

Localizing Periodicity in Time Series and Videos

Giorgos Karvounas^{1, 2}

gkarv@ics.forth.gr

Iason Oikonomidis¹

oikonom@ics.forth.gr

Antonis A. Argyros^{1, 3}

argyros@ics.forth.gr

¹ Institute of Computer Science, FORTH,
Heraklion, Crete, Greece

² Technological Educational Institute of
Crete, Department of Informatics
Engineering, Greece

³ Computer Science Department,
University of Crete, Greece

Abstract

Periodicity detection is a problem that has received a lot of attention, thus several important tools exist to analyse purely periodic signals. However, in many real world scenarios (time series, videos of human activities, etc) periodic signals appear in the context of non-periodic ones. In this work we propose a method that, given a time series representing a periodic signal that has a non-periodic prefix and tail, estimates the start, the end and the period of the periodic part of the signal. We formulate this as an optimization problem that is solved based on evolutionary optimization techniques. Quantitative experiments on synthetic data demonstrate that the proposed method is successful in localizing the periodic part of a signal and exhibits robustness in the presence of noisy measurements. Also, it does so even when the periodic part of the signal is too short compared to its non-periodic prefix and tail. We also provide quantitative and qualitative results obtained from the application of the proposed method to the problem of unsupervised localization and segmentation of periodic activities in real world videos.

1 Introduction

Periodic patterns and motions are ubiquitous in both natural and man-made environments [8]. Common periodic signals include the undulatory motion of biological organisms as well as the repetitive motions of man-made machines. Thus, the detection and the characterization of such periodic patterns has been a topic addressed in several disciplines such as signal processing, pattern analysis, image processing and computer vision.

Several well established tools and techniques such as the Fourier Transform [10], auto-correlation [2] and wavelets [11] can be used to analyse purely periodic signals. However, in many real life scenarios, periodic signals appear as segments of larger signals containing non-periodic parts. For example, consider the scenario of a sitting human who stands up, performs a repetitive/periodic motion like walking, hand waiving, etc and then sits down again. It is also common that the periodic part of the signal constitutes a small part of the whole signal. The detection of such periodic parts of the motion or signal, along with their characterization (i.e., the estimation of the period and temporal extent of the detected part)

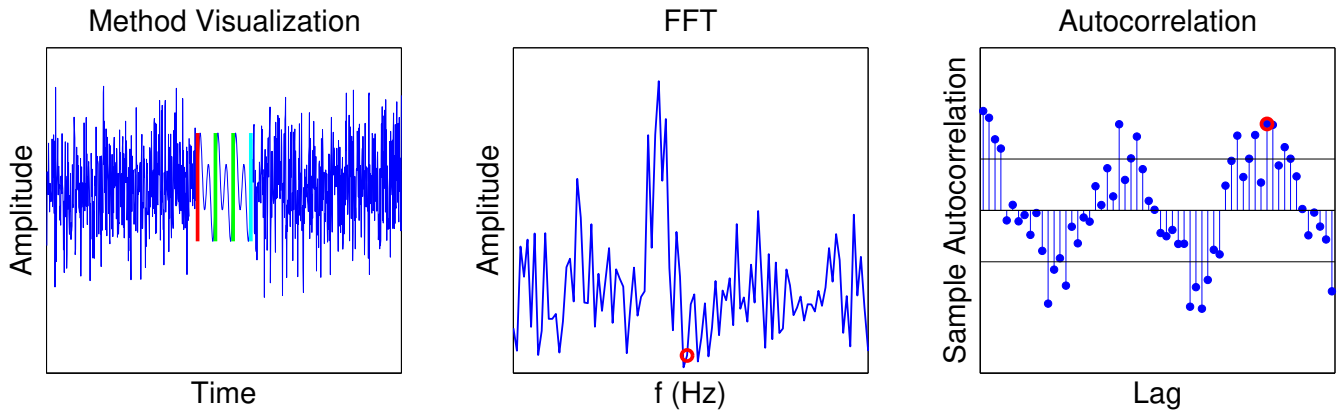


Figure 1: The proposed method (left) detects the start (red delimiter), the end (cyan delimiter) and the period length (green delimiters) of a periodic signal that appears within a non-periodic one. Because of the extended non periodic prefix and tail, the Fourier Transform (center) and the autocorrelation (right) cannot cope with this problem in a robust and computationally efficient way. The red circle in the center and right plots corresponds to the ground truth regarding the periodic signal. The Fourier-based and autocorrelation-based responses exhibit spurious peaks, different than that of the ground truth.

are subproblems of a challenging problem that cannot be easily addressed by existing approaches and techniques.

