Visual Quasi-Periodicity

E. Pogalin[†]   A.W.M. Smeulders[‡]   A.H.C. Thean[†]

[†]Signal Processing Group
TNO Science & Industry, The Netherlands
{erik.pogalin, andrew.thean}@tno.nl

[‡]Intelligent Systems Lab Amsterdam
University of Amsterdam, The Netherlands
a.w.m.smeulders@uva.nl

Abstract

Periodicity is at the core of the recognition of many actions. This paper takes the following steps to detect and measure periodicity. 1) We establish a conceptual framework of classifying periodicity in 10 essential cases, the most important of which are flashing (of a traffic light), pulsing (of an anemone), swinging (of wings), spinning (of a swimmer), turning (of a conductor), shuttling (of a brush), drifting (of an escalator) and thrusting (of a kangaroo). 2) We present an algorithm to detect all cases by the one and the same algorithm. It tracks the object independent of the object's appearance, then performs probabilistic PCA and spectral analysis followed by detection and frequency measurement. The method shows good performance with fixed parameters for examples of all above cases assembled from the Internet. 3) Application of the method, completely unaltered, to a random half hour of CNN news has led to an 80% score.

1. Introduction

Many natural objects exhibit periodicity in their motion. Examples include running athlete, flapping wings, nodding, swimming, a beating heart and ironing. The notion of periodicity has been used in Computer Vision for object detection or segmentation [17, 18], tracking [7, 22] and classification [1, 10, 19]. In sports like speed skating and swimming, periodicity gives important measure on how efficiently an athlete moves [24]. This paper develops a model of the 10 essential cases of visual periodicity. Furthermore, a method is devised to robustly analyze these different cases, by evaluating the object appearance over time.

Periodicity in a sequence can be caused by regularity in the motion of the object. Alternatively, it can be caused by the regularity in appearance of the object over time. Or, it can be caused by the motion of the regular pattern on the object. These categories are divided further until we arrive at the 10 fundamental cases. This categorization is important as a different model of computation may result in the strive to good detection results.

We present a novel approach to periodic motion analysis that is inspired by the method developed in [34]. The essential observation of which this paper rests is that the portions of the image that vary (periodically) together also constitute a part of one physical entity. Hence, avoiding explicit and computationally expensive segmentation, we aim for regions of maximum extend for which regions vary coherently. By using Principal Component Analysis (PCA) over time sequence after removal of the translational component one can detect co-varying parts of the image. We prefer to use probabilistic PCA (pPCA) as it delivers a more robust estimate [30] and there is an incremental implementation available (ipPCA) [20]. This approach avoids the burden of many of the past methods.

The paper is organized as follows. Section 2 discusses the classification of visual periodicity types mentioned earlier in more details. We review the related work in section 3. Section 4 describes the proposed periodicity analysis method. In section 5, various experimental results of our method are compared to the results from our own implementation of a recent method [26]. Finally, in section 6 we draw some conclusions.

2. Visual periodicity

Visual periodicity in the perception of an object may result from the appearance variation A_O of the object (which includes both intensity pattern f_O and shape s_O), the perceived path $\mathbf{p}_O$ of the object, or by the motion of a regular pattern on the object. From these general groups we distinguish 10 fundamental cases of periodicity. This categorization we aim for encompasses many more cases than in [11], which focusses on the perceived motion path only. Apart from the two groups described above, the classification of the periodicity cases is based on several other modes. Firstly, the motion can either move forwardly, or create an oscillating pattern, i.e. moving back and forth between two states. Secondly, the flow field created by the variation can be continuous or intermittent. This only holds for the forward motion, since an oscillation will always has moments where the motion stops briefly to change direction. Thirdly,

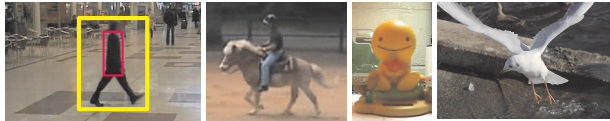(a) *Flash* case: railroad sign, bike lamp, car indicator, sirens

(b) *Pulse* case: anemone, accordeon, spring, heart

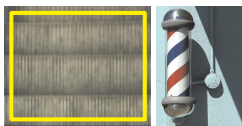(c) *Inflate* case   (d) *Turn* case: conducting

(e) *Swing* case: walking, galloping, toy nodding, flapping

(f) *Spin* case: swimming, spinning disc, windmill, spinning top

(g) *Shuttle* case: brushing, dribbling, sawing, ironing

(h) *Drift* case: escalator, bar-   (i) *Thrust* case: jumping kangaroo, inter-
ber's pole   mittent conveyor

Figure 1. Several examples of visual periodicity grouped by case. The annotated images will be used for evaluation. The smaller rectangle indicates the part that was used for tracking; the larger one is the input for the periodicity detector. The *Intensify* case is missing since no representative example could be found. See Table 1 for the complete list.

the periodicity can be induced either by the object motion or the motion of the surrounding background. In other words, one can make a recording of some child on a swing, or one can record the background while fixing the child in the center of the screen. Our periodicity detector will treat both cases similarly (It would even generate the same frequency of periodicity if the camera would be fixed to the swing and record the background pattern). Several images of real-life periodicity for every case are depicted in Figure 1.

