



Synchronization of oscillations for machine perception of gaits

Jeffrey E. Boyd

Department of Computer Science, University of Calgary, Calgary, Canada T2N 1N4

Received 18 March 2003; accepted 27 April 2004

Available online 6 August 2004

Abstract

Substantial evidence supports a relationship between gait perception and gait synthesis. Furthermore, passive mechanical systems demonstrate that the jointed leg systems of humans have innate oscillations that form a gait. These observations suggest that systems may perceive gaits by synchronizing an internal oscillating model to observed oscillations. We present such a system in this paper that uses phase-locked loops to synchronize an internal oscillator with oscillations from a video source. Arrays of phase-locked loops, called video phase-locked loops, synchronize a system with the oscillations in pixel intensities. We then test the perception of the resulting synchronized-oscillator model in various gait recognition tasks. Tools based on Procrustes analysis and directional statistics provide the computational mechanism to compare patterns of oscillations. We discuss the possibility of an alternative model for motion perception based on synchronization with the transient oscillations of temporal band-pass filters that is consistent with other proposed models for human perception. Synchronization of a kinematic model to oscillations also suggests a path to bridge the gap between the model-free and model-based domains.

© 2004 Published by Elsevier Inc.

Keywords: Gait analysis; Gait recognition; Gait perception; Motion analysis; Oscillating motion; Phase-locked loops; Optical flow

E-mail address: boyd@cpsc.ucalgary.ca.

1. Introduction

Perception of gaits is important for two reasons. First, people continually generate gaits using an innate ability to produce the necessary coordination of motions. Second, people continually view human gaits as part of their daily interactions. Perception of gaits thus forms an important part of how people perceive the actions of others.

Evidence for a relationship between gait perception and gait synthesis exists. Berntenthal and Pinto [8] present evidence gleaned from psychological experiments. Fazio et al. [31] indicate, based on PET studies, that the same parts of the brain that are used to perform an action are also used in the perception of similar actions. Cohen et al. [20] observed that while humans can easily recognize human motion, they have more difficulty recognizing animal motion. Cohen et al. explain this observation by suggesting that humans rely on the same mechanisms that they use to generate their own gait to perceive the gaits of others.

Passive mechanical systems that exhibit gaits suggest that walking is an innate part of the kinetic structure of the human body. For example, McGeer [49,50] describes two-dimensional passive mechanical models that walk like humans. The models are two-dimensional because they can move only in the sagittal plane. More recently, Coleman and Ruina [22], Garcia et al. [34], and Collins et al. [24] describe passive mechanical machines that walk in three dimensions. Collectively, these systems demonstrate that the jointed leg systems of humans have an innate ability to walk with little or no external control. The role of control in a gait is to deal with deviations from the innate gait made necessary by variations in terrain or the need to change directions. Laszlo et al. [43] describe such a control system that synthesizes gaits for computer graphics.

Based on this evidence we form the following two hypotheses related to the perception of gaits.

1. People have an internal gait model that is used to synthesize their own gait. This model is a combination of a person's own kinematic structure that has an innate ability to walk and a control system that produces variations of the gait as needed.
2. Humans use this internal gait model to perceive the gaits of others.

Although the evidence cited above supports these hypotheses, we cannot know if they are true. However, if they are true, then the hypotheses suggest a way to build machine systems that perceive gaits by synchronizing an external stimulus with an internal gait model. We present such a system in this paper. Phase-locked loops (PLLs) are control systems that synchronize the oscillations of an internal oscillator with oscillations from an external source. We use arrays of phase-locked loops, called video phase-locked loops (vPLLs), introduced by Boyd [13], to synchronize a system with the oscillations in pixel intensities that occur when viewing a gait or other oscillating stimulus.

We then use the information extracted from the synchronized internal oscillators to recognize gaits as well as perform other tasks related to gait perception, including identifying viewpoint, variations in gaits, and individual gaits (Section 4). While ma-

chine gait recognition is a useful task, it is not sufficient to completely justify the use of synchronized oscillations. In Section 5 we discuss potential connections between synchronizing oscillations using vPLLs, and the perception of motion in general, and the possibility of connecting model-based and model-free representations of human motion.

1.1. Oscillations and gait perception

A sensitivity to oscillations is a fundamental part of human vision. Watson [69] documents the temporal band-pass filter properties of the human eye with the center of the pass band at approximately 10Hz. Geisler and Albrecht [35] also measured this band-pass effect at the neurological level in monkeys and cats. The band-pass filter not only attenuates steady-state oscillations outside of its passband, but also produces a damped oscillation in response to transient inputs. Watson and Ahumada [70], and Adelson and Bergen [1] incorporate this transient response in their models of motion perception. Recent work by Lee and Blake [45], Usher and Donnelly [65], Sekuler and Bennett [58], Farid and Adelson [30], Farid [29], and Shadlen and Movshon [59] addresses the role of oscillations and synchronization in perceptual grouping, referred to as temporal synchrony. Ng and Gong [53] have applied concepts of temporal synchrony to visual surveillance.

Not only are oscillations an important part of human motion perception in general, but they play a critical role in the perception of gaits. Bertenthal and Pinto [8] identify *frequency entrainment* (component motions of the gait sharing a common frequency), and *phase locking* (constant phase/timing relationships among the component motions) as oscillation-related properties of a gait that are essential for human perception.

The importance of phase in gaits appears in medical literature. Murray et al. [52] show a detailed analysis of the gaits for several *normal* humans. The analysis focuses on the timing of joint and limb movements throughout the cycle of the gait. In subsequent work, Murray [51] shows that these sinusoidal variations change in phase for subjects that have abnormal gaits due to some physical affliction. The changes in the gaits are manifest in the timing patterns.

Systems that synthesize gaits also show the importance of phase. For example, Laszlo et al. [43] describe a limit-cycle control system for animation of walking figures. Varying phases in the limit cycles alters the type of synthesized gait. Unuma et al. [64] describe a Fourier-based system for gait synthesis. Their system derives Fourier coefficients for periodic joint angles. The phase component of the Fourier series plays a significant role in the synthesis of the varying gaits.

1.2. Evaluation of recognition

We make the distinction between gait recognition and people recognition (biometrics) using gaits. In this paper we cite people recognition rates in the results. We do this partly to conform to common practice, but also because we lack a better way of evaluating machine perception. Just as psychologists measure recognition rates when

they study human motion perception from moving light displays, we too look at recognition rates to verify that a system or model exhibits desired properties. Consequently, we suggest that errors in people recognition be treated carefully. For example, if I wish to analyze a gait to evaluate the effectiveness of an orthotic insole, then an excellent biometric measure that is invariant to footwear is unlikely to be useful.

Recognition rates are extremely sensitive to context, including sampling method and sample size. For example, it is possible to inflate a recognition rate by using a small sample. Cumulative match characteristics (CMC) also depend on sample size. Recognition rates may also vary with the classifier used.

In this paper we use the F statistic that is the basis of analysis of variance (ANOVA) to address the issue of sample size. Appendix B describes an extrapolation of ANOVA to get F_{WW} , a descriptive statistic useful for periodic signals like those often found in gait analysis. Bobick and Johnson [12] characterize gait recognition systems in a similar manner by computing the *expected confusion*, $E[A] = \frac{\sigma_i}{\sigma_p}$, (related to confusion matrices) using: where σ_i is the standard deviation for a group of individuals, and σ_p is the standard deviation for the population. From $E[A]$ they predict recognition rates. $E[A]$ captures the same properties of the data as the F statistic. In fact, $F = MS_B/MS_W$, where MS_B and MS_W are the *between* and *within* group mean square errors. In the context of gait recognition, MS_B and MS_W are unbiased estimates of σ_p^2 and σ_i^2 . Therefore, $E[A] = 1/\sqrt{F}$ when the sample size is infinite. Alternatively, if one views MS_B as the energy in the signal of interest, then MS_W is noise energy and F is the signal-to-noise ratio.

ANOVA evaluates the significance of a particular F with respect to the null hypothesis that all groups are identical, for a given sample size. However, we have observed that features measured by a system that can recognize tend to have a large F , and the probability of the null hypothesis, $P(H_0)$, is infinitesimal. Thus, while we report $P(H_0)$, we use F as a descriptive statistic. The value of this statistic can be gleaned from recognition rate versus sample size plots in Bhanu and Han [10]. Although it is not reported in these terms, individual curves in the plots characterize recognition rates over a broad range of sample sizes for a constant F .

2. Gait analysis systems

Machine vision systems for human motion analysis fall into two broad categories, model-free and model-based, depending on whether or not the system uses a kinematic model. This section reviews existing techniques in both categories.

