International Journal of Computer Vision, 25, 1–23 (1997)

View-Invariant Analysis of Cyclic Motion

STEVEN M. SEITZ, CHARLES R. DYER

seitz@cs.wisc.edu, dyer@cs.wisc.edu

Department of Computer Sciences, University of Wisconsin, Madison, WI 53706

Received August 17, 1995. Accepted August 6, 1996.

Abstract.

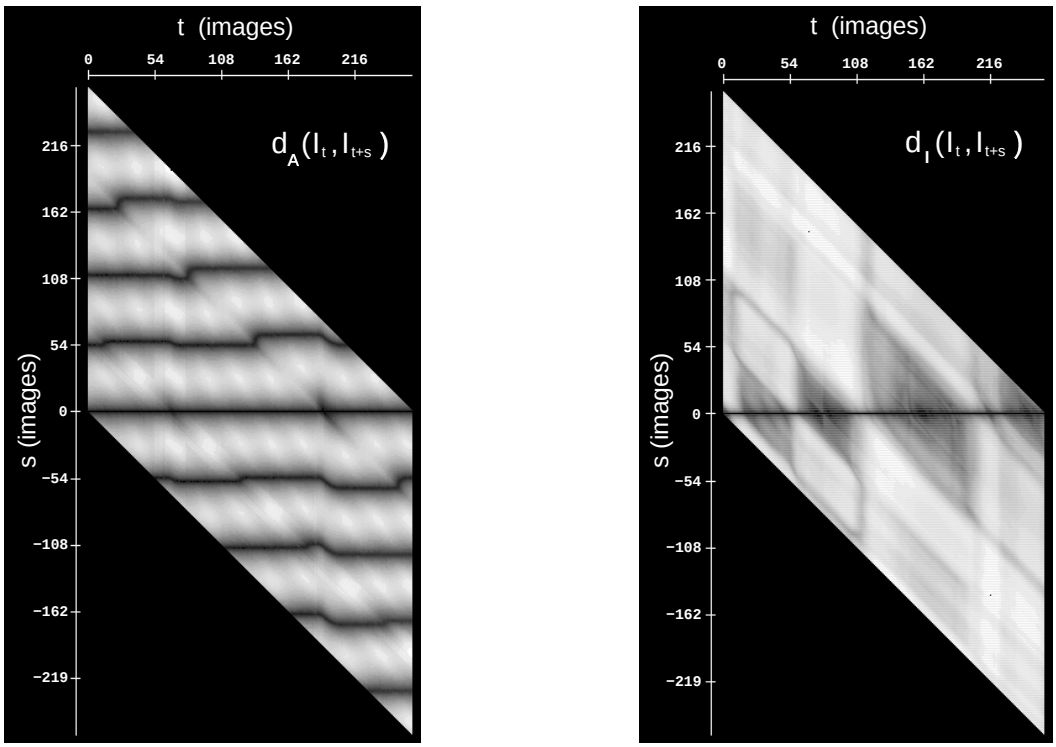
This paper presents a general framework for image-based analysis of 3D repeating motions that addresses two limitations in the state of the art. First, the assumption that a motion be perfectly even from one cycle to the next is relaxed. Real repeating motions tend not to be perfectly even, i.e., the length of a cycle varies through time because of physically important changes in the scene. A generalization of *period* is defined for repeating motions that makes this temporal variation explicit. This representation, called the period trace, is compact and purely temporal, describing the evolution of an object or scene without reference to spatial quantities such as position or velocity. Second, the requirement that the observer be stationary is removed. Observer motion complicates image analysis because an object that undergoes a 3D repeating motion will generally not produce a repeating sequence of images. Using principles of affine invariance, we derive necessary and sufficient conditions for an image sequence to be the projection of a 3D repeating motion, accounting for changes in viewpoint and other camera parameters. Unlike previous work in visual invariance, however, our approach is applicable to objects and scenes whose motion is highly *non-rigid*. Experiments on real image sequences demonstrate how the approach may be used to detect several types of purely temporal motion features, relating to motion trends and irregularities. Applications to athletic and medical motion analysis are discussed.

1. Introduction

Repeating motions abound in nature as well as in the man-made world. Examples include the motions of a heart beating, an athlete running, and a wheel rotating. Repeating motions have an intrinsic temporal structure that describes how a motion varies from one cycle to the next. The temporal structure is invisible in any single image, but becomes apparent when images are compared and correlations are analyzed in the temporal domain. Fig. 1 shows temporal image correlation plots for a repeating and a non-repeating motion. Observe that the repeating sequence (a) exhibits a pattern of dark contours corresponding to pairs of images with high correlation. In this paper we show that this pattern can be analytically modeled to recover meaningful information about motion trends and irregularities in the underlying scene.

Analyzing 3D repeating motions from image sequences is challenging because changes in viewpoint affect the apparent motion. In particular, the projection of a repeating motion with a moving camera will generally not produce a repeating sequence of images. In order to determine salient features of the underlying motion in the scene, the repeating component must be isolated. We describe a technique for analyzing repeating motions that is invariant with respect to (1) changes in the position, orientation, and scale of the moving object(s), and (2) changes in viewpoint and camera parameters such as focal length and aspect ratio. The key observation is that effects due to both (1) and (2) can be modeled by affine transformations so we can make use of principles of affine invariance (Koenderink and Van Doorn, 1991; Tomasi and Kanade, 1992; Shapiro et al., 1994).

Although many real motions are intrinsically repeating, few are perfectly periodic. For instance, a walker's stride frequency may vary visibly from one cycle to the next and a heart may



(a) **(b)**

Fig. 1. Temporal correlation plot for a repeating motion (a) shows a pattern of dark contours. The specific form of the contours encodes the motion's variation from one cycle to the next. No such pattern arises from a non-repeating motion (b). Dark regions in plot (b) correspond to intervals with little or no motion. The plots were formed by comparing every pair of images in the sequence and plotting correlation values as intensity (light – low correlation, dark – high correlation). The image comparison functions, $d_{\mathcal{A}}$ and d_I, are described later in this paper.

beat slower or faster according to changing activity levels. We use the term *cyclic* to describe motions that repeat but lack a constant period. An analog of period can be defined for cyclic motions, which we call *instantaneous period*, that gives cycle length as a function of time. In applications such as cardiac image analysis the instantaneous period can be used to extract individual cycles and to obtain correspondences between cycles, providing a means of tracking changes in a heart over time.

Meaningful information about irregularities can be inferred by comparing different cycles of a cyclic motion. If two cycles are of different lengths, one is irregular with respect to the other. The exact point in time where an irregularity occurs can be determined by considering the derivative of cycle length (which we define in Section 3.4). Furthermore, the nature of the irregularity can be recovered, indicating how the motion varies locally relative to the norm.

Detection of irregularity information has applications in athletic and medical motion analysis. By analyzing the variance of the period in several cycles of a runner's gait or a swimmer's stroke, anomalies can be located providing feedback on specific areas which may need improvement (Perry, 1990). Medical imagery can be analyzed using the same techniques, for instance to detect unevenness in a heart's beating motion. Finally, a motion's periodic variance provides a purely motion-based signature that can be used to detect certain qualities of a motion based on the pattern of irregularities that the motion exhibits.

The remainder of the paper is structured as follows. Section 2 discusses related work on periodic motion analysis. Section 3 formally defines cyclic motion and introduces the notion of *period trace*. Section 4 introduces several motion features

that can be derived from the period trace. Section 5 describes our method for affine-invariant image matching. Section 6 presents a method for automatically recovering a motion's period trace using an optimization approach. Section 7 describes experiments on real image sequences.

2. Related Work

Several researchers have investigated ways of measuring periodicity information from image sequences. Allmen and Dyer (1990) described an approach for detecting periodicity under orthographic projection. They used the curvature scale space of point trajectories to detect repeating patterns of curvature maxima and hence infer a period. Polana and Nelson (1993) presented a method for detecting periodic motions using Fourier transforms of several point trajectories. In theory, the period of the motion could be detected as well by averaging the fundamental frequencies of the point trajectories, although the authors indicated that determining the period in this way was unreliable. Tsai et al. (1994) described a similar technique, using Fourier transforms of curvature values, where the period was determined from a single point trajectory.

None of these approaches are appropriate for analyzing repeating motions that lack a constant period; in particular, they cannot be used to detect or locate motion irregularities. Furthermore, they rely on detecting periodicity in 2D and fail to account for changes in camera viewpoint and reference frame. For example, consider the motion of a person running along an arbitrary path. The runner's motion is periodic in a 3D reference frame that moves with the runner. With a stationary camera, however, the motion of points on the runner will not project to periodic image paths, due to the runner's constantly changing attitude relative to the camera. Therefore, methods based on detecting periodic image paths will fail to detect many 3D repeating motions.

Motion information has proven useful in a number of related recognition problems. Johansson's pioneering work on moving light displays (MLD's) (1973) demonstrated that human motions such as walking and running can be recognized solely based on the trajectories of a small number of bright spots attached to different parts of the body. Since then, motion-based recognition has become an active area of study within computer vision. These efforts include identification of pedestrians (Hogg, 1983; Rohr, 1993), MLD-based motion recognition (Goddard, 1989), hand gesture recognition (Darrell and Pentland, 1993; Davis and Shah, 1994), interpretation of facial expressions (Yacoob and Davis, 1994; Essa and Pentland, 1994), and temporal textures (Polana and Nelson, 1992). An overview of work in this area, including a more thorough list of references, is available in a recent survey article (Cedras and Shah, 1995) and an upcoming book (Shah and Jain).

Cyclic motion analysis is unique among motion-based recognition approaches in that it does not require any type of object- or motion-specific model. Other motion-based recognition techniques (Hogg, 1983; Rohr, 1993; Darrell and Pentland, 1993; Yacoob and Davis, 1994; Essa and Pentland, 1994; Polana and Nelson, 1992) require *a priori* models of the underlying object and/or the motion in the scene, although these models can potentially be learned (Darrell and Pentland, 1993; Baumberg and Hogg, 1994; Bobick and Wilson, 1995). In contrast, periodicity is a *universal* motion characteristic that can be detected and described without knowledge of the underlying object and without reference to a previous instance of the motion. Therefore, the techniques in this paper can be used to interpret image sequences of completely unfamiliar objects and motions.

In Seitz and Dyer (1994a) we introduced a match function that permits view-invariant comparison of images and removes the restriction of a stationary observer. In Seitz and Dyer (1994b) we described a technique for analyzing irregular cyclic motions using the *period trace*. Here we present a unified framework for the analysis of periodic and cyclic motions from image sequences based on the period trace and view-invariant image matching.

3. Motion Classes

Most existing representations of motion describe how a set of intrinsic parameters of a specific object model change over time. Examples include feature trajectories (Gould and Shah, 1989), spatiotemporal surfaces (Baker and Bolles, 1989),

joint angle paths (Goddard, 1989), and evolving physically-based representations (Terzopoulos et al., 1988). In many applications, however, we are interested in the temporal evolution of an object or a scene and not its instantaneous shape. Consequently, in this section we introduce an alternative, purely temporal approach that is independent of an object's spatial properties. Our temporal motion representation, called the *period trace*, describes how the period changes throughout the course of a motion and contains a wealth of useful motion information. Because no assumptions are made about the spatial structure or representation of the scene, the period trace can be recovered from any of several standard motion representations, including (Gould and Shah, 1989; Baker and Bolles, 1989; Goddard, 1989; Terzopoulos et al., 1988), and be used to analyze both rigidly and non-rigidly moving objects and scenes. To motivate the period trace, we first define what it means for a motion to be *periodic* and, more generally, *cyclic*.