In this paper we present a new method to solve exactly this problem. Given a univariate or multivariate time series representing a periodic signal that has a non-periodic prefix and tail (see Figure 1), our goal is to be able to estimate the start of the periodic signal, its end and its period length. Furthermore, we want to achieve this in a totally unsupervised manner. In that direction we formulate an appropriate objective function that is minimised by employing Particle Swarm Optimization (PSO) [14]. PSO belongs to the class of evolutionary optimization algorithms that mimic the process of natural selection. The employed objective function is parametrized by the begin and end of the periodic segment within the signal, as well as the period to be estimated. As shown by experiments on synthetic data, the proposed method is robust to noise. Noise tolerance is attributed to both the robustness of the chosen optimization technique and also to the appropriate design of the employed objective function.

Our target application is to detect and characterize periodic events and activities in monocular videos. Towards this end, we also propose a method that, based on motion information, automatically extracts several time series from videos. The analysis of these time series based on the proposed periodicity detection method leads to the spatio-temporal localization of periodic events as well as to the estimation of their frequency. The resulting method has only a small number of tunable parameters, is totally unsupervised and can detect short periodic events occurring in the context of extended non-periodic activities. Several experiments show its effectiveness and accuracy in real world scenarios. From a computational point of view, the method delivers near-optimal results at a computational cost that is several orders of magnitude less compared to the one required by the naive exhaustive search over the parameters of the problem.

2 Related Work

Physics was the first scientific area to deal with problems related to periodicity detection. Furthermore, many fields in computer science such as computer vision, databases and digital signal processing have benefited from and improved upon the relevant knowledge.

Signal processing methods: Firstly, back in the 19th century, Fourier [10] introduced the Fourier Transform. Cooley and Tukey in [7] made it efficient to use the Fourier Transform in digital signals by introducing the Fast Fourier Transform (FFT) that decomposes a signal into its constituent frequencies. This transform is not suitable for our problem because the non-periodic prefix and tail of the system introduce several frequencies, therefore the unique frequency of the periodic part of the signal or its harmonics cannot be distinguished. Additionally, the periodic part of the signal cannot be localized. Autocorrelation [2] is a method to measure the similarity between values in a signal and is more tolerant to non-stationary periodic signals. Wavelets, introduced by Grossmann et al. [11] can tackle the Fourier transform locality issue, but their application requires manual initialization of several parameters. All of the above methods have low tolerance to noise.

Data mining: Finding sequential patterns has also received a lot of attention in the area of data mining. Agrawal et al. [1] introduced the first rule for pattern mining in databases and, two years later, proposed mining of often occurring sub-sequences. Elfeky et al. [9] proposed the WARP algorithm for the detection of reoccurring, same or similar transactions in databases. Han et al. [12, 13] also proposed a method for mining single or multiple periodic patterns in databases. Data mining is out of the scope of this work, but many real-world problems can be addressed using such techniques.

Vision-based tracking of periodic motions: Seitz et al. [22] present an algorithm to estimate cyclic and periodic motions based on the Kolmogorov-Smirnov test. The method is affine invariant regarding the observation viewpoint. The output of the method is the estimated period of the motion, as well as a certainty score. Polana and Nelson [21] devise an extension of the Fourier formula to detect periodicity. Visual features are extracted by tracking the objects and then spatially aligning the frames using as guide the centroid of each object. Cutler and Davis in [8] address the problem of periodicity detection for both the case of stationary and non-stationary periodic signals. For the case of stationary signals, this can be achieved by a Fourier Transform followed by a Hanning filter. However, for the non-stationary case, Short-Time Fourier Transform is employed to better handle the shifting spectrum. As in [21], the objects are tracked and aligned before the periodicity analysis.

Human gait analysis: Urtasun et al. [25] use the anthropomorphic walker [15] for the application of person tracking. The anthropomorphic walker is a physics model describing bipedal locomotion. By detecting the frequency of collision with the ground, the model can create a strong prior for the next move of the legs. On the other hand, Collins et al. [6] successfully applied gait cycle analysis to identify humans by extracting landmark poses.

Motif Detection: The detection of repeating sub-sequences (termed motifs) in time series, is also a problem relevant to periodicity detection. Buhler et al. [3] propose an algorithm based on random projections to detect motifs. Chiu et al. [4] devise an adaptive version of the random projection algorithm [3]. They group input samples into clusters that they call symbols, achieving low computational complexity. In a method closely related to the proposed one, Serra and Arcos [23] propose the SWARMMOTIF, an evolutionary algorithm for the problem of motif detection. Their method can find motifs with a prefix dissimilarity. An important issue of the algorithm is the maximum number of motifs that can be detected in a time series. This parameter is decided upon the initialization of the optimization algorithm.