2.1. Model-free

A variety of shape descriptions have seen use in gait analysis. Little and Boyd [46] use optical flow to identify the moving regions in a gait image sequence. They then describe the shape of the moving region with a set of scalar features that oscillate with the gait. The system extracts the relative phases of the scalar oscillations and

forms a phase feature vector that they use to identify individual gaits. Shutler and Nixon [61] extended this concept using Zernike velocity moments to compute shape descriptions over an entire sequence, rather than on a frame by frame basis. Lee and Grimson [44] partition the moving figure into seven regions, then analyze shape variations in each region. They apply the resulting feature vector to person and gender recognition. From a sequence of images, Davis and Bobick [28] compute motion energy images (MEI) and motion-history images (MHI) that indicate where motion is occurring and how recently the motion occurred. They describe the shape of the moving regions with a set of Hu moments, which they in turn use to recognize patterns of motion, such as various aerobic exercises. Cutler and Davis [26] identify periodicities in a vector of image intensities for a tracked object. As a periodic motion goes through its cycles, various frames are similar to others. The periodic behavior due to the various self-similarities allows the system to distinguish between human, animal, and mechanical motion. More recently Ben-Abdelkader et al. [4,7] extended this method to people recognition. Vega and Sarkar [56] measure spatial relationships among moving edge pixels, which they encode into histograms. They then compare the histograms to recognize individuals.

Several existing methods focus on the use of a silhouette of a moving figure. Baumberg and Hogg [2] describe a method that extracts the silhouette of a moving figure. They extend the concept by treating changes in shape with a vibration model [3]. More recently, Kale et al. [41], Tassone et al. [63], Wang et al. [67,68], and Hayfron-Acquah et al. [38] use a variety of metrics derived from points on the outline of a person to recognize individual people, and various walking and running gaits. Collins et al. [23] recognize individuals by comparing the shapes of silhouettes at the both the double-support (two feet on the ground) and mid-stride phase of a gait using a spatial cross-correlation.

Observations of a tracked gait produce oscillations in pixels. For example Polana and Nelson [55] examine oscillations in the magnitude of the optical flow in a sequence containing periodic motion. They compute a coarse resolution (four by four) flow magnitude image at six points in the period of the motion. From this they form a 96-element vector that is used to recognize a broad range of periodic motions. Also, Liu and Picard [47] examine oscillations in pixel intensity for a gait sequence using fast Fourier transforms (FFT). Their analysis identifies the amplitude of the fundamental frequency of the gait.

2.2. *Model-based*

Oscillations in a kinematic model occur in the joint angles and also in the positions of the limbs. Fujiyoshi and Lipton [33] use a very simplified kinematic model they call a star skeleton. While they estimate the skeleton on a frame-by-frame basis, the skeleton reveals period limb motion. Bissacco et al. [11] use Bregler's method [17] to extract joint angle information from a motion sequence. They then compute an auto-regressive moving-average (ARMA) model of the joint movement which they in turn use as a feature vector. Their method recognizes different types of gaits such as running, walking, or walking a stair case. Tanawongsuwan and Bobick [62] use

joint angle trajectories derived from a motion capture system. They synchronize trajectories to a common reference point in the gait and then re-sample so that all subjects have the same sample size. Trajectories for the various joints combine to form a feature vector used to recognize individual gaits. Cunado et al. [25] measure the angle motion of a subject's thigh. Fourier coefficients extracted from the temporal sequence of angles recognizes act as features for recognition of individuals.

Bobick and Johnson [12,40] describe a system that uses static parameters derived from a gait such as stride and torso length for biometric recognition. Similarly, Ben-Abdelkader et al. [5,6] measure height and stride characteristics, also with application to biometrics. Davis [27] uses stride length and frequency to distinguish between gaits of adults and children. Although these methods make minimal use of the dynamic characteristics of the gait, they can recognize individuals.

Several methods do not exploit the periodic nature of gaits, but take the more general approach of estimating a series of poses that may or may not be periodic. These methods include work by Hunter et al. [39], Rowley and Rehg [57], Wachter and Nagel [66], Wren et al. [71], and Bregler and Malik [18].

3. Synchronization with image oscillations using vPLLs

Video phase-locked loops are a mechanism that synchronizes an array of oscillators to intensity oscillations in an image. This section describes their operation and design.

3.1. Phase-locked loops

Fig. 1A shows a block diagram for a basic PLL. PLLs are found primarily in communications systems where they are used to lock on carrier signals. Its basic components are a phase detector (PD), a low-pass (LP) loop filter, and a voltage-controlled oscillator (VCO). The literature on PLLs is extensive. Best [9] provides an excellent introduction to PLLs.

To understand a PLL, suppose we initially know nothing about the frequency of the VCO output. The phase detector produces a *phase difference* signal, $u_d(t)$, that is proportional to the difference between the phases of $u_1(t)$ and $u_2(t)$. If u_1 and u_2 oscillate at different frequencies, u_d will increase or decrease at a constant rate. The loop filter removes high-frequency components of u_d to obtain the filtered phase difference signal, $u_f(t)$. u_f feeds back to the VCO, changing the frequency of VCO oscillations until the frequencies of u_1 and u_2 are equal. At this point the phase difference will remain constant, and the loop reaches a steady state and is said to be locked. One can think of a PLL as a *control system* in which the phase of u_1 controls the phase of u_2 . The PLL has two outputs: the VCO output, u_2 , and the LP loop filter output, u_f . Fig. 1B shows the linear relationship between the VCO control signal, u_f , and the frequency it oscillates at. K_0 is the slope of the line, while the intercept, $\omega_0 = 2\pi f_0$, is the center frequency of the oscillator. Fig. 1C shows the linear relationship between the phase-detector output, u_d , and

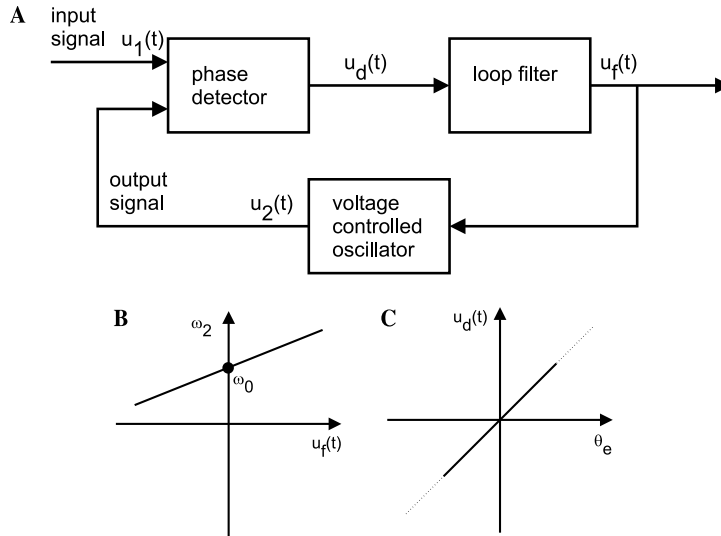


Fig. 1. The basic phase-locked loop: (A) a block diagram of the loop, (B) the transfer function of the voltage-controlled oscillator, and (C) the transfer function of the phase detector. The loop filter is usually a low-pass filter.

the phase difference between u_1 and u_2 , θ_e . The slope is K_d and the intercept is zero. Best gives a set of PLL design equations summarized in Table 1.

There are many ways to implement PLLs. For the video PLLs described here we have opted for a software implementation using quantized signals sampled at discrete, uniform time intervals, referred to as an all-software PLL. This allows us to use a digital-averaging PD, Fig. 2. The oscillator produces both an in-phase and a quadrature signal (labeled I and Q, respectively), each of which is multiplied with the input signal. The loop filter is brought inside the phase detector and averages or smooths the in-phase and quadrature products. A division and inverse tangent operation give the phase angle between u_1 and u_2 as required. The digital-averaging PD is not sensitive to the magnitude of u_1 (as is the case with other PDs) giving $K_d = 1$ always. The in-phase and quadrature products can also be used to derive the magnitude of u_1 by computing $\sqrt{X^2 + Y^2}$.