3.1. Periodic Motions

We define a *motion* $\mathbf{M}(t)$ to be a function whose value at time t is the instantaneous configuration of a continuously-moving object or scene[1]. We call a motion $\mathbf{P}$ *periodic* if it repeats with period p, i.e.,

$$\mathbf{P}(t+p) = \mathbf{P}(t) \tag{1}$$

for some constant $p > 0$ and all times t in a given time domain. The smallest such constant p is the *period* and the set $\mathbf{P}_{t_0} = \{\mathbf{P}(t) \mid t_0 \leq t < t_0 + p\}$ is called the *cycle* beginning at time t_0.

3.2. Cyclic Motions

The notion of periodicity defined by Eq. (1) has a very restrictive temporal constraint, namely, that the motion is perfectly regular from one cycle to the next. This constraint is relaxed with the introduction of a *period-warping* function, ϕ, as follows:

DEFINITION 1: *A motion* $\mathbf{C}$ *is called* **cyclic** *if*

$$\mathbf{C}(\phi(t)) = \mathbf{C}(t) \tag{2}$$

for all times t *in a given time domain and some increasing continuous function* ϕ *satisfying* $\phi(t) > t$. *A function* ϕ *satisfying these properties is called* $\mathbf{C}$-warping.

Intuitively, $\phi(t)$ corresponds to the start of the next cycle after the cycle beginning at time t. The increasing condition on ϕ ensures that a cyclic motion is order preserving, i.e., $t_1 < t_2$ implies $\phi(t_1) < \phi(t_2)$. Notice that all periodic motions are cyclic and that a cyclic motion is periodic when Eq. (2) is satisfied for $\phi(t) = t + p$, with p the period. Although not all cyclic motions are periodic, any cyclic motion can be *warped* into a periodic motion by appropriate temporal transformations (see Section 4.4).

We emphasize that Definition 1 describes in a very general sense what it means for a *motion* to be cyclic, as opposed to a signal or a texture. The terms *cyclic* and *periodic* have been used previously to mean different things in different contexts. Some authors (Allmen and Dyer, 1990; Tsai et al., 1994) have used the term cyclic to describe strictly periodic phenomenon. On the other hand, a somewhat broader notion of cyclic is often used to characterize one-dimensional signals that vary both in frequency and amplitude. The advantage of Eq. (2) is that it applies equally to one-, two-, and three-dimensional motions, and is expressed purely with respect to temporal quantities.

3.3. Match Functions

Examples of cyclic motions include a rotating wheel, a spinning top, a beating heart, waving gestures, and an athlete running in place. However, many familiar motions that we intuitively characterize as being repeating do not satisfy Definition 1. For instance, consider a runner who moves along an arbitrary path. The runner's motion can be decomposed into a cyclic component (running in place) and a component consisting of whole-body movement along the path[2].

Definition 1 can be generalized to describe partially-cyclic motions, such as running, by re-

placing equality with equivalence under a certain class of transformations. Towards this end, we introduce the concept of a *match function*, $d : P(\mathbf{C}) \Rightarrow \Re$, a non-negative real-valued function on the power set $P(\mathbf{C})$ consisting of all subsets of configurations of a motion $\mathbf{C}$. Match functions can be used to identify configurations that belong to a given equivalence class, i.e., $d(\{\mathbf{C}(t_1), \mathbf{C}(t_2), \ldots, \mathbf{C}(t_k)\}) = 0$ if and only if $\mathbf{C}(t_1) \equiv \mathbf{C}(t_2) \equiv \ldots \equiv \mathbf{C}(t_k)$. Various forms of invariance can be achieved by using appropriate match functions. In particular, a view-invariant match function is presented in Section 5. In addition to incorporating forms of invariance, match functions also provide a natural way of incorporating different methods for comparing images. This framework incorporates existing metrics such as frame difference, optical flow, Hausdorff distance (Huttenlocher et al., 1993), and recent rigidity tests (Seitz and Dyer, 1994a; Soatto and Perona, 1994; McReynolds and Lowe, 1995). Eq. (2) is generalized to incorporate a match function d as follows:

$$d(\{\mathbf{C}(\phi^i(t)) \mid i \in \mathcal{Z}\}) = 0 \tag{3}$$

where

$$\phi^i(t) = (\overbrace{\phi \circ \phi \circ \ldots \circ \phi}^{i})(t)$$

3.4. The Period Trace

A $\mathbf{C}$-warping function ϕ contains information concerning the temporal variation of a cyclic motion $\mathbf{C}$. However, Definition 1 does not determine a unique $\mathbf{C}$-warping ϕ for a given cyclic motion $\mathbf{C}$. For instance, if $\mathbf{C}$ is periodic with period p, $\phi(t) = t + kp$ for any positive integer k satisfies Definition 1. We therefore introduce the notion of *instantaneous period*:

DEFINITION 2: *Let* $\mathbf{C}$ *be a cyclic motion. Let*

$$\phi_1 = \text{pointwise-infimum } \{\phi \ \mathbf{C}\text{-warping}\}$$

Define $\tau_1(t) = \phi_1(t) - t$. τ_1 *is called the* **instantaneous period** *of* $\mathbf{C}$.

The fact that ϕ_1 is $\mathbf{C}$-warping follows from the continuity of $\mathbf{C}$: Let $\mathcal{A} = \{\phi \ \mathbf{C}\text{-}warping\}$ and let $\{\psi_i\}_{i=1}^{\infty}$ be a subset of $\mathcal{A}$ converging to ϕ_1 at t. Then $\mathbf{C}(\psi_i(t)) = \mathbf{C}(t)$ for each i, so $\lim_{i\to\infty} \mathbf{C}(\psi_i(t)) = \mathbf{C}(t)$. But $\lim_{i\to\infty} \mathbf{C}(\psi_i(t)) = \mathbf{C}(\phi_1(t))$ by continuity of $\mathbf{C}$. Hence $\mathbf{C}(\phi_1(t)) = \mathbf{C}(t)$ for all values of t, i.e., ϕ_1 is $\mathbf{C}$-warping.

Intuitively, $\tau_1(t)$ is the length of the cycle beginning at time t. For instance, if $\mathbf{C}$ is periodic, τ_1 is the period. From τ_1, several useful quantities are computable, including τ_n, the instantaneous combined length of the next n cycles. Accordingly, let $\phi_n = \phi^n$ for integers $n > 0$, where product implies functional composition. Then the *nth instantaneous period*, τ_n, is defined as

$$\tau_n(t) = \phi_n(t) - t \tag{4}$$

In addition to being continuous, τ_n has the following property:

$$\frac{\tau_n(t+h) - \tau_n(t)}{h} > -1 \tag{5}$$

which follows from Eq. (4) and the increasing condition on ϕ_1. If τ_n is differentiable, Eq. (5) is equivalent to the condition $\tau_n' > -1$.

The inverse functions exist and are defined as

$$\begin{aligned} \phi_{-n} &= (\phi_n)^{-1} \\ \tau_{-n}(t) &= \phi_{-n}(t) - t \end{aligned}$$

for $n > 0$. $\tau_{-n}(t)$ has the intuitive interpretation as the combined length of the previous n cycles ending at time t. Its value is always negative.

We refer to the set of functions $\{\tau_n \mid n \neq 0\}$ as the *period trace* of a cyclic motion $\mathbf{C}$. For instance, the period trace of a periodic motion is a set of constant functions: $\tau_n = np$, where p is the period, as shown in Fig. 2. The period trace is a comprehensive map of all the cycles in a repeating motion and identifies all corresponding configurations in different cycles. These properties make the period trace a useful tool for comparing different cycles and tracking changes in a repeating motion over time.

4. Trends and Irregularities

The period trace of a cyclic motion describes how the motion varies from one cycle to the next. Consequently, many important attributes of a motion can be computed directly from the period trace, without reference to the spatial structure of the underlying scene. In this section we present several features that quantify temporal trends and irregularities in cyclic motions.

4.1. Local Features

Because the period trace reflects the cumulative motion of n cycles, it is not obvious that local temporal information can be derived from it. In this section we show that local temporal irregularities can indeed be computed from derivatives of the instantaneous periods. Any interval $[t, t+h]$ in which τ_n is not constant, for some n, is said to be an *irregular interval*. The quantity $\tau_n(t+h) - \tau_n(t)$ gives the cumulative change in the nth period, and $\frac{\tau_n(t+h)-\tau_n(t)}{h}$ is the mean rate of change of τ_n in the interval $[t, t+h]$. Letting $h \to 0$, the mean rate of change converges to the instantaneous rate of change, $\tau_n'(t)$. Henceforth, we denote τ_n' as the one-sided derivative defined by $\tau_n'(t) = \lim_{h \to 0} \frac{\tau_n(t+h)-\tau_n(t)}{h}$.

Values of t for which $\tau_n'(t) \neq 0$ are *irregular points*, i.e., points where the motion is faster or slower at $\phi_n(t)$ than at t. Moreover, τ_n' provides both the sign and magnitude of the irregularity. For instance, $\tau_1'(t) = 1$ indicates that the period at time t is increasing at a rate of 1 unit, i.e., motion is faster at t than at $\phi_1(t)$ by 1 unit. Because the motion is faster at the beginning of the cycle than at the end, it must slow down during the course of the cycle, thereby causing a net increase in the period.

Generally, not all points t where $\tau_n'(t) \neq 0$, for some $n \neq 0$, correspond to times where the motion is globally uneven (see Fig. 2). A nonzero value of $\tau_n'(t)$ only indicates that motion is irregular at time t *relative* to time $\phi_n(t)$. A point t where the motion is globally irregular satisfies $\tau_n'(t) \neq 0$, for all $n \neq 0$ (see Fig. 2 (f)).

The total first-order irregularity of a cyclic motion $\mathbf{C}$ at a point in time t can be defined as

$$irreg_{\mathbf{C}}(t) = mean\ \{|\tau_n'(t)| \mid n \neq 0\} \qquad (6)$$

Higher order irregularities may also be relevant in certain situations. For instance, consider a jogger whose stride frequency is steadily decreasing. Due to the changes in speed, τ_n' will be nonzero and second-order irregularities may be more interesting, i.e., where $\tau_n''(t) \neq 0$. Unless otherwise qualified, the term *irregularity* refers to a first-order irregularity.

4.2. Global Features

The period trace can also be used to compute various global features of a cyclic motion. These include

- period
- cyclic acceleration
- shortest and longest cycle
- regular and irregular points, cycles, and intervals

The simplest of these is the period. If $\tau_1'(t)$ is uniformly 0 then τ_1 is constant and corresponds to the period. Similarly, if $\tau_1' = c$, where $c > 0$, then the period is uniformly increasing at a rate of c. In the case of a decelerating jogger, τ_1' gives the change in speed in terms of the rate of change of the stride period. The mean value of τ_1' gives the mean rate of increase of the period, which we call the *average cyclic acceleration*.

In addition to trends such as the period and cyclic acceleration, the period trace can be used to compute globally significant points, cycles, or intervals in a motion. For instance, a longest (shortest) cycle is given by any interval $[t, \phi_1(t))$ for which $\tau_1(t)$ is maximal (minimal).

