Recognition of Human Periodic Motion - a Frequency Domain Approach

Baihua Li   and   Horst Holstein

Department of Computer Science

University of Wales, Aberystwyth, SY23 3DB, Wales UK

bbl00@aber.ac.uk   hoh@aber.ac.uk

Abstract

We present a frequency domain analysis technique for modelling and recognizing human periodic movements from moving light displays (MLDs). We model periodic motions by motion templates, that consist of a set of feature power vectors extracted from unidentified vertical component trajectories of feature points. Motion recognition is carried out in the frequency domain, by comparing an observed motion template with pre-stored templates. This method contrasts with common spatio-temporal approaches. The proposed method is demonstrated by some examples of human periodic motion recognition in MLDs.

1. Introduction

The perception of high-level motion is crucial in a variety of applications including "smart" surveillance, human-machine interaction, and medical studies. The recognition of articulated periodic movements, such as performed by humans, is a routine function of the human visual system. It has been demonstrated extensively by Johansson's Moving Light Displays (MLDs) [7]. The MLDs illustrate that humans can perceive complex human movements solely on the reduced motion cues from a small number of unstructured moving dots on a subject.

In the computer vision literature, there are two kinds of theories about motion recognition [4]. One is structure-based recognition. Perception depends on identification of human kinematic structure from motion cues, such as joint angle [6] [3]. The other theory, motion-based recognition, deals with direct use of motion information over time for recognition, such as suggested by MLDs. The latter avoids the complex vision problem of structure recovery.

In this paper we propose motion-based (non-structural) recognition of periodic motions using the frequency domain. The rest of the paper is organized as follows: Section 2 reviews related work on cyclic motion recognition. Section 3 describes the frequency recognition approach in details. Section 4 provides experimental results on recognition of human cyclic motion. We discuss and conclude our work in sections 5 and 6.

2. Related work

In related non-structural approaches, human motion, specifically walking, has been studied extensively using various spatio-temporal cues from images or MLDs. Polana and Nelson [8] propose a method for detecting periodic motion using Fourier transforms on several point trajectories. They indicated that in principle, the period of the movement could be inferred from averaging the fundamental frequencies of the point trajectories. Tsai et al. [9] apply a similar method to the curvature of one trajectory. An autocorrelation is performed to enhance self-similarity within the curvature function. The Fourier Transform is finally used to detect the presence of a cycle and its period. Fujiyoshi and Lipton [5] generate a "star" skeleton from the object boundary. They apply Fourier analysis to the skeleton for detecting periodic motion. Then they utilize both posture and motion cycle of the "star" skeleton to recognize activities such as walking and running. Recent work by Boyd and Little [2], using global shape-of-motion features derived from MLD images, has shown that it is possible to recognize individual people from their gaits, by non-structural means.

In these approaches, Fourier transforms are usually used to detect or recognize periodicity. The detected periodicity is used to assist recognition. For spatio-temporal domain approaches, in order to deal with the problems of human motion irregularities or change in speed, techniques such as scale space or Dynamic Time Warping, considered computationally expensive, are usually used to match portions of scale space to find repeated patterns from identified curves.

The direct application of frequency domain analysis for motion recognition has received much less attention in the literature. We are pursuing the interpretation and recognition of human periodic motion using a pure frequency domain approach. We believe that in common periodic move-

ments, such as walking, running and skipping, motion characteristics exist not only in the spatio-temporal domain, but also in the frequency domain. They are preserved even in the reduced information of kinematic data from MLDs, allowing motion recognition without knowledge of underlying structure.

3. Method

All human kinematic data used in our investigation are acquired from a commercial 3D marker-based optical motion capture (MoCap) system, the Vicon 512. The system provides 3D coordinates of unidentified trajectories of markers attached to feature points on a subject, in the manner of a 3D-MLD system. Feature points used in our experiments are at locations of head, torso and on each of shoulder, elbow, wrist, hip, knee, ankle and toe. The high quality captured data allows us to apply a frequency domain approach directly for motion recognition. In this study, we use the vertical-components, which are the Z-coordinates of 3D-MLD trajectories, as the only motion cue to be analyzed. In Fig.2 and Fig.3, we have indicated feature point identity for clarity, but this information is not used in the identification process. We find that cues from the unidentified tracks suffice to discriminate between a number of simple periodic human activities.