Table 1
Design equations for PLLs [9], based on linear PLLs with passive lag loop filters

Parameter	Description	Equation
ω_0	Oscillator center frequency	
τ_1, τ_2	Loop filter time constants	
ω_n	Natural frequency of PLL	$\omega_n = \sqrt{\frac{K_0 K_d}{\tau_1 + \tau_2}}$
ζ	Damping factor	$\zeta = \frac{\omega_n}{2} \sqrt{\tau_2 + \frac{1}{K_0 K_d}}$
$\Delta\omega_H$	Hold range	$\Delta\omega_H = K_0 K_d$
$\Delta\omega_L$	Lock range	$\Delta\omega_L \approx 2\zeta\omega_n$
T_L	Lock-in time	$T_L \approx \frac{2\pi}{\omega_n}$

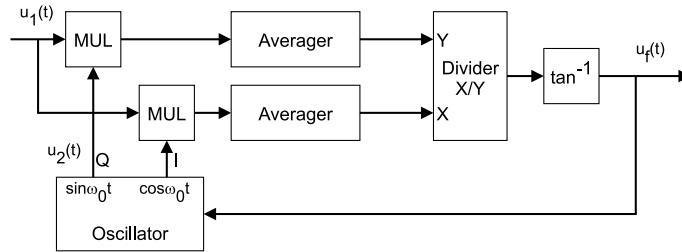


Fig. 2. Phase-locked loop using a digital-averaging phase detector.

3.2. Video phase-locked loops

A video PLL consists of a set of independent PLLs, each operating at a single pixel site. The intensity value of a pixel taken over time gives a temporal sequence. Pixel intensities are non-negative so that when observing an oscillating motion the temporal sequence consists of a constant background value with a sinusoid (or perhaps several sinusoids) superimposed on top. Since the PLL locks on sinusoids only, the video PLL uses a temporal band-pass filter to preprocess the signal to remove the background and pass only the sinusoid. This temporal band-pass filter is consistent with the temporal band-pass properties of human vision [70].

At first it might appear the independent operation of the PLLs would be troublesome. It turns out that this is not the case, at least for gait analysis, for several reasons. The band-pass filter allows sinusoids only in a given frequency band to pass through to the PLL. Also, the PLL can lock only within a given frequency band (the PLL and band-pass filter must have matching frequency bands for this reason). This combination restricts the range of frequencies that the PLL can lock. Given that a gait has only a fundamental frequency and harmonics [15] the individual PLLs are likely to all lock to the same frequency, since the band-pass filter and PLL locking range can only lock on a single frequency component of the gait. Still it may be possible that other frequency components show up in the background of a gait. The Procrustes shape analysis (see Appendix A) sees these stray frequencies as noise that obscures a signal, then recovers the signal by averaging over time. Thus, there is little need to couple the individual PLLs.

This small innovation has a significant impact on machine perception of gaits and other oscillatory motion. Video PLLs can perform frequency entrainment and phase locking quickly. Furthermore, they extract sufficient information to enable recognition of gaits and other oscillatory motions.

3.3. Signals in the phase-locked loop

Within a video PLL there is an abundance of signals useful in the perception of oscillating motion. The output of the loop, u_f , is a phase error, but also gives an instantaneous estimate of the frequency of oscillation in a pixel. The relationship be-

tween u_f and the frequency is given by the VCO transfer function, Fig. 1B. As stated previously, for a gait, all pixels for the walker will lock to the same frequency.

With all oscillators locked on the same frequency, the only difference between oscillators is phase. The oscillator produces its output by computing the phase of the oscillation at each time interval, and from there, computing the in-phase and quadrature components of u_2 . The relative phases of the oscillators throughout the image give the relative phases of the intensity oscillations.

The use of a digital-averaging PD has the added benefit of computing both the phase difference between u_1 and u_2 , and the magnitude of u_1 . The magnitude indicates the amplitude, or strength, of the locked signal. By combining the magnitude with the phase, we get a phasor (phase vector, see Appendix A) at each pixel site. The combination of phasors over several pixels forms a *phasor configuration* that rotates at the entrained frequency. We can use the configuration to classify and recognize oscillatory motion.

Fig. 3 shows phasor configurations for a sample gait sequence. While a temporal gait sequence results in a single phasor configuration, the figure shows the configurations expanded in time to highlight the temporal relationships in pixel intensities. Each configuration is first computed from an input image sequence at full resolution, then down sampled and centered in a 21-by-21 *pixel* square. Most of the configuration is background with approximately 150 pixels covering the moving figure.

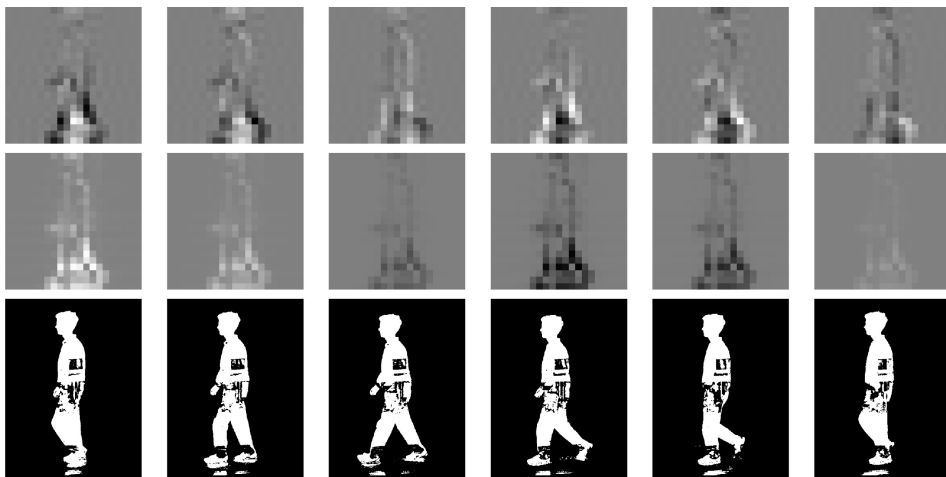


Fig. 3. Example phase configurations for a gait sequence. Each configuration is a 21-by-21 *pixel* array of complex phasors, each phasor representing the magnitude and relative phase of a sinusoid. The top row shows the phase configuration as images expanded in time through one complete cycle. The middle row shows the phase configuration as images, with the relative phases forced to zero, also expanded in time. In the configurations, gray is zero, dark is negative, and bright is positive. The bottom row shows the silhouette frames for the gait sequence corresponding to the above configurations.

3.4. Alternatives to phase-locked loops

We use PLLs to synchronize internal oscillators with the oscillations of an external source. However, we could do that with any system that can measure the frequency and phase of a sinusoidal signal. The most obvious alternative is the discrete Fourier transform (DFT). In the manner described by Liu and Picard [47], one can capture and store n frames, do an n -element DFT for each pixel, identify the fundamental frequency, and record the amplitude and phase at that frequency. If $n = 2^r$, where $r \in \mathbb{I}^+$, then using the fast Fourier transform algorithm we can compute the DFTs in $O(n \lg n)$ time, using $O(n)$ storage. We reject this for three reasons.

First, the DFT is fundamentally more complex. It uses n quadrature oscillators to estimate a single sinusoid, with a frequency resolution of $1/n$ (normalized to the sample rate). A PLL uses a single oscillator and feedback in a closed-loop control system to accurately lock to the correct frequency.

Second, the PLL operates as a causal system. We can turn a non-causal DFT system into a real, causal system by incorporating an n -frame delay, then scanning back in time. In contrast, a PLL can track phase continuously throughout a gait cycle. For a subject with an asymmetric gait, the vPLL phase configuration will track the changes in alternate left and right strides. A DFT can only look at steady-state properties of a signal.

Last, causality and continuous operation mean that a PLL can respond to transient events. While this may not be essential for biometric gait analysis, it appears to be an important part of human motion perception, as shown by Watson and Ahumada [70] and Adelson and Bergen [1].

4. Application to gait recognition

To evaluate the gait recognition capabilities of vPLLs we employ the system shown in Fig. 4. The system applies a vPLL to a gait image sequence, producing a sequence of phase configurations. The Procrustes mean (Appendix A) of a subset of the frames gives a more stable phase configuration that represents the sequence (where sufficient frames are available). A database of exemplars is processed the same way. The Procrustes distance (Appendix A) is used to compare the test sequence phase configuration to the exemplars. We use a nearest-neighbor (NN) recognition system that selects the exemplar with the shortest Procrustes distance to the test configuration. Although it is possible to employ methods other than NN, we restrict ourselves to NN for simplicity and uniformity in the evaluation.

We test the vPLLs using two databases: the Carnegie Mellon University (CMU) Motion of Body (Mobo) [37], and the University of Southampton [60] gait databases.