Periodicity in the intensity pattern is described in its simplest case by a repetitive intensity variation, either intrisically or by external illumination, with the smallest value of $\tau > \epsilon$ as the periodicity:

$$f_O(t + \tau) = f_O(t) \qquad (1)$$

This is object periodicity due to oscillation regardless of the position $\mathbf{x}$, we call it the *Flash* case. Other words that describes this case are blinking and shutter. The second case is the *Intensify* case, when the intensity becomes stronger (or weaker) with periodic pauses. Since the continuous version of the forward intensity increase does not cause periodicity, we only have two cases for this group.
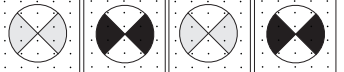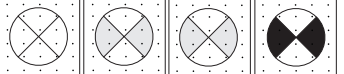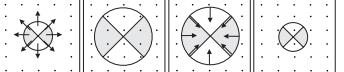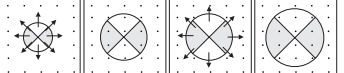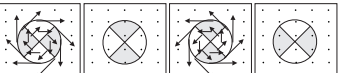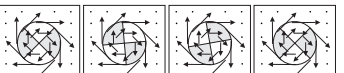
An essentially different case is when the object rotates with an axis perpendicular to the image plane and pointing towards the viewer. We arrive at the *Swing*, *Spin* and *Turn* case. These three cases are caused either by texture variation, usually whole body rotation with the axis going through the object's center of gravity, or by both texture and shape, when the axis is not passing the center of gravity. A few examples for *Swing* motion are: walking, nodding and hand waving. The *Spin* case, the forward and continuous rotation, includes spinning, rolling (without the translation) and swimming. The *Turn* case is similar to *Spin* but with periodic pausing. The *Turn* case is sometimes seen in manufacturing but otherwise rare.
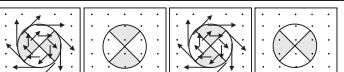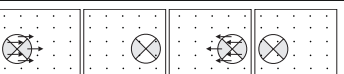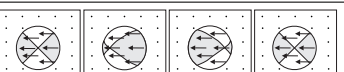
Due to lack of space, the complete discussion on other cases will be included elsewhere. For the remaining of the paper, the *Intensify* and *Inflate* case will not be considered as they are very rare in real situations. The *Turn* case is excluded since it appears very similar to *Spin* case.

3. Related work

The majority of the papers on periodicity analysis in the literature can be categorized as appearance-based. Object tracking is done first to align the image frames. The less common method is to detect periodicity in shape deformation [13, 16]. Generally the shape-based approach is limited only for the cases when the apparent object shape is varying, i.e. the *Pulse* case and a subset of the *Swing* and *Turn* cases. The more recent approaches are focussed on analyzing intensity variations to detect periodicity [5, 6, 10, 18, 29]. The well-cited method of Cutler and Davis [10] calculates the correlation of successive frames, resulting in a similarity plot over time. Ran et al. [26] improved the results of [10] by applying hypothesis testing (based on periodogram maximizer) on individual pixels, to check whether a given pixel is periodic. Only the periodic pixels are then used to

Table 1. Classification of visual periodicity.

	Type	Periodicity abstraction	Examples
Intensity var.	**Flash** (Oscillating)		Fig. 1(a)
	Intensify (Forward intermittent)		None
Deformation	**Pulse** (Oscillating)		Fig. 1(b)
	Inflate (Forward intermittent)		Fig. 1(c)
Rotation	**Swing** (Oscillating)		Fig. 1(e)
	Spin (Forward continuous)		Fig. 1(f)
	Turn (Forward intermittent)		Fig. 1(d)
Translation	**Shuttle** (Oscillating)		Fig. 1(g)
	Drift[†] (Forward continuous)		Fig. 1(h)
	Thrust (Forward intermittent)		Fig. 1(i)

[†]Note: For a periodic appearance, this requires repetition of objects as in a rolling staircase.

estimate the dominant period in the sequence. Next, vPLL (video Phase-Locked Loops) approach [4] is used on the edge images to improve the result, but this second part is only specific for gait analysis. We implemented the first part of Ran's method and will use it for performance comparison. This method however, is based on two assumptions: the motion must be quasi-periodic and no DC component (i.e. shift of intensity value) is present in the pixel variations. Pixel correlations methods described above are sensitive to noise. We propose the use of pPCA before detecting the period to reduce the noise sensitivity. In [6] an online filtering scheme to estimate temporal frequency of quasi-periodic object motion is presented. The filter response gives the location and a measure on how strong the periodicity is at a certain pre-selected tuning frequency. However, the frequency range must be determined manually. Our method requires no such a priori knowledge. Finally, a method to find multiple periodic motions over static background without the need for object tracking is proposed in

[5]. By projecting the intensity in one direction and observing the repetitive patterns in the time-frequency distribution of the video, this method is able to detect periodicity even if it the motion is superimposed to translation. Summation in one spatial axis, however, might eliminate the periodic intensity we are trying to detect. Despite the minor drawbacks discussed above, appearance-based method is the most suitable approach to detect periodicity in all fundamental cases.

