The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

View-Invariant Analysis of Periodic Motion

Evan Ribnick and Nikolaos Papanikolopoulos

Abstract— **Periodicity has been recognized as an important cue for tasks like activity recognition and gait analysis. However, most existing techniques analyze periodic motions only in image coordinates, making them very dependent on the viewing angle. In this paper we propose a new technique for reconstructing periodic point trajectories in 3D given only their apparent trajectories in image coordinates from a single stationary camera. We show that this reconstruction is possible without performing a costly gradient descent-type optimization, and is based only on a single SVD. This new algorithm is shown to accurately reconstruct natural human motions, allowing them to be compared in 3D world coordinates, independent of the angle from which they were originally viewed.**

I. INTRODUCTION

Periodic motion is quite common in our everyday experience, and one of the most frequent and interesting examples arises from natural human motions such as walking and running. It has been recognized as an important cue in the literature by researchers interested in various areas, such as activity recognition and gait analysis, among others. However, since monocular systems are far more commonplace than multi-camera observations, existing techniques for analyzing periodicities are largely image-based. This typically implies a lack of view-invariance, since the same motion can have a drastically different appearance in the image if viewed from a different angle.

Because they are not view-invariant, motion/action classifiers or mappings that are learned solely from image coordinate observations typically do not generalize well to different cameras or viewing angles. In general, a new classifier or mapping must be learned separately for every viewing angle (e.g., [7]). As such, one way to circumvent this need is to first infer characteristics of the motion in 3D, and then any subsequent analysis is performed in world coordinates, independent of the viewing angle from which the motion was originally imaged. This allows one to develop general classifiers or mappings, which theoretically can be used for any captured sequence. Of course, in the case of any arbitrary motion it may not be possible to reconstruct the object's trajectory in 3D given only its appearance in image coordinates. However, when more is known about the motion of the object in the world, this additional information can sometimes be used to adequately constrain the reconstruction problem.

E. Ribnick is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN ribnick@cs.umn.edu

N. Papanikolopoulos is with the Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN npapas@cs.umn.edu

In this paper we explore the idea of reconstructing periodic point trajectories in 3D given only their appearances in image coordinates, along with the known physical constraints on a motion that is periodic in the world. Broadly speaking, the goals of this paper are twofold: (i) to show that such reconstructions are both possible to obtain and accurate in realistic settings, and (ii) that they contain useful information. We demonstrate that this information can be used to analyze and compare natural human motions, filmed from completely different viewing angles. The focus here is on human motion, because it is the type of periodic motion for which the richest set of applications exist. In addition, it is more challenging to reconstruct and analyze than mechanical motions, since it often includes significant deviations from pure periodicity.

We demonstrated in [15] that it is possible in most cases to reconstruct periodic motions in 3D. That work was the first published attempt to perform reconstruction of periodic motion (to the best of our knowledge), which represents a significant proof-of-concept, but it left several questions unanswered. Here we present a new formulation of the periodic motion reconstruction problem, and attempt to address several of the remaining issues. Some of the advantages of this formulation are:

- we assume only that the observed motion is periodic in the world, and do not require any additional constraints for reconstruction,
- estimation does not require a costly gradient descent-type optimization, but only a single SVD,
- extrinsic camera parameters do not need to be known, and
- reconstruction can be performed from observation of only two periods of motion.

Additionally, the formulation presented here has an intuitive multi-view geometry analogy, from which significant insight can be obtained regarding the nature of the problem and the information that is required to solve it. Our formulation is presented in the context of a more complete system, in which points of interest are tracked automatically, and a novel heuristic for automatically estimating the period of motion is presented. It is illustrated that accurate and useful reconstructions are possible to obtain in a very automated fashion, and that they are robust to the noise and errors inherent in real scenarios.

II. RELATED WORK

As mentioned above, we demonstrated in [15] that it is possible in most cases to reconstruct periodic motions in 3D given their image trajectories. That formulation is based on minimizing the squared reprojection error, and requires

978-1-4244-3804-4/09/$25.00 ©2009 IEEE

one additional constraint in order to perform the estimation. It requires that at least three periods of motion must be observed in order to perform reconstruction, and tracking and period estimation are performed manually.