4.1. CMU Mobo database

The CMU Mobo database [37] contains gait sequences for 25 subjects. Six cameras record each subject performing four different *activities* while walking on a tread-

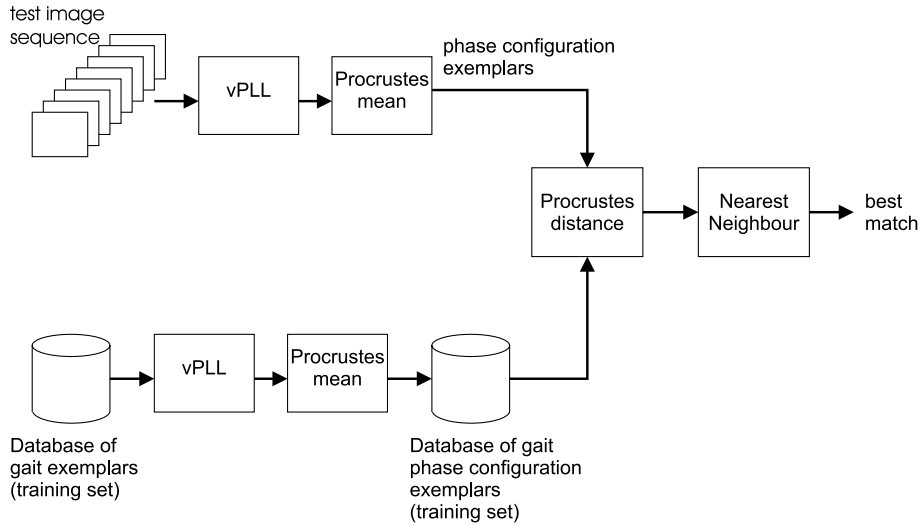


Fig. 4. Nearest-neighbor exemplar recognition system for gait phasor configurations.

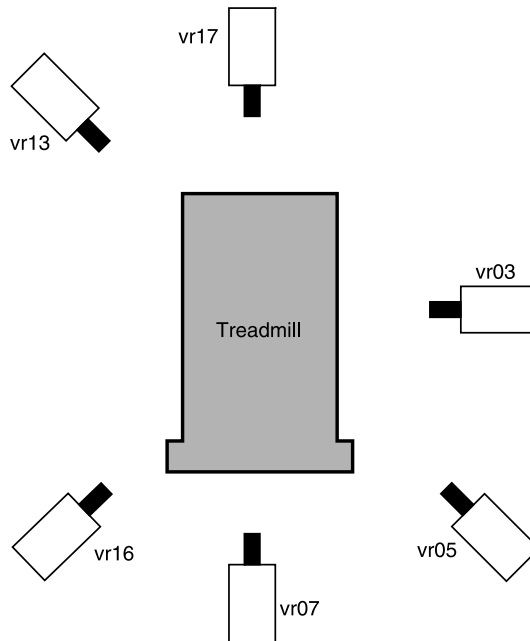


Fig. 5. Camera placement for CMU Mobo database.

mill. Fig. 5 shows the camera viewing angles, labeled *vr03*, *vr05*, *vr07*, *vr13*, *vr16*, and *vr17*. The four activities are a slow walk, a fast walk, walking up an incline, and walking while carrying a ball. Each sequence is about 11 s long sampled at 30 frames

per second. Thus the database contains variations in subject, viewing angle, and activity. The database provides full color images of the sequences for each subject–view–activity combination, as well as a binary foreground–background image that separates the subject from the background.

Clothing and background affects a vPLL in two ways. First, the contrast between the clothing and the background determines the magnitude of the observed oscillations. Second, a change in polarity of the contrast, i.e., switching from white-on-black to black-on-white, results in a 180° phase shift. We take advantage of the segmented binary images in the Mobo database to eliminate these effects. Ignoring segmentation errors, each subject appears as a white silhouette on a black background. The multiple factors in the database allow one to test recognition of viewing angles and activity in addition to recognizing individuals.

Since each sequence is approximately 11 s long, we compute two or three phase configurations from each sequence by finding the Procrustes mean for frames 101 through 140, 201 through 240, and, where the sequence is long enough, 301 through 340. Table 2 gives the design parameters for the vPLL used to process the sequences. We subsample the region of the walking figure in the vPLL output to get phasor configurations that are 21-by-21 *pixels*, as shown in Fig. 3. Thus, the database contains multiple exemplars for each person–view–activity combination, although these exemplars are taken only a few seconds apart and do not necessarily represent the true variation possible for a particular individual.

To see what is gained by using phase information, we also test the recognition system using phase configurations in which the phase is forced to zero. The result is a configuration that indicates where there are oscillations, but without timing information. These zero-phase images (examples in middle row of Fig. 3) are analogous to the motion energy images described by Davis and Bobick [28] and could be called *oscillation energy images*.

In our evaluation, we look at two figures of merit. The first is $F_{\text{WW}} = MS_{\text{B}}/MS_{\text{W}}$, the ratio of the mean square error between and within groups, given in second Table of Appendix B, and its associated probability. The second is the recognition rate based on the confusion matrix produced using a k -fold cross-validation [21].

Viewing angle. The ANOVA for viewing angle yields $MS_{\text{B}} = 90.7$, $MS_{\text{W}} = 0.6$, and $F_{\text{WW}} = 159$. Such a high value for F_{WW} corresponds to $P(H_0) < 0.001$ within the precision of our computations. It is reasonable to conclude that the configurations for at least one viewing angle are different than the others. This observation

Table 2
vPLL design parameters

Symbol	Definition	Value
Ω_1	Lower cut-off for band pass	0.0375
Ω_2	Upper cut-off for band pass	0.075
ζ	Damping factor	0.7
T_{L}	Locking time	40 frames
f_0	VCO center frequency	0.05625

Frequencies are normalized to sample rate.

is borne out by the recognition rates. Table 3 shows the confusion matrix resulting from a k -fold cross-validation test of viewing angle recognition. Sequences for the same person in the test image are dropped from the training set, requiring that correct matches be due to different people viewed from the same angle. The table indicates a recognition rate of 98.7%. We can make things even harder for the recognition system by biasing the test and dropping sequences from the training set that match the test sequence activity. This means that correct matches can only occur between different people doing different things, but viewed by the same camera. Table 4 shows the confusion matrix for this test. The recognition rate is 94.2%.

The zero-phase configurations ANOVA yields $MS_B = 40.0$, $MS_W = 0.3$, and $F_{WW} = 139$, which also gives $P(H_0) < 0.001$. Recognition rates for the k -fold cross-validation test were 96.3 and 92.9% corresponding to the two cases above.

We can conclude that vPLLs can recognize the viewing angle extremely well. Use of the phase information corresponds to a modest improvement in recognition rate.

Activity. When computing F_{WW} for activity recognition with the Mobo database, it is essential to account for variations in viewing angle. There are two approaches one can take: compute a separate F_{WW} for each of the six viewing angles, or compute F_{WW} where there are 24 categories corresponding to all the combinations of viewing angles and activities. Table 5 shows the results for both approaches.

The result in the last column of Table 5 represents a combination of the sensitivity to viewing angle combined with any additional sensitivity to activity. Therefore, it is not surprising that F_{WW} is so large. The tests for the individual viewing angles isolate

Table 3

The confusion matrix showing results of cross-validation test of viewpoint recognition in Mobo database

	vr07	vr13	vr17	vr03	vr16	vr05
vr07	286	0	3	0	0	0
vr13	0	274	0	0	6	0
vr17	4	0	282	1	0	2
vr03	0	1	0	287	0	0
vr16	0	5	0	0	284	0
vr05	0	0	0	0	0	289

The test excludes matches of sequences for the same subject.

Table 4

The confusion matrix showing results of cross-validation test of viewpoint recognition in Mobo database

	vr07	vr13	vr17	vr03	vr16	vr05
vr07	284	0	5	0	0	0
vr13	0	247	0	0	33	0
vr17	7	0	278	1	1	2
vr03	0	1	0	287	0	0
vr16	0	46	0	0	243	0
vr05	0	0	0	4	0	285

The test excludes matches of sequences for the same subject or same activity.

gles, and one for all the data combined. Summaries of F_{WW} for both with-phase and zero-phase configurations are tabulated in Tables 8 and 9.

We test the recognition rate of individual people from the Mobo database under three conditions:

1. using the full training set,
2. dropping sequences with the same activity as the test sequence, and
3. dropping sequences with the same activity and same person as the test sequence.