Highly irregular points can be found by maximizing Eq. (6). Similarly, the irregularity of a cycle is measured by integrating Eq. (6) over the entire cycle:

$$irreg(\mathbf{C}_t) = \frac{1}{\tau_1(t)} \int_t^{\phi_1(t)} irreg_{\mathbf{C}}(s)\ ds$$

Another useful feature is the median[3] cumulative irregularity of an interval:

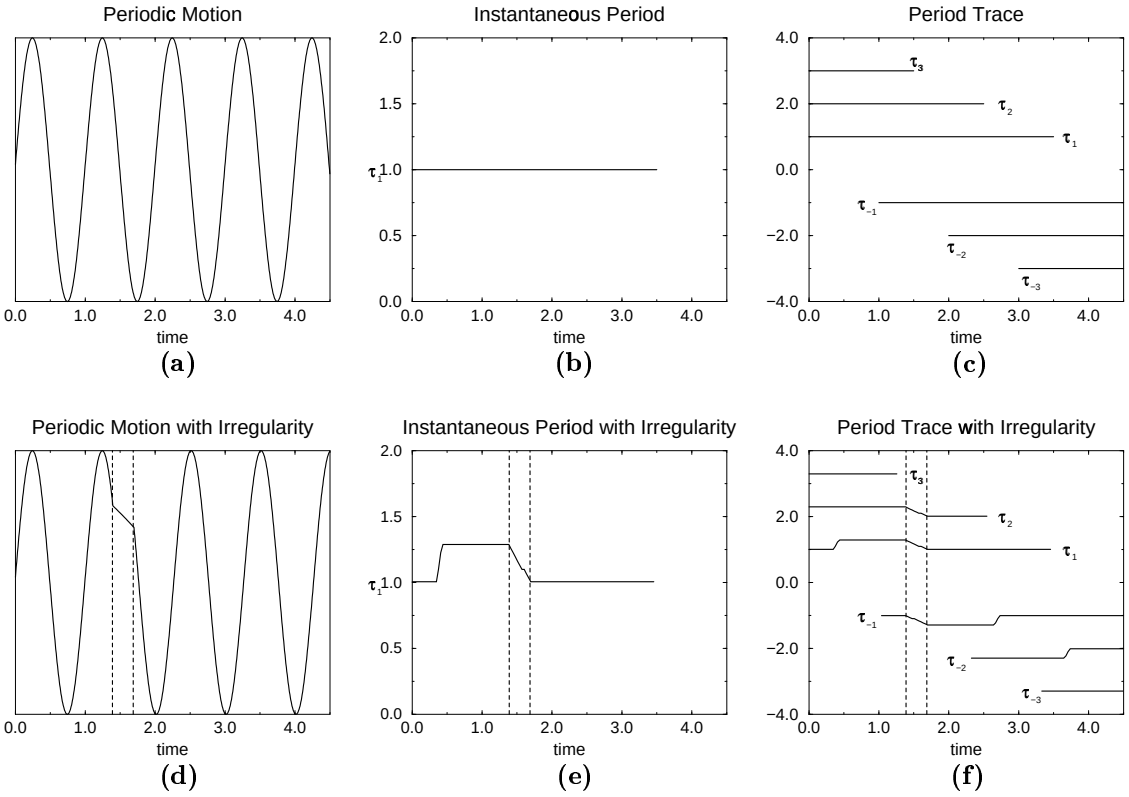


Fig. 2. Effects of an irregularity on the instantaneous period. A periodic sinusoid signal (a) has a constant period (b) and a period trace consisting of multiples of the period (c). A small irregularity is introduced (d) by temporarily slowing by a factor of 10. The irregularity shows up in two places in the instantaneous period (e), where a ramp occurs once when the irregularity enters the current cycle and once where it leaves. The ambiguity is resolved using the period trace (f) where the ramps "line-up" only at the true irregularity (i.e., in a small interval around time 1.5).

$$irreg_{\mathbf{C}}(t_1, t_2) = median\ \{|\tau_n(t_2) - \tau_n(t_1)| \ \mid \ n \neq 0\}$$

4.3. Motion Signatures

Some types of cyclic motion produce telltale temporal signatures. Examples include periodic motions and motions with uniform cyclic acceleration, both of which have instantaneous periods that are linear. The form of a motion's period trace may also indicate something about the distribution of irregularities. Cyclic motions that contain isolated irregularities have very distinct signatures (see Figs. 2, 6, and 8); the instantaneous period of a motion of this sort is piecewise-constant except at isolated intervals whose locations and extents correspond to the irregularities.

The locations of irregularities in the period trace provide other important qualitative clues about the behavior of a moving object. Any cyclic motion containing at least one irregularity has a period trace with ramps in various places, e.g., see Fig. 2 (f). There are two categories of ramps; those that line-up, recurring at the same place in each cycle, and isolated ramps that don't recur. Generally, irregular motions will have both types of ramps, but certain motions have only ramps that line-up. These are motions where irregularities occur at roughly the same point in each cycle. For instance, a heart may contract at a different rate in each cycle, resulting in a series of ramps that line-up at each contraction. A similar phenomenon is shown in Fig. 8.

4.4. Measuring Relative Speed and Removing Irregularities

The period trace provides detailed information about cycles and changes in speed but does not explicitly give instantaneous speed as a function of time. Consider the motion of a jogger whose stride frequency varies through time. Suppose we wish to describe these variations not as changes in stride frequency, but in terms of differences in speed relative to the norm. For instance, we might say that the jogger is currently running twice as fast as normal (e.g., 15 mph versus 7.5 mph). This can be achieved by recovering a *time-warping* function that maps a perfectly regular, e.g. periodic, motion into the uneven motion of the runner.

Any cyclic motion can be warped into a periodic motion by *flattening* its period trace. This flattening transformation and its inverse are the time-warping functions we're after. The first step is to find a function that maps a cyclic motion into a periodic motion. In other words, given a cyclic motion, $\mathbf{C}(t)$, we wish to find a reparameterization $s = \sigma(t)$ such that $\tilde{\tau}_1(s)$ is constant, where

$$\tilde{\tau}_1(s) = (\sigma \circ \phi_1 \circ \sigma^{-1})(s) - s$$

The result is that $\mathbf{C}(s)$ is periodic, satisfying Eq. (1).

For any cyclic motion, $\mathbf{C}(t)$, there exist innumerable reparameterizations, $s = \sigma(t)$, such that $\mathbf{C}(s)$ is periodic. A natural choice is one that synchronizes the motion with respect to a particular cycle. The result is that the reference cycle remains intact and all other cycles are aligned to match the reference cycle. This approach is attractive because a cycle may be chosen that satisfies a particular criterion and the whole motion will be rectified to conform to the criterion. For instance, irregularities may be removed by choosing a *regular* reference cycle, using the techniques of Section 4.2. Similarly, choosing a cycle that has a particular anomaly will result in a motion with a corresponding anomaly in every cycle.

Let $\mathbf{C}(t)$ be a cyclic motion and choose a cycle, $\mathbf{C}_{t_0}$, starting at time t_0. Define time warping function σ and its inverse as

$$\sigma(t) = \phi_{-n}(t) + n\tau_1(t_0) \quad \text{for } t \in [\phi_n(t_0), \phi_{n+1}(t_0))$$

$$\sigma^{-1}(s) = \phi_n(s - n\tau_1(t_0)) \quad \text{for } \sigma^{-1}(s) \in [\phi_n(t_0), \phi_{n+1}(t_0))$$

CLAIM: $\mathbf{C}(s)$ *is periodic, where* $s = \sigma(t)$

Proof: Fix a value of s. Suppose, without loss of generality, that $\sigma^{-1}(s) \in [\phi_n(t_0), \phi_{n+1}(t_0))$. Then

$$\begin{aligned}
\tilde{\tau}_1(s) &= (\sigma \circ \phi_1 \circ \phi_n)(s - n\tau_1(t_0)) - s \\
&= (\phi_{-n-1} \circ \phi_1 \circ \phi_n)(s - n\tau_1(t_0)) \\
&\qquad +(n+1)\tau_1(t_0) - s \\
&= s - n\tau_1(t_0) + (n+1)\tau_1(t_0) - s \\
&= \tau_1(t_0)
\end{aligned}$$

Therefore $\tilde{\tau}_1(s)$ is constant so $\mathbf{C}(s)$ is periodic with period $\tau_1(t_0)$. ■

Assuming a constant frame rate, a cyclic image sequence represents a motion sampled uniformly in the t domain. Rectification involves selecting an image subsequence corresponding to a uniform sampling in the s domain. Uniform s sampling ensures that the new sequence is periodic when played back at a constant frame rate. The rectified sequence may be constructed using σ^{-1} as follows: given an initial image sequence $I_1, I_2, \ldots, I_m$, the rectified sequence is

$$I_{\sigma^{-1}(1)}, I_{\sigma^{-1}(2)}, \ldots, I_{\sigma^{-1}(k)}$$

where $k = \lfloor \sigma(m) \rfloor$. The effect is that all irregularities are removed and there is no variation in speed from one cycle to the next.

Whereas σ^{-1} can be used for making a sequence periodic, σ is useful for describing how a cyclic motion deviates from periodicity. The derivative of σ gives the factor by which motion is locally faster with respect to the reference cycle. Contrast this with τ_1', which gives the local change in cycle length as a function of time; $\frac{d\sigma}{dt}$ gives relative speed directly, but requires a reference cycle. For instance, $\frac{d\sigma}{dt}(t_1) = 2$ indicates that the motion at time t_1 is twice as fast as normal. In Section 7 these techniques are used to measure relative speed and to demonstrate the rectification process for a real image sequence.

5. Affine Invariance

Three-dimensional motion analysis from image sequences requires isolating salient 3D motion characteristics from artifacts arising from the projection process. For instance, consider the problem of measuring the stride frequency of an athlete filmed while running along a path. Both the relative movement of the runners' limbs and his change in attitude relative to the camera contribute to the motion in the image sequence, but only the former is relevant to the stride frequency. Equivalently, the athlete may be running in place while the camera moves. We would like to be able to determine the stride frequency in a way that is not affected by changes in viewpoint. This is accomplished using a view-invariant match function that equates a set of images if and only if they represent views of an object in the same configuration.

5.1. *Projected Match Criterion*

Formally, we want to determine if a set of images *match*, i.e., correspond to views of the same 3D configuration of an object or a scene. Our formulation assumes an *affine* camera model (Koenderink and Van Doorn, 1991; Mundy and Zisserman, 1992; Shapiro et al., 1994) which is a generalization of orthographic, weak perspective, and paraperspective projections. An image sequence is represented as a sequence of $2 \times n$ matrices $\mathbf{I}_1, \ldots, \mathbf{I}_m$ where each column is the instantaneous position of one of n image features. It is assumed that corresponding columns of $\mathbf{I}_t$ and $\mathbf{I}_s$ represent projections of the same scene point. Hence, the correspondence between features in different images is assumed to be known. An object or scene is represented as a $3 \times n$ matrix $\mathbf{S}$ with columns corresponding to the 3D coordinates of features in an object-centered affine reference frame. Without loss of generality, assume that each row of $\mathbf{I}_t$ has zero mean; choose the origin of each image to be the centroid of the feature points. Under an affine projection model a set of images, $\Gamma = \{\mathbf{I}_{t_1}, \ldots, \mathbf{I}_{t_k}\}$, match if and only if

$$\mathbf{I}_{t_i} = \mathbf{\Pi}_{t_i}\mathbf{S} \tag{7}$$

for some fixed $3 \times n$ matrix $\mathbf{S}$, $i = 1, \ldots, k$, and 2×3 matrices $\mathbf{\Pi}_{t_1}, \ldots, \mathbf{\Pi}_{t_k}$.