Belongie and Wills [2] propose a technique for estimating the structure of an object that is undergoing periodic motion. They consider snapshots of the object separated by one period in time, and treat them as multiple views of the same object. These multiple views are used to perform geometric inference using techniques from multiview geometry. Note that this is fundamentally differently than what we aim to achieve in this paper, since our goal is to estimate the trajectory of an object in 3D, and not the structure of the object itself. In addition, the technique in [2] only makes use of one image per period of motion for their purposes, while we use samples from every image during the observed trajectory in order to obtain complete motion information.

Other work related to periodic motion has focussed mainly on detection and analysis in image coordinates, and does not explicitly consider the 3D information that is embedded in the periodicity. In some cases, the object of interest is first tracked, and translational motion removed before processing. Several techniques (e.g., [3], [5], [11]–[14]) use some type of Fourier analysis of pixel intensities or appearances to detect, segment, and represent periodic motions. Frequency-domain techniques have also been used in the activity classification and recognition tasks [8], [10], [17].

Some of the other existing work does not rely on Fourier techniques to detect and analyze periodic motion. Seitz and Dyer [16] present a framework for analyzing cyclic motions that deviate from pure periodicity. In [1], cyclic motions are detected as repetitions on surfaces carved out by a moving object in xyt space. Periodicity has also been used to recognize and classify human gestures [4], human facial expressions, and bird movements [6].

III. Periodic Motion Reconstruction

We now describe in more detail the problem we are trying to solve, and introduce the mathematical formulation. For now it is assumed that a point of interest has been tracked in image coordinates, and that this point corresponds to a point in the world on an object that is undergoing periodic motion with a known temporal period. For example, in this paper we consider points on the hands and feet of moving people. Later we will describe how the track and period estimate are obtained automatically.

Periodic motion here is defined as any movement that is periodic in velocity (in 3D world coordinates):

$$\mathbf{v}(t+nT) = \mathbf{v}(t), \tag{1}$$

for any integer n, where $\mathbf{v} \triangleq (\dot{X}, \dot{Y}, \dot{Z})$ and T is the period. In terms of the 3D position of the point, we have:

$$\mathbf{p}(t+T) = \mathbf{p}(t) + \boldsymbol{\Delta}_{\mathbf{p}_T}, \tag{2}$$

where $\boldsymbol{\Delta}_{\mathbf{p}_T} \triangleq (\Delta_{X_T}, \Delta_{Y_T}, \Delta_{Z_T})$ is the displacement per period of the point, which is constant over any period of length T.

Since samples are taken at discrete times determined by the video frame rate, we represent times using discrete indices of the form t_k^i. This represents the time of the k-th sample in the i-th period. From equation (2) we can then arrive at the expression:

$$\begin{pmatrix} X_k^i \\ Y_k^i \\ Z_k^i \end{pmatrix} = \begin{pmatrix} X_k^0 \\ Y_k^0 \\ Z_k^0 \end{pmatrix} + i \begin{pmatrix} \Delta_{X_T} \\ \Delta_{Y_T} \\ \Delta_{Z_T} \end{pmatrix}, \tag{3}$$

for any integers k and i, where t_k^0 is the time of sample k in the zeroth period.

In order to see how samples from the periodic trajectory are projected into image coordinates, we rely on the pin-hole camera model to obtain equations for the image point (u_k^i, v_k^i) in pixel coordinates:

$$\begin{pmatrix} u_k^i \\ v_k^i \end{pmatrix} = \frac{1}{Z_k^i} \begin{pmatrix} f_x & 0 \\ 0 & f_y \end{pmatrix} \begin{pmatrix} X_k^i \\ Y_k^i \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix}, \tag{4}$$

where we have placed the origin of the world coordinate system at the camera center, with the Z-axis parallel to the camera's optical axis. The quantities f_x, f_y, c_x, and c_y are intrinsic parameters of the camera representing the focal length and image plane center in pixel units. Equation (4) can be expanded using (3) as follows:

$$\begin{pmatrix} u_k^i \\ v_k^i \end{pmatrix} = \frac{1}{Z_k^0 + i\Delta_{Z_T}} \begin{pmatrix} f_x & 0 \\ 0 & f_y \end{pmatrix} \begin{pmatrix} X_k^0 + i\Delta_{X_T} \\ Y_k^0 + i\Delta_{Y_T} \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix}. \tag{5}$$

