Beta Iterative Synchronization: An Algorithm for Structural Signal Averaging

DAMIAN MIRECKI\textsuperscript{1}, TOMASZ MELCER\textsuperscript{2}, CEZARY SIeluŻYCKI\textsuperscript{2}, and MONIKA E. DANIELEWSKA\textsuperscript{2}

\textsuperscript{1}Department of Computer Science, Faculty of Fundamental Problems of Technology, Wrocław University of Science and Technology, Wrocław, Poland
\textsuperscript{2}Department of Biomedical Engineering, Faculty of Fundamental Problems of Technology, Wrocław University of Science and Technology, Wrocław, Poland

Corresponding author: Damian Mirecki (e-mail: dmirecki.dmm@gmail.com).

This research was funded by the National Centre for Research and Development grant number LIDER/074/L-6/14/NCBR/2015 and was supported in part by PLGrid Infrastructure.

ABSTRACT Many biomedical signals can be considered as sets of repetitions due to the occurrence of a repetitive pattern of features. Those features, however, are characterized by some jitter, which often renders standard arithmetic averaging inadequate. Examples include electrocardiogram, ocular pulse or corneal pulse, series of evoked potentials in electro- and magnetoencephalography, among many others. We propose a new approach to structural averaging of such signals. We use the family of beta cumulative distribution functions as a set of candidates for time warping function in order to synchronize repetitions, and then apply a variant of the Procrustes method to find the average signal. For both synthetic and real data, we provide a comparison where we challenge the Dynamic Time Warping (DTW) method and present both theoretical and practical advantages of our algorithm. As an illustrative real-data example we address corneal pulse waveforms with their dicrotic valleys as the feature of interest. The detection of the dicrotic valleys turned out more reproducible than in the case of DTW while maintaining similar classification performance and having fewer parameters. The proposed method for structural averaging provides effective estimation in the case of the analyzed signals. The method can readily be extended to other biomedical signals characterized by repetitive feature patterns.

INDEX TERMS Beta distribution, Dynamic Time Warping (DTW), iterative synchronization, ocular dicrotism, signal averaging.

I. INTRODUCTION

The motivation for this work was the commonly acknowledged problem of slight changes in frequency of appearances of repetitive patterns in many biomedical signals. Examples of such signals include electrocardiogram (ECG), blood pulse (BPL), ocular pulse (OP) with the related corneal pulse (CP), and series of evoked potentials in electro- (EEG) and magnetoencephalography (MEG). It is known that simple arithmetic averaging of the cycles of such signals in order to, for example, increase the signal-to-noise ratio (SNR) or to find the best representative set of features, results in losing important information [1]. This is caused by the fact that those features are either lost or at best smeared out in the process of averaging due to their intrinsic temporal jitter from cycle to cycle (ECG, OP, CP) or from trial to trial when repetitive stimulation is used in EEG or MEG in order to find estimates of the so-called evoked brain responses.

To avoid this and other problems associated with simple averaging, a common approach is to model each repetition of a cycle, or each new trial, as a new realization of a subject-specific random process defined by not only some kind of an expected, typical or average signal shape common to all realizations, but also some assumed signal variations that can plausibly occur in different realizations, e.g., signal dilation, compression, shifts, amplification, modulation, partial concealment, etc. In this view, a set of realizations of a single random process can be used to estimate the average signal shape using a class of methods broadly called signal averaging or structural average estimation. Then, the morphological...
parameters of the recovered waveform are used for further analysis.

Often, as an intermediate step, a set of repetitions is marked to pair up similar features that occur at different times in different repetitions. This procedure is called signal matching.

Many methods for signal matching and structural averaging have been proposed. A popular approach is to use a parametrized class of continuous functions to represent time transformations. Examples include affine transforms [2]–[4] and splines [1]. Flexible non-parametric methods such as Self-Modelling Non-linear Regression [3], [5] and methods based on functional analysis [1] have also been studied. A different class of algorithms is based on pairing up samples of discrete signals, e.g., Dynamic Time Warping (DTW) [6] and its variants [7]. In cases where a method is based on synchronizing a set of realizations using a pairwise distance measure between signals, the Procrustes method may be used to approach a solution that minimizes the distance from the estimated average to each realization [1]. A related approach to warping uses a beta distribution to reparametrize the input space for Bayesian optimization [8].

Signal variations that a given method can model define its inductive bias [9], in this case, what estimates are considered plausible and how flexibly a method can fit an estimate to the sample repetitions. Simple methods may not be flexible enough, that is, they may fail to represent signal variations considered in a given domain (e.g., affine transforms being unable to represent compression of only part of a signal). On the other hand, over-flexible methods (e.g. DTW) may find solutions that can be implausible (e.g. exaggerated dilations of local features). By this reasoning, an adaptive method that can be tuned for a specific application is better than a general method [10]. Other common flaws of signal averaging methods include difficulties with identifiability (lack of a unique solution), problems with convergence or computational complexity [1], [3], [11].