Methods that analyze motion variation can be divided further into two types: those that find point correspondences [1, 2, 11, 17, 19, 27, 32] and those that analyze motion field variation [8, 23, 31]. Among the first group, the method described in [17] is most interesting. Periodic motion detection is considered here as a sequence alignment, where an image sequence is matched to itself over one or more periods. This assumption combines both tracking and periodicity detector in the same optimization step. Generally, point correspondences are difficult to achieve when the images of the object of interest are of low resolution, and if the background as well as the foreground are strongly varying in appearance. In the second group, optic flow method is used to estimate the motion field. In [8], SVD is used as replacement for FFT in estimating the period of a dynamic texture. It is reported that SVD is superior to FFT on quasi-periodic data, but data windowing can be applied to FFT to cope with quasi-periodicity. Also, just as in [6] the SVD method described here needs to test a range of period lengths to find the best estimate. Both groups of the motion-based approach cannot detect periodicity in the *Flash* case, because point correspondences and motion field are not available there. Furthermore, the *Spin* and *Drift* case produce a constant flow field, so the second group will have difficulties to detect the periodicity in those cases. In general, appearance-based method such as ours can detect periodicity in a broader range of cases.

PCA has been used before to model object appearance [9, 33]. The use of PCA to analyze periodic motion has also been independently proposed [12], but for the purpose of reconstructing 3D animation. Functional PCA has been used to represent human gait motion [22]. Two other papers [14, 25] used PCA to learn a model for behavior and action classification. In this paper, we use pPCA instead, which provides better robustness to outliers and missing data [30].

In conclusion, the appearance-based method is more generally applicable than the motion-based approach, since it can detect periodicity in more cases. However, none of the methods described above has considered the object of interest as one physical entity. Usually, an ensemble of estimation from parts of the images are produced, and the best period estimation that represents the periodicity is chosen. This would make the methods susceptible to noise by alignment error, occlusion or changing illumination. We will address this issue by applying pPCA on the data.
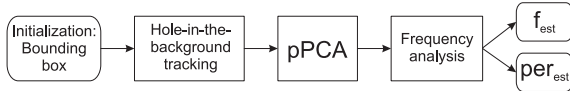
Figure 2. Block diagram of our method. The input is two bounding boxes that is determined manually (See Figure 1). The output is a frequency estimate and the degree of periodicity of the sequence.

4. Periodicity analysis

Our approach is appearance-based compatible to [6, 26] discussed above. Just as most of the approaches, tracking is performed first to align the object in the center of the frame. The key novelty is that by using PCA, the input data that are spatially correlated are grouped together. If the object of interest causes a periodic variation in intensity and shape, PCA models the variation with one or more unobserved variables. Unlike previous approaches, PCA considers these pixels as one physical entity, not as individually varying pixel. Hence the method is compatible with all of the above types and at the same time robust to noise.

Figure 2 shows the block diagram of our method. The process is initialized manually by giving two bounding boxes as input (See Figure 1). The first smaller one is used for tracking. After the image is aligned, we enlarge the region of interest with the second box and pass the cropped sequence to the periodicity detector. At the heart of our method, we use pPCA to detect the maximum spatially coherent changes over time in the object's appearance without the need of segmentation. The online version [20] might be considered if online computation is needed. In the final output we describe the periodicity of the input data by two indicators: the estimated period and the degree of periodicity. Note that the first part of Ran et al. [26] is similar to our method, except that no PCA calculation is done and no degree of periodicity estimation is given.

4.1. The hole-in-the-background tracking

We have selected the method in [21] to track the object of interest. This method makes no assumptions on the appearance of the object other than its distinction from the background. So even if the object's appearance changes drastically over time due to its periodicity or otherwise, the tracker is still able to lock on to the object.

The tracking result plays an essential role in the performance of the periodicity detector. Better tracking allows the periodic components to be modeled by the first few principal components. If severe misalignment occurs during tracking, the first few components would be dominated by a non-periodic variation instead.