Thus, we see that it is possible to express the projection of any sample projected into image coordinates, (u_k^i, v_k^i), as a function of the corresponding 3D point in the zeroth period (X_k^0, Y_k^0, Z_k^0) and the inter-period displacement $(\Delta_{X_T}, \Delta_{Y_T}, \Delta_{Z_T})$.

A. Estimation

Rearranging the terms in Equation (5), we can obtain expressions for X_k^0 and Y_k^0 in terms of estimates of Z_k^0 and $(\Delta_{X_T}, \Delta_{Y_T}, \Delta_{Z_T})$:

$${}^{i}\hat{X}_k^0 = \frac{u_k^i - c_x}{f_x}(\hat{Z}_k^0 + i\hat{\Delta}_{Z_T}) - i\hat{\Delta}_{X_T} \tag{6}$$

$${}^{i}\hat{Y}_k^0 = \frac{v_k^i - c_y}{f_y}(\hat{Z}_k^0 + i\hat{\Delta}_{Z_T}) - i\hat{\Delta}_{Y_T}, \tag{7}$$

where " $\hat{}$ " denotes that a quantity is an estimate, and ${}^{i}\hat{X}_k^0$ and ${}^{i}\hat{Y}_k^0$ are approximations of X_k^0 and Y_k^0 based on the estimates and the image-coordinate samples of period i. Such equations can be formed for each sample $k = 0, 1, ...N-1$ and each period $i = 0, 1, ...M-1$. Note that, from these relations, it is clear that the parameters we wish to estimate, Z_k^0, $k = 0, 1, ...N-1$ and $(\Delta_{X_T}, \Delta_{Y_T}, \Delta_{Z_T})$, completely characterize the trajectory of the point in 3D.

Ideally ${}^{i_1}\hat{X}_k^0 = {}^{i_2}\hat{X}_k^0$ and ${}^{i_1}\hat{Y}_k^0 = {}^{i_2}\hat{Y}_k^0$ for any sample k and any pair of periods i_1 and i_2. Therefore, making use of (6) and (7), we can obtain a pair of equations as follows:

$$\begin{aligned} {}^{i_1}\hat{X}_k^0 - {}^{i_2}\hat{X}_k^0 = \frac{u_k^{i_1} - u_k^{i_2}}{f_x}\hat{Z}_k^0 + (i_2 - i_1)\hat{\Delta}_{X_T} \\ + \frac{i_1 u_k^{i_1} - i_2 u_k^{i_2} + c_x(i_2 - i_1)}{f_x}\hat{\Delta}_{Z_T} = 0 \end{aligned} \tag{8}$$

and

$$
{}^{i_1}\hat{Y}_k^0 - {}^{i_2}\hat{Y}_k^0 = \frac{v_k^{i_1} - v_k^{i_2}}{f_y}\hat{Z}_k^0 + (i_2 - i_1)\hat{\Delta}_{Y_T} + \frac{i_1 v_k^{i_1} - i_2 v_k^{i_2} + c_y(i_2 - i_1)}{f_y}\hat{\Delta}_{Z_T} = 0. \quad (9)
$$

Two equations of the form (8) and (9) can be obtained for every sample k, for every pair of periods i_1 and i_2. This results in a total of $2N\begin{pmatrix}M\\2\end{pmatrix}$, where M is the number of periods, and N is the number of samples from each period. If we stack all these equations together in matrix form, the result is the overconstrained homogeneous linear system:

$$
A\mathbf{X} = \mathbf{0}, \quad (10)
$$

where $A \in \mathbb{R}^{\left(2N\begin{pmatrix}M\\2\end{pmatrix}\right)\times(N+3)}$ is the coefficient matrix, and $\mathbf{X} \in \mathbb{R}^{N+3}$ is the vector of the parameters we wish to estimate:

$$
\mathbf{X} = \begin{pmatrix}\hat{Z}_0^0 & \hat{Z}_1^0 & \hat{Z}_2^0 & \dots & \hat{Z}_{N-1}^0 & \hat{\Delta}_{X_T} & \hat{\Delta}_{Y_T} & \hat{\Delta}_{Z_T}\end{pmatrix}^T. \quad (11)
$$