3.1. Power spectral analysis of periodic movement

We apply spectral analysis to a vertical-component trajectory $Z_{i(n)}$ in frames $n = 1 \ldots N$, of each feature point $i = 1 \ldots I$. The decomposition is expressed by

$$Z_{i(n)} = \frac{1}{2}a_{i(0)} + \sum_{k=1}^{N/2} a_{i(k)}\cos(2\pi k\Delta f) + b_{i(k)}\sin(2\pi k\Delta f)$$

(1)

where f_s is the sample rate and $\Delta f = f_s/N$ denotes the frequency resolution. The DC component $\frac{1}{2}a_{i(0)}$ denotes the average vertical position of this point relative to the origin on the ground during motion. When the sequence length N is an exact power of two, we adopt the Fast Fourier Transform (FFT) algorithm which is more efficient than the raw Discrete Fourier Transform (DFT). The power spectrum for the kth frequency multiple of feature trajectory i is

$$P_{i(k)} = a_{i(k)}^2 + b_{i(k)}^2, \quad k = 1, 2 \ldots N/2$$

(2)

An example of such a power spectrum for the toe is shown in Fig.1, in which N=1024, $f_s = 60Hz$, and the gait-cycle frequency is about 1Hz. From a number of experiments, we observed the dominant power of normal human
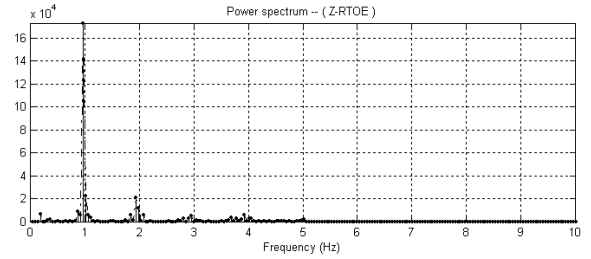

Figure 1. Power spectrum for the toe of a walking person

periodic movements occupy only a narrow bandwidth, with an upper limit less than 10Hz, similar to results in [1]. The power spectrum distribution shows clustering around some frequencies related to the fundamental activity cycle, such as the gait cycle in walking and running, and its multiples. The magnitude of the spectral component is time shift invariant. It retains its spectral envelope regardless of where in time the FFT is performed. For the same kind of motion in different subjects, the spectral distribution is largely similar, hinting at the motion nature, differences being related to variation in individual speed and amplitude.

To obtain speed-invariant representation, we normalize these spectra to the fundamental activity cycle or generalized gait cycle (Gc). To achieve an accurate Gc, we sum power spectra within a band-limited frequency $[0.4 \sim 3Hz]$. The frequency corresponding to the maximum power magnitude in the summed-spectrum,

$$P_{Gc} = \max_{0.4Hz < k < 3Hz}\left\{\sum_i P_{i(k)}\right\}$$

(3)

is regarded as the activity cycle, or generalized gait-cycle.
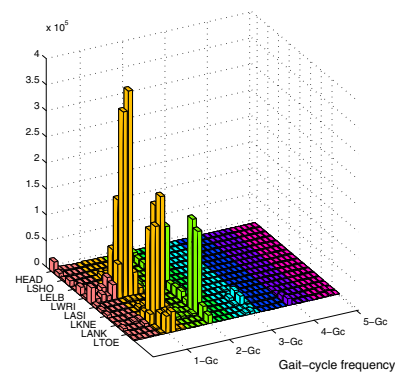

Figure 2. The whole body power spectra of a walking person (scaled by Gc)

The cycle frequency is subsequently used to normalize the frequency axis from Hz to Gc. Fig.2 shows an example

of Gc-scaled whole body power spectra of a walking person. We observe that power components are reflected by the active body parts. This is to be expected since they undergo different motions.

3.2. Feature power vector and motion template

In order to achieve an adequate frequency resolution $\Delta f = f_s/N$, long sequences are required. Moreover, a uniform frequency resolution between different trials is needed in order to make a direct spectral comparison feasible. However, in practice, a random length sequence N may contain only a few cycles and result in an imprecise spectrum and different frequency resolution.