4.2. pPCA for periodicity detection

Let $X = [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_N]$ be a $D \times N$ matrix that represents the input video datacube, with N the number of image frames and D the number of pixels in one frame. The rows
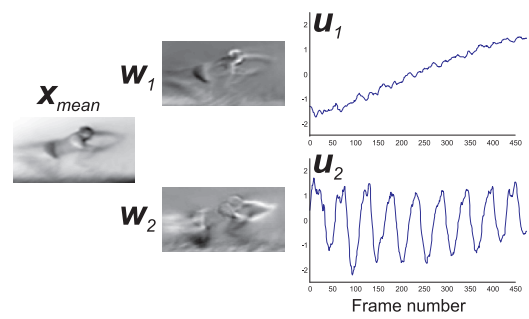

Figure 3. An example of the pPCA analysis: Visualization of the pPCA analysis of a swimmer sequence. Only two components are shown here. The first component models the gradual shift of the swimmer. The second component is dominated by the motion of his arms and legs.

of an aligned image frame are concatenated to form the vector $\mathbf{x}_n$. The optimal linear reconstruction $\hat{X}$ of the datacube is given by:

$$\hat{X} = WU + \overline{X} \qquad (2)$$

with $D \times Q$ matrix $W = [\mathbf{w}_1\ \mathbf{w}_2\ \ldots\ \mathbf{w}_Q]$ as the set of orthonormal basis vectors ($Q \ll D$) onto which the retained variance under projection is maximal, $Q \times N$ matrix $U = [\mathbf{u}_1\ \mathbf{u}_2\ \ldots\ \mathbf{u}_N]$ as a set of Q-dimensional vectors of unobserved variables and $\overline{X}$ the set of mean vectors $\overline{\mathbf{x}}$. The vectors $\mathbf{w}_q$ are also called the eigen images since each represents an image. Some textbooks also refer to the vectors $\mathbf{u}_n$ as principal components. The amount of variation contained in each component is indicated in pPCA by the eigenvalues $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_d)$ of the covariance matrix S of the input data X:

$$S = V\Lambda V^T \qquad (3)$$

which is calculated by Eigenvalue decomposition. The dimension Q is selected by setting the maximum percentage of retained variance we want to have in the reconstructed matrix $\hat{X}$. See [30] for the complete derivation of pPCA and the estimation of W and U in equation (2).

Figure 3 shows a selection of the pPCA output of a sequence with a swimmer. As we can see, the eigen image $\mathbf{w}_q$ accounts for the motion variation, and the change in the time is captured in the unobserved variable $\mathbf{u}_q^T$, in time axis direction. Furthermore, if the periodic and non-periodic variations are not correlated, PCA breaks down the data into either periodic or non-periodic components and describes these variations separately. The reorganization of the data by PCA is quite effective since each unobserved variable vector $\mathbf{u}_q^T$ provides a compact representation of the periodic variation in a group of spatially-correlated pixels, which we will use for spectrum estimation.

4.3. Frequency analysis

In principle, there are two approaches to estimate the power spectrum of a time series: parametric and non-

parametric methods [15]. Non-parametric methods are based on the idea of estimating the autocorrelation function of a random process, and taking the Fourier transform to estimate the power spectrum. The periodogram and its modifications belongs to this class, which need no a priori information about the random process. Parametric methods, on the other hand, need some knowledge on the generation of the random process. Some examples of methods that belong to this class are Burg Algorithm, Modified Covariance method and MUSIC [15]. Although parametric methods can give a higher resolution estimate, in our case knowledge about the input data is in most cases not available. For this reason, we have chosen the modified periodogram of the non-parametric class, which provides a better estimate than plain periodogram with only 1 additional parameter, the window type $w(n)$ to be used [15]:

$$P_q(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} w(n)x(n) \exp(-jn2\pi f) \right|^2 \quad (4)$$

Spectrum estimation is performed for each principal component vector $\mathbf{u}_q^T$ from the pPCA. By weighing the spectra $P_q(f)$ with the relative percentages λ_q^* of the retained variance and summing them together, a spectrum is obtained:

$$\overline{P}(f) = \sum_{q=1}^{Q} \lambda_q^* P_q(f), \quad \lambda_q^* = \frac{\lambda_q}{\sum_{d=1}^{D} \lambda_d} \quad (5)$$

To detect the dominant frequency component in the spectrum $\overline{P}(f)$, we use a similar approach to the one used in [18]. Local maxima detection using a dilation operator is used to detect peaks in the spectrum. Local minima are also detected to define the peak's supports (see Figure 4). The peaks with a frequency lower than $\frac{f_s}{N}$ are discarded, with f_s being the sampling rate of the video and N the frame length. This value is the frequency resolution of the periodogram and the smallest frequency value that can be detected. Afterwards, starting from the lowest found frequency to the highest, each peak is checked against the others for its harmonicity, i.e. if its frequency can be expressed as a linear combination of the existing fundamental frequencies. We require that a fundamental frequency have a higher peak than its harmonics and a tolerance of $\frac{f_s}{N}$ is used in the matching process. Since multiple fundamentals may exist, we select the one group k with the highest total energy to represent the dominant frequency component in the data. The total energy is the sum of the area between the left and right supports $E(.)$ of the fundamental frequency peak f_k^0 and its harmonics f_k^i:

$$f_{est} = \underset{f_k^0}{\mathrm{argmax}} \left\{ E\left(f_k^0\right) + \sum_i E\left(f_k^i\right) \right\} \quad (6)$$