The following work introduces a new method of structural average estimation. The method is based on a pairwise synchronization using a cumulative distribution function from the beta family to represent time warping. A variant of the generalized Procrustes analysis [12] is then used to compute the estimate of a typical signal. We study the method and compare it to the widely used DTW algorithm using a data set of corneal pulsation signals as an illustrative study.

II. METHODS

A. PRELIMINARIES

The considered setting for the proposed signal averaging is based on the following assumptions.

1) A certain repetitive physical process is measured at a sampling interval that allows recognition of features of each repetition; this results in a digital signal that has repetitive features that are visually recognizable (e.g., CP: a single 10-second recording at 400 Hz may contain 8 to 12 full repetitions of a heart cycle, subsequent cycles have visibly similar signal shapes).
2) Repetition of features occur at roughly constant pace; each repetition interval may have slightly different length (e.g., CP: typical cycle length variability under measurement conditions is around 5%).
3) Each repetition interval exhibits inherent variability in the form of noisy amplitude, compression or dilation of parts of the signal (often, but not always, accompanying the cycle length variability). The variability is in most cases not significant enough to make it difficult for an expert to recognize signal features, as in the case of a cardiologist evaluating the ECG signal morphology. It is also rare, though possible, that a feature disappears from some repetitions (e.g., a missing QRS complex in ECG).
4) There are long-term trends (e.g., breathing induces slow head movements that are visible as trends over many repetitions; EEG, MEG, CP).
5) A small percentage of repetitions are outliers. The outliers are easy to identify and can be removed before signal averaging (e.g., blinking significantly obscures parts of recordings, but it is easy to detect; EEG, MEG, CP).
6) Identification of starting and ending points of repetitions may be imprecise.

We present a method for obtaining an average signal shape that takes the above assumptions into account. The method considers the morphology of the cumulative distribution function of a beta distribution (CDF_B) as a means for time warping between the time domain of the estimated average and that of the collected realizations of a process. In this view, we do not exploit the probabilistic aspects of CDF_B.

B. SIGNALS MATCHING

Let us denote the analyzed multi-repetition signal by s, where a repetition is understood as a part of s between two consecutive time instants that correspond to events that start cycles. In the following sections the signal s will be considered as a real-valued function.

Consider a repetition f within signal s and a reference g modeling the average repetition from s. Synchronization of f and g is a transformation of f by local temporal changes in a way that morphological properties of f and g appear in the same time. Formally, let us consider continuous functions f, g : [0; 1] → ℝ corresponding to the repetition and the reference. The goal is to find a time transformation function h that minimizes the error function:

$$\delta_c(f, g; h) = \int_0^1 \left( f(h(t)) - g(t) \right)^2 dt,$$  \hspace{1cm} (1)

where subscript c stands for continuous. We will therefore seek $\arg \min_h \delta_c$. 

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/.
C. BETA DISTRIBUTION

In principle, any reasonable function can be chosen for \( h \). In this paper we consider using cumulative distribution functions from the beta distribution family (CDF\(_\text{B} \)) for \( h \):

\[
h \in \{ \text{CDF}\(_\text{B} \)(; \alpha, \beta) : \alpha, \beta \in (0; \infty) \}.
\]

(2)

The CDF\(_\text{B} \) is equal to the regularized (note the denominator) incomplete (note that \( t \leq 1 \)) beta function

\[
\text{CDF}\(_\text{B} \)(t; \alpha, \beta) = \frac{\int_0^t u^{\alpha-1}(1-u)^{\beta-1}du}{B(\alpha, \beta)} \quad t \in [0, 1],
\]

(3)

where

\[
B(\alpha, \beta) = \int_0^1 v^{\alpha-1}(1-v)^{\beta-1}dv
\]

(4)

is the complete beta function.

The beta distribution has a few properties [13] useful for time synchronization:

1) CDF\(_\text{B} \) : [0; 1] \( \xrightarrow{1:1} \) [0; 1] (all points of both repetitions are synchronized),
2) CDF\(_\text{B} \) is continuous on [0; 1] (no time skips),
3) CDF\(_\text{B} \) is nondecreasing (no going back in time),
4) the family of CDF\(_\text{B} \) has only two parameters, \( \alpha \) and \( \beta \),
5) for \( \alpha = 1 \) and \( \beta = 1 \) the distribution is uniform on [0; 1] (no time shifts at all),
6) \( \alpha \) and \( \beta \) can be interpreted as controlling relative compression of the initial and final part of the signal.

The proposed signal averaging method aims at a sufficient flexibility of the time warping function on the one hand (property 6), and at a parsimonious representation of that function with possibly few parameters on the other hand (property 4). The latter is advantageous because a large number of parameters can render a method prone to noise via, e.g., over-fitting. Fig. 1 demonstrates time transformations with the original time spanning along abscissae and the transformed time spanned along ordinates for several values of \( \alpha \) and \( \beta \).

Let us emphasize that in this paper, CDF\(_\text{B} \) serves exclusively as a time warping function, with no probabilistic interpretation. \( t \) denotes time which is subject to transformation by either squeezing (compression) or stretching (dilation) in order to provide flexible mutual matching of the analyzed signals in time.