To mitigate this problem, we extract a set of powers from the spectrum of each trajectory, regarded as a feature power vector $\vec{\nu}_i$, as shown in Eq.4,

$$\nu_{i(0)} = DC_i$$

$$\nu_{i(n)}|_{n=1,2...5} = \sum_{k \in W_n} P_{i(k)}$$

$$\nu_{i(6)} = \sum_{k \le 6Gc, k \notin W_{1,2...5}} P_{i(k)} \qquad (4)$$

where $\Delta Gc = \Delta f/Gc$ is the Gc-scaled frequency resolution. The first component $\nu_{i(0)}$ of $\vec{\nu}_i$ of each feature point i is its DC component. For $\nu_{i(1)}$ to $\nu_{i(5)}$, we utilize a shift sum-window $W_{1,2...5}$ to sum three power components around each of the 1st-Gc, 2nd-Gc ... 5th-Gc frequencies respectively. The last item $\nu_{i(6)}$ is used to compensate small power components lying up to the $6th$ Gc that have not yet been taken into account. Using feature power vectors also reduces the number of frequency components to be considered for matching.

We stack I feature power vectors together as a motion template, $V = \{\vec{\nu}_i \mid i \le I\}$. The template can be viewed as an $I \times 7$ array. Each column in the motion template is scaled relative to the maximum value in this column, to normalize the power amplitude. An averaging technique is used to smooth the small differences of individuals to generate a standard motion template. Some examples of motion templates are shown in Fig.3. In these examples, we find high-frequency feature power $\nu_{i(4)}$ and $\nu_{i(5)}$ are very small relative to the power on the 1st-Gc, 2nd-Gc and 3rd-Gc, so we set them to zero.

3.3. Motion recognition

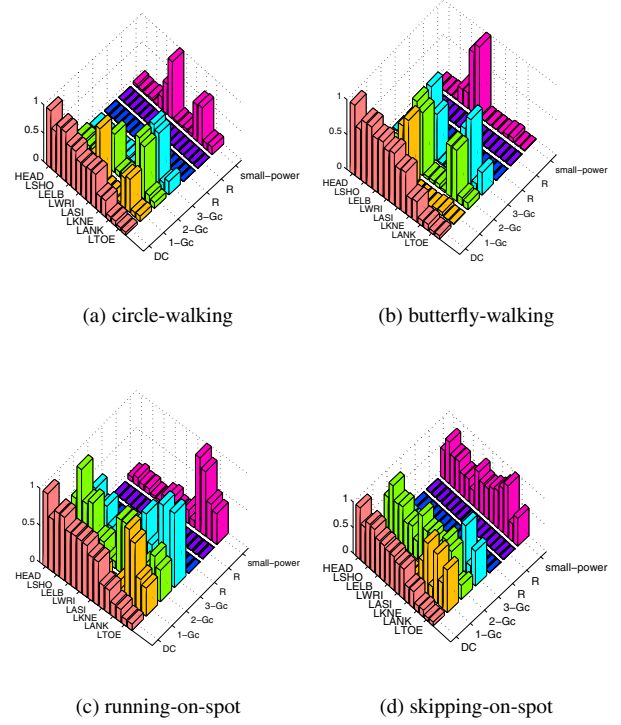Recognition entails finding the best match for an observed motion template among pre-stored motion templates.

(a) circle-walking   (b) butterfly-walking

(c) running-on-spot   (d) skipping-on-spot

Figure 3. Examples of motion templates

We apply the same algorithm to an observed motion to create its motion template $U = \{\vec{\mu}_j \mid j \le J\}$, with J feature power vectors. We use an $I \times J$ match matrix $M = \{m_{(i,j)} \mid i \le I, j \le J\}$, see Eq.5, to measure the weighted-difference between the ith motion vector $\vec{\nu}_i$ in a model template and jth motion vector $\vec{\mu}_j$ in the observed template,

$$m_{(i,j)} = \sum_{n=0}^{6} |\nu_{i(n)} - \mu_{j(n)}| \, \omega_n \qquad (5)$$

where $\vec{\omega} = [0.4, 0.25, 0.25, 0.05, 0, 0, 0.05]$ is the weight vector. The best match of point j should be the point corresponding to the minimum element $d_j = \min_i\{m_{(i,j)}\}$ in a row of M. Subsequently, whole motion similarity $\mathcal{S}$ is measured according to all best matches as shown in Eq.6.

$$\mathcal{S} = 1 - \frac{\sum_{j=1}^{J} d_j}{J} \qquad (6)$$

The motion with maximum similarity $\mathcal{S}$ for all the searched templates is taken to indicate recognition. The feature points of the observed motion can be an adequate subset ($J \le I$) of the standard database templates.