The literature review shows that there is no method that is both computationally efficient and robust to noise for the problem of recovering spatial and temporal parameters of periodic signals that appear in the context of non periodic ones. To the best of our knowledge, in this work we propose the first such method.

3 The Proposed Method

The core of the proposed framework is a method that, given a univariate time series containing a periodic part, detects the start, the end and the period length of that part (Section 3.1). However, several phenomena can be represented more effectively as multivariate time series (e.g., motion capture data representing the joint angles of a human body in motion as a function of time). We consider multivariate time series as a set of synchronized, univariate time series. We apply the core periodicity detection method to each of them. Then, we employ a simple yet effective method to aggregate partial results towards characterizing the periodicity of the event that is represented with the multivariate time series (Section 3.2). Finally, we show how to transform an input video to time series which are analysed by the proposed techniques in order to detect and characterize periodic events in real world videos in a totally unsupervised manner (Section 3.3).

3.1 Periodicity detection in a univariate time series

We consider a univariate time series $\mathbf{x} = \langle x_1, x_2, \dots, x_N \rangle$. We assume that this time series is periodic between times b and e . The period of that part of the signal is l . Thus, between b and e , the signal consists of $n = \lfloor (e - b)/l \rfloor$ repetitions of a certain motif. Given the time series $\mathbf{x}$ and no other information, our goal is to estimate b , e and l . We formulate the task as an optimization problem in a search space defined by b , e and l .

3.1.1 Objective function

The role of the defined objective function is to quantify the quality of a candidate solution (b, e, l) . Given such a candidate solution the time series is segmented into a prefix $\langle x_1, x_2, \dots, x_{b-1} \rangle$, a tail $\langle x_{e+1}, x_{e+2}, \dots, x_N \rangle$ and an in-between part $\langle x_b, x_{b+1}, \dots, x_e \rangle$, supposedly consisting of n repetitions of a segment of length l . If the solution (b, e, l) is correct, then the n segments

$$s_i = \langle x_{b+l \cdot (i-1)}, \dots, x_{b+l \cdot i-1} \rangle, i \in \{1 \dots n\}, \quad (1)$$

each of length l , should be very similar to each other. We quantify the total dissimilarity of these segments as the mean squared error among all pairs of segments:

$$\epsilon_s(l) = \frac{1}{n \cdot l} \sum_{i=1}^n \sum_{j=i+1}^n \|s_i - s_j\|_2^2, \quad (2)$$

where $\|\cdot\|_2^2$ denotes the squared L_2 norm. Note that whenever $[(e - b) \bmod l] \neq 0$, the last segment s_n is not entirely within the bounds b and e . Therefore, special care is taken to properly handle the comparison of this last segment to the others. Essentially, the same computation as Equation (2) is performed, but with appropriate limit values. Specifically,

the length of the last segment is $l - [(e - b) \bmod l]$. Therefore, only this part of this segment is compared to the corresponding parts of the rest of the segments.

The sought solution (b, e, l) to our problem minimizes $\varepsilon_s(l)$. However, $\varepsilon_s(l)$ is also minimized if parts of the periodic signal are integrated to the non-periodic prefix or tail. Thus, another term is incorporated to the objective function to favour solutions with larger temporal extent. We achieve this by defining:

$$\varepsilon_t(b, e, l, \mathbf{x}) = \alpha \cdot \varepsilon_s(l) + \frac{1}{e - b}. \quad (3)$$

In Equation (3), $\alpha = 0.1$ is an experimentally determined weight factor. The intuition behind the use of the second, additive term is that, as $1/(e - b)$ diminishes for large values of $e - b$, the optimization tends to favour solutions that include a larger number of samples. Without this term, the method tends to prefer very small temporal extents for the periodic signal since it is probable to find a good score in a few samples.

The special handling of the last segment results in an objective function surface that is asymmetric around a given start and end. To alleviate this issue we compute Equation (3) for the input signal, as well as for its time-reversed version, appropriately transforming the boundaries. Thus, the objective function $O(b, e, l, \mathbf{x})$ guiding the optimization is defined as

$$O(b, e, l, \mathbf{x}) = \varepsilon_t(b, e, l, \mathbf{x}) + \varepsilon_t(L(\mathbf{x}) - e, L(\mathbf{x}) - b, l, r(\mathbf{x})), \quad (4)$$

where $L(\mathbf{x})$ denotes the length of signal $\mathbf{x}$ and $r(\mathbf{x})$ represents the time-reversed signal $\mathbf{x}$.