Since the system (10) is overconstrained, the estimation is cast as a homogeneous linear least-squares problem in which we aim to minimize $\|A\mathbf{X}\|_2$ subject to $\|\mathbf{X}\|_2 = 1$. Note that ideally the solution $\mathbf{X}$ must lie in the nullspace of A, and that the nullity of A is one. This implies that it is possible to obtain a solution only up to a scaling factor. We will demonstrate later that this is sufficient for many activity analysis tasks. This minimization can be performed efficiently using one Singular Value Decomposition (SVD), where $A = U\Sigma V^T$, and the minimizer $\mathbf{X}^*$ is the last column of V [9].

Finally, once the parameters $\mathbf{X}$ are estimated, the trajectory can be reconstructed using (6) and (7). We estimate X_k^0 by taking the mean of the estimates ${}^{i}\hat{X}_k^0$ over all periods i, and similarly for Y_k^0, for all samples k. Points from subsequent periods can then easily be reconstructed from the estimates of (X_k^0, Y_k^0, Z_k^0) using the relation (3), $k = 1, 2, ..., N-1$.

B. Multi-View Geometry Analogy

An interesting observation is that the problem described above is mathematically analogous to multi-view reconstruction. Since it is assumed that each period of the trajectory is identical in world coordinates, observation of M periods from a single camera is mathematically equivalent to observing a single period from M cameras. Additionally, this implies that the motion between these "virtual cameras" is purely translational, and that the translation from camera i to camera $i+1$ is $(-\Delta_{X_T}, -\Delta_{Y_T}, -\Delta_{Z_T})$. To see this, simply reinterpret all equations presented thus far, taking k to be the index of a sample in the one observed period, and i to be the index of the virtual camera view. Under this analogy, equation (5) can be interpreted as expressing the pixel coordinates of sample k in the i-th camera view, as a function of the 3D position of sample k in the coordinate frame fixed at camera $i = 0$, and the translation between the camera views.

Recognizing this equivalence allows us to understand the current problem in a quite intuitive manner, and to draw some simple conclusions regarding its solvability based on well-known results from multi-view geometry. First, we point out that when the motion is periodic in *position*, this implies that the translation between the virtual cameras, $-(\Delta_{X_T}, \Delta_{Y_T}, \Delta_{Z_T})$, is zero. In this case, it is not possible to reconstruct the periodic trajectory, since all of the virtual cameras are collocated. Indeed, it is possible to arrive at the same conclusion by observing that the system of equations of the form (8) and (9) becomes ill-posed when the inter-period displacement goes to zero.

Second, the mathematical equivalence between our problem and multi-view reconstruction also assures us that only two periods of motion need to be observed in order to estimate the trajectory in 3D. This is equivalent to binocular stereo, where we have two virtual cameras observing a single period of motion.

Finally, since the motion between these virtual cameras is known to be purely translational, it is guaranteed that the only other case in which the problem becomes degenerate is when the 3D samples in the first period of motion happen to be coplanar with the optical centers of both virtual cameras [9]. This can only happen in the unlikely case when all points on the periodic trajectory lie in a plane, and the camera views the trajectory from a point on this plane.

C. Tracking Points of Interest

Until now it has been assumed that the trajectory of the point in image coordinates is already known. However, in the experiments here we have used an automatic tracker in order to demonstrate that reconstruction of periodic motion is still possible when the image coordinate samples contain significant noise. Note that any arbitrary algorithm which can accurately track a point of interest in image coordinates could be used, and that the choice of tracking algorithm is purely one of convenience based on the particular application.

For the results presented here we have used a manually initialized marker-based tracker, in which the subject wears a brightly colored patch that can be easily identified in the image. In most cases, it is attached near the subject's ankle. Colors are analyzed in normalized RGB space, which is known to be illumination-invariant for Lambertian surfaces. A simple Bayesian classifier is learned from the manual initialization, in which the marker color is modeled as a multivariate Gaussian in normalized RGB space. This classifier is then used to identify the colored marker in subsequent images.