Define the measurement matrix $\mathbf{M}_\Gamma$ of a set of images to be the concatenation of the image measurements:

$$\mathbf{M}_\Gamma = \begin{bmatrix} \mathbf{I}_{t_1} \\ \vdots \\ \mathbf{I}_{t_k} \end{bmatrix}$$

If the images Γ match, then by Eq. (7), we can express $\mathbf{M}_\Gamma$ as follows:

$$\mathbf{M}_\Gamma = \begin{bmatrix} \mathbf{\Pi}_{t_1} \\ \vdots \\ \mathbf{\Pi}_{t_k} \end{bmatrix} \mathbf{S} \tag{8}$$

Therefore $\mathbf{M}_\Gamma$ is the product of two matrices, each having rank at most 3. It follows that $\mathbf{M}_\Gamma$ has rank 3 or less. This is the Rank Theorem, due to Tomasi and Kanade (1992), and generalized for an affine camera. Conversely, any measurement matrix of rank 3 or less can be decomposed as in Eq. (8) using singular value decomposition (Tomasi and Kanade, 1992). Therefore, we have the following:

GENERALIZED RANK THEOREM: *A set of images, Γ, match if and only if $\mathbf{M}_\Gamma$ is of rank at most 3.*

Under orthographic projection (the case considered by Tomasi and Kanade), the rank condition alone is not sufficient to determine that a set of images match. However, it is both necessary and sufficient under the more general affine camera model. Hence, the Generalized Rank Theorem can be used to detect if a set of images could have been produced by a 3D cyclic motion. Accordingly, let $\mathbf{I}_1, \ldots, \mathbf{I}_m$ be a sequence of images, let ϕ be an integer-valued function, and denote $\Gamma_t^\phi = \{\mathbf{I}_{\phi^i(t)} \mid 1 \leq \phi^i(t) \leq m\}$. The following result is a consequence of the Generalized Rank Theorem and Eq. (2):

PROJECTED MATCH CRITERION: *An image sequence, $\mathbf{I}_1, \ldots, \mathbf{I}_m$, is the affine projection of a 3D cyclic motion if and only if there exists an increasing function satisfying $\phi(t) > t$ such that $rank(\mathbf{M}_{\Gamma_t^\phi}) \leq 3$ for $t = 1, \ldots, m$.*

Such a function ϕ is called **I**-warping. The Projected Match Criterion suggests a provably-correct way of detecting cyclic motions: check all possible **I**-warping functions. Under an assumption of periodicity, this brute-force strategy is feasible (Seitz and Dyer, 1994a) but exhaustive search is too costly in general. A more efficient method is described in Section 6.

5.2. Affine-Invariant Image Matching

Because rank measurements are highly sensitive to numerical errors, a more robust measure of match quality is needed. We can characterize the residual error of a sequence of images by the amount by which features must be perturbed in order to make the images match. Define the match error of a set of images as follows:

$$dist_{\mathcal{A}}(\Gamma) = min\ \{\|\mathbf{E}\|_{rms} \mid rank(\mathbf{M}_\Gamma + \mathbf{E}) \leq 3\}$$

$\|\mathbf{E}\|_{rms}$ is the root-mean-squared norm of the matrix $\mathbf{E}$ defined by $\|\mathbf{E}\|_{rms} = \sqrt{\frac{1}{2kn}\sum_{i,j}\mathbf{E}_{ij}^2}$. Alternatively, $dist_{\mathcal{A}}$ can be expressed in terms of the singular values of $\mathbf{M}_\Gamma$:

THEOREM: $dist_{\mathcal{A}}(\Gamma) = \sqrt{\frac{1}{2kn}\sum_{i=4}^{n}\sigma_i^2}$ *where* $\sigma_1 \ldots \sigma_n$ *are the singular values of* Γ.

Proof: Singular value decomposition gives

$$\mathbf{M}_\Gamma = \mathbf{U}\mathbf{\Sigma}\mathbf{V}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices and the singular values of $\mathbf{M}_\Gamma$ appear along the diagonal of $\mathbf{\Sigma}$ (a $2k \times n$ diagonal matrix). The above equation can be rewritten as

$$\mathbf{M}_\Gamma = \mathbf{U}\tilde{\mathbf{\Sigma}}\mathbf{V} + \mathbf{U}\mathbf{\Sigma}'\mathbf{V} \tag{9}$$

where $\tilde{\mathbf{\Sigma}}$ and $\mathbf{\Sigma}'$ are structured as

$$\tilde{\mathbf{\Sigma}} = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \mathbf{0} \\ & \sigma_3 & \\ \mathbf{0} & & \mathbf{0} \end{bmatrix} \quad \mathbf{\Sigma}' = \begin{bmatrix} \mathbf{0} & & \mathbf{0} \\ & \sigma_4 & \\ \mathbf{0} & \ddots & \\ & & \sigma_n \end{bmatrix}$$

Eq. (9) can be expressed as

$$\mathbf{M}_\Gamma = \tilde{\mathbf{M}}_\Gamma + \mathbf{M}'_\Gamma$$

where $\tilde{\mathbf{M}}_\Gamma = \mathbf{U}\tilde{\mathbf{\Sigma}}\mathbf{V}$ and $\mathbf{M}'_\Gamma = \mathbf{U}\mathbf{\Sigma}'\mathbf{V}$.

$\tilde{\mathbf{M}}_\Gamma$ is the optimal (in an RMS sense) rank-3 approximation of $\mathbf{M}_\Gamma$ (Stewart, 1973). Hence, $\mathbf{M}'_\Gamma$ is the minimal perturbation of $\mathbf{M}_\Gamma$ that produces a match. Therefore, $dist_{\mathcal{A}}(\Gamma) = \|\mathbf{M}'_\Gamma\|_{rms}$. The latter term is just $\|\mathbf{\Sigma}'\|_{rms}$ since $\mathbf{U}$ and $\mathbf{V}$ are orthogonal, and the result follows. ■

The measure $dist_{\mathcal{A}}$ gives the average amount (in pixels) necessary to additively perturb the coordinates of each image feature in order to produce a set of matching images. In the case where there are only two images we abbreviate $dist_{\mathcal{A}}(\{\mathbf{I}_s, \mathbf{I}_t\})$ as $dist_{\mathcal{A}}(\mathbf{I}_s, \mathbf{I}_t)$. $dist_{\mathcal{A}}$ has the following properties:

- $dist_{\mathcal{A}}(\Gamma) = 0$ if and only if the images, Γ, match exactly.
- $dist_{\mathcal{A}}$ is well-behaved with respect to noise (see, for example, Stewart (1973)).
- $dist_{\mathcal{A}}$ is defined in image coordinates and can be directly related to measurement errors.
- $dist_{\mathcal{A}}$ is always zero when less than five features are considered.
- $dist_{\mathcal{A}}$ may be non-zero when the features are co-planar. Therefore co-planarity of the feature set does not cause problems, in contrast to related techniques (Koenderink and Van Doorn, 1991; Tomasi and Kanade, 1992; Shapiro et al., 1994).
- For n features, $dist_{\mathcal{A}}(\mathbf{I}_s, \mathbf{I}_t) = \frac{\sigma_4}{2\sqrt{n}}$ and the evaluation cost is $O(n)$ arithmetic operations. For m images, the evaluation cost is the smaller of $O(nm^2)$ and $O(mn^2)$.

5.3. A Simplified Match Function

The match function $dist_{\mathcal{A}}$ is appropriate when a large number of features are continuously visible throughout an image sequence. For real world applications however, this is seldom the case due to factors such as noise and occlusion. Therefore, we use a simplified match function that computes pairwise correlations between images that have a suitable number of features in common:

$$d_{\mathcal{A}}(\Gamma) = mean\ \{dist_{\mathcal{A}}(\mathbf{I}_{t_i}, \mathbf{I}_{t_j}) \mid 1 \leq i < j \leq k, \\ \mathbf{I}_{t_i} \text{ and } \mathbf{I}_{t_j} \text{ share at least 5 features}\}$$

Note, however, that $d_{\mathcal{A}}(\Gamma) = 0$ is only a necessary condition for a set of images to match, whereas $dist_{\mathcal{A}}(\Gamma) = 0$ is both necessary and sufficient.

6. Recovering the Period Trace

In order to detect cyclic motion features such as those introduced in Section 4, we must be able to recover functions $\tau_n(t)$ for each integer $n \neq 0$ from a cyclic motion $\mathbf{C}$. Because the complete period trace can be computed from τ_1, it is sufficient to determine τ_1. For periodic motions, τ_1 is constant and can be determined using unconstrained scalar optimization methods. More generally, the recovery of τ_1 is posed as a constrained functional minimization problem.

6.1. The Periodic Case

If a cyclic motion is known to have a certain form, this information may be used to help recover its period trace. Particularly, if an *a priori* motion model is available, e.g., linear or piecewise constant, the model can ameliorate the task of determining τ_1 by reducing the search space and making the recovery procedure less sensitive to noise. Existing techniques for recovering periodicity information (Allmen and Dyer, 1990, Polana and Nelson, 1993; Seitz and Dyer 1994a; Tsai et al., 1995) implicitly use this strategy, exploiting the assumption of a constant period. The case where such a model is not available is treated in the next section.

The instantaneous periods of a periodic motion are constant and correspond to the multiples of the period. Therefore, the period trace of a periodic motion can be found by optimizing some function over the set of candidate periods. Towards this end, we introduce a tool for assessing the significance of a candidate period or period trace.

A Significance Measure In the presence of noise and quantization effects it is unlikely that corresponding configurations of a cyclic motion will match precisely, yielding the zero match scores predicted by Eq. (3). Matching configurations can nevertheless be detected by minimizing a well-behaved match function d. However, it is still necessary to evaluate the significance of a solution in order to discriminate between a cyclic and a non-cyclic motion. For this purpose we use a statistical measure based on the Kolmogorov-Smirnov (K-S) test (Press et al., 1988) .

The quality of a computed period trace can be assessed by considering the problem within the context of sampling theory. The operation of comparing two configurations $\mathbf{C}(s)$ and $\mathbf{C}(t)$ can be thought of as taking a random point sample from *match space*:

$$\mathcal{M} = \{d(\mathbf{C}(s), \mathbf{C}(t)) \mid s, t \in \Re, s < t\}$$

A period trace $\pi = \{\tau_n \mid n \neq 0\}$ determines a specific sample $\mathcal{S}_\pi \subset \mathcal{M}$ given by

$$\mathcal{S}_\pi = \{d(\mathbf{C}(t), \mathbf{C}(t + \tau_n(t))) \mid n > 0\}$$

For example, a period of p corresponds to the sample:

$$\mathcal{S}_p = \{d(\mathbf{C}(t), \mathbf{C}(t + np)) \mid n > 0\}$$

$\mathcal{S}_\pi$ contains the match scores of precisely the configurations that correspond, under the hypothesis that $\mathbf{C}$ is cyclic with period trace π. If this hypothesis is correct, the distribution of $\mathcal{S}_\pi$ should differ significantly from that of $\mathcal{M}$, since corresponding configurations minimize d by assumption. The K-S test gives the probability, or *P-value*, of a random sample of $\mathcal{M}$ matching the cumulative histogram of $\mathcal{S}_\pi$, i.e., the probability that the motion is not cyclic.