3.1.2 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a stochastic method that performs optimization by iteratively improving a candidate solution with regard to a given measure of quality (objective function). PSO has been applied successfully to solve a number of computer vision problems such as object detection [24], head pose estimation [20], 3D hand tracking [18, 19] and 3D tracking of hands in interaction with objects [16]. PSO maintains a population of candidate solutions, called particles, and moves these particles in the search space according to mathematical formulae governing the particle's position p and velocity V . The movement of each particle p_i is influenced by its local best known position P_i , and simultaneously guided towards the globally best known position G in the search-space. Both these positions are updated as better positions are found by other particles. The update to the k -th generation is described by:

$$V_{i,k} = r_1 c_1 (P_i - p_{i,k-1}) + r_2 c_2 (G - p_{i,k-1}) + \omega V_{i,k-1} \quad (5)$$

$$p_{i,k} = p_{i,k-1} + V_{i,k}, \quad (6)$$

where $p_{i,k}$ and $V_{i,k}$, respectively, denote the position and velocity of the particle p_i at the k -th generation, r_i are samples of the uniform distribution $U(0, 1)$, and c_1 , c_2 and ω are parameters controlling the convergence speed of PSO. The particles are allowed to move within per-defined ranges along each dimension of the search space. To enforce this constraint, whenever it is violated the respective velocity $V_{i,k}$ is reduce up to the point that the constraint is again satisfied. These steps are followed iteratively, until a fixed upper bound of generations is reached. Regarding the parameters c_1 , c_2 and w , we follow the guidelines proposed in [5]. Specifically, $c_1 = 2.8$, $c_2 = 1.3$ and $\omega = 2 / \left| 2 - \psi - \sqrt{\psi^2 - 4\psi} \right|$, where $\psi = c_1 + c_2$.

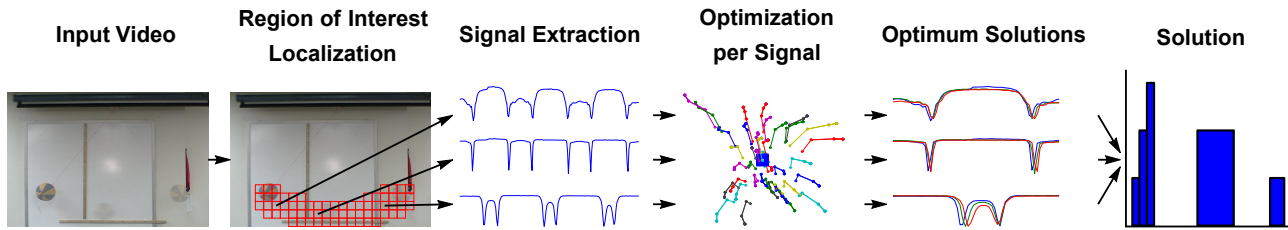


Figure 2: Illustration of the basic steps of the proposed method. From left to right: pre-processing of input video to yield a set of time series, estimation of the start, end and period length of each of them and, finally, aggregation of results through weighted voting.

PSO can handle large search spaces and noisy, multi-modal objective functions. It is suitable for our problem since the objective function exhibits multiple minima. For our problem at hand, we use a standard variant of PSO which is termed canonical PSO [5] to minimize the objective function of Equation (4) over candidate solutions (b, e, l) . Canonical PSO optimizes real-valued parameters within a pre-specified cuboid of the search-space. Our problem has integer-valued parameters, therefore we resort to rounding the inputs to the closest integer. Furthermore, since parameters b and e signify the begin and end of the periodic pattern, it should also hold that $e > b$. We enforce this constraint by returning a very high objective function value whenever it is violated.

3.2 Periodicity detection in a multivariate time series

The result of PSO is an estimated period length l as well as the start b and the end e of the periodic part within a single input signal. Given a K -dimensional multivariate times series, we consider this as a collection of K univariate time series and seek for the triplet (b, e, l) suggested by all of them. Towards this end, each univariate time series is individually processed, resulting in a set of candidate solution triplets (l_k, b_k, e_k) , $k \in \{1..K\}$. To come up with the single triplet describing the whole set, we resort to computing the weighted median of all the estimated periods. As weights, we employ the variances of the values of the time series. This forces signals with more information to contribute more to the final estimation.