D. DISCRETE SIGNALS

Let us define the standardized domain of a discrete-time repetition \( c \), whose length is equal to \( T \) samples, as

\[
T = \left\{ 0, \frac{1}{T-1}, \frac{2}{T-1}, \ldots, \frac{T-2}{T-1}, \frac{T-1}{T-1} \right\}
\]

(5)

and the standardized function, denoted by \( \sigma \)

\[
\sigma : T \rightarrow \mathbb{R},
\]

\[
\sigma(t) = c ((T - 1) t + 1) .
\]

(6)

(7)

So far only continuous functions (signals) \( f, g, \delta_c \) were considered. However, signals analyzed in practice are discrete. Therefore, it is necessary to introduce a discrete version of \( \delta_c \). Let us consider standardized functions \( \sigma_f \) and \( \sigma_g \) and assume \( n_f = n_g = n \) for the time moments. Now, the discrete version of (1) can be introduced:

\[
\delta_d(\sigma_f, \sigma_g; h) = \sum_{t \in T} \left( \sigma_f(h(t)) - \sigma_g(t) \right)^2.
\]

(8)

Note that \( h \) is a continuous function, whereas \( \sigma_f \) is a discrete-time signal. Therefore, we must interpolate \( \sigma_f \) in the new discrete time points after the transformation \( h \). For (8), we approximate \( \sigma_f \) using linear interpolation.

We also attempted to use kriging [14] instead of linear interpolation for computing the values of \( \sigma_f \) in (8). However, we found no differences in either the optimal values of the loss function or the performance of other experiments described hereafter, and hence settled on linear interpolation as faster to compute.

E. CHOOSING PARAMETERS FOR CDF\(_\text{B} \)

The goal is to find the values of the two CDF\(_\text{B} \) parameters, \( \alpha \) and \( \beta \), that minimize \( \delta_d \). In this way, the problem of synchronization is reduced to an optimization problem.

The transformations applied to the signals should stay within boundaries plausible to the field of application. The error function \( \delta_d \) is difficult to optimize (see Fig. 2 for an example). Therefore, in order to minimize \( \delta_d \) we employ a brute-force grid search, i.e., select plausible values of \( \alpha \) and \( \beta \) by domain-specific prior knowledge and try all combinations of these values.
F. ITERATIVE SYNCHRONIZATION FOR FINDING THE REPRESENTATIVE

Let \( C = \{c_1, c_2, \ldots, c_k\} \) be the set of \( k \) repetitions, and assume that for \( j \in \{1, \ldots, k\} \) each \( c_j \) is defined on the same time domain with time instants \( t \in \{1, \ldots, w\} \). We will also define \( m = \{c_1, \ldots, c_k\} \), i.e., the arithmetic mean across all repetitions.

Having a method that enables synchronization of a repetition with a reference (see (8)), we propose an iterative synchronization algorithm for deriving an estimate of a typical repetition (BIS, standing for Beta Iterative Synchronization) from the set \( C \); see Algorithm 1. The BIS approach resembles the generalized Procrustes analysis [1]. BIS performs three steps in a loop:

1) computes the arithmetic mean \( m(t) \) across the set \( C' \) of repetitions (line 7),
2) performs synchronization between \( c_j \) and \( m \) to find \( c'_j \), i.e. the transformed repetition (lines 8–11),
3) computes the total difference between \( c'_j \) and \( m \) (line 13).

The algorithm stops when the difference between errors in two consecutive loop executions is smaller than some constant \( \varepsilon \). In our experiments we set the stopping criterion to \( \varepsilon = 0.05 \). As shown in Appendix, BIS always converges monotonically to a constant.

III. RESULTS

A. SYNTHETIC SIGNALS

We compare the proposed method with its analogue based on DTW instead of CDFB, on a set of synthetic signals with known characteristics.

The two methods differ in the form of line 9 (CDFB versus DTW) in Algorithm 1, the rest of the function remaining unchanged. This analogue version would be called DTW-IS, for iterative synchronization employing DTW. Synthetic signals have been prepared according to the following procedure. Firstly, we set the number of repetitions \( k = 10 \). Each repetition consists of 100 samples; e.g., samples \( x[501], \ldots, x[600] \) compose the sixth repetition. A signal is constructed as:

\[
x[\tau] = \sum_{i=1}^{k} N(\tau - 10(i - 1); \mu_i, \max(\sigma_i, 10))
\]

\[
+ \max(\theta_i, 0) \sum_{i=1}^{k} N(\tau - 10(i - 1); \mu_i, \max(\sigma_i, 0))
\]

\[
+ a\tau^2 + b\tau + \varepsilon[\tau]
\]

for \( \tau \in \{1, \ldots, 1000\} \), where \( N(\tau; \mu, \sigma) \) is the value of the probability density function of a Gaussian distribution with expected value \( \mu \) and standard deviation \( \sigma \). \( \varepsilon \) is either an uncorrelated or correlated noise, the latter being created from an uncorrelated noise by means of a correlating window of the width \( w = 10 \). For each repetition \( i \in \{1, \ldots, k\} \), the characteristics of a generated signal are sampled according to the following equations:

\[
\theta_i \sim \mathcal{N}(0.5, 0.5), \quad \mu_1 \sim \mathcal{N}(25, 10), \quad \mu_2 \sim \mathcal{N}(70, 10),
\]

\[
\sigma_1 \sim \mathcal{N}(15, 5), \quad \sigma_2 \sim \mathcal{N}(15, 5),
\]

\[
a \sim \mathcal{N}(0, 10^{-5}), \quad b \sim \mathcal{N}(0, 2 \cdot 10^{-8}),
\]

where \( \mathcal{N}(\mu, \sigma) \) is a Gaussian distribution with mean \( \mu \) and standard deviation \( \sigma \). An example of such a signal is shown in Fig. 3.