D. Period Estimation

We also propose a new heuristic for estimating the period of motion (or equivalently, the dominant frequency), which has been found to perform reliably on real data. A key observation here is that, even though the periodic trajectory has been warped by the perspective projection into image coordinates (and as such is no longer periodic in the image), much of the *temporal* frequency content of the original world coordinate signal can still be recovered. This is not always immediately obvious from the appearance of the trajectory in the image.

Once a point undergoing periodic motion has been tracked, its image coordinate trajectory can be thought of as a curve through u-v-t-space, where, u and v are image coordinates, and t is time. If one were to plot the time derivatives of the $u(t)$ and $v(t)$ signals separately, it might be observed that one of these signals appears to be "more periodic" than the other. In fact, these are just two examples of projections of the $\dot{u}$-$\dot{v}$-t trajectory onto planes parallel to the t-axis. We estimate the period by projecting the $\dot{u}$-$\dot{v}$-t trajectory onto several planes parallel to the t-axis at evenly spaced rotations. For each projection, the Fast Fourier Transform (FFT) is computed, and the dominant spectral component is identified. The median of these dominant frequencies over all projections is then taken to be the inverse of the period of motion.

IV. EXPERIMENTS

In order to demonstrate the effectiveness of the proposed technique for reconstructing periodic trajectories in 3D, we perform several experiments with videos of real human motion. We focus on natural human motion because it is quite challenging, in that there can be significant deviations from perfect periodicity. Additionally, human motion analysis is an area of interest in the computer vision community, and has a rich set of potential applications. In the following experiments we attempt to show that the reconstructions are accurate (§IV-C), that they are view-invariant (§IV-A), and that useful information is embedded in them, which can be used to make inferences about the motion of a human subject (§IV-B).

The tracking algorithm outlined in §III-C was used to track points of interest in all experiments, except the one in §IV-C, where we have used an existing dataset for comparison. In all cases where our tracking algorithm was used, the marker was fixed near the subject's ankle. The period estimation technique presented in §III-D was used in all of the following experiments.

A. View-Invariance

In this experiment we analyzed videos of humans walking in a straight line. The videos contained 34 walking sequences from two different people, filmed simultaneously from two very different viewing angles. Each trajectory was reconstructed separately from both views. The purpose of this experiment is to show that the two reconstructions are very similar for each sequence, implying view-invariance.

Since reconstruction is accurate only up to a scaling factor, and the extrinsic camera parameters are unknown, it was necessary to first project the reconstructions onto a standard set of basis vectors in order to compare them. New axes were computed for any given sequence using Principal Component Analysis (PCA) on the reconstructed