In practice $\mathbf{C}(t)$ may be sampled both in space and time so that the exact distribution of $\mathcal{M}$ is not known. For an image sequence, $\Gamma = \{\mathbf{I}_1, \ldots, \mathbf{I}_k\}$, $\mathcal{M}$ is approximated by $\mathcal{M}' = \{d(\mathbf{I}_s, \mathbf{I}_t) \mid 1 \leq s < t \leq k\}$. For details on how to calculate the K-S statistic, consult (Press et al., 1988).

The P-value of a candidate period trace π is evaluated by applying the K-S test to $\mathcal{S}_\pi$ and $\mathcal{M}'$. Since the objective is to find a period trace that minimizes d, it is necessary to assess P-values of period traces only for which

$$mean(\mathcal{S}_\pi) < mean(\mathcal{M}') \tag{10}$$

Candidate period traces for which this inequality is violated are normalized by assigning P-values of

1. The P-value of a motion is defined to be the minimum P-value of all possible period traces.

In practice, a motion's exact P-value is never computed; the P-value of the period trace obtained by the optimization procedure provides an upper bound that is generally sufficient to differentiate cyclic from non-cyclic motions. We have found this test to be quite robust; real cyclic motions yield P-values that are exceedingly small, on the order of 10^{-20} or smaller, whereas non-cyclic motions have P-values at or near 1. In fact, every one of the non-cyclic motions we evaluated had a P-value of 1. The reason is that image sequences are decidedly non-random; adjacent images (i.e., $\mathbf{I}_t$ and $\mathbf{I}_{t+1}$) tend to be similar, resulting in a natural bias towards small periods. $\mathcal{M}'$ contains match scores for *all* adjacent frames and it is therefore unlikely that Eq. (10) will be satisfied for any period trace π chosen at random. As a result, most period traces will have P-values of 1. Note that this bias helps to differentiate cyclic from non-cyclic motions since the latter tend to have P-values of 1. For cyclic images sequences there is an implicit assumption that the instantaneous period is uniformly greater than 1. In other words, the frame rate (images per second) should be greater than the frequency (cycles per second), or else the period trace cannot be detected.

Detecting Periods Using the K-S Test Using the K-S test it is straightforward both to calculate the probability that a motion is periodic and to compute the most likely period. Given a motion $\mathbf{C}(t)$ defined on the domain $[t_1, t_2]$, evaluate the K-S test for every candidate period $\tilde{p} \leq \frac{t_2 - t_1}{2}$. The value of $\tilde{p}$ with lowest P-value constitutes the most likely period and the corresponding P-value gives an upper bound on the probability that the motion is not periodic.

A potential complication arises in discriminating between the true period and its multiples. If p is the true period, note that $S_{kp} \subset S_p$ for any positive integer multiple k of p. Hence, it is possible that $mean(S_{kp}) \leq mean(S_p)$, for some value of k. However, the K-S test assigns greater significance to larger samples, adding a built-in bias in favor of smaller periods. Because S_{kp} constitutes a smaller sample than S_p, the latter is automatically assigned more significance, resulting in a smaller P-value. Consequently, the P-value of the true period will tend to be more significant than those of its multiples due to the natural bias of the K-S test. This phenomenon can be seen in Fig. 7, where the true period has a much smaller P-value than its multiple.

6.2. The Cyclic Case

When nothing is known *a priori* about a cyclic motion, the recovery of its period trace can be posed as a constrained functional optimization problem. Because the complete period trace can be computed from τ_1, it is sufficient to determine $\tau_1(t)$ for each value of t. Given a cyclic motion $\mathbf{C}$, we seek to minimize $\tau_1(t)$ subject to the following constraints:
1. $\mathbf{C}(t) = \mathbf{C}(\phi^i_{\tau_1}(t))$ for $t \in \Re$, $i \in \mathcal{Z}$
2. τ_1 is continuous
3. $\tau_1(t) > 0,\ t \in \Re$
4. $\tau_1'(t) > -1,\ t \in \Re$

where $\phi_{\tau_1}(t) = \tau_1(t) + t$.

For discrete data sampled on the interval $[t_1, t_2]$, the problem is simplified by introducing a match function, d, (see Section 3.3) and formulating an energy function[4] to be minimized. It is assumed that d is well-behaved near zeros so that near-matching configurations show up as local minima of d. The energy function, $E(\tau_1)$, has two terms to enforce constraints 1 and 2, weighted by a scalar factor α:

$$\begin{aligned} E_1(\tau_1, t) &= d(\{\mathbf{C}(\phi^i_{\tau_1}(t)) \mid t_1 \leq \phi^i_{\tau_1}(t) \leq t_2, \\ &\qquad\qquad i \in \mathcal{Z}\}) \\ E_2(\tau_1, t) &= |\tau_1'(t)| \\ E(\tau_1) &= \sum_{t=t_1}^{t_2} [E_1(\tau_1, t) + \alpha E_2(\tau_1, t)] \end{aligned} \tag{11}$$

Constraints 3 and 4 are local and easily enforced, although the strict inequality in constraint 4 is relaxed (since τ_1' may be arbitrarily close to -1)[5].

We propose a multiscale *snake* algorithm adapted from (Kass et al., 1988; Williams and Shah, 1992) for recovering the period trace by iteratively minimizing the energy function E. The snake is initialized to a rough estimate of τ_1 and is incrementally refined so as to reach a state of locally minimal energy. In each iteration a value

of t is selected and $\tau_1(t)$ is adjusted by 0, 1, or -1 unit so as to globally decrease $E(\tau_1)$ while checking that constraints 3 and 4 are satisfied. A pass consists of iterating over all successive values of t in the domain of τ_1. The algorithm performs repeated passes until E converges. Convergence is guaranteed since energy decreases monotonically in each iteration. To avoid problems with local minima, the optimization approach is repeated at three scales of increasing resolution using the output of each stage as the initialization for the next.

Initialization is performed by setting τ_1 to the constant function that minimizes the K-S test, as described in Section 6.1. In combination with the multiscale approach, we have found this simple strategy to provide adequate initialization for the optimization procedure to converge correctly for the sequences that we tested. However, this strategy is designed for near-periodic sequences; highly irregular cyclic sequences may require a more sophisticated initialization procedure, perhaps incorporating user interaction, as in (Kass et al., 1988).

6.3. The Recovery Algorithm

The following sequence of operations is used to compute the period trace and its P-value from a sequence of images $\mathbf{I}_1, \ldots, \mathbf{I}_m$.

- *Step 1.* Choose a match function d for comparing images.
- *Step 2.* Compute P-values for each candidate period $\tilde{p} = 2, \ldots, \frac{m-1}{2}$. The candidate p with smallest P-value is the most likely period. The associated P-value gives the probability that the motion is not periodic.
- *Step 3.* Apply energy minimization passes until the snake converges at a succession of finer scales. The snake is initialized to $\tau_1 = p$ at the coarsest scale. Upon convergence, the new snake configuration is used to initialize energy minimization at the next finer scale.
- *Step 4.* Obtain the P-value of the derived period trace using the K-S test, giving the probability that the motion is not cyclic.

As noted in Section 6.1, this algorithm cannot detect periods of one frame or less. Such motions will generally be found to be *non-cyclic* with a P-value of 1.

7. Experiments

Three image sequences were used to evaluate the performance of the optimization methods described in Section 6 and the detection of temporal motion features. In addition, three non-cyclic sequences were used as a control set to test the discrimination accuracy of the algorithm. The first two sequences illustrate the use of match functions for view-invariant recovery of irregularity information. The third experiment uses X-ray data of a beating heart to show that the period trace can be calculated from medical imagery without preprocessing and can be used for temporal segmentation and analysis. Selected images from all six sequences are shown in Fig. 3. All three control sequences yielded P-values of 1, indicating with great probability that they are not cyclic. In each case, the period minimizing $mean(S_p)$ was $p = 1$ frame. More experiments can be found in Seitz and Dyer (1997).

7.1. Walking Sequence

A human subject was filmed walking in an arc subtending about 70 degrees. The camera was manually rotated to keep the subject in view, and the focal length was slowly changed during filming. To aid software feature detection and tracking, reflective markers were placed in areas which were visible for the duration of the sequence (i.e., right arm, right leg, mid torso, and head). We used the method described in Section 6.1 to determine the period with smallest P-value. The affine-invariant match function, $d_{\mathcal{A}}$, of Section 5 was used for comparing images. Fig. 3 shows that the algorithm successfully detected matching images; images 1, 38, and 112 are views of the subject at roughly the same point in her walk cycle. The algorithm detected a period of 37 with a P-value of 10^{-18}. The P-value of the detected period clearly beats out all other candidates (see Fig. 5).

Because features on only one side of the torso were tracked, the feature set was nearly planar.

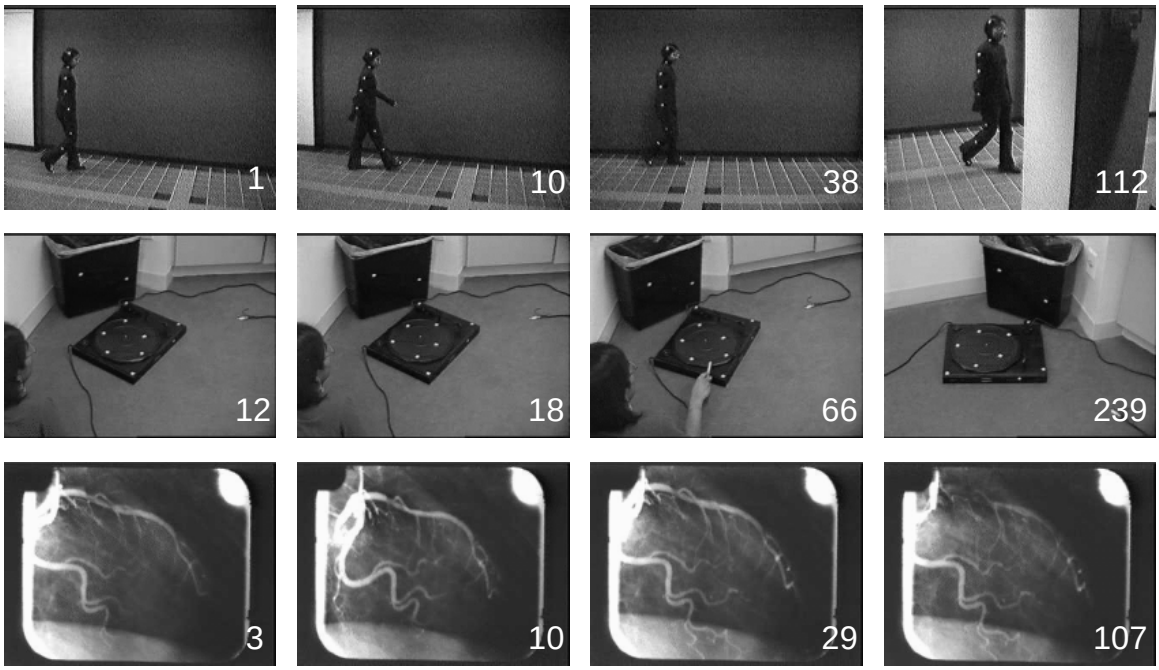


Fig. 3. Selected images from image sequences of cyclic motions. Top: Human subject walking. Middle: Rotating turntable. Bottom: X-ray images of a beating heart. The first, third, and fourth images in each set match and these correspondences were computed automatically from the period trace.