We computed the Euclidean distance between the result of the structural signal averaging method and an ideal signal \( y \) defined by (11) in order to assess the quality of the match. The smaller the distance the better the match.

Algorithm 1: The iterative synchronization algorithm for deriving the representative repetition (BIS) from the set of repetitions of a multi-repetition signal.

1: function BIS(C)
2: \( C = \{c_1, \ldots, c_k\} \) is the set of repetitions.
3: \( e_1 \leftarrow \infty \)
4: \( e_2 \leftarrow \infty \)
5: \( C' \leftarrow C \)
6: repeat
7: \( m \leftarrow \overline{C'} \)
8: for \( j = 1 \) to \( k \) do
9: \( \alpha, \beta \leftarrow \text{argmin}_{\alpha, \beta} \delta(c_j; m; h, \alpha, \beta) \)
10: \( c'_j \leftarrow h(c_j; \alpha, \beta) \)
11: end for
12: \( e_1 \leftarrow e_2 \)
13: \( e_2 \leftarrow \sum_{j=1}^{k} \delta(c'_j, m) \)
14: until \( e_1 - e_2 \leq \varepsilon \)
15: return \( m \) is the representative.
16: end function
B. CORNEAL PULSE SIGNALS

Corneal pulsation is defined as temporal variations in corneal expansion due to the choroidal blood volume pulsation and changes in intraocular pressure [16]–[18]. In 2014, utilizing a new non-contact ultrasonic technique allowed, for the first time, registering a characteristic double-peak shape in corneal pulse (CP) signal for one heart cycle, named the ocular dicrotic pulse (ODP) [19]. This newly observed phenomenon, termed ocular dicrotism, was observed in more than 70% of healthy subjects over 50 years old and was found to increase with age [19], [20]. In addition, it was shown that with advancing age, the incidence of ocular dicrotism was higher in glaucomatous eyes than in normal eyes [20].

To aid the study of ocular dicrotism, an automated method for detecting the presence of ODP is sought. The approach of using statistics based on the frequency domain to detect features in a periodic-like signal [21] is hampered by the non-stationarity of the CP signal caused by heart rate variability and the respiratory sinus arrhythmia [22]. Previous attempts of overcoming this problem include using a Dynamic Time Warping algorithm to synchronize repetitions [21] and using a wavelet transform to analyze the signal without explicitly modeling repetitions [23].

In [23] it was shown that by using the wavelet transform it is possible to perform detection of the ODP signal from the CP signal alone, i.e., without the need to acquire additional cardiovascular signals. For this purpose, agglomeration of a set of features derived from the Continuous Wavelet Transform representation of the CP signal resulted in a reliable scheme that achieved detection rates reaching 93%. However, the collected features were not physically interpretable.

The illustrative data set considered in this work consists of signals that were obtained from a 10-second acquisition at the sampling frequency of 400 Hz. We have used retrospective data from the study approved by the Ethics Committee of Wrocław Medical University (decision No. KB 503/2011) and adhered to the Tenets of the Declaration of Helsinki. In that study, the corneal pulse (CP) signal was measured synchronously with the electrocardiogram (ECG), where the latter was used exclusively to segment the CP signal into repetitions based on the R peaks. Each CP signal was classified manually by an expert as dicrotic or non-dicrotic. The data set consists of 458 pairs of signals (CP, ECG) measured from 85 individuals. There were 375 dicrotic and 63 non-dicrotic recordings. In the remaining 20 signals, the dicrotism could not be determined unambiguously by means of visual inspection by an expert.

A few steps of preprocessing are performed. First, all analyzed signals are detrended with a 1st-degree polynomial and then detrended signals are filtered from 0.5 Hz to 20 Hz. This filtering aims at removing frequencies below 0.5 Hz related to breathing modulation and those that may contaminate the morphology of the QRS complex of the ECG signal. Repetitions that contain eye blink artifacts are removed. Finally, discrete derivatives of repetitions are used for synchronization as in [1].
Below, we present a few methods that were developed and used to compare the BIS, DTW-IS, and DTW algorithms. All three algorithms require to define a meta-parameter—the threshold (depth of the dicrotic valley) that is used in the classification criterion. In a previous publication [19] a 3-dB threshold was used. Unfortunately, a threshold defined this way is sensitive to the amplitude of an entire signal. For example, two CP signals can both be dicrotic but if one has significantly smaller variance the 3-dB criterion may not give good results because two signals with significantly different norms are not comparable in the framework of this criterion. To tackle this problem we propose to first normalize the L2 norm of each signal to 1. After that, we can measure the dicrotic valley depth in decibels.