4. Experimental Results

All experiments were performed using real motion capture data from MoCap system Vicon 512, $f_s = 60Hz$, $N \approx 350 \sim 1300$. Some small gaps raised by occlusion were interpolated during pre-processing. The proposed algorithm was tested on recognition of some normal cyclic motions, such as walking-on-spot, circle-walking, butterfly-walking (walking with waving hands up and down), running-on-spot, circle-running, skipping-on-spot. Recognition indicated by motion similarity parameter $\mathcal{S}$ is given in Table 1. The results are averages of a group of ten subjects that consisted of males and females, with ages from 5 to 60 years. The highest column entries, high-lighted, occur when the observed row activity matches the model column activity.

Table 1. Recognition of human periodic movements

motion template	circle -walking	running on-spot	skipping on-spot
circle-walking	**0.98**	0.85	0.84
walking-on-spot	0.94	0.89	0.85
butterfly-walking	0.92	0.86	0.86
running-on-spot	0.86	**0.97**	0.95
circle-running	0.89	0.95	0.93
skipping-on-spot	0.85	0.94	**0.97**

From the results above, we observe that the proposed frequency analysis approach is able to classify cyclic motions solely on unidentified vertical-component trajectories. Because the algorithm utilizes a whole-body power spectral analysis, detailed recognition is realized by comparison of the movement on each feature point. The different classes of movements, such as walking and running, have larger differences in the similarity parameter $\mathcal{S}$ due to the considerable difference among feature power vectors; while the similar movements, such as running-on-spot and circle-running, can also be distinguished, because the magnitudes of power spectra on left limbs usually differ from that on right limbs in circular activities.

5. Discussion

Our experiments assumed availability only of unidentified vertical-component trajectories obtained from 3D-MLDs. Normalization both on frequency and power magnitude allows recognition to be carried out in frequency domain for a wide range of subjects. We can combine this approach with motion periodicity assisted recognition $\mathcal{S}' = \{\mathcal{S}, \xi_{Gc}\}$, where ξ_{Gc} is a Gc-related recognition factor, that presents a more precise recognition.

In practice, for shorter sequences $N < 512$, to guarantee a desired frequency resolution, we suggest using the whole length for Fourier transform, rather than cutting it to a length of an exact power of two. In this case, the FFT is reduced to the raw Discrete Fourier transform (DFT). This is a tradeoff between frequency resolution and computational cost. Nevertheless, because each motion template contains a small number of parameters by using feature power vector analysis, and because we also only use one standard template for indexing each kind of motion, a short matching time and small database are required. The method proved to be computationally efficient.

6. Conclusions

We have described a pure frequency domain technique for human periodic motion recognition. We address a motion-based approach using unidentified vertical-component trajectories from 3D-MLDs. The experimental results indicate that the inherent characteristics of periodic movements exist not only in the spatio-temporal domain but also in the frequency domain. Frequency domain features hint motion nature which can be used to classify different periodic activities, and even discriminate similar movements within the same class.

Acknowledgments:

References

[1] C. Angeloni, P. O. Riley, and D. E. Krebs. Frequency content of whole body gait kinematic data. *IEEE Transaction on Rehabilitation Engineering*, 2(1):40–46, March 1994.

[2] J. Boyd and J. Little. Global versus structured interpretation of motion: moving light displays. In *Proceedings of the Computer Vision and Pattern Recognition*, 1997.

[3] L. Campbell and A. Bobick. Recognition of human body motion using phase space constraints. In *Proc. of International Conference on Computer Vision*, pages 624–630, Cambridge, 1995.

[4] C. Cedras and M. Shah. Motion-based recognition: A survey. *IVC*, 13(2):129–155, March 1995.

[5] H. Fujiyoshi and A. Lipton. Real-time human motion analysis by image skeletonization. In *Proceedings of the IEEE Workshop on Applications of Computer Vison*, 1998.

[6] N. H. Goddard. *The Perception of Articulated Motion: Recognizing Moving Light Displays*. PhD thesis, University of Rochester, 1992.

[7] G. Johansson. Visual motion perception. *Scientific American*, 232(6):75–80, 85–88, 1975.

[8] R. Polana and R. Nelson. Detecting activities. In *DARPA93*, pages 569–574, 1993.

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