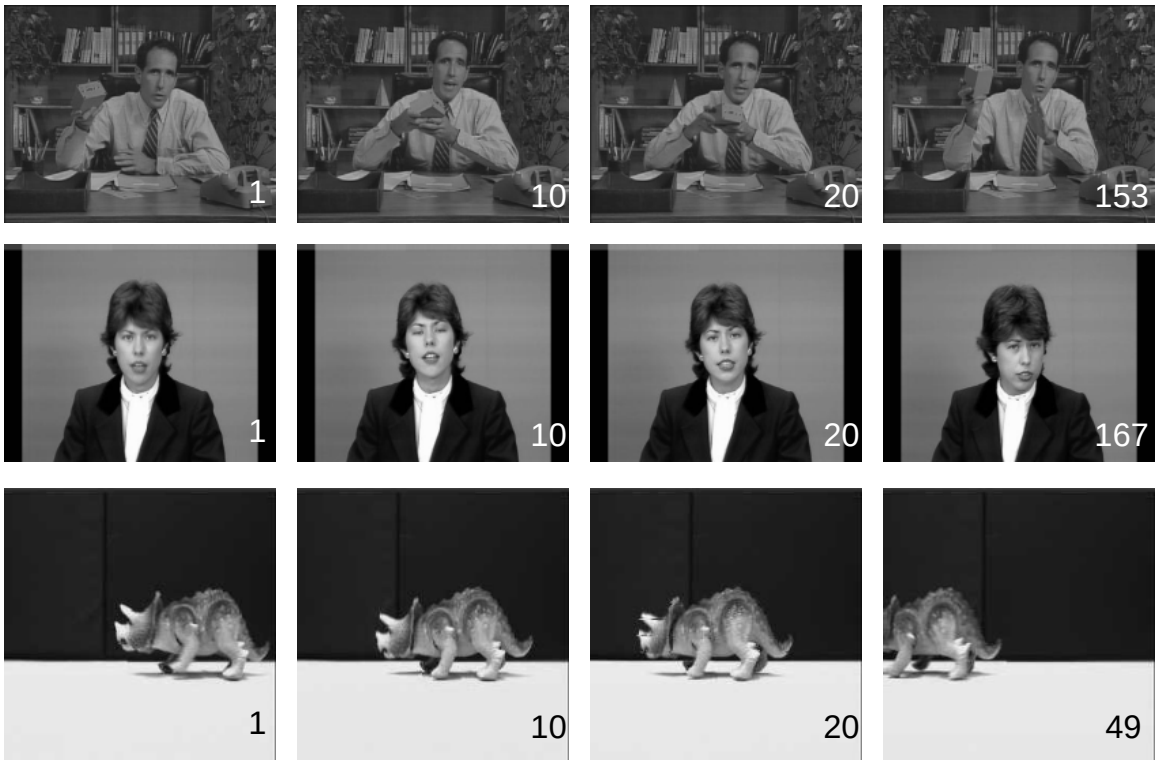


Fig. 4. Selected images from three non-repeating image sequences: *salesman* (top), *claire* (center), and *toy* (bottom). All three sequences yielded P-values of 1, indicating with great probability that the motions are non-cyclic.

Co-planarity of the feature set is known to be problematic for many shape-from-motion algorithms (Koenderink and Van Doorn, 1991; Tomasi and Kanade, 1992; Shapiro et al., 1994). Note, however, that it does not cause a problem for our algorithm because an explicit 3D scene representation is never computed. Note also that the entire motion is nearly affine; the features in each image can be roughly approximated by a horizontal shear and/or reflection of the features in the first image. In fact, we have found this property to be true of other human locomotory motions such as running, skipping and jumping. Since the match function d_A equates all images that are related by a 3D affine transformation, purely affine motions appear stationary, and period detection fails. The periods of nearly-affine motions, however, can be reliably detected with our approach, but the resulting P-values tend to be less significant.

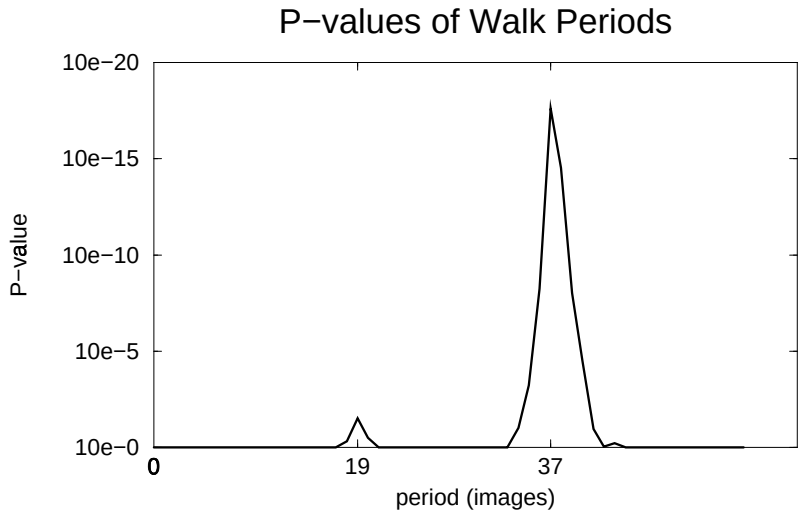


Fig. 5. Period P-values of a walking human. The period with the highest level of significance is 37, with a P-value of 10^{-18}.

7.2. Turntable Motion Sequence

A rotating turntable was filmed using a moving hand-held video camera. Reflective markers were placed both on the turntable and elsewhere in the static scene. Note that although the motion of the turntable is rigid, the entire scene moves non-rigidly. Custom software was used to track the markers as the camera moved about the scene and the turntable simultaneously rotated. Twice the turntable was briefly touched to temporarily slow the rotation and artificially produce an irregularity. Images were compared using the affine-invariant match function, $d_{\mathcal{A}}$, of Section 5.

Fig. 6 shows the recovered period trace. The figure illustrates that the optimization process successfully located low-energy "valleys" in the match space, corresponding to the dark contours in the temporal correlation plot. The two regions with the highest irregularity values correspond to the two brief intervals in which the turntable was touched. The width of each interval indicates how long the turntable was touched and the irregularity value determines the extent to which the rotation was slowed. Notice that there are several ramps throughout the period trace but all coin-

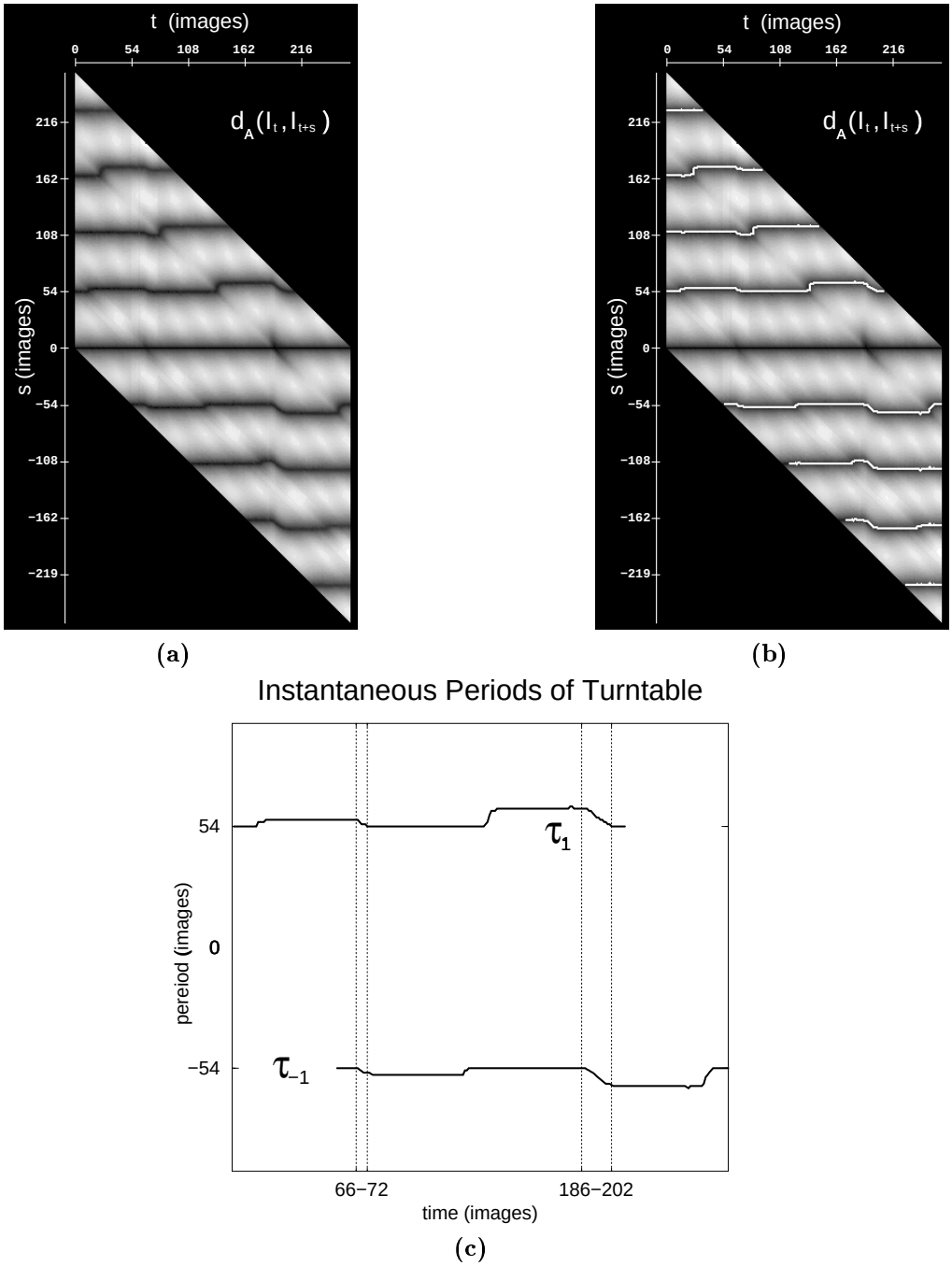


Fig. 6. Period trace of the turntable motion sequence. (a) Temporal correlation plot of $d_{\mathcal{A}}(I_t, I_{t+s})$ with values shown as intensity (white – low correlation, black – high correlation). (b) The recovered period trace (white) is superimposed on the correlation plot. (c) A graph of τ_1 and τ_{-1} with the two most irregular intervals marked. Notice that discontinuities appear in various places but "line up" only at actual motion irregularities. The ground truth turntable frequency is $33\frac{1}{3}$ revolutions per minute, or 54 frames per revolution at NTSC video rate.