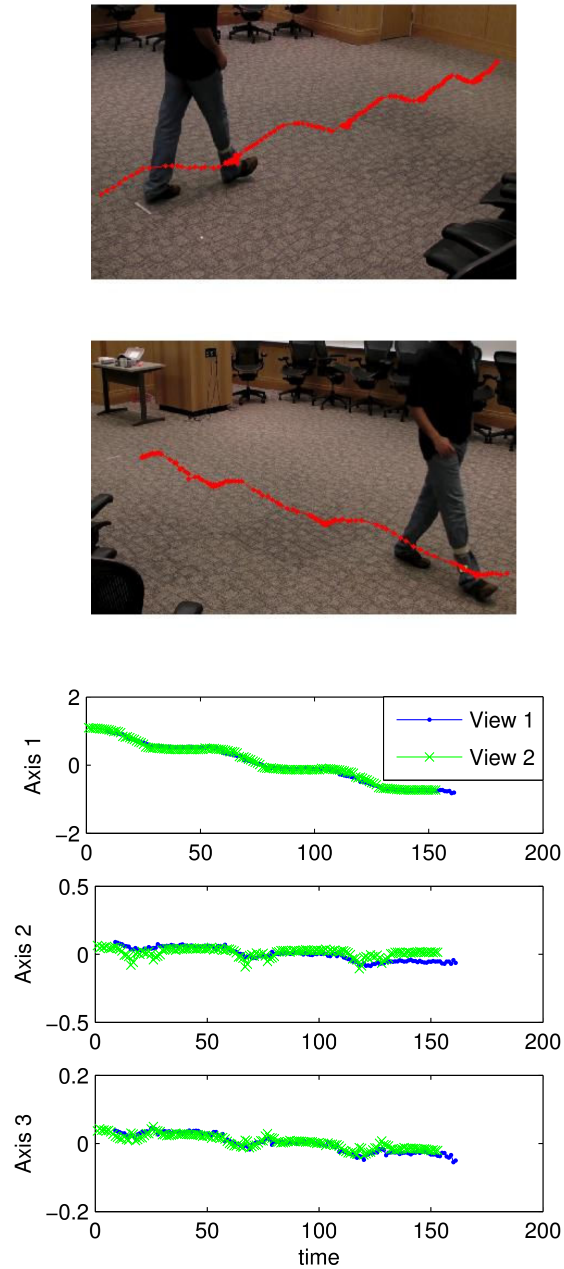

Fig. 1. Example of view-invariant reconstruction. (top) Two views of the same trajectory. (bottom) Both reconstructions plotted together, reprojected onto the axes computed using PCA. Each dimension is plotted against time. Three periods are shown.

3D datapoints from one period of motion. The datapoints were then projected onto the new axes, and scaled. Finally, we compare two reconstructions by aligning the principal components, and computing the mean Euclidean distance between corresponding samples in this new space. Note that the new axes computed with PCA do not necessarily correspond to any standard XYZ axes in the world.

An example of a single walking sequence, reconstructed independently from two views, is shown in Figure 1. Notice that even though the trajectory of the tracked point is the

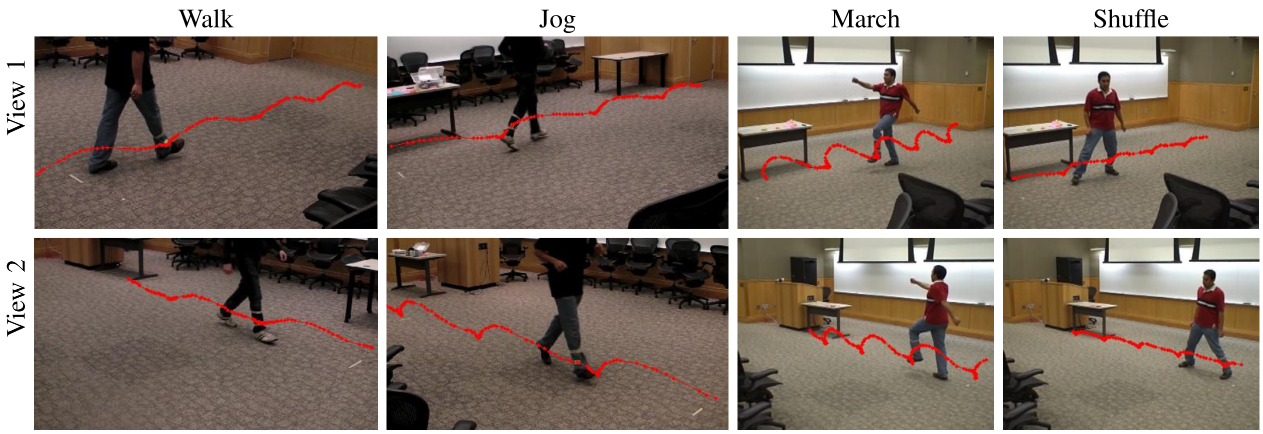

Fig. 2. Examples of the four activities, from two different views.

same in both cases, its appearance in image coordinates changes dramatically with the viewing angle. Figure 1 (bottom) shows plots of the reconstructions from the two views, superimposed. Note that they are plotted on the new axes, as described above, where the first axis is the principal component. Three periods of the reconstructions are shown, plotted against time on the horizontal axis. As can be seen, the two reconstructions match each other very well, even over multiple periods.