For the purpose of choosing a proper threshold (dicrotic valley depth) for the three compared algorithms, receiver operating characteristic curves (ROC) [25] were plotted (see Fig. 8 and Fig. 9). To compare BIS, DTW-IS, and DTW, two thresholds were chosen for each of the three algorithms: 1) one that maximizes TPR (true positive rate) for FPR (false positive rate) not greater than 0.05, 2) another one, which

For example, (12a) forbids strong warping of the point at 1/6th of the signal. The set of $\alpha$ and $\beta$ that satisfies the above inequalities forms a kite-like shape that is symmetric with respect to the line $\alpha = \beta$ (see Fig. 7).

\[
\begin{align*}
\text{CDF}_B \left( \frac{1}{6}; \alpha, \beta \right) &< \frac{1}{3}, \quad (12a) \\
\text{CDF}_B \left( \frac{5}{6}; \alpha, \beta \right) &> \frac{2}{3}, \quad (12b) \\
\text{CDF}_B \left( \frac{1}{3}; \alpha, \beta \right) &> \frac{1}{6}, \quad (12c) \\
\text{CDF}_B \left( \frac{2}{3}; \alpha, \beta \right) &< \frac{5}{6}. \quad (12d)
\end{align*}
\]
TABLE 1. TPR and FPR for selected thresholds.

<table>
<thead>
<tr>
<th></th>
<th>TPR</th>
<th>FPR</th>
<th>threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIS0.05</td>
<td>0.356</td>
<td>0.039</td>
<td>2.7</td>
</tr>
<tr>
<td>DTW-IS0.05</td>
<td>0.34</td>
<td>0.044</td>
<td>2.46</td>
</tr>
<tr>
<td>DTW0.05</td>
<td>0.359</td>
<td>0.033</td>
<td>2.8</td>
</tr>
<tr>
<td>BISY</td>
<td>0.847</td>
<td>0.294</td>
<td>1.5</td>
</tr>
<tr>
<td>DTW-ISY</td>
<td>0.798</td>
<td>0.267</td>
<td>1.26</td>
</tr>
<tr>
<td>DTWY</td>
<td>0.847</td>
<td>0.233</td>
<td>1.18</td>
</tr>
</tbody>
</table>

maximizes the popular Youden’s index [26]. The algorithms with so-obtained thresholds will be denoted as BIS0.05, DTW-IS0.05, and DTW0.05 for the former, and BISY, DTW-ISY, and DTWY for the latter case. The obtained thresholds with appropriate values of TPR and FPR are shown in Table 1.

In our experiments, BIS usually converges (see line 14 in Algorithm 1) after about seven iterations.

C. DICROTIC VALLEY DISPERSION MEASURE

A good synchronization method should be robust in the sense of insensitivity to a particular selection of repetitions of a signal. This robustness was measured by the following routine. Suppose there is a signal \( s_i \) divided into repetitions \( C_i = \{c_1, \ldots, c_{k_i}\} \). Let us define \( C_{i-j} = C_i \setminus \{c_j\} \). For each \( C_{i-1}, \ldots, C_{i-k_i} \) one can obtain representatives \( r_{i-1}, \ldots, r_{i-k_i} \) and measure the dicrotic valley depth for each representative. In the absence of a dicrotic waveform, the depth is considered to be 0. Let us denote the set of dicrotic valley depths as \( D_i \). Let also \( z_i = \max D_i - \min D_i \) be the difference between the depth of the deepest and shallowest valley. Then, for a set of \( n \) signals one will obtain the set of dispersion values \( Z = \{z_1, \ldots, z_n\} \). We can then calculate the mean and standard deviation over \( Z \). The lower the mean the more robust a method of detection is. A comparison between BIS, DTW-IS, and DTW is presented in Table 2 and the histograms of values in Fig. 10. The BIS method has lower mean and standard deviation than both DTW-IS and DTW, which indicates that BIS is more robust. Thus, the other two algorithms are more sensitive to small perturbations.

D. CLASSIFICATION DISCREPANCY MEASURE

Similarly to the strategy presented in Sec. III-C the representatives \( r_{i-1}, \ldots, r_{i-k_i} \) are obtained. Then, each of such representatives is classified as dicrotic or non-dicrotic. The efficiency of a method can also be measured by counting the ratio of the number of representatives that were classified as dicrotic and the \( k_i \) number. Let us define

\[
u_i = \left\lfloor \frac{|\{j : r_j \text{ was classified as dicrotic}\}|}{k_i} \right\rfloor.
\]  

(13)

The interpretation of the \( u_i \) value is as follows. If \( u_i = 0 \) then all representatives \( r_{i-1}, \ldots, r_{i-k_i} \) were classified as non-dicrotic, whereas if \( u_i = 1 \) than all representatives were classified as dicrotic. The best possible option is that \( u_i \) is

![FIGURE 7](image)

FIGURE 7. Parameter space (shaded area) that satisfies constraints (12).