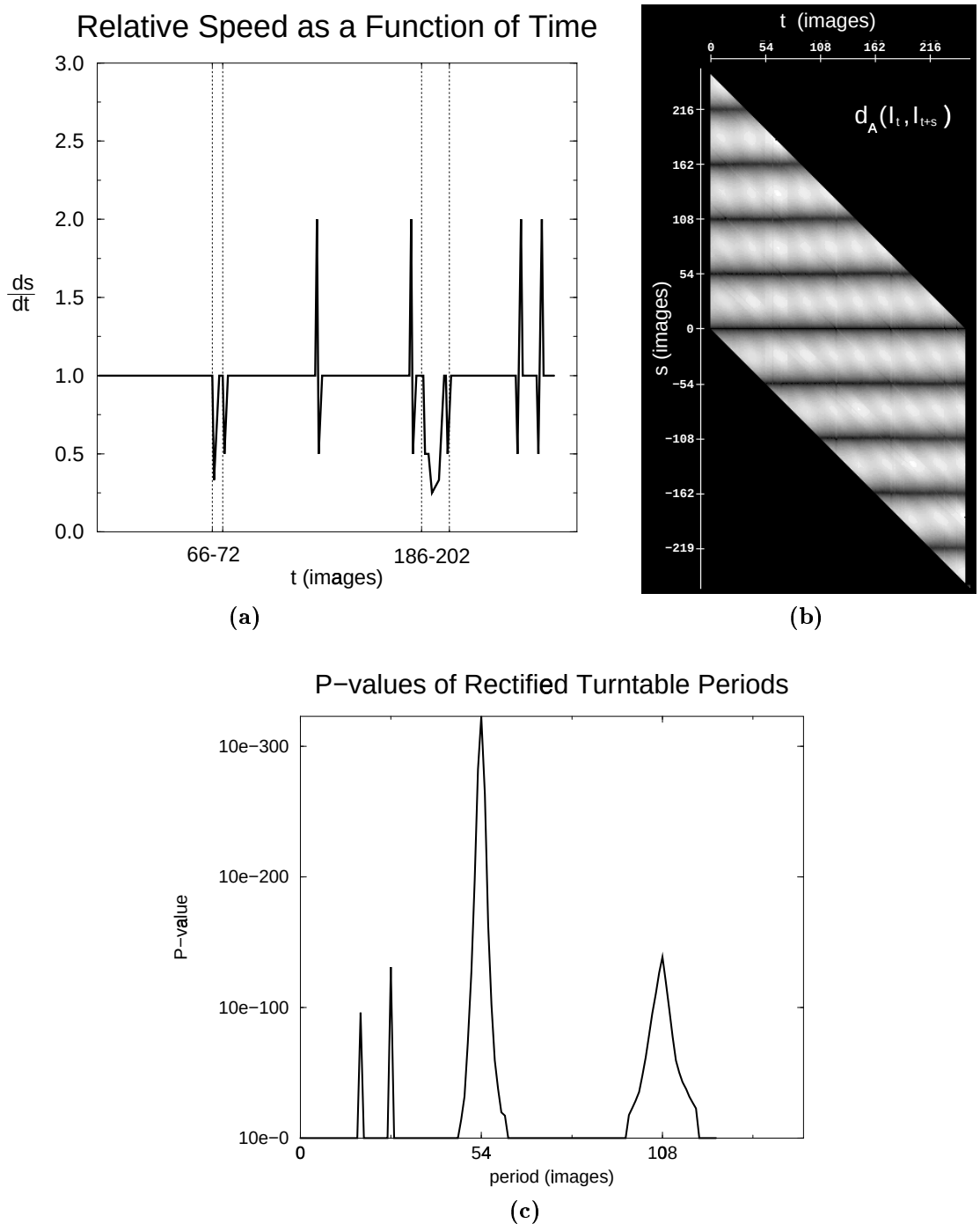


Fig. 7. Image sequence rectification using the period trace. The turntable sequence was rectified, producing a periodic sequence with irregularities automatically removed. (a) Speed relative to the first cycle. Touching the turntable slowed the rotational speed, producing momentary reductions in $\frac{ds}{dt}$. Other fluctuations are present but are not significant since the local variations in speed cancel. (b) Temporal correlation plot of the rectified sequence. Observe that the period trace (dark lines) consists of constant functions, indicative of a periodic motion. (c) Period P-values for the rectified sequence. The period with smallest P-value is 54, exactly matching the ground truth of $33\frac{1}{3}$ rpm at NTSC video rate.

cide only in the intervals where the turntable was touched. Image 66 in Fig. 3 is an image from a detected irregular interval, showing the turntable being touched.

The rectification procedure of Section 4.4 was applied to the image sequence, producing a periodic image sequence (see Fig. 7) comprised of a subset of the images in the original turntable sequence. Notice that the period trace has been *flattened* and no longer contains signs of irregularities. The first complete cycle in the image sequence was automatically detected and used as the reference cycle. In the rectified sequence, the turntable rotates at a constant rate of 54 frames per revolution, exactly matching the ground truth rate of $33\frac{1}{3}$ rpm at NTSC video rate. The turntable sequence was found to be cyclic with a P-value of 10^{-229}. The rectified sequence was determined periodic with a P-value smaller than 10^{-324}, the smallest number representable with an 8-byte double float.

The graph at left in Fig. 7 shows instantaneous speed relative to the reference cycle. The slowing of the turntable is reflected by a reduction in speed in the two intervals in which the turntable was touched. Notice that there are four other momentary fluctuations in speed. These fluctuations, however, are not significant because the local changes in speed cancel out. For example, the rotational speed is twice as fast as normal in frame 126, and half as fast as normal in frame 127. These reciprocal variations exactly cancel each other out, producing no net change in cycle length. This characteristic pattern is common in graphs of this type and is indicative of a small error in the recovered period trace rather than a motion irregularity in the scene. In contrast, motion irregularities generally leave variations that do not cancel, resulting in a net increase or decrease in cycle length during the time in which the irregularity occurred (e.g., frames 66-72 and 186-202 in Fig. 7).

7.3. Heart X-ray Motion Sequence

We also tested our recovery method on a beating heart motion using a sequence of X-ray images obtained through coronary angiography. A few images from the sequence are shown in Fig. 3. Because of the difficulty of automatically segmenting these images, we chose to match the raw data directly, using the grayscale image intensities as the underlying spatial representation. The match function was just the sum of absolute pixel intensity differences between image pairs:

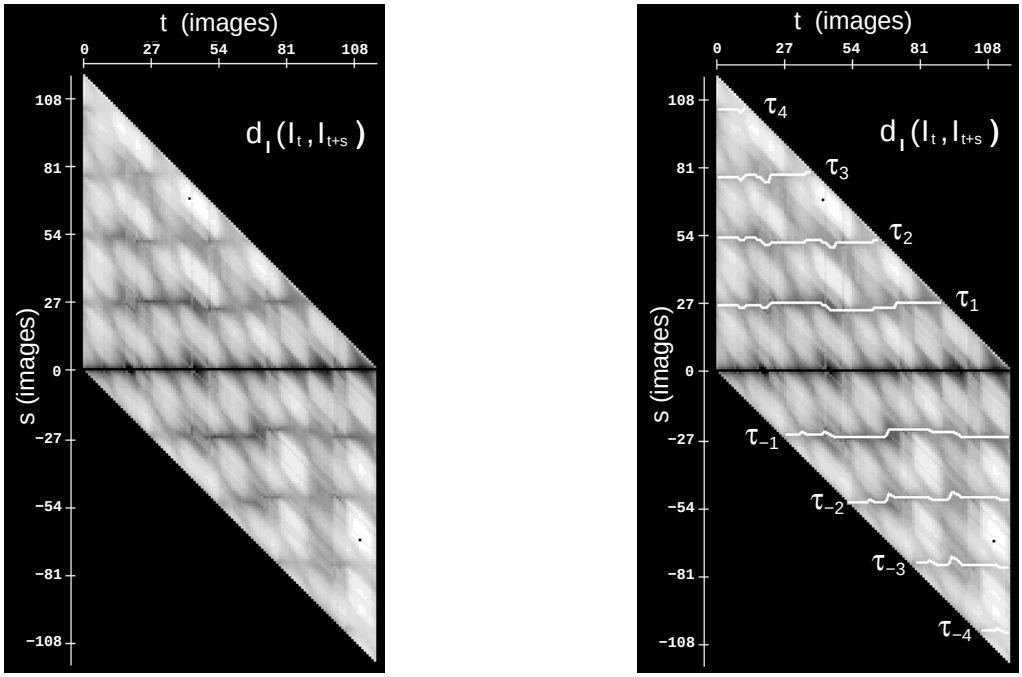
$$d_I(\mathbf{I}_s, \mathbf{I}_t) = \|\mathbf{I}_t - \mathbf{I}_s\|_1$$

Due to the poor quality of the images, the pattern of dark contours in the temporal correlation plot is weaker than in the turntable plot. It was distinct enough, however, for the algorithm to find a period trace with a P-value of 10^{-51}, indicating that the motion is cyclic with very high probability.

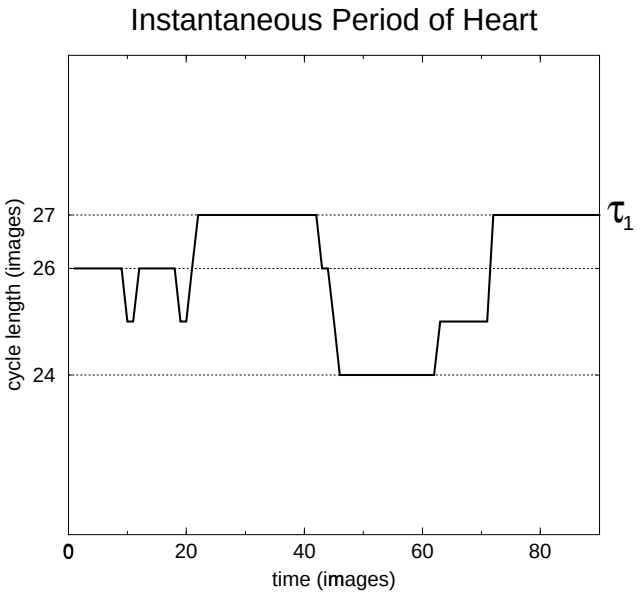
The τ_1 curve shown in Fig. 8 is qualitatively similar to an electrocardiogram (ECG), which can also be used to detect heart periods, although different information is conveyed. An ECG measures electric activity of the heart so period information is not explicit, in contrast to the period trace which gives cycle length as a function of time.

Fig. 8 (c) shows that the period changed from 26 frames to 27 to 24 and back to 27, but remained roughly constant for the duration of each cycle. A period trace of this form is indicative of a motion with isolated irregularities. In addition, observe that the ramps in the period trace (b) all tend to line up, indicating that irregularities occurred in roughly the same place in each cycle. The locations of these ramps coincide with intervals in which the heart is at rest, indicating that irregularities occurred between contractions as the heart relaxed, pausing for slightly different lengths of time in different cycles.

In addition to temporal irregularities, real motions can contain *spatial* irregularities where the set of physical configurations changes somewhat from one cycle to the next. Although the period trace is a purely temporal representation, it can be used in conjunction with other methods to detect spatial variations between cycles. For example, in order to visualize spatial changes in the heart sequence, we generated a sequence of difference images comparing corresponding frames from different cycles. Fig. 9 shows sets of these difference images corresponding to one interval in which the heart is undergoing rapid spatial changes and another in which it is nearly stationary. Pixels



(a) **(b)**



(c)

Fig. 8. Period trace of the heart X-ray motion sequence. (a) Temporal correlation plot of the image sequence. (b) Period trace superimposed on correlation plot. (c) Enlargement of the instantaneous period. Notice that the instantaneous period is roughly piecewise-constant, corresponding to the predicted model of a motion with isolated irregularities. Qualitatively, the motion is seen to be relatively even except in certain short intervals (near frames 20, 40, and 70), suggesting that the period variation is due to uneven motion solely in these short intervals.