Table 10 summarizes the k -fold cross-validation recognition rates for both with-phase and zero-phase configurations. Condition one allows matches with exemplars taken from the same sequence but at different points in time. We would expect this to be fairly easy, and it is, with a recognition rate of 90.5%. Condition two excludes matches to the same activity and thus prevents matching to the same sequence. At the same time, condition two also prevents matching to another person doing the same activity, which makes it easier than condition three where it is possible to match a different person doing the same activity. Condition two yields a recognition rate of 59.7% and condition three yields 47.3%. Recognition rates for the zero-phase configurations are lower under all conditions. Thus both F_{WW} and cross-validation tests indicate an advantage in using the phase information.

Note that case three is particularly difficult. Consider a test sequence where the subject is walking on an incline. To register a correct match, the system must find

Table 8
ANOVA results for people recognition

	vr03	vr05	vr07	vr13	vr16	vr17	All
MS_B	1.77	1.66	3.16	1.86	1.68	3.68	3.41
MS_W	0.32	0.49	0.36	0.41	0.41	0.43	0.80
F_{WW}	5.5	3.4	8.7	4.5	4.5	8.6	4.28
$P(H_0)$	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Table 9
ANOVA results for people recognition with zero-phase data

	vr03	vr05	vr07	vr13	vr16	vr17	All
MS_B	0.99	0.82	1.22	1.02	0.94	1.41	1.57
MS_W	0.19	0.26	0.16	0.25	0.25	0.19	0.39
F_{WW}	5.1	3.1	7.7	4.2	3.7	7.5	4.05
$P(H_0)$	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Table 10
Summary of people-recognition results

Test condition	With phase (%)	Zero phase (%)
All	90.5	88.8
Drop same activity	59.7	55.4
Drop same person–activity	47.3	43.0

a good match to a sequence for the same person, but not on an incline, while avoiding potential matches to similar subjects that are on an incline. The ANOVA clearly indicates that the configurations have a strong dependence on individual gaits.

Biometric benchmarks. Creators of the Mobo database recommend a set of benchmarks for evaluation of biometric algorithms with their data. We report a subset of these benchmarks here. The benchmarks partition the data into probe and gallery sets, and report recognition results in the form of CMC plots [54] to investigate variates of interest for biometrics. We deviate from the Mobo recommendations only in specific selection of images within a sequence to accommodate the continuous operation of the vPLL.

Fig. 6A shows the results when recognizing a person from data within the same sequence from camera vr03. Fig. 6B shows results for recognizing a person from data within the same sequence, always for a slow walk, but from camera angles vr03, vr05 or vr07. Fig. 6C shows the results for recognizing an individual doing a fast walk or walking with a ball from a gallery of slow walks, from camera vr07.

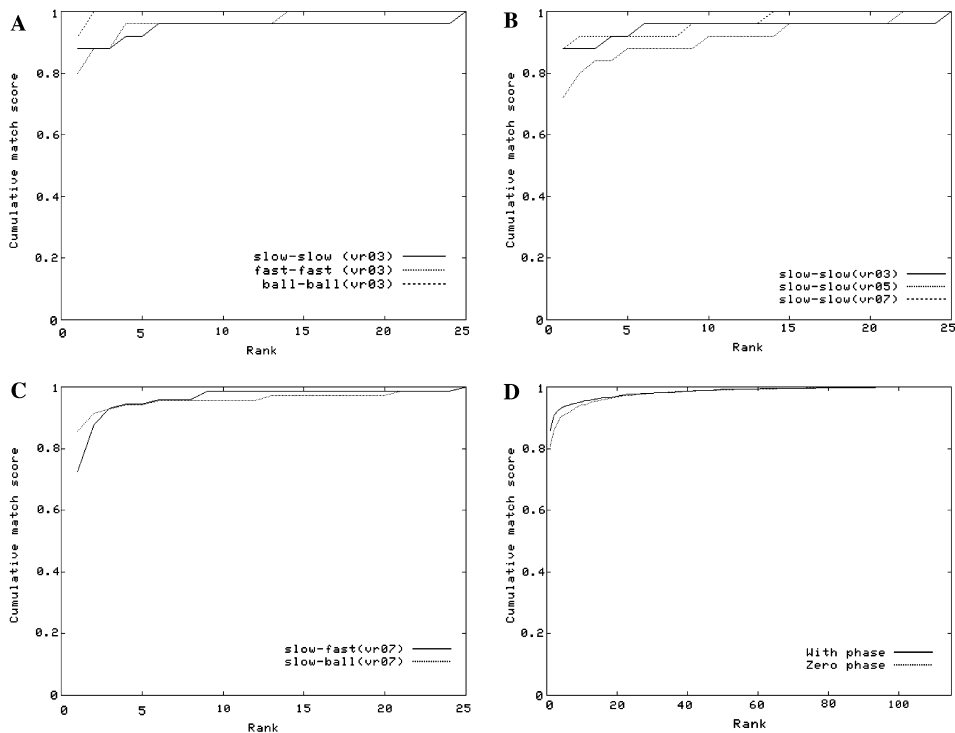


Fig. 6. CMC curves for a subset of recommend biometric benchmarks on the Mobo and Southampton databases: (A) Mobo within sequence, varying activity, vr03, (B) Mobo within sequence, varying camera, slow walk, (C) Mobo match from different sequence, vr07, and (D) k -fold cross-validation on the Southampton database. Mobo results are for with-phase only.

4.2. University of Southampton database

Shutler et al. [60] provide a gait database with 115 subjects acquired at the University of Southampton. There are several sequences acquired for each subject from multiple camera views. We restricted our tests to the subset of the database using a single camera aimed perpendicular to the subject's motion. All subjects walked in front of a green backdrop to facilitate chroma-key segmentation. The sequences are approximately evenly split between toward-the-left (L) and toward-the-right (R) motion. We reverse the horizontal axis for the R sequences so that the orientation of the subject matches for all sequences. The Southampton database does not contain forced variations in a subject's gait such (like the *activities* in Mobo), so we can only test people recognition.

As with the Mobo database, we register the silhouettes of the sequence, compute phase configurations, then subsample and center the data to a common size of 21-by-21 *pixels*. The Southampton database sequences are significantly shorter than those in the Mobo database, typically 31 frames or fewer. For this reason, we omit the Procrustes averaging step, reduce T_L for the PLLs (shorter lock time), and use only the last frame of vPLL output for the phase configuration.

The omission of the Procrustes averaging and shorter lock time increases measurement errors, MS_W , thereby reducing F_{WW} and recognition rates. Whereas in the Mobo trials, Procrustes averaging blurs asymmetries in a gait, in our Southampton trials an asymmetric gait can appear to be different depending on whether the sequence ends with a left- or right-foot swing. This happens because the vPLL tracks phase continuously.

For with-phase configurations, we measure $MS_B = 2.09$, $MS_W = 0.197$, and $F_{WW} = 10.6$, and for zero-phase configurations we get $MS_B = 1.21$, $MS_W = 0.104$, and $F_{WW} = 10.4$. The corresponding recognition rates based on k -fold cross-validation are 85.8% with phase, and 80.6% for zero-phase. Fig. 6D shows the corresponding CMC plots.

There is a strong preference to match L sequences to L and R sequences to R. If we bias the test to force recognition to match only sequences of opposite orientation, the recognition rate drops to 39%. This suggests either a strong sensitivity to asymmetries in the gaits, or that the camera was not aimed precisely perpendicular to the motion.

5. Discussion

5.1. Recognition rates and F_{WW}

The results in Section 4 present a mixed view of recognition capabilities of the vPLL-based system with numbers ranging from 47 to 86% for recognition of individual gaits. To shed light on the issue we offer the following observations.

We used a nearest-neighbor classifier for all tests. Other classifiers may give different rates.

Randomization of a sample is essential to external validity of an experiment [19]. George Gallup based his polling success on the realization that sample bias is far more critical than sample size. Unfortunately, we do not know of any databases containing a randomized sample of human gaits, and it would not be trivial to assemble one. The ratio of male to female subjects in both the Mobo (23/2) and Southampton (91/24) databases suggests that neither is a random sample of the general population.

F_{WW} provides a useful descriptive statistic for recognition. The values of MS_B indicate that the inclusion of phase information leads to larger variation between views, activities, and individuals. However, MS_W indicates that the inclusion of phase increases variability in measurements, i.e., noise. In all but one trial, the net effect of including phase was to reduce F_{WW} and improve recognition.

In one anomalous case, activity recognition in Mobo, inclusion of phase *reduces* F_{WW} , but recognition still improves. There are only four activities present in the Mobo database, so the number of degrees of freedom in this test is small. Given that F_{WW} and recognition rates are statistics, and that we have only three degrees of freedom in our estimate of MS_B , neither F_{WW} nor the recognition rate is very accurate. However, the values of F_{WW} appear consistent across all camera views, so we conclude that for activity recognition, the signal gained by including phase was negated by increased measurement error, and that the improved recognition rate is a statistical anomaly.

