TMD, and the rest are healthy subjects. The symbol $(*)$ in Table 1 indicates statistical significant difference with p-value < 0.05. As the results show, the $\hat{p}$ of healthy individuals are all significant at the 5% level. Therefore, we conclude that the detected

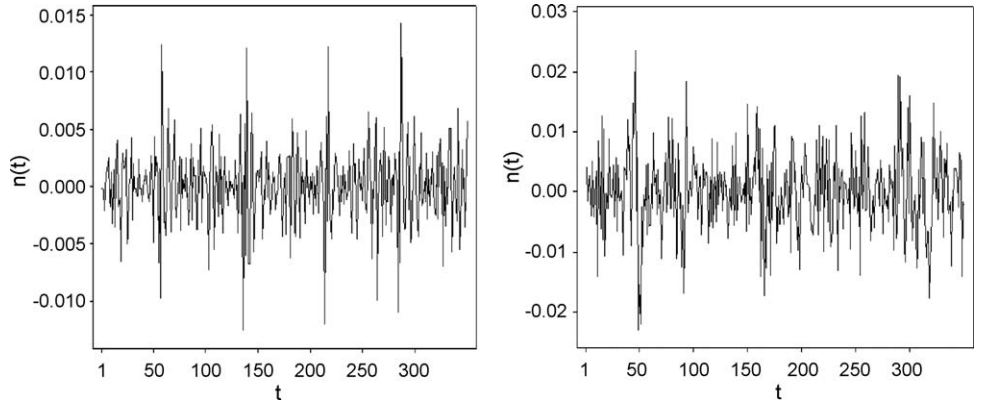

Fig. 10. Noise series (eigentriples 6–40) for an individual with TMD (in the left) and that of a healthy individual (in the right).

Table 1
The results of the binomial test.

N	Subject	Z	$\hat{p}$
1	TMD	8	1.00
2	TMD	8	1.00
3	TMD	7	0.88
4	TMD	6	0.75
5	Healthy	2	0.25*
6	Healthy	2	0.25*
7	Healthy	1	0.12*
8	Healthy	1	0.12*
9	Healthy	1	0.12*
10	Healthy	0	0.00*
11	Healthy	0	0.00*
12	Healthy	0	0.00*
13	Healthy	0	0.00*
14	Healthy	0	0.00*

* Indicates statistical significant difference with p-value <0.05.

peaks from noise series of healthy individual are not related to TMD.

4. Conclusions

In this paper we have established a method based on the SSA technique for detection of useful information in a noise series, and applied it as a method for TMD detection. The motion data of markers placed on the points of special interest on the faces of several subjects extracted and analysed using the proposed method.

The results illustrated that, in addition to time series analysis and forecasting, the SSA technique can be successfully used as a detection method, or at least for extraction of the main signal, and a filtering technique for processing of the biomedical time series with irregular behavior such as the data for TMD analysis.

The results also show that we can use the extracted noise for detection. In fact, we extract a useful pattern for classification of the individuals with TMD and healthy individuals using the noise series. Therefore, it is established that the extracted noise contains useful information which can be used for detection.

Acknowledgments

The authors would like to thank the editor and the referees for their constructive comments which have led to substantial improvements in this paper. The authors are grateful to Anatoly Zhigljavsky (Cardiff University), Vladimir Nekrutkin and Nina Golyandina (University of St. Petersburg) for useful suggestions.