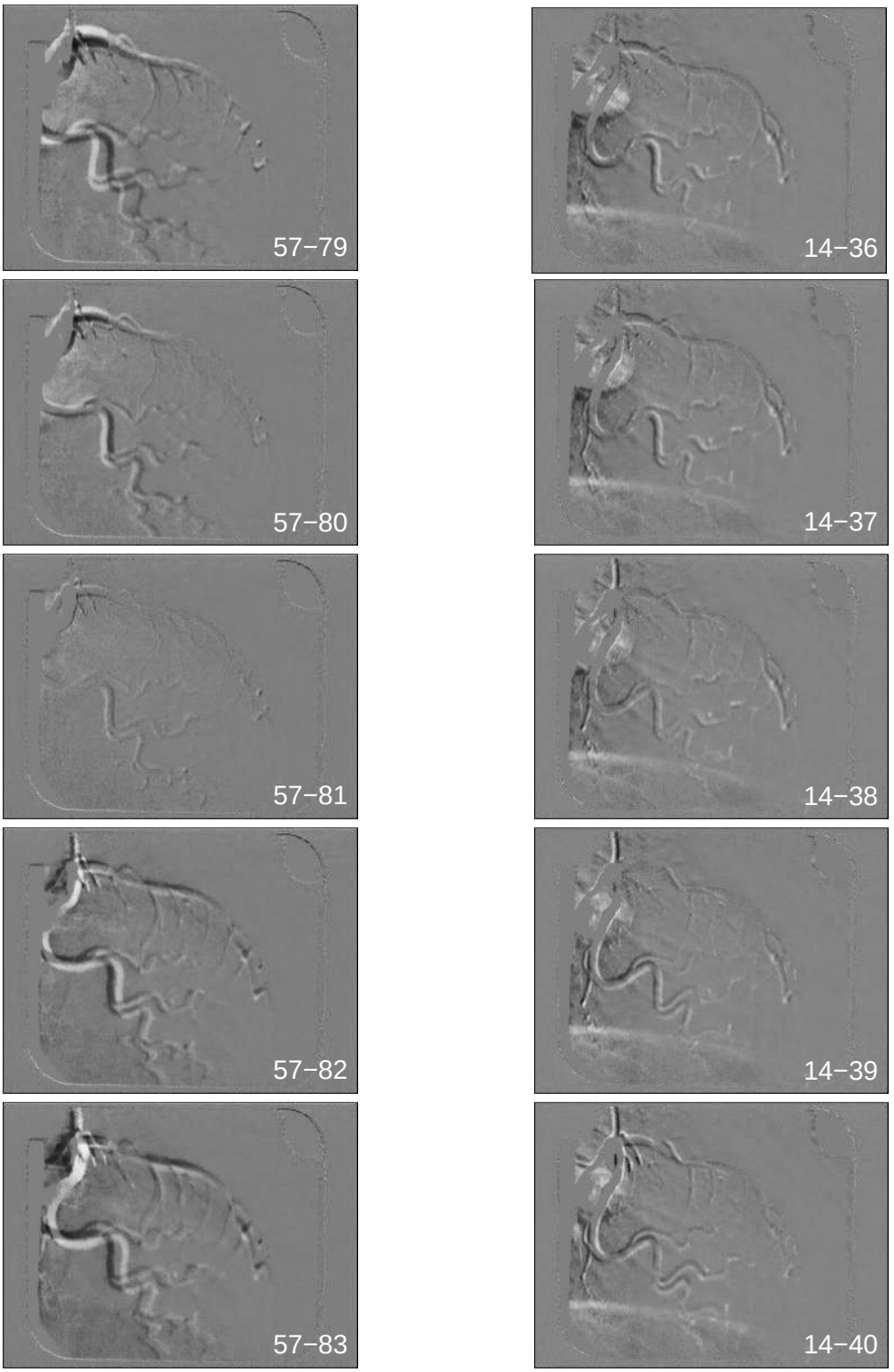


Fig. 9. Visualizing spatial variations in the heart X-ray sequence. Difference images comparing corresponding frames from different cycles where a heart is moving quickly (left) and slowly (right). Frames 57 and 81 correspond, as do frames 14 and 38. Similar pixels appear as neutral gray. Note that changes are much more rapid at left than at right. Light and dark regions in image 14-38 reveal significant spatial differences, indicating that the heart passed through slightly different configurations in the two cycles.

with higher (lower) intensity in the first image appear lighter (darker) in the difference image. Similar pixels appear as neutral gray. These difference images provide simultaneous visualization of both instantaneous velocity and spatial irregularities between cycles. For instance, the left column of Fig. 9 shows that the configuration of the heart is changing very rapidly relative to the frame rate. However, the overall match is visually quite good between images 57 and 81, indicating that the heart is passing through nearly the same configuration at this point in the two cycles. In contrast, the difference images in the right column of Fig. 9 have similar gray-values almost everywhere, indicating an interval where the heart is nearly stationary. Images 14 and 38 yielded the locally best match score, yet the difference image shows several areas of disparity, revealing that the heart paused in slightly different configurations in the two cycles. The patterns of light and dark regions provide a way of visualizing spatial irregularities between cycles. Optical flow techniques could be used to gain a more quantitative measure of these changes, provided that the changes are small. These images illustrate that the period trace can be used in conjunction with other methods to visualize spatial as well as temporal variations in cyclic motions.

8. Conclusion

This paper presented two main contributions to the field of motion analysis. First, we introduced a novel motion representation, called the *period trace*, that provides a complete description of temporal variations in a cyclic motion, and can be used to detect motion trends and irregularities. Unlike previous motion representations, the period trace is purely temporal, describing the evolution of an object or scene without reference to spatial quantities such as position or velocity. Second, we derived necessary and sufficient conditions for determining whether an uncalibrated image sequence is the projection of a 3D cyclic motion in the scene, using principles of affine invariance. In contrast to most previous work on visual invariance, however, our approach is applicable to objects and scenes that undergo *non-rigid* cyclic motions.

Our approach makes use of a novel affine-invariant match function, $d_{\mathcal{A}}$, that permits view-invariant image comparisons, enabling analysis of images taken with a moving camera. Because $d_{\mathcal{A}}$ is invariant to 3D affine transformations of the scene, the technique is also applicable to motions, like an athlete running, that are not strictly cyclic but have a cyclic component. The algorithm is robust with respect to measurement errors and performs well for a wide range of image sequences.

Acknowledgements

The authors would like to thank Dr. Michael Van Lysel at the University of Wisconsin–Madison Medical School for providing the X-ray images used in this paper. The support of the National Science Foundation under Grant Nos. IRI-9220782 and CDA-9222948 is gratefully acknowledged.

Notes

1. Although time is implicit, t could represent another quantity.
2. Others have referred to this decomposition in terms of "relative" and "common" motion.
3. We have found *median* to be more robust than *mean* in this context.
4. Optimization based directly on the K-S test leads to numerical problems, since P-values can be exceedingly small close to the true period trace.
5. This modification poses minor invertibility problems for $\phi^n_{\tau_1}$.

References

M. Allmen and C. R. Dyer. Cyclic motion detection using spatiotemporal surfaces and curves. In *Proc. 10th Int. Conf. on Pattern Recognition*, pages 365–370, 1990.

H. H. Baker and R. C. Bolles. Generalizing epipolar-plane image analysis on the spatiotemporal surface. *Int. J. of Computer Vision*, 3(1):33–49, 1989.

A. M. Baumberg and D. C. Hogg. An efficient method for contour tracking using active shape models. In *Proc. Workshop on Motion of Non-Rigid and Articulated Objects*, pages 194–199, 1994.

A. F. Bobick and A. D. Wilson. A state-based technique for the summarization and recognition of gesture. In *Proc. Fifth Int. Conf. on Computer Vision*, pages 382–388, 1995.

C. Cedras and M. Shah. Motion-based recognition: A survey. *Image and Vision Computing*, 13(2):129–155, 1995.

T. Darrell and A. Pentland. Space-time gestures. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 335–340, 1993.

J. W. Davis and M. Shah. Visual gesture recognition. *IEE Proc. Vision, Image and Signal Processing*, 141(2):101–106, 1994.

I. A. Essa and A. Pentland. A vision system for observing and extracting facial action parameters. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 76–83, 1994.

N. H. Goddard. The interpretation of visual motion: Recognizing moving light displays. In *Proc. IEEE Workshop on Visual Motion*, pages 212–220, 1989.

K. Gould and M. Shah. The trajectory primal sketch: A multi-scale scheme for representing motion characteristics. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 79–85, 1989.

D. Hogg. Model-based vision: A program to see a walking person. *Image and Vision Computing*, 1(1):5–20, 1983.

D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge. Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Analysis and Machine Intell.*, 15(9):850–863, 1993.

G. Johansson. Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14(2):201–211, 1973.

M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. of Computer Vision*, 1(4):321–331, 1988.

J. J. Koenderink and A. J. van Doorn. Affine structure from motion. *J. Opt. Soc. Am. A*, 8:377–385, 1991.

D. P. McReynolds and D. G. Lowe. Rigidity checking of 3D point correspondences under perspective projection. In *Proc. Fifth Int. Conf. on Computer Vision*, pages 945–950, 1995.

J. L. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision*. MIT Press, Cambridge, MA, 1992.

T. S. Perry. Biomechanically engineered athletes. *IEEE Spectrum*, pages 43–44, April 1990.

R. Polana and R. Nelson. Recognition of motion from temporal texture. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 129–134, 1992.

R. Polana and R. Nelson. Detecting activities. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 2–7, 1993.

W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, Cambridge, MA, 1988.

K. Rohr. Incremental recognition of pedestrians from image sequences. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 8–13, 1993.

S. M. Seitz and C. R. Dyer. Affine invariant detection of periodic motion. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 970–975, 1994.

S. M. Seitz and C. R. Dyer. Detecting irregularities in cyclic motion. In *Proc. Workshop on Motion of Non-Rigid and Articulated Objects*, pages 178–185, 1994.

S. M. Seitz and C. R. Dyer. Cyclic motion analysis using the period trace. In *Motion-Based Recognition (M. Shah and R. Jain, Eds.)*, pages 61–85. Kluwer Academic Publishers, Boston, MA, 1997.

M. Shah and R. Jain. *Motion-Based Recognition*. Kluwer Academic Publishers, Boston, MA, 1997.

L. Shapiro, A. Zisserman, and M. Brady. Motion from point matches using affine epipolar geometry. In *Proc. Third European Conf. on Computer Vision*, pages 73–84, 1994.

S. Soatto and P. Perona. Three dimensional transparent structure segmentation and multiple 3D motion estimation from monocular perspective image sequences. In *Proc. Workshop on Motion of Non-Rigid and Articulated Objects*, pages 228–235, 1994.

G. W. Stewart. *Introduction to Matrix Computations*. Academic Press, New York, NY, 1973.

D. Terzopoulos, A. Witkin, and M. Kass. Constraints on deformable models: Recovering 3D shape and nonrigid motion. In *Proc. AAAI Conf.*, pages 91–122, 1988.

C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: A factorization method. *Int. J. of Computer Vision*, 9(2):137–154, 1992.

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

D. J. Williams and M. Shah. A fast algorithm for active contours and curvature information. *CVGIP: Image Understanding*, 1(55):14–26, 1992.

Y. Yacoob and L. Davis. Computing spatio-temporal representations of human faces. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 70–75, 1994.