Similar analysis was performed for each of the 34 motion sequences, and the distance between each corresponding pair of reconstructions was computed – the so-called inter-view reconstruction error. Since the reconstructions are on an arbitrary scale, the distance between them is computed as a fraction of the total displacement over the three periods. We found that the average inter-view reconstruction error was 8.7% of the total displacement, which shows that the corresponding reconstructions from different views match each other quite well. This will be further explored below.

B. Activity Classification

Reconstruction quality and view-invariance are further examined through the task of activity classification. In this experiment we show that the reconstructed trajectory of a single point on the body (near the ankle in this case) can be used to distinguish between several activities. Here we analyzed videos containing two people performing several instances of the following four activities: walking, jogging, marching, and sideways shuffling. As before, motion sequences were filmed from two very different viewing angles. Examples of these four activities are shown in Figure 2 from two different views.

It is important to note that the aim of this paper is *not* to propose a new method for activity recognition. Rather, we choose the task of activity classification only to demonstrate that it is possible to accurately reconstruct periodic motions in 3D from a single camera, independent of viewing angle, and to show that useful information is embedded in these reconstructions.

The dataset used in this experiment consisted of 32 sequences, including four instances of each activity from each of the two views (see Figure 2 for examples). These sequences contain motions from two different people. We employed an *opposite-view* nearest-neighbor classification scheme, in which a reconstructed trajectory from one view is assigned an activity label based on its nearest-neighbor from the reconstructed trajectories from the other view. Distances between trajectories were computed using the same PCA-based axis-alignment and scaling as described above in §IV-A.

TABLE I

CONFUSION MATRIX FOR OPPOSITE-VIEW ACTIVITY CLASSIFICATION.

		Classification			
		Walk	Jog	March	Shuffle
Actual	Walk	8	0	0	0
	Jog	0	8	0	0
	March	3	0	5	0
	Shuffle	0	0	0	8

Results from the opposite-view activity classification are shown in Table I in the form of a confusion matrix. Overall classification accuracy is 90.6%. Notice that the walking, jogging, and shuffling sequences are classified perfectly. Some of the march sequences were erroneously classified as walks since the motions are quite similar in cases when the marching action is not performed well. Reconstruction is also sensitive to errors in period estimation.

C. Reconstruction Accuracy

So far we have demonstrated that it is possible to reconstruct periodic motions in 3D irrespective of the viewing angle, and that these reconstructions can be used to make inferences about the human subject. In this final experiment, we examine the absolute accuracy of the reconstruction. For the sake of comparison, the data used in this experiment was the same used in [15]. The video contains footage of a walking person. Importantly, we also have the ground-truth trajectory of a point on the person's hand in 3D, which was captured simultaneously with a commercial motion capture system. An image from the sequence is shown in Figure 3, with the track overlaid.

For this experiment our tracking algorithm was not used – rather, tracking information was provided with the dataset. However, our automatic period estimation was used. Reconstruction of the periodic trajectory was performed from a single view and compared with the 3D ground-truth. The reconstruction and ground-truth were matched and error computed as described earlier using PCA, except that in this case we scale the reconstruction to be the same scale as the ground-truth. The aligned reconstructed and ground-truth trajectories are shown in Figure 4 over three periods, on the reprojected axes. Note that in this case, the units of the vertical axes are mm. As can be seen, the reconstruction matches quite closely with the ground-truth trajectory. For this trajectory, the average reconstruction error is $5.92cm$ over three periods. This is quite similar to the reconstruction error reported for this sequence in [15]. Additionally, in our case the error in the stride length estimation (i.e., the magnitude of the inter-period displacement) was $7.9mm$.

V. Conclusions and Future Work

We have presented a new algorithm for reconstructing periodic motions in 3D from a single stationary view, and loosened many of the assumptions of the previous technique.

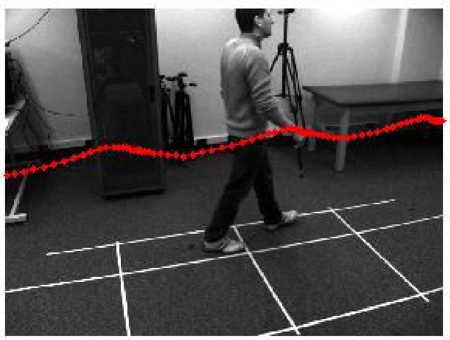

Fig. 3. An image from the walking sequence.