5.2. An alternative model for motion perception

In gait analysis a vPLL synchronizes to a steady-state oscillation. However, when presented with a transient change in pixel intensity, a vPLL will lock to the transient oscillations of the band-pass filter. Boyd and Little [16] showed that it is possible to estimate translational motion by comparing the detected phases of the transient oscillations in adjacent pixels. This can be understood by comparing the operation of a PLL temporal filters used by Fleet and Jepson [32] to compute phase-based optical flow. However, Fleet and Jepson require a bank of filters tuned to separate velocities to cover a range of possible inputs. A PLL uses feedback to track phase and frequency, and thus requires only one filter.

Watson and Ahumada [70], and Adelson and Bergen [1] describe models for human motion perception that, among other things, depend on the transient response of temporal band-pass filters in the human vision system. It is interesting to note that both the vPLL system and the human models can detect motion from band-pass transients. Although we have not investigated the possibility of a vPLL-like model for human vision, the idea bears consideration.

5.3. Synchronization with kinematic models

While it is possible to divide human motion analysis methods into model-free and model-based categories, it is clear that neither approach can be considered complete in the sense that it can solve all vision problems related to human motion. It is evident that both approaches are necessary at least some of the time. The concept of

synchronization suggests a path to connect the model-free and model-based domains. If we assume that a kinematic model has an innate gait, as demonstrated by passive mechanical walkers [22,24,34,49,50], and that the model's gait can be altered by an external control system, as demonstrated by Laszlo et al. [43], then synchronization of the model gait to oscillations in a model-free representation of the gait is possible. Boyd [14] describes a system that synchronizes a simple, two-dimensional kinematic model to gait sequences. Yam et al. [72] describe a system that synchronizes a parametric oscillating gait model with an image sequence, then uses the parameters to recognize gaits. Much of the synchronization process could be done directly from vPLL-based phase configurations.

6. Conclusions

In Section 1 we pose the hypotheses that humans have a built-in gait model and that humans perceive motion by using that model. Inspired by these hypotheses, we created a system that synchronizes an array of vPLL oscillators to oscillations in an image sequence. The magnitude and phase of the synchronized oscillations gives feature vectors that can distinguish various properties of a gait such as viewing angle and activity, and also recognize individual gaits.

The ability to synchronize oscillations suggests an alternative model for motion perception based on measuring the phases of the transient oscillations of temporal band-pass filters. Synchronization of a kinematic model to oscillations in model-free representations suggests a path to bridge the gap between the model-free and model-based domains.

Acknowledgments

The Mobo [37] and Southampton [60] databases are valuable assets to the community of gait analysis researchers. The author thanks Jianbo Shi and Ralph Gross of the Robotics Institute at Carnegie Mellon University for use of Mobo, and Mike Grant and Mark Nixon at the University of Southampton for use of their data. The work presented here was supported, in part, by the Natural Science and Engineering Research Council of Canada.

Appendix A. Directional statistics

A phasor allows us to represent a sinusoid as a complex number of the form $V = Ae^{i\phi}$, where A is the amplitude and ϕ is the phase of the sinusoid. The implication is that we can treat sinusoids as vectors in the complex number plane, with a magnitude and direction. Consequently, we can analyze sinusoids derived from a gait by treating them as directional data. Procrustes shape analysis [42] is a method in directional statistics [48] that can summarize (by finding means) and compare

(using distance measures) shapes. By virtue of using phasors, we can represent a phase or timing pattern in a gait as a set of directions, which is mathematically equivalent to a shape, making Procrustes analysis a useful tool for analyzing the phasor patterns that emerge from gaits.

The following is a summary based on Mardia and Jupp [48]. Describe a shape in two dimensions using a vector of k complex numbers, $\mathbf{z} = [z_1, z_2, \dots, z_k]^T$, called a configuration. Two configurations, \mathbf{z}_1 and \mathbf{z}_2 , represent the same shape if by a combination of translation, scaling, and rotation, their configurations are equal, i.e.,

$$\mathbf{z}_1 = \alpha \mathbf{1}_k + \beta \mathbf{z}_2, \quad \alpha, \beta \in \mathcal{C}$$

$$\beta = |\beta| e^{i\angle\beta},$$

as shown in Figs. 7A and B. That is, $\alpha \mathbf{1}_k$ translates \mathbf{z}_2 , and $|\beta|$ and $\angle\beta$ scale and rotate \mathbf{z}_2 . It is convenient to center shapes by defining the centered configuration $\mathbf{u} = [u_1, u_2, \dots, u_k]^T$, $u_i = z_i - \bar{z}$, and $\bar{z} = \sum_{i=1}^k z_i/k$. From here we can define the rotation-invariant *full Procrustes distance* between two configurations, $d_F(\mathbf{u}_1, \mathbf{u}_2)$ as:

$$d_F(\mathbf{u}_1, \mathbf{u}_2) = 1 - \frac{|\mathbf{u}_1^* \mathbf{u}_2|^2}{\|\mathbf{u}_1\|^2 \|\mathbf{u}_2\|^2}, \quad (\text{A.1})$$

where \mathbf{u}^* denotes the transpose conjugate of \mathbf{u} . Note that $0 \leq d_F \leq 1$. The Procrustes distance allows us to compare two shapes independent of position, scale, and

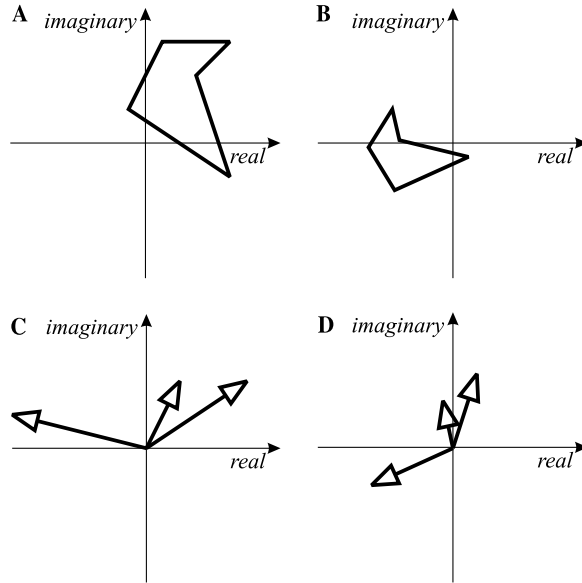


Fig. 7. Procrustes analysis applied to shape and phase configurations. In the conventional application, a shape is represented by a vector of complex vertices. The shape in (A) is the same as the shape (B) because one is a translated, scaled, and rotated version of the other. A phasor configuration is also a vector of complex numbers. The configuration in (C) is the same as that in (D) because one is a rotated and scaled version of the other. Rotation is always about the origin so translation is omitted.

rotation. To find the *Procrustes mean shape*, $\hat{\mu}$, which is an estimate of a mean, μ , based on a sample of n shapes, we compute the matrix

$$\mathbf{S}_u = \sum_{j=1}^n (\mathbf{u}_j \mathbf{u}_j^*) / (\mathbf{u}_j^* \mathbf{u}_j). \quad (\text{A.2})$$

The *Procrustes mean shape* is the dominant eigenvector of \mathbf{S}_u , i.e., the eigenvector that corresponds to the greatest eigenvalue of \mathbf{S}_u . Note that the power method allows fast computation of the dominant eigenvalue and vector without forming \mathbf{S}_u directly [36].

A vector of complex phasors, or a phasor configuration, is equivalent to a shape configuration, as illustrated in Figs. 7C and D. Shapes are invariant through translation, scaling, and rotation. Translational invariance is achieved by using the centered configuration \mathbf{u} . When using phasors the issue of translation becomes irrelevant. All phasors rotate about the origin, $0 + 0i$, at the entrained frequency, and the configurations, \mathbf{z} , are already centered, i.e., $\mathbf{z} = \mathbf{u}$.