![FIGURE 8](image)

FIGURE 8. ROC curve that compares BIS, DTW-IS and DTW methods. Solid, dashed and dotted are for BIS, DTW-IS and DTW, respectively.

![FIGURE 9](image)

FIGURE 9. Enlargement of ROC curve that exposes low FPR values fragment.
very close to 0 or 1. For unification, the following value will be used instead of \( u_k \)

$$q_i = \frac{1}{2} - \left| \frac{1}{2} - u_k \right|,$$

(14)

where \( q_i \) indicates the discrepancy measure in the classification of the representative. If \( q_i = 0 \) then the classification method assigns the same label to all representatives. If \( q_i = 0.5 \) then half of representatives were classified as dicrotic and half as non-dicrotic. The histograms of values \( q_i \) for \( i \in \{1, \ldots, n\} \) for all considered methods are demonstrated in Fig. 11, whereas arithmetic mean and standard deviation of those values are shown in Table 3. For the threshold that maximizes TPR while keeping FPR below 0.05, the results indicate that DTW\(_{0.05}\) has lower mean and standard deviation than DTW-IS\(_{0.05}\) or BIS\(_{0.05}\). However, for the other threshold, both DTW\(_Y\) and DTW-IS\(_Y\) have these values larger than those of BIS\(_Y\). This shows the dependence of the actual robustness on the selected type of threshold.

E. WITHIN-PERSON CLASSIFICATION CONSISTENCY

There are on average three CP signals for each eye and person in the data set considered in this paper. The signals were grouped with respect to eye and person and it was measured how often the compared methods classify signals from the same group to the same label. For a given method (e.g., BIS) and a given group (e.g., the left eye of a given individual) the results will be considered as consistent if the method classified all such signals as dicrotic or non-dicrotic. Otherwise, the result is denoted as inconsistent. We count the number of groups for which the results are consistent. Table 4 demonstrates the results. Again, the consistency depends on the selected type of threshold.

<table>
<thead>
<tr>
<th>Method</th>
<th>Consistent groups of results</th>
<th>Inconsistent groups of results</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIS(_{0.05})</td>
<td>67 (45%)</td>
<td>81 (55%)</td>
</tr>
<tr>
<td>DTW-IS(_{0.05})</td>
<td>77 (52%)</td>
<td>71 (48%)</td>
</tr>
<tr>
<td>DTW(_{0.05})</td>
<td>59 (40%)</td>
<td>89 (60%)</td>
</tr>
<tr>
<td>BIS(_Y)</td>
<td>61 (41%)</td>
<td>87 (59%)</td>
</tr>
<tr>
<td>DTW-IS(_Y)</td>
<td>55 (37%)</td>
<td>93 (63%)</td>
</tr>
<tr>
<td>DTW(_Y)</td>
<td>41 (28%)</td>
<td>107 (72%)</td>
</tr>
</tbody>
</table>
IV. DISCUSSION

In our simulations, synthetic signals were designed to have two clearly visible peaks in each repetition. Realizations differed by variable location, magnitude and width of these peaks, as well as a slight nonlinear trend and noise, both characteristic to biological signals. The goal was to recover the shape of the two peaks. Results of these simulations show that the BIS method has competitive accuracy of recovering the original signal, as measured by the Euclidean distance. This is despite BIS having significantly less parameters and offering less flexibility than DTW-IS. In some cases, such as uncorrelated noise with low signal-to-noise ratio, BIS additionally has lower standard deviation than DTW-IS.

We also note as an observation on implementation quality that the BIS method is easier to parallelize and make perform well. In fact, the brute-force search for the best parameters \( \alpha \) and \( \beta \) to synchronize two signals is an “embarrassingly parallel” problem [27]. Also, for each set of parameters, only \( O(n) \) steps is performed, as opposed to \( O(n^2) \) steps for an unconstrained DTW-IS algorithm, where \( n \) is the length of each repetition in samples. Performance also scales linearly with the number of repetitions and the number of iterations necessary for convergence, for both DTW-IS and BIS. Additionally, while the BIS algorithm provides a tunable trade-off between speed and accuracy in the form of the grid size used for the search for the best \( \alpha, \beta \) parameters, our practice suggests that there are no significant benefits in increasing it above the grid size suggested in sec. II-E in applications such as the ocular dicrotism, and therefore it can be considered a constant within a specific field of use.

In the illustrative analysis presented in this paper, we have sought to improve on the previous procedure used to detect ocular dicrotism by providing a better structural averaging method. A discrete, very flexible method of DTW exhibited an undesired behavior of distorting the signal shape by overextending or skipping large parts of a signal [19]. We have identified excess flexibility as the probable cause of this problem. The proposed new method was aimed at correcting this shortcoming. It takes advantage of domain-specific knowledge to carefully constrain flexibility.

Despite having only two parameters, the new method has been found to recover sufficient signal shape information to exceed the ODP detection performance of the DTW-based method. In addition, the results are comparable to the procedure based on the wavelet method [23] while being simpler and more interpretable.