As long as there is some minor peaks present in the spectrum $\overline{P}(f)$ the above method may still give a frequency
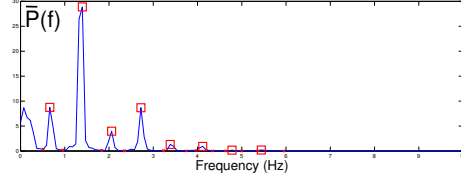

Figure 4. The weighted spectrum $\overline{P}(f)$ with the detected spectral harmonics peaks ($\square$) and their supports ($\times$).

estimate, even if there is no periodicity in the data. It is therefore important to separate those cases by measuring the strength of the periodic component. Ideally, low value should be assigned for non-periodic motion, e.g. translational movement. To this end, we propose to compare the energy of all peaks found in $\overline{P}(f)$ with the total energy:

$$per_{est} = \frac{\sum_K E_\Delta(f_k)}{\sum_f \overline{P}(f)} \quad (7)$$

with M as the number of peaks detected and $E_\Delta(f_k)$ as the area of a triangle formed by the peak and its left and right support (see Figure 4). Note that the peak supports should have zero energy for the spectrum of periodic signal. By only using the triangle area for the nominator in eq. (7), we assign a lower per_{est} value for quasi-periodic signal.

For pure periodic motion, the above analysis is sufficient. However, for quasi-periodic motion, we use data windowing. The window length depends on the rate of the instantaneous frequency changes and is given by [3, Ch. 2.3]:

$$\Delta = \sqrt{2} \left| \frac{df_i(t)}{dt} \right|^{-1/2} \quad (8)$$

The equations (2), (5)-(7) give the core of our method, leading to the periodic frequency f_{est} in eq. (6). The algorithm is applied with the following parameters: pPCA retained variance 75%, Hanning window for periodogram smoothing, dilation operator with size 7 to find frequency peaks and peak difference threshold of 0.1 to ignore small peaks. It is applied to all test data without alteration.

5. Experimental results

We designed two experiments to evaluate our method's performance. In the first experiment, seven short sequences are used to assess the accuracy of the period estimation. A 30 minutes CNN news sequence is used in the second experiment to evaluate the degree of periodicity measure.

For accuracy evaluation we picked seven sequences from the Internet, each representing a fundamental case described in section 2. Our dataset, as summarized in Table 2, represents more cases than the ones discussed previously [e.g. 6, 10, 26]. All sequences are converted to grayscale prior

Table 2. Overview of the dataset per case (See also Figure 1). The dataset size is given as the frame size and the number of frames.

Dataset name	Size	Ground truth period
	(w × h × #frames)	mean ± std (fr/cl)
Flash - railroad sign	111 × 127 × 152	20.07 ± 0.93
Pulse - anemone	80 × 64 × 160	18.88 ± 0.99
Swing - walking	96 × 172 × 118	12.20 ± 0.45
Spin - swimming	136 × 72 × 474	53.94 ± 0.56
Shuttle - brushing	142 × 120 × 150	13.57 ± 0.79
Drift - escalator	88 × 72 × 250	21.27 ± 0.47
Thrust - kangaroo	96 × 66 × 200	18.00 ± 0.50
Spinning disc	244 × 208 × 250	5.00 - 20.00

Table 3. The quantitative results of our method applied on the datasets in Table 2, compared to the results from our improved implementation of Ran's method. #PC is the number of principal components used to reach 75% retained variance. ([†]average value)

Dataset	Our method			Improved Ran method
	#PC	Estimated period	per_{est}	Estimated period
		$p_{est} \pm \epsilon_p$ (fr/cl)		$p_{est} \pm \epsilon_p$ (fr/cl)
Flash	1	20.08 ± 1.32	0.99	20.08 ± 1.32
Pulse	5	18.79 ± 1.10	0.20	18.96 ± 1.10
Swing	15	12.49 ± 0.66	0.28	12.49 ± 0.66
Spin	33	**53.89** ± 3.06	0.50	55.35 ± 3.06
Shuttle	11	**13.65** ± 0.62	0.78	14.22 ± 0.62
Drift	5	21.33 ± 0.91	0.99	21.33 ± 0.91
Thrust	16	**17.96** ± 0.81	0.41	18.62 ± 0.81
Sp. disc	9	4.82 - 20.48	0.73[†]	-

to the processing. As some objects have translational motion, tracking is applied on all sequences except for the motion of regular pattern case like the escalator sequence. The appearance variation in each sequence after tracking is stationary and quasi-periodic. The spinning disc sequence is added to evaluate the performance of our method against non-stationary periodicity. In this sequence the disc rotation speed reduces slowly toward the end of the sequence. The ground truth of all sequences are determined by visual inspection (represented in frames/cycle).