A final step is the handling of the case where our solution is in fact a multiple of the basic frequency (called the fundamental). This is trivially handled by exhaustive, yet efficient search. Keeping the estimated boundaries b and e , we vary the period length parameter l . Starting from the estimated period l_e , we successively try the period values $\lfloor l_e/i \rfloor$ for $i = 1, 2, \dots, l_e/2$. For each such value we compute the average objective function value and check if it improves upon the previous average value. We terminate the search as soon as the tested period value does not improve upon the previous one. Regarding the boundaries, we compute the overall begin b and end e as the minimum of all b_k , and the maximum of all e_k over solutions with index k that resulted in the dominant period length l .

It has to be noted that we intentionally handle a multivariate time series as a collection of K univariate time series. This is because in practical situations, it is not guaranteed that all K input signals (e.g., all joint angles of the human body or all time-varying image patches of a video, see also Section 4) exhibit some periodicity. The adopted voting strategy handles this issue in an effective and natural way.

3.3 From real world videos to time series

The proposed method as described in sections 3.1 and 3.2, is generic, in the sense that it can be applied to any univariate or multivariate time series. In this section, we describe how a video can be processed to give rise to time series that can feed periodicity detection, resulting in the detection and characterization of periodic events in arbitrary videos.

Assuming a static camera, objects that undergo periodic motion result in periodic fluctuations of the brightness values of image points in time. Therefore, the intensity value of each and every such point could form a time series to be analysed by the proposed algorithm. However, apart from being unnecessarily complex, this would also be very sensitive to noise. Thus, instead of processing individual pixels, we split the input video in tiles of size 30×30 , and select the ones that exhibit large intensity variation over time.

Initially, we transform every frame from RGB to gray scale and then de-noise it by applying a 9×9 Gaussian filter of $\sigma = 2$. After noise removal, we compute the median image MI of all the video frames by computing the median value of each pixel intensity over time. This serves as the background image, with most parts of the moving objects being removed. By subtracting each frame from MI , we create a new video that mostly contains moving objects. The resulting video is then split into the aforementioned 30×30 tiles, and the ones that exhibit significant motion are selected for further processing. The motion threshold is adaptive, computed as the sharpest increase in the histogram of the whole video motion values. Finally, a time series is defined for each of the remaining tiles, representing the time evolution of its average intensity. Figure 2 summarizes the steps of the proposed method for detecting periodic activities in videos.

4 Experimental Results

In this section we present the results we obtained using the method of Section 3. We first evaluate the performance of the method on synthetically generated sequences, determining appropriate parameters for PSO. Given these parameters, we evaluate the performance of the method under the presence of varying amounts of noise. We conclude the section with results on real-world videos.

Computational budget vs accuracy: We evaluated the performance of the proposed method for periodicity detection in univariate time series based on synthetic data. Through this process we investigate the performance of the method in varying configurations of the computational budget of PSO. This investigation allows us to balance between computational cost and accuracy of the estimated values.

The runtime of PSO is determined by the product of the total number of particles times the maximum number of generations. We consider PSO budgets consisting of a number of particles (from 5 to 100 in steps of 5) running for a number of generations (from 5 to 100 in steps of 5). Thus, a total of 400 different budget combinations are considered.

We investigate the effect of the computational budget allocated to PSO using as input synthetic signals with known ground truth. Specifically, we generate 100 signals containing periodic parts of either $l = 5$, $l = 20$ or $l = 50$ samples per period. Figure 1 (left) shows a sample of these signals. Each periodic part has values ranging between -1 and 1 and it is prepended and appended with 100 samples drawn from the uniform random distribution $U(-1, 1)$. For each such signal and PSO parametrization, we apply our method 10 times to factor out the stochastic nature of PSO. From these 10 runs, we retain the median values for

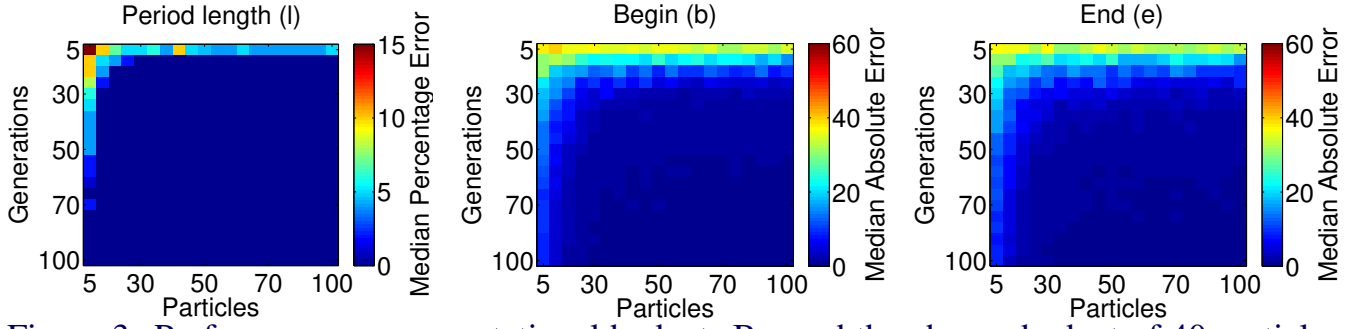


Figure 3: Performance vs computational budget. Beyond the chosen budget of 40 particles and 40 generations the performance gain is small.