Appendix B. Analysis of variance on directional data and phase configurations

Mardia and Jupp [48] describe ANOVA for highly concentrated directional data, summarized in first Table. The analysis applies to q groups of directional data, $\theta_{i1}, \dots, \theta_{in}$, $i = 1, \dots, q$. Each sample has a resultant length, R_i , where

$$\begin{aligned} C_i &= \sum_{j=1}^{n_i} \cos \theta_{ij}, & S_i &= \sum_{j=1}^{n_i} \sin \theta_{ij}, \\ C &= \sum_{i=1}^q C_i, & S &= \sum_{i=1}^q S_i, \\ R_i^2 &= C_i^2 + S_i^2, & R^2 &= C^2 + S^2. \end{aligned} \quad (\text{B.1})$$

F_{WW} , as given in first Table, is approximately distributed as $F_{q-1, n-q}$. A correction factor is applied to F_{WW} if the data are not sufficiently concentrated on one part of the unit circle.

We extrapolate directional ANOVA using Procrustes analysis for application to phase configurations. The eigenvalue decomposition of the matrix \mathbf{S}_u (Eq. (A.2)) yields the Procrustes mean shape, $\hat{\mu}$, the eigenvector that corresponds to the largest eigenvalue, $\hat{\lambda}$. When a sample of configurations have identical shapes, $\hat{\lambda} = n$, where n is the sample size. For orthogonal shapes all the eigenvalues are one. Therefore, $\hat{\lambda}$ gives a measure of variation in a shape sample that is analogous to R in Eq. (B.1), and we define the ANOVA procedure shown in the second Table that is analogous to first Table. $\hat{\lambda}$ is the largest eigenvector for the entire sample while $\hat{\lambda}_i$ is the largest eigenvector for group i .

Using Monte Carlo simulation, we have determined that for configurations of two phasors, F_{WW} under the null hypothesis is distributed approximately as $F_{q-1, n-q}$. This is to be expected since the situation is equivalent to basic directional ANOVA. However, we observed that the degrees of freedom added by increasing the number

of phasors in the configuration yield a narrower distribution than $F_{q-1, n-q}$. Lacking an analytical model, we use $F_{q-1, n-1}$ to give an upper bound on $P(H_0)$ and use Monte Carlo simulation to get a lower bound. We expect the truth to lie somewhere in between since gait configurations contain many zeros that do not contribute additional degrees of freedom. Computation time limits our Monte Carlo probability estimates to a precision of 1/1000.

ANOVA table for highly concentrated directional data. df is degrees of freedom, SS is sum of squares, MS_B is mean square error between groups, and MS_W is mean square error within groups

Source	df	SS	Mean square	F
Between samples	$q - 1$	$\sum_{i=1}^q R_i - R$	$MS_B = (\sum_{i=1}^q R_i - R)/(q - 1)$	$F_{WW} = MS_B/MS_W$
Within samples	$n - q$	$n - \sum_{i=1}^q R_i$	$MS_W = (n - \sum_{i=1}^q R_i)/(n - q)$	
Total	$n - 1$	$n - R$		

ANOVA table for highly concentrated shape configuration data

Source	df	SS	Mean square	F
Between samples	$q - 1$	$\sum_{i=1}^q \hat{\lambda}_i - \hat{\lambda}$	$MS_B = (\sum_{i=1}^q \hat{\lambda}_i - \hat{\lambda})/(q - 1)$	$F_{WW} = MS_B/MS_W$
Within samples	$n - q$	$n - \sum_{i=1}^q \hat{\lambda}_i$	$MS_W = (n - \sum_{i=1}^q \hat{\lambda}_i)/(n - q)$	
Total	$n - 1$	$n - \hat{\lambda}$		

References

- [1] E.H. Adelson, J.R. Bergen, Spatiotemporal energy models for the perception of motion, *J. Opt. Soc. Am. A* 2 (2) (1985) 284–299.
- [2] A.M. Baumberg, D.C. Hogg, Learning flexible models from image sequences, Technical Report 93.36, University of Leeds School of Computer Studies October, 1993.
- [3] A.M. Baumberg, D.C. Hogg, Learning spatiotemporal models from training examples, in: *British Machine Vision Conference*, Birmingham, September, 1995.
- [4] C. Ben-Abdelkader, R. Cutler, L. Davis, Motion-based recognition of people in eigengait space, in: *Automatic Face and Gesture Recognition 2002*, Washington, DC, May, 2002 pp. 254–259.
- [5] C. Ben-Abdelkader, R. Cutler, L. Davis, Person identification using automatic height and stride estimation, in: *16th International Conference on Pattern Recognition*, Quebec, August 2002, pp. 377–380.
- [6] C. Ben-Abdelkader, R. Cutler, L. Davis, View-invariant estimation of height and stride for gait recognition, in: *Workshop on Biometric Authentication (BIOMET)*, in association with ECCV, 2002.

- [7] C. Ben-Abdelkader, R. Cutler, H. Nanda, L. Davis, Eigengait: motion-based recognition of people using image self-similarity, in: *Audio- and Video-Based Biometric Person Authentication*, Halmstad, Sweden, June 2001, pp. 284–294.
- [8] B.I. Bertenthal, J. Pinto, Complementary processes in the perception and production of human movements, in: L.B. Smith, E. Thelen (Eds.), *A Dynamic Systems Approach to Development: Applications*, MIT Press, Cambridge, MA, 1993, pp. 209–239.
- [9] R.E. Best, *Phase-locked Loops Design, Simulation and Applications*, McGraw-Hill, New York, 1999.
- [10] B. Bhanu, J. Han, Bayesian-based performance prediction for gait recognition, in: *IEEE Workshop on Motion and Video Computing*, Orlando, Florida, December 2002, pp. 145–150.
- [11] A. Bissacco, A. Chiuso, Y. Ma, S. Soatto, Recognition of human gaits, in: *Computer Vision and Pattern Recognition 2001*, vol. II, Kauai, HI, December 2001, pp. 52–57.
- [12] A. Bobick, A. Johnson, Gait recognition using static activity-specific parameters, in: *Computer Vision and Pattern Recognition 2001*, vol. I, Kauai, HI, December 2001, pp. 423–430.
- [13] J.E. Boyd, Video phase-locked loops in gait recognition, in: *International Conference on Computer Vision*, Vancouver BC, 2001, pp. 696–703.
- [14] J.E. Boyd, Oscillating models for perception of human motion, Technical Report 2003-718-21, Department of Computer Science, University of Calgary, 2003.
- [15] J.E. Boyd, J.J. Little, Phase in model-free perception of gait, in: *IEEE Workshop on Human Motion*, Austin TX, December 2000.
- [16] J.E. Boyd, J.J. Little, Motion from transient oscillations, in: *Computer Vision and Pattern Recognition 2001*, Kauai HI, December 2001. Technical Sketeches on CD-ROM.
- [17] C. Bregler, Learning and recognizing human dynamics in video sequences, in: *Computer Vision and Pattern Recognition 1997*, San Juan, Puerto Rico, June 1997, pp. 568–574.
- [18] C. Bregler, J. Malik, Tracking people with twists and exponential maps, in: *Computer Vision and Pattern Recognition 1998*, Santa Barbara, June 1998.
- [19] D.T. Campbell, J.C. Stanley, *Experimental and Quasi-Experimental Designs for Research*, Houghton Mifflin, Boston, 1963.
- [20] L. Cohen, T.F. Shipley, E. Marshark, K. Taht, D. Aster, Detecting animals in point-light displays, in: *22nd Annual Meeting of the Cognitive Science Society*, Philadelphia, PA, August 2000, p. 70.
- [21] P.R. Cohen, *Empirical Methods for Artificial Intelligence*, MIT Press, Cambridge, MA, 1995.
- [22] M.J. Coleman, A. Ruina, An uncontrolled toy that can walk but cannot stand still, *Phys. Rev. Lett.* 80 (16) (1998) 3658–3661.
- [23] R.T. Collins, R. Bross, J. Shi, Silhouette-based human identification from body shape and gait, in: *Automatic Face and Gesture Recognition 2002*, Washington DC, May 2002, pp. 351–356.
- [24] S.H. Collins, M. Wisse, A. Ruina, A 3-d passive-dynamic walking robot with two legs and knees, *Int. J. Robot. Res.* (2001), in press.
- [25] D. Cunado, M.S. Nixon, J.N. Carter, Automatic extraction and description of human gait models for recognition purposes, *Comput. Vis. Image Understand.* 90 (1) (2003) 1–41.
- [26] R. Cutler, L. Davis, Robust periodic motion and motion symmetry detection, in: *Computer Vision and Pattern Recognition 2000*, 2000.
- [27] J.W. Davis, Visual categorization of children and adult walking styles, in: *Audio- and Video-Based Biometric Person Authentication*, Halmstad, Sweden, June 2001 pp. 295–300.
- [28] J.W. Davis, A.F. Bobick, The representation and recognition of human movement using temporal templates, in: *IEEE Computer Vision and Pattern Recognition*, June 1997, pp. 928–934.
- [29] H. Farid, Temporal synchrony in perceptual grouping: a critique, *Trends Cognit. Sci.* 6 (7) (2002) 284–288.
- [30] H. Farid, E.H. Adelson, Synchrony does not promote grouping in temporally structured displays, *Nat. Neurosci.* 4 (9) (2001) 875–876.
- [31] F. Fazio, D. Perani, N.A. Borghese, S. Ferrari, F. Lo Monaco, V. Blasi, M.C. Gilardi, How can virtual reality be? a pet study of motor actions observation, in: *Human Brain Mapping Conference*, 1999.
- [32] D.J. Fleet, A.D. Jepson, Computation of component image velocity from local phase information, *Int. J. Comput. Vis.* 5 (1) (1990) 77–104.