The results of our method applied to the sequences described above are summarized in Figure 5 and Table 3. To improve the frequency detection, zero-padding with $N_{FFT} = 2048$ is applied on the modified periodogram. The theoretical uncertainty measure of the estimation results is derived from the frequency resolution of the original sequence length $\epsilon_f = f_s/N$:

$$\epsilon_p = \left|\frac{dp}{df}\right| \epsilon_f = \left(\frac{1}{2f_{est}^2}\right)\left(\frac{f_s}{N}\right) = \frac{f_s}{2Nf_{est}^2} \qquad (9)$$

As expected, on videos with cluttered background (the last four) more components are needed to capture 75% of the variations. We observe that the first few components captured the periodic variation, even if the tracking is not perfect. This is illustrated in Figure 5(b) where the camera is shifted in the middle of the sequence, and in Figure 5(d) where a gradual camera viewpoint change is observed. This effect is also represented by the lower degree of periodicity per_{est} of those cases. In the case of non-stationarity (spinning disc), the windowing method is able to track the gradual deceleration of the rotation. In general, when compared to the ground truth, an error of less than 0.5 frame/cycle is achieved.

We have implemented and improved the first part of Ran et al.'s method [26] and applied it to the same dataset. Although not described in the original paper, we also applied the same amount of zero-padding as in our method to perform correct comparison. Ran's method is based on two important assumptions: the motion must be quasi-periodic and no DC component (i.e. shift of intensity value) is present in the pixel variations. The estimation results of Ran's method is included in Table 3 after ignoring the first few values of

the periodogram, just as in our method. Without this step, the DC peak would have been selected.

As we can see, the estimation of both our method and our improved Ran implementation is comparable since both methods used periodogram-based frequency analysis (Hence the same uncertainties). The difference lies in where it is applied. Ran's method applied periodogram analysis on every pixel, while we applied it a lot less frequent; only on the Q principal components. Our method seems to perform better (given in bold in Table 3) for complex data such as the *Spin*, *Shuttle* and *Thrust* cases, owing to PCA who separates spatially co-varying parts of the image.

The algorithm was also applied on a randomly selected 30 minutes sequence of CNN news from TRECVID 2005 dataset [28] without any parameters adjustment. The goal here is to detect general periodicity in the full frames, so unlike the first experiment tracking is not applied. Visual inspection yields 69 instances of periodicity in the whole sequence. Detecting periodicity in this sequence is very challenging as the periodic motion is often a tiny fraction of the whole frame and only a short number of cycles is present in each instance. A window of 4 seconds is used, but no further training, tuning or testing of the algorithm was done. We consider $per_{est} > 0.1$ as a positive detection. With this non-adaptive whole screen algorithm, we detected 58 events of periodicity (80% of 69). The remaining 20% were too small where the periodic component is overpowered by non-periodic variation. False detection is invoked in 13 instances when non-periodic abrupt changes produce ringing in the frequency spectrum. These figures compare very favorably to common object detector rates. Figure 6 shows a 4 minutes interval of the sequence. Events A, B and D are people shaking their heads. The weak response of event C shows a *Drift* case when the camera pans 3 standing boys. Event E shows some flashing as quick camera shot changes occur during advertisement. And event F is the flashing of a white patch.

(a) Flash - railroad sign   (b) Pulse - anemone   (c) Swing - walking   (d) Spin - swimming

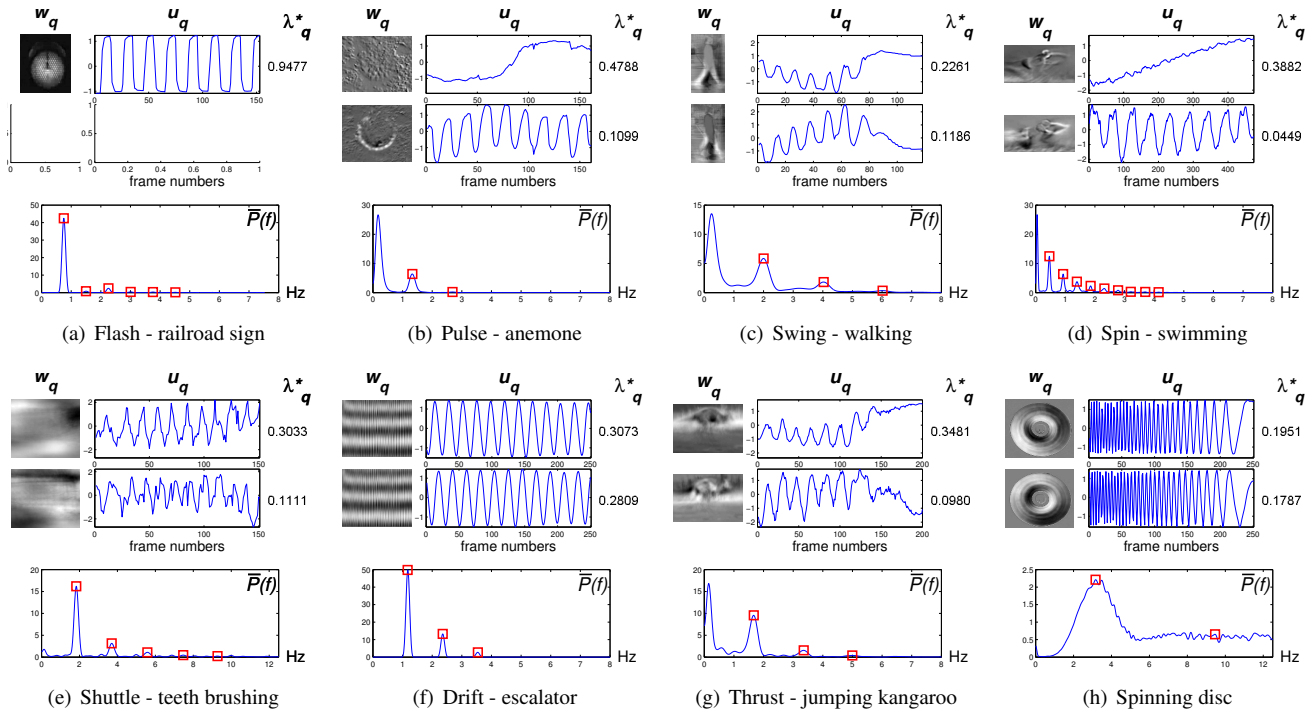(e) Shuttle - teeth brushing   (f) Drift - escalator   (g) Thrust - jumping kangaroo   (h) Spinning disc