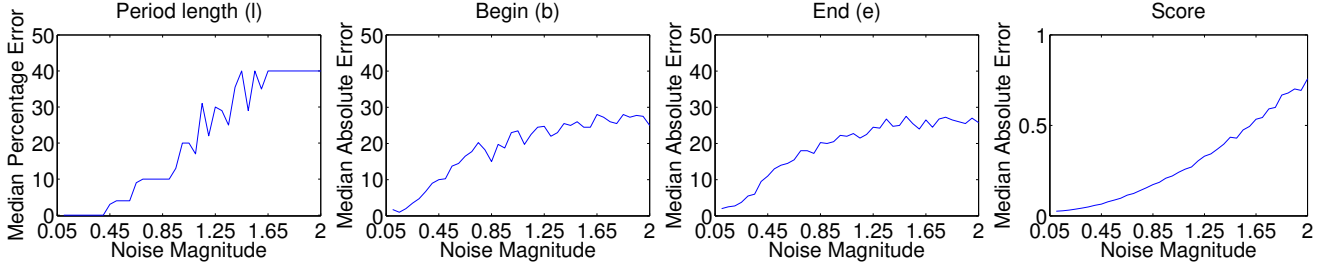


Figure 4: Performance in the presence of additive noise.

the estimated parameters b , e and l .

For a certain signal, consider that the ground truth parameters are (b_g, e_g, l_g) and that some PSO parametrization estimated the solution (b_e, e_e, l_e) . The defined error metrics are $m_l = |((l_e + l_g/2) \bmod l_g) - l_g/2|/l_g$ (as %, treating harmonics as correct), $m_b = |b_g - b_e|$ and $m_e = |e_g - e_e|$. Figure 3 (left to right) plots the median m_l , m_b , m_e over all signals and PSO parametrizations. Evidently there is a direct correlation between the PSO budget and the performance of the method. In practice, after 40 particles and 40 generations, the performance gain is disproportionate to the extra computational budget. Thus, for the remaining of the experiments we keep this PSO budget. Interestingly, the individual plots of Figure 3 for each of the three classes of signals for $l = 5$, $l = 20$, $l = 50$ (omitted due to space limitations) reveal that PSO requires more budget in case that the periodic part of the signal constitutes a smaller part of the whole signal. Thus, in case that there are known statistics about the input signals, the accuracy/computational budget trade-off can be tuned appropriately.

Noise tolerance: We evaluate the performance of our method in the presence of noise. We employ the same synthetic dataset as before adding to each sample random noise drawn from a Gaussian distribution. The variance of the Gaussian samples is in the range $[0.05..2.0]$ in steps of 0.05. The computed performance metrics m_l , m_b and m_e are shown in Figure 4 as functions of the noise level. It can be verified that the method can tolerate noise levels up to almost twice the amplitude of the original signal. It is interesting to note that, although the estimation error flattens after the noise level of 1, the actual objective function score keeps increasing, reflecting the noisier input signal. Figure 5 shows indicative results for the 7 activities. More qualitative results are provided in the supplementary material.

Experiments on the MHAD dataset: We experimented with the MHAD dataset [17] that captures human activities that are repeatedly performed in front of a camera system. Interestingly, this dataset features MoCap data corresponding to 93 joint angles of the human body captured by an optical motion capture system at 480Hz. We thus performed two different experiments based on this dataset:

(a) **Periodicity detection based on motion capture data:** We down-sampled the motion

Activity	frames	(a) Motion Capture			(b) RGB video		
		l	m_b	m_e	l	m_b	m_e
Jumping	174	28	0	2	27	3	0
Jumping jacks	194	31	0	0	30	1	0
Bending	429	71	0	9	70	0	1
Punching	204	31	12	1	29	9	1
Waive two hands	238	39	0	4	38	0	0
Waive one hand	248	45	0	7	43	0	5
Clapping	131	22	3	0	21	14	1

Table 1: Summary of results for periodicity detection on the MHAD dataset for the cases of employing motion capture data and RGB videos. See text for details.