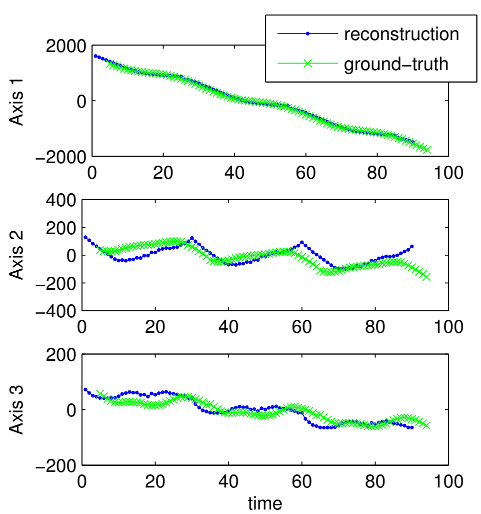

Fig. 4. Reconstruction of hand trajectory, plotted together with ground-truth. Here the vertical axes are on the mm scale. Three periods are shown.

The proposed formulation also has an intuitive multi-view geometry interpretation, from which important conclusions can be drawn. It was demonstrated that these reconstructions are accurate, view-invariant, and contain useful information. We focussed on natural human motion, and illustrated that the reconstructed trajectory of a single point on the body can be used to classify various activities.

There are several aspects that we plan to investigate further in the future. It would be interesting to explore the structure of the manifold on which different periodic human motions reside. Additionally, we plan to develop reconstruction techniques which explicitly model fluctuations in instantaneous frequency.

VI. Acknowledgements

This material is based upon work supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office under contract #911NF-08-1-0463 (Proposal 55111-CI), and the National Science Foundation through grants #IIS-0219863, #CNS-0224363, #CNS-0324864, #CNS-0420836, #IIP-0443945, #IIP-0726109, and #CNS-0708344.

References

[1] M. Allmen and C.R. Dyer. Cyclic motion detection using spatiotemporal surfaces and curves. *Proc. Int'l. Conf. Pattern Recognition*, 1990.

[2] S. Belongie and J. Wills. Structure from periodic motion. *Proc. Int'l. Workshop on Spatial Coherence for Visual Motion Analysis*, 2004.

[3] A. Briassouli and N. Ahuja. Extraction and analysis of multiple periodic motions in video sequences. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 29(7):1244–1261, July 2007.

[4] C.J. Cohen, L. Conway, and D. Koditschek. Dynamical system representation, generation, and recognition of basic oscillatory motion gestures. *Proc. Int'l. Conf. Automatic Face and Gesture Recognition*, 1996.

[5] R. Cutler and L.S. Davis. Robust real-time periodic motion detection, analysis, and applications. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(8):781–796, August 2000.

[6] J. Davis, A. Bobick, and W. Richards. Categorical representation and recognition of oscillatory motion patterns. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2000.

[7] A. Elgammal and C.-S. Lee. Inferring 3d body pose from silhouettes using activity manifold learning. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2004.

[8] H. Fujiyoshi, A.J. Lipton, and T. Kanade. Real-time human motion analysis by image skeletonization. *IEICE Trans. Information and Systems*, E87-D(1):113–120, 2004.

[9] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd ed., 2003.

[10] J. Little and J. Boyd. Describing motion for recognition. *Proc. Int'l. Symposium on Computer Vision*, 1995.

[11] F. Liu and R.W. Picard. Finding periodicity in space and time. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1998.

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

[13] X. Orriols and X. Binefa. Classifying periodic motions in video sequences. *Proc. Int'l. Conf. Image Processing*, 2003.

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

[15] E. Ribnick and N. Papanikolopoulos. Estimating 3d trajectories of periodic motions from stationary monocular views. *Proc. European Conf. Computer Vision*, 2008.

[16] S.M. Seitz and C.R. Dyer. View-invariant analysis of cyclic motion. *Int'l. J. Computer Vision*, 25:1–23, 1997.

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