Figure 5. The results of the proposed approach applied on the datasets in Table 2. Only the first 2 components are visualized here. For each component, the eigen image $\mathbf{w}_q$ and the principal component in time axis direction $\mathbf{u}_q^T$ is plotted, along with the percentage of retained variance λ_q^*. The weighted spectrum of all components is shown at the bottom of each figure.
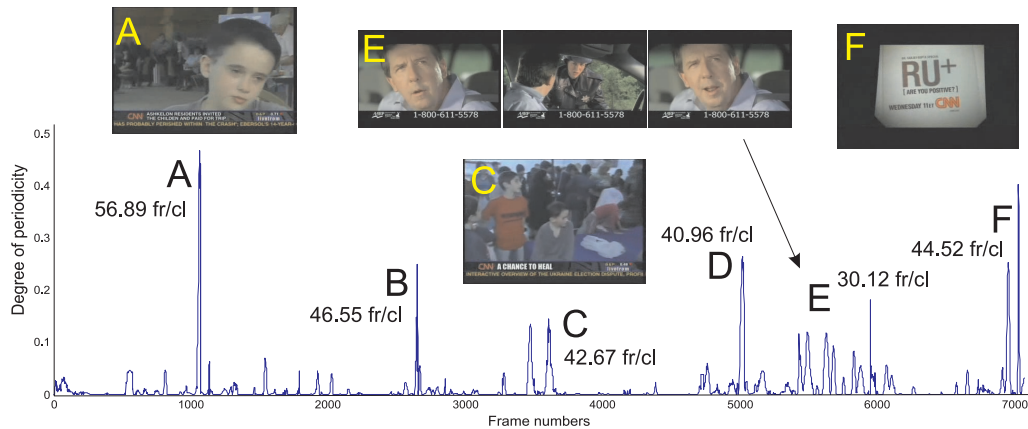

Figure 6. Results of the degree of periodicity estimation applied to a 4 minutes segment of CNN news sequence.

6. Conclusions

Visual periodicity is an important cue for many applications such as detection of gait and performance measurement of top athletes. This paper takes the following steps to investigate the topic:

1. We propose a conceptual framework for the categorization of visual periodicity in 10 fundamental cases.
2. We present a model to detect all these types of periodicity by one and the same algorithm. The method is based on tracking the hole in the background and on probabilistic Principal Component Analysis (pPCA). By spectral analysis on the principal components we estimate the frequency and the degree of periodicity. Due to the nature of PCA to group spatially coherent changes over time, our method has performed favorably to a recently developed method [26], especially on cluttered sequences.
3. The method, which is tuned on the small data set of 10 cases, is also applied without any alteration on a random half hour of CNN news to detect periodicity successfully.

We have good reasons to assume that periodicity is an important indicator for the attention arousal as illustrated in Figure 6. In a sequel study we will evaluate the attention of periodicity as detected by our algorithm in a quantitative manner.

Acknowledgement

References

[1] A. B. Albu, R. Bergevin, and S. Quirion. Generic temporal segmentation of cyclic human motion. *Pattern Recognition*, 41(1):6–21, 2008.

[2] M. Allmen and C. R. Dyer. Cyclic motion detection using spatiotemporal surfaces and curves. *Proc. IEEE Int. Conf. Pattern Recog.*, 1:365–370, 1990.

[3] B. Boashash. *Time frequency signal analysis and processing: A comprehensive reference*. Elsevier, Amsterdam, 2003.

[4] J. E. Boyd. Synchronization of oscillations for machine perception of gaits. *Comp. Vision Image Underst.*, 96:35–59, 2004.

[5] A. Briassouli and N. Ahuja. Extraction and analysis of multiple periodic motions in video sequences. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(7):1244–1261, 2007.