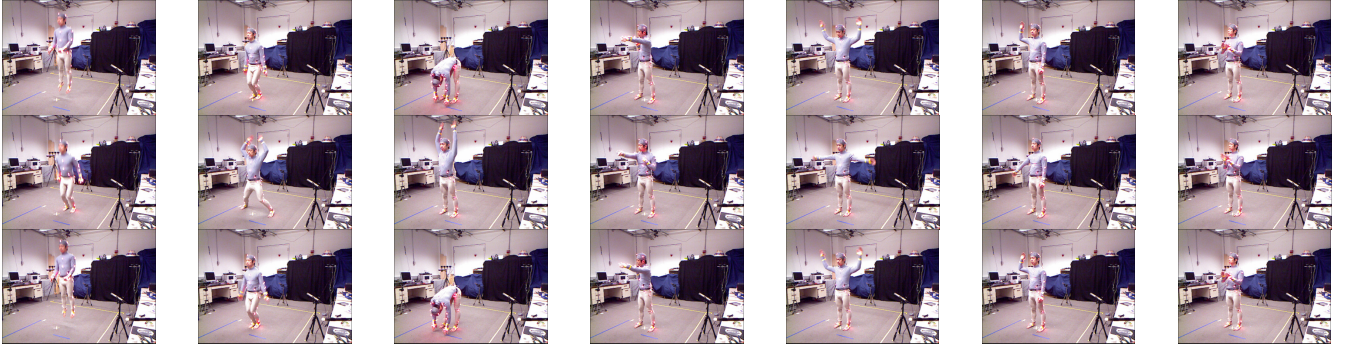


Figure 5: First row: a hand-picked frame from each of the 7 MHAD activities. Second row: frame at half estimated period apart. Third row: frame one estimated period apart.

capture data to 30 Hz. $K = 93$ time series were defined and fed to our method.

- (b) **Periodicity detection based on video data:** We performed periodicity detection based on time series produced as described in Section 3.3 based on the 30Hz RGB videos.

We experimented with 7 MHAD activities. For each of them, Table 1 shows the estimated parameter l and the error metrics m_b , m_e for cases (a) and (b) above. Both approaches provide very similar results that are in agreement with the ground truth.

Experiments on real-world sequences: Using the method described in Section 3.3, we extract signals from four input videos. The 1st (*pendulum*, 192 frames) shows three swings of a pendulum. The 2nd (*turntable*, 157 frames) shows a spinning turntable. The 3rd (*rope*, 106 frames) captures a person pulling a rope. Finally, the 4th video (*machine*, 836 frames) shows the periodic process of milling an object in a production line. Our method estimated the period lengths of the observed activities (63, 52, 35 and 158 frames, respectively) within 3 frames from the manually determined ground truth. Indicative frames are shown in Figure 6. The similarity of frames at a temporal distance of l frames indicates the accuracy of period length estimation. Sample qualitative results are provided at <https://youtu.be/2oa0Y9znH6g>.

5 Summary

We presented a method to detect and characterize periodic signals that appear as parts of non-periodic ones. In the heart of the proposed approach lies a stochastic optimization algorithm that estimates the start, the end and the period length of the periodic part of a univariate time



Figure 6: Indicative results from experiments in four videos (pendulum, turntable, rope, machine). A pair of frames is selected from each video, a random one (left) and one at a temporal distance equal to the period length computed by the proposed method (right).

series. We also propose a method for aggregating individual results obtained from several such time series. Finally, we demonstrated an effective application of the proposed method for localizing (both spatially and temporally) periodic activities in videos. The method is totally unsupervised and has a small number of tunable parameters. PSO is shown to require as many as 40 particles evolving in 40 generations, thus 1,600 objective function evaluations. This should be contrasted to the exhaustive search on the 3-dimensional space of possible starts, ends and period lengths. Given that each of them is in the order of the length of the time series/video, this means that the proposed approach provides accurate results, while requiring orders of magnitude less computations compared to the brute-force, exhaustive search. An extension of the current work could be to handle multiple periodic phenomena within the same input. A simple way to handle this would be to apply the current method multiple times, omitting each time the detected periodic part.

Acknowledgments

This work has been partially supported by the EC H2020 project ACANTO.