References

[1] C. McNeill, Temporomandibular Disorders: Guidelines for Classification, Assessment and Management, Quintessence, Chicago, 1993, pp. 1–38.
[2] J.P. Okeson, Temporomandibular disorders in the medical practice, J. Family Pract. 43 (4) (1996) 347.
[3] http://www.daemen.edu/offices/grants/Images/TMJ.JPG.
[4] S.F. Dworkin, L. LeResche, Research diagnostic criteria for temporomandibular disorders: review, criteria, examinations and specifications, critique, J. Craniomandib. Disord. 6 (4) (1992) 301–355.
[5] J.P. Okeson (Ed.), Orofacial Pain. Guidelines for Assessment, Diagnosis, and Management, Quintessence Publishing Co., Inc., Chicago, 1996, pp. 116–117.
[6] R.J. Wilding, M. Shaikh, Muscle activity and jaw movements as predictors of chewing performance, J. Orofac. Pain 11 (1) (1997) 24–36.
[7] L.V. Christensen, M.N. Rassouli, Experimental occlusal interferences. Part V. Mandibular rotations versus hemimandibular translations, J. Oral Rehabil. 22 (12) (1995) 865–876.
[8] C.C. Took, S. Sanei, J. Chambers, S. Dunne, Underdetermined blind source separation of temporomandibular joint sounds, IEEE Trans. Biomed. Eng. 53 (2006) 2123–2126.
[9] M. Ghodsi, S. Sanei, Y. Hicks, T. Lee, S. Dunne, Detection of temporomandibular disorder from facial pattern, in: Proceedings of the 15th International Conference on Digital Signal Processing, Cardiff, UK, 2007.
[10] M. Ghodsi, S. Sanei, Y. Hicks, T. Lee, S. Dunne, A facial pattern recognition approach for detection of temporomandibular disorder, in: Proceedings of the EUSIPCO, Poznan, Poland, 2007.
[11] R. Wilding, A. Lewin, A computer analysis of normal humanmasticatory movements recorded with a sirognathograph, Arch. Oral Biol. 36 (1) (1997) 65–75.
[12] T. Kuwahara, R.W. Bessette, T. Maruyama, Chewing pattern analysis in TMD patients with unilateral internal and bilateral derangement, J. Cranio. 13 (3) (1995) 167–172.
[13] T. Kuwahara, C. Yoshioka, H. Ogawa, T. Maruyama, Effect of malocclusion on mandibular movement during speech, Int. J. Prosthodont. 7 (3) (1994) 264–270.
[14] D.S. Broomhead, J.P. Huke, R. Muldoon, Linear filters and non-linear systems, J. R. Stat. Soc. Ser. B: Methodol. 54 (1992) 373–382.
[15] T. Schreiber, P. Grassberger, A simple noise reduction method for real data, Phys. Lett. A 160 (1991) 411–418.
[16] J.D. Farmer, J. Sidorowich, Optimal shadowing and noise reduction, Physica D 47 (1991) 373–392.
[17] R. Vautard, P. Yiou, M. Ghil, Singular-spectrum analysis: a toolkit for short, noisy chaotic signals, Physica D 58 (1992) 95–126.
[18] T. Sauer, A noise reduction method for signals from nonlinear systems, Physica D 58 (1992) 193–201.
[19] F. Takens, Detecting strange attractors in turbulence, Lect. Notes Math. 898 (1981) 365–381.
[20] D.S. Broomhead, G.P. King, Extracting qualitative dynamics from experimental data, Physica D 20 (1986) 217–236.
[21] E.J. Kostelich, T. Schreiber, Noise reduction in chaotic timeseries data: a survey of common methods, Phys. Rev. E 48 (1993) 1752–1763.
[22] D.S. Broomhead, J.P. Huke, R. Jones, Signals in chaos: a method for the cancellation of deterministic noise from discrete signals, Physica D 38 (1995) 423–432.
[23] S.H. Jensen, P.C. Hansen, S.D. Hansen, J. Sorensen, Reduction of broadband noise in speech by truncated QSVD, IEEE Trans. Speech Audio Process. 3 (1995) 439–448.
[24] Y. Ephraim, H.L. Van Trees, A signal subspace approach for speech enhancement, IEEE Trans. Speech Audio Process. 3 (1995) 251–256.
[25] U. Mittal, N. Phamdo, Signal/noise KLT based approach for enhancing speech degraded by colored noise, IEEE Trans. Speech Audio Process. 8 (2001) 159–167.
[26] A. Rezayee, S. Gazor, An adaptive KLT approach for speech enhancement, IEEE Trans. Speech Audio Process. 9 (2001) 87–95.
[27] Y. Hu, P. Loizou, A generalized subspace approach for enhancing speech corrupted by colored noise, IEEE Trans. Speech Audio Process. 11 (2003) 334–341.
[28] H. Lev-Ari, Y. Ephraim, Extension of the signal subspace speech enhancement approach to colored noise, IEEE Signal Process. Lett. 10 (2003) 104–106.
[29] F. Jabloun, B. Champagne, Incorporating the human hearing properties in the signal subspace approach for speech enhancement, IEEE Trans. Speech Audio Process. 11 (2003) 700–708.
[30] N. Golyandina, V. Nekrutkin, A. Zhigljavsky, Analysis of Time Series Structure: SSA and Related Techniques, Chapman & Hall/CRC, New York/London, 2001.
[31] F.J. Alonsoa, J.M. Del Castilloa, P. Pintado, Application of singular spectrum analysis to the smoothing of raw kinematic signals, J. Biomech. 38 (2005) 1085–1092.
[32] W.C.A. Pereira, S.L. Bridal, A. Coron, P. Laugier, Singular spectrum analysis applied to backscattered ultrasound signals from in vitro human cancellous bone specimens, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 51 (3) (2004) 302–312.
[33] A.R. Teixeira, A.M. Tome, E.W. Lang, P. Gruber, A. Martins da Silva, On the use of clustering and local singular spectrum analysis to remove ocular artifacts from electroencephalograms, in: Proceeding of the International Joint Conference on Neural Networks, Montreal, Canada, 2005.
[34] D. Danilov, A. Zhigljavsky (Eds.), Principal Components of Time Series: The 'Caterpillar' Method, University of St. Petersburg, St. Petersburg, Russian, 1997.
[35] R. Vautard, P. Yiou, M. Ghil, Singular-spectrum analysis: a toolkit for short, noisy chaotic signals, Physica D 58 (1992) 95–126.
[36] V.G. Moskvina, A. Zhigljavsky, An algorithm based on singular spectrum analysis for change-point detection, Commun. Stat.: Simul. Comput. 32 (2003) 319–352.
[37] H. Hassani, Singular spectrum analysis: methodology and comparison, J. Data Sci. 5 (1) (2007) 239–257.
[38] H. Hassani, S. Heravi, A. Zhigljavsky, Forecasting European industrial production with singular spectrum analysis, Int. J. Forecasting., Forthcoming, 75 (2008) 1131–1142.
[39] H. Hassani, A. Dionisio, M. Ghodsi, The effect of noise reduction in measuring the linear and nonlinear dependency of financial markets, Nonlinear Anal.: Real World Applications, Forthcoming.
[40] H. Hassani, A. Zhigljavsky, Singular spectrum analysis: methodology and application to economics data, J. System Sci. Complexity, Forthcoming.
[41] G.H. Golub, C.F. van Loan, Matrix Computations, third edition, The Johns Hopkins University Press, Baltimore, MD, 1996.
[42] A. Mohamad-Djafari (Ed.), Bayesian inference for inverse problems, in: Proceedings of the SPIE/International Society for Optical Engineering, San Diego, CA, 1998, pp. 128–135.
[43] H.M. Wadsworth, Jr. (Ed.), Handbook of Statistical Methods for Engineers and Scientists, McGraw-Hill, NY, 1990.