[6] G. J. Burghouts and J.-M. Geusebroek. Quasi-periodic spatio-temporal filtering. *IEEE Trans. Image Process.*, 15(6):1572–1582, 2006.

[7] C. Chang, R. Ansari, and A. Khokar. Efficient tracking of cyclic human motion by component motion. *IEEE Signal Proc. Lett.*, 11(12):941–944, 2004.

[8] D. Chetverikov and S. Fazekas. On motion periodicity of dynamic textures. *Proc. British Machine Vision Conf.*, 1:167–176, 2006.

[9] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *Proc. European Conf. Comp. Vision*, 2:484–498, 1998.

[10] R. Cutler and L. S. Davis. Robust real-time periodic motion detection, analysis and applications. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(8):781–796, 2000.

[11] J. Davis, A. Bobick, and W. Richards. Categorical representation and recognition of oscillatory motion patterns. *Proc. IEEE Conf. Comp. Vision Pattern Recog.*, 1:628–635, 2000.

[12] L. Favreau, L. Reveret, C. Depraz, and M.-P. Cani. Animal gaits from video: Comparative studies. *Graphical Models*, 68(2):212–234, 2006.

[13] H. Fujiyoshi and A. J. Lipton. Real-time human motion analysis by image skeletonization. *Proc. IEEE Workshop Applicat. Comp. Vision*, pages 15–21, 1998.

[14] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky. Behavior classification by eigendecomposition of periodic motions. *Pattern Recognition*, 38(7):1033–1043, 2005.

[15] M. H. Hayes. *Statistical Digital Signal Processing and Modeling*. John Wiley & Sons, New York, 1996.

[16] B. Heisele and C. Wöhler. Motion-based recognition of pedestrians. *Proc. IEEE Int. Conf. Pattern Recog.*, 2:1325–1330, 1998.

[17] I. Laptev, S. J. Belongie, P. Perez, and J. Wills. Periodic motion detection and segmentation via approximate sequence alignment. *Proc. IEEE Int. Conf. Comp. Vision*, 1:816–823, 2005.

[18] F. Liu and R. W. Picard. Finding periodicity in space and time. *Proc. IEEE Int. Conf. Comp. Vision*, pages 376–383, 1998.

[19] C. M. Lu and N. J. Ferrier. Repetitive motion analysis: Segmentation and event classification. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(2):258–263, 2004.

[20] H. T. Nguyen, Q. Jie, and A. W. M. Smeulders. Spatio-temporal context for robust multi-target tracking. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(1):52–64, 2007.

[21] H. T. Nguyen and A. W. M. Smeulders. Robust tracking using foreground-background texture discrimination. *Int. J. Comp. Vision*, 68(3):277–294, 2006.

[22] D. Ormoneit, M. J. Black, T. Hastie, and H. Kjellström. Representing cyclic human motion using functional analysis. *Image and Vision Computing*, 23(14):1264–1276, 2005.

[23] R. Polana and R. C. Nelson. Detection and recognition of periodic, nonrigid motion. *Int. J. Comp. Vision*, 23(3):261–282, 1997.

[24] F. Potdevin, B. Bril, M. Sidney, and P. Pelayo. Stroke frequency and arm coordination in front crawl swimming. *Int. J. Sports Med.*, 27(3):193–198, 2006.

[25] M. M. Rahman and S. Ishikawa. Human motion recognition using an eigenspace. *Pattern Recog. Lett.*, 26(6):687–697, 2005.

[26] Y. Ran, I. Weiss, Q. Zheng, and L. S. Davis. Pedestrian detection via periodic motion analysis. *Int. J. Comp. Vision*, 71(2):143–160, 2007.

[27] S. M. Seitz and C. R. Dyer. View-invariant analysis of cyclic motion. *Int. J. Comp. Vision*, 25(3):1–23, 1997.

[28] A. Smeaton, P. Over, and W. Kraaij. Evaluation campaigns and TRECvid. In *Proc. 8th ACM Int. Workshop Multimedia Information Retrieval*, pages 321–330, 2006.

[29] A. Thangali and S. Sclaroff. Periodic motion detection and estimation via space-time sampling. *Proc. IEEE Workshop Motion and Video Computing*, 2:176–182, 2005.

[30] M. E. Tipping and C. M. Bishop. Probabilistic principal component analysis. *J. of Royal Stat. Society, Series B*, 61(3):611–622, 1999.

[31] X. Tong, L. Duan, C. Xu, Q. Tian, and H. Lu. Local motion analysis and its application in video based swimming style recognition. *Proc. IEEE Int. Conf. Pattern Recog.*, 2:1258–1261, 2006.

[32] P. Tsai, M. Shah, K. Keiter, and T. Kasparis. Cyclic motion detection for motion based recognition. *Pattern Recognition*, 27(12):1591–1603, 1994.

[33] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neurosci.*, 3(1):71–86, 1991.

[34] Z. Zivkovic and J. Verbeek. Transformation invariant component analysis for binary images. *Proc. IEEE Conf. Comp. Vision Pattern Recog.*, pages 254–259, 2006.