References

- [1] Rakesh Agrawal and Ramakrishnan Srikant. Mining sequential patterns. In *Data Engineering, 1995. Proceedings of the Eleventh International Conference on*, pages 3–14. IEEE, 1995.
- [2] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
- [3] Jeremy Buhler and Martin Tompa. Finding motifs using random projections. *Journal of computational biology: a journal of computational molecular cell biology*, 9:225–242, 2002.
- [4] Bill Chiu, Eamonn Keogh, and Stefano Lonardi. Probabilistic discovery of time series

- motifs. *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining KDD 03*, 304:493, 2003.
- [5] Maurice Clerc and James Kennedy. The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *Evolutionary Computation, IEEE Transactions on*, 6(1):58–73, 2002.
- [6] Robert T. Collins, Ralph Gross, and Jianbo Shi. Silhouette-based human identification from body shape and gait. *Proceedings - 5th IEEE International Conference on Automatic Face Gesture Recognition, FGR 2002*, pages 366–371, 2002.
- [7] James W Cooley and John W Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965.
- [8] Ross Cutler and Larry S. Davis. Robust real-time periodic motion detection, analysis and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 2000.
- [9] Mohamed G Elfeky, Walid G Aref, and Ahmed K Elmagarmid. Periodicity detection in time series databases. *Knowledge and Data Engineering, IEEE Transactions on*, 17, 2005.
- [10] Jean Baptiste Joseph Fourier. *The analytical theory of heat*. The University Press, 1878.
- [11] Jean Grossmann, Alexander and Morlet. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM journal on mathematical analysis*, 15: 723–736, 1984.
- [12] Jiawei Han, Wan Gong, and Yiwena Yin. Mining segment-wise periodic patterns in time-related databases. In *Proceeding of the ACM SIGKDD international conference on knowledge discovery in databases*, 1998.
- [13] Jiawei Han Jiawei Han, Guozhu Dong Guozhu Dong, and Yiwen Yin Yiwen Yin. Efficient mining of partial periodic patterns in time series database. In *Proceedings 15th International Conference on Data Engineering (Cat. No.99CB36337)*, pages 1–10, 1999.
- [14] James Kennedy. Particle swarm optimization. In *Encyclopedia of machine learning*, pages 760–766. Springer, 2011.
- [15] A D Kuo. A simple model of bipedal walking predicts the preferred speed-step length relationship. *Journal of biomechanical engineering*, 123(3), 2001.
- [16] Nikolaos Kyriazis and Antonis A Argyros. Scalable 3d tracking of multiple interacting objects. In *IEEE Computer Vision and Pattern Recognition (CVPR 2014)*, pages 3430–3437, Columbus, Ohio, USA, June 2014. IEEE.
- [17] Ferda Ofli, Rizwan Chaudhry, Gregorij Kurillo, Rene Vidal, and Ruzena Bajcsy. Berkeley MHAD: A comprehensive Multimodal Human Action Database. pages 53–60, 2013.

- [18] Iason Oikonomidis, Nikolaos Kyriazis, and Antonis A Argyros. Efficient model-based 3d tracking of hand articulations using kinect. In *British Machine Vision Conference (BMVC 2011)*, volume 1, pages 1–11, Dundee, UK, 2011. BMVA Press, BMVA.
- [19] Iason Oikonomidis, Nikolaos Kyriazis, and Antonis A Argyros. Tracking the articulated motion of two strongly interacting hands. In *IEEE Computer Vision and Pattern Recognition (CVPR 2012)*, pages 1862–1869, Providence, Rhode Island, USA, June 2012. IEEE.
- [20] Pashalis Padeleris, Xenophon Zabulis, and Antonis A Argyros. Head pose estimation on depth data based on particle swarm optimization. In *IEEE Computer Vision and Pattern Recognition Workshops (CVPRW 2012)*, pages 42–49, Providence, Rhode Island, USA, June 2012. IEEE.
- [21] Ramprasad Polana and Randal C Nelson. Detection and recognition of periodic, non-rigid motion. *International Journal of Computer Vision*, 23(3):261–282, 1997.
- [22] Steven M Seitz and Charles R Dyer. View-invariant analysis of cyclic motion. *International Journal of Computer Vision*, 25:1–23, 1997.
- [23] Joan Serrà and Josep Luis Arcos. Particle swarm optimization for time series motif discovery. *Knowledge-Based Systems*, 92, 2016.
- [24] Stefanos Stefanou and Antonis A Argyros. Efficient scale and rotation invariant object detection based on hogs and evolutionary optimization techniques. In *Advances in Visual Computing (ISVC 2012)*, pages 220–229, Rethymnon, Crete, Greece, July 2012. Springer.
- [25] Raquel Urtasun, David J. Fleet, and Pascal Fua. 3d people tracking with gaussian process dynamical models. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1, 2006.