The previous work [19] established a threshold of 3 dB for recognizing a dicrotic pulse when using a structural averaging method based on DTW. This threshold has been found to be extreme when applied to the new method. After manual inspection we found that the flexibility of the DTW-based method allowed the structural average to repeatedly match samples that represent peaks and dicrotic valleys in one signal to an unrelated part of another signal, resulting in exaggerated amplitude differences. That, in turn, led to a high value of the former threshold.

TABLE 5. Statistical measures.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>( F_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIS0.05</td>
<td>0.4</td>
<td>0.98</td>
<td>0.31</td>
<td>0.47</td>
</tr>
<tr>
<td>DTW-IS0.05</td>
<td>0.41</td>
<td>0.98</td>
<td>0.51</td>
<td>0.48</td>
</tr>
<tr>
<td>BIS0.05</td>
<td>0.4</td>
<td>0.99</td>
<td>0.31</td>
<td>0.47</td>
</tr>
<tr>
<td>BISY</td>
<td>0.78</td>
<td>0.95</td>
<td>0.78</td>
<td>0.86</td>
</tr>
<tr>
<td>DTW-ISY</td>
<td>0.76</td>
<td>0.96</td>
<td>0.75</td>
<td>0.84</td>
</tr>
<tr>
<td>DTWY</td>
<td>0.78</td>
<td>0.98</td>
<td>0.76</td>
<td>0.86</td>
</tr>
</tbody>
</table>

F. STATISTICAL MEASURES

One way to visualize the efficiency of a method is to draw confusion matrices (see Fig. 12). A few measures were derived from the confusion matrices (see Table 5). Here again, the advantage of BIS over the other two methods depends on the selected type of threshold.
The results from the reproducibility experiments show that the new method is much more stable in results than the method based on DTW: a single patient’s set of signals is more consistently classified as either ODP or non-ODP. We suspect that for the DTW-based method this problem was a result of frequent wrong synchronization of signals stemming from overflexibility in scenarios with poor SNRs.

In order to check whether the improved performance of BIS over DTW stemmed from the constrained flexibility of the time transform itself or from the iterative synchronization, we introduced and evaluated a third method, named DTW-IS, that was identical to BIS with the exception that the time transformation was replaced by the traditional DTW algorithm. Our results show that the differences in performance between the three tested methods depend on the threshold parameter. These findings reveal that the best choice of averaging method may depend on the desired false positive or false negative ratios.

With any signal processing procedure care must be taken in order not to lose relevant information. The illustrative example shows that BIS is capable of preserving as much information as DTW and DTW-IS (see, e.g., the mutual agreement of the statistical measures in Table 5).

We considered adding a mechanism to answer observation 6 from Sect. II-A, imprecision in estimating boundaries of repetitions, by allowing the procedure to synchronize consecutive triplets of repetitions, as opposed to just a single repetition, in order to detect ODP in the middle repetition of the triplet. In such a case, for the CP signal, the points that represent boundaries of the middle repetition would also be synchronized by the beta CDF function instead of taking their location straight from the ECG signal. However, we found no difference in classification performance between the basic version of the method and the triplet-based method.

In the context of the algorithm presented here, the two parameters, $\alpha$ and $\beta$, can be interpreted as controlling relative time compression or dilation between a single repetition and of the method and the triplet-based method. The values of $\alpha = \beta = 1$ represent no change. This observation has proved useful in testing our implementation of the algorithm.

The choice of a structural average estimation method turns out to have non-trivial impact on the effectiveness of analysis of signals with many repetitions. Overflexibility of the traditional DTW may have a detrimental effect on the procedure. Therefore, when synchronizing repetitions, it is a good practice to use a tool that can be adapted to average out the types of signal variations considered nonessential for the task at hand.

V. CONCLUSION
We have designed a structural average estimation method that takes into account modes of variability specific to signals with repetitive feature patterns. In simulations, BIS performed comparably to DTW, but with lower standard deviation of estimates and requiring less computation. When compared to DTW used on CP signals, BIS performs better when used to estimate depths of dicrotic valleys: the estimates are more precise, and within-signal as well as within-person standard deviation of estimates is lower. As a result, more reliable detection of the phenomenon of ocular dicrotism is achieved. The method can potentially be applied to other repetitive signals such as ECG, EEG or MEG if characteristics of particular data sets justifies this choice. In that case, one should carefully constrain the range of possible values of the $\alpha$ and $\beta$ parameters.

ACKNOWLEDGMENT
Damian Mirecki and Tomasz Melcer contributed equally to this work. We thank Prof. D. Robert Iskander for his fruitful comments on an earlier version of the manuscript.

APPENDIX CONVERGENCE OF BIS

Theorem 1. BIS converges monotonically to a constant.