- [33] H. Fujiyoshi, A.J. Lipton, Real-time human motion analysis by image skeletonization, in: DARPA Image Understanding Workshop, Monterey, California, November 1998.
- [34] M. Garcia, A. Chatterjee, A. Ruina, M. Coleman, The simplest walking model: stability, complexity, and scaling, *ASME J. Biomech. Eng.* 120 (2) (1998) 281–288.
- [35] W.S. Geisler, D.G. Albrecht, Visual cortex neurons in monkeys and cats: detection, discrimination and identification, *Vis. Neurosci.* 14 (1997) 897–919.
- [36] G.H. Golub, C.F. Van, Loan, *Matrix Computations*, second ed., The Johns Hopkins University Press, Baltimore, 1989.
- [37] R. Gross, J. Shi, The cmu motion of body (mobo) database. Technical Report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, June 2001.
- [38] J.B. Hayfron-Acquah, M.S. Nixon, J.N. Carter, Automatic gait recognition by symmetry analysis, *Pattern Recognit. Lett.* 24 (2003) 2175–2183.
- [39] E.A. Hunter, P.H. Kelly, R.C. Jain, Estimation of articulated motion using kinematically constrained mixture densities, in: *Nonrigid and Articulated Motion Workshop*, San Juan, Puerto Rico, June 1997.
- [40] A.Y. Johnson, A.F. Bobick, A multi-view method for gait recognition using static body parameters, in: *Audio- and Video-Based Biometric Person Authentication*, Halmstad, Sweden, June 2001, pp. 301–311.
- [41] A. Kale, A.N. Rajagopalan, N. Cuntoor, V. Kruger, Gait-based recognition of humans using continuous hmms, in: *Automatic Face and Gesture Recognition 2002*, Washington, DC, May 2002, pp. 321–326.
- [42] J.T. Kent, New directions in shape analysis, in: K.V. Mardia (Ed.), *The Art of Statistical Science: A Tribute to G. S. Watson*, Wiley, Chichester, 1992, pp. 115–127.
- [43] J. Laszlo, M. van de Panne, E. Fiume, Limit cycle control and its application to the animation of balancing and walking, in: *SIGGRAPH 96*, 1996, pp. 155–162.
- [44] L. Lee, W.E.L. Grimson, Gait analysis for recognition and classification, in: *Automatic Face and Gesture Recognition 2002*, Washington, DC, May 2002, pp. 148–155.
- [45] S. Lee, R. Blake, Visual form created solely from temporal structure, *Science* 284 (1999) 1165–1168.
- [46] J.J. Little, J.E. Boyd, Recognizing people by their gait: the shape of motion, *Videre* 1 (2) (1998) 1–32.
- [47] F. Liu, R.W. Picard, Finding periodicity in space and time, in: *Int. Conf. Computer Vision*, 1998.
- [48] K.V. Mardia, P.E. Jupp, *Directional Statistics*, Wiley, Chichester, 2000.
- [49] T. McGeer, Passive dynamic walking, *Int. J. Robot. Res.* 9 (2) (1990) 62–82.
- [50] T. McGeer, Passive walking with knees, in: *IEEE Int. Conf. Robotics and Automation*, 1990, pp. 1640–1645.
- [51] M.P. Murray, Gait as a total pattern of movement, *Am. J. Phys. Med.* 16 (1) (1967) 290–332.
- [52] M.P. Murray, A. Bernard, R.C. Kory, Walking patterns of normal men, *J. Bone Joint Surg. A* 46 (2) (1964) 335–359.
- [53] J. Ng, S. Gong, On the binding mechanism of synchronized visual events, in: *IEEE Workshop on Motion and Video Computing*, Orlando, Florida, December 2002, pp. 112–117.
- [54] P.J. Phillips, S. Sarkar, I. Robledo, P. Brother, K. Bowyer, Baseline results for the challenge problem of human id using gait analysis, in: *Automatic Face and Gesture Recognition 2002*, Washington, DC, May 2002, pp. 130–135.
- [55] R. Polana, R. Nelson, Detection and recognition of periodic, nonrigid motion, *Int. J. Comput. Vis.* 23 (3) (1997) 261–282.
- [56] I. Robledo, S. Sarkar, Experiments on gait analysis by exploiting nonstationarity in the distribution of feature relationships, in: *Int. Conf. Pattern Recognition*, vol. 1, Quebec Canada, 2002, pp. 1–4.
- [57] H.A. Rowley, J.M. Rehg, Analyzing articulated motion using expectation-maximization, in: *Computer Vision and Pattern Recognition 97*, San Juan, Puerto Rico, June 1997, pp. 935–941.
- [58] A.B. Sekuler, P.J. Bennett, Generalized common fate: grouping by common luminance changes, *Psychol. Sci.* 12 (6) (2001) 437–444.
- [59] M.N. Shadlen, J.A. Movshon, Synchrony unbound: a critical evaluation of the temporal binding hypothesis, *Neuron* 24 (1999) 67–77.
- [60] J.D. Shutler, M.G. Grant, M.S. Nixon, J.N. Carter, On a large sequence-based human gait database, in: *Fourth Int. Conf. Recent Advances in Soft Computing*, Nottingham, UK, 2002, pp. 66–71.

- [61] J.D. Shutler, M.S. Nixon, Zernike velocity moments for description and recognition of moving shapes, in: *British Machine Vision Conference 2001*, page Session 8: Modelling Behaviour, Manchester, UK, July 2001.
- [62] R. Tanawongsuwan, A. Bobick, Gait recognition from time-normalized joint-angle trajectories in the walking plane, in: *Computer Vision and Pattern Recognition 2001*, vol. II, Kauai HI, December 2001, pp. 726–731.
- [63] E. Tassone, G. West, S. Venkatesh, Temporal pdms for gait classification, in: *Int. Conf. Pattern Recognition*, vol. 2, Quebec, Canada, 2002, pp. 1065–1068.
- [64] M. Unuma, K. Anjyo, R. Takeuchi, Fourier principles for emotion-based human figure animation, in: *SIGGRAPH 95*, 1995, pp. 91–96.
- [65] M. Usher, N. Donnelly, Visual synchrony affects binding and segmentation in perception, *Nature* 394 (1998) 179–182.
- [66] S. Wachter, H.-H. Nagel, Tracking of persons in monocular image sequences, in: *Nonrigid and Articulated Motion Workshop*, San Juan, Puerto Rico, June 1997.
- [67] L. Wang, W. Hu, T. Tan, A new attempt to gait-based human identification, in: *Int. Conf. Pattern Recognition*, vol. 1, Quebec, Canada, 2002, pp. 115–118.
- [68] L. Wang, H. Ning, T. Tan, W. Hu, Fusion of static and dynamic body biometrics for gait recognition, in: *IEEE Int. Conf. Computer Vision*, Nice, France, 2003, pp. 1449–1454.
- [69] A.B. Watson, Temporal sensitivity, in: K.R. Boff, L. Kaufman, J.P. Thomas (Eds.), *Handbook of Perception and Human Performance*, vol. 1, Wiley, New York, 1986 (Chapter 6).
- [70] A.B. Watson, A.J. Ahumada, Model of human visual-motion sensing, *J. Opt. Soc. Am. A* 2 (2) (1985) 322–342.
- [71] C.R. Wren, A. Azarbayenjani, T. Darrell, A.P. Pentland, Pfinder: real-time tracking of the human body, *IEEE Trans. Pattern Anal. Mach. Intell.* 19 (7) (1997) 780–785.
- [72] C.-Y. Yam, M.S. Nixon, J.N. Carter, Extended model-based automatic gait recognition of walking running, in: *Audio- and Video-Based Biometric Person Authentication*, Halmstad, Sweden, June 2001, pp. 278–283.