Proof. Assume that the signal $s$ is composed of repetitions $c_1, \ldots, c_k$. For the $l$th iteration of BIS, let $m^{l+1} = (c_1^l, \ldots, c_k^l)$, i.e. the arithmetic mean across repetitions $c_j^l$, where $j \in \{1, \ldots, k\}$. Also, let $h_j^{l+2}$ be the transformed (see Sect. II-B) function $c_j$ that minimizes $\delta_d(c_j, m^{l+1}; h^{l+2})$. Fix $j$ and $l$, since $h_j^{l+2}$ minimizes the error $\delta(c_j, m^{l+1}; h^{l+2})$, we have

$$\delta_d(c_j, m^{l+1}; h^{l+2}) \leq \delta(c_j, m^{l+1}; h^{l+1})$$

and equivalently

$$\sum_{\tau \in T} (c_j^{l+2}(\tau) - m^{l+1}(\tau))^2 \leq \sum_{\tau \in T} (c_j^{l+1}(\tau) - m^{l+1}(\tau))^2.$$  

Since (16) holds for all $j$, we have

$$\sum_{j \in \{1, \ldots, k\}} \sum_{\tau \in T} (c_j^{l+2}(\tau) - m^{l+1}(\tau))^2 \leq \sum_{j \in \{1, \ldots, k\}} \sum_{\tau \in T} (c_j^{l+1}(\tau) - m^{l+1}(\tau))^2.$$  

Let us evaluate the right-hand side of (17):

$$\sum_{j \in \{1, \ldots, k\}} \sum_{\tau \in T} (c_j(h_j^{l+1}(\tau)) - m^{l+1}(\tau))^2 =$$

$$= \sum_{\tau \in T} \sum_{j \in \{1, \ldots, k\}} (c_j(h_j^{l+1}(\tau)) - m^{l+1}(\tau))^2 \leq \sum_{\tau \in T} \sum_{j \in \{1, \ldots, k\}} (c_j(h_j^{l+1}(\tau)) - m^{l}(\tau))^2 =$$

$$= \sum_{j \in \{1, \ldots, k\}} \sum_{\tau \in T} (c_j(h_j^{l+1}(\tau)) - m^{l}(\tau))^2.$$  

10 VOLUME X, 2018
Inequality (18) stems from the fact that the arithmetic mean minimizes \[ \sum_{j \in \{1, \ldots, k\}} \left( e_j h_j^{l+2}(\tau) - m^{l+1}(\tau) \right)^2 \leq \sum_{j \in \{1, \ldots, k\}} \sum_{\tau \in T} \left( e_j h_j^{l+1}(\tau) - m^{l}(\tau) \right)^2 \]

\[ \sum_{j \in \{1, \ldots, k\}} \delta_d(e_j, m^{l+1}, h_j^{l+2}) \leq \sum_{j \in \{1, \ldots, k\}} \delta_d(e_j, m^{l}, h_j^{l+1}). \]

(19)

Therefore, the error considered in iteration \( l + 1 \) is not greater than the error in iteration \( l \). Thus, the sequence of errors is monotonic and bounded, and therefore the algorithm converges to a constant.

REFERENCES


DAMIANKIREKIREceived the M.Sc. degree in computer science, with a specialization in algorithms from the Wroclaw University of Science and Technology, Wroclaw, Poland, in 2017. He has commercial experience in the field of computer vision and machine learning. His current interests include image and signal processing, analysis of biomedical data and machine learning.

TOMASZ MELCER received the M.Sc. degree in computer science from the Wroclaw University of Science and Technology (WUST), Wroclaw, Poland. He has commercial experience in the field of software for scientific and laboratory applications: cancer biomarker research and image processing for bacteria identification, among others. Currently working in the Department of Biomedical Engineering at WUST, he is a member of the Biomedical Signal Processing Group. His current research interests include applications of modern machine learning techniques to diagnostics of eye diseases and developing statistical methods tailored to analysis of biomedical data.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2877918, IEEE Access
CEZARY SIELUZYCKI received the M.Sc. degree in technical physics from Wrocław University of Technology in 1999 and the Ph.D. degree in biocybernetics and biomedical engineering from Warsaw University of Technology in 2003. From 2000 to 2001, he was a Socrates–Erasmus fellow at the VU University Medical Center of the Vrije Universiteit Amsterdam. From 2004 to 2013, he worked as a Research Fellow in the Special Laboratory for Noninvasive Brain Imaging of the Leibniz Institute for Neurobiology Magdeburg. After that, he joined the Control of Normal and Abnormal Movements Team in the ICM Brain and Spine Institute, Pierre and Marie Curie University (Paris VI, Sorbonne). In 2015, he started working at the Wrocław University of Science and Technology, first in the Department of Computational Intelligence and now in the Biomedical Signal Processing Group of the Department of Biomedical Engineering. His research interests include biomedical signal processing with emphasis on magnetoencephalography applied to neuroscience of hearing and movement, and recently also applications of the Microsoft Kinect to studying motor learning.

MONIKA E. DANIELEWSKA received the M.Sc. degree in technical physics, with a specialization in biomedical engineering from the Wrocław University of Science and Technology, Wrocław, Poland, in 2008, and the Ph.D. degree in biocybernetics and biomedical engineering from the Silesian University of Technology, Gliwice, Poland, in 2013. Her current research interests include biomedical signal processing and its application to ocular instrumentation, particularly related to early diagnosis of glaucoma. Dr Danielewska is a member of the Association for Research in Vision and Ophthalmology.