3D Reconstruction of Periodic Human Walking Trajectories Based on Single View*

ZHANG Jun, HUANG Weiqiang and LUO Tingjin

(Department of Information System and Management, National University of Defense Technology, Changsha 410073, China)

Abstract — Periodicity of a moving body is one of important characteristics in activity monitoring. We present a method to estimate the trajectory of human gait in 3D space from a single camera by exploring the periodicity of the human movement. Geometric constraints are established to characterize the periodic motion, which are invariant under viewing geometry variations. According to our analysis, these geometric constraints reduce the overall computational complexity of evaluating periodicity. Meanwhile, with the help of these geometric constraints, we develop a novel method to reconstruct the 3D motion trajectory from a single camera. Experimental results demonstrate the accuracy and robustness of the algorithm.

Key words — Periodic motion, 3D reconstruction, Virtual view, Gait analysis.

I. Introduction

Gait analysis plays an important role in human activity monitoring\textsuperscript{1–2}. Unlike other popular biometrics based on face recognition or fingerprint analysis which only works in a short distance, gait analysis allows remote extraction of biometric features of the human activities from a relatively long distance. This advantage of gait analysis has been applied in many areas such as security and health monitoring. The first and enabling step in gait analysis is to extract the 3D motion trajectory of the person\textsuperscript{3–5}.

Existing methods based on multiple views for analyzing the human motions are greatly dependent upon the image information from more than one camera monitoring the same object. However, since the mono-camera systems which provide almost no overlap of the mentoring region for each camera are much commonly used in real world than multi-camera systems, the multi-view based techniques are likely to result in a lack of view-invariance since the same motion can have very different appearances in the image if viewing from different angles.

Assuming the periodicity of the human moving motion, it is possible to interpret several periodic movements of the object obtained from a single view and reconstruct the real motion trajectory of the object in 3D that used to be only determined from multiple-view techniques. Ribnick et al.\textsuperscript{6} prove the possibility and applicability of this thought. They derive the equations of the trajectories of the motion and also propose two optimization algorithms to realize the 3D reconstruction. Compared with Ribnick’s algorithms which focus on optimizing the trajectory equations, Troje et al.\textsuperscript{7} focus their work on the periodic movement reconstruction by using gait analysis. They use Fourier method to analyze human movements and acquired a concise expression of motion. A linear model with low dimensions is trained by a group of samples and the data of a 2D image was projected to the model using the least squares method. Belongie et al.\textsuperscript{8} further develop a method of reconstructing the object structure based on the periodic movements. Their method extracts the contour of the object from one image frame for each movement period and acquires the depth information of the object based on the binocular vision. Allmen et al.\textsuperscript{9} use the space partition method to reconstruct the motion. The movements obtained from the different movement periods can be regarded as equivalent to the motion in the same periodic but taken from different view angles on the camera plane. By extending the line between the projecting center and the contour in the imaging plane in each movement period, a cone shape is primarily obtained. Then the basic structures of the object can be determined from the intersections among the cones obtained from several movement periods.

This paper relies on the fact that the problem can be shown to be mathematically equivalent to reconstructing a single object from multiple views under certain conditions. This allows us to draw on the rich body of existing theory in the domain of multi-view geometry and 3D reconstruction. Currently, most algorithms used to analyze the human walking motion are based on the 2D coordinates which are largely view-angle-dependents. The periodic nature of the human movement could be used to reconstruct the trajectory is ignored. In this paper, gait recognition is used as an example. The features of the gait extracted from a single view are and can only be applied to the classifier in accordance to the same angle of view. The reliance on angle of view though limits the application of these features in the gait recognition. This limitation can be avoided if the features of the gait is extracted in a 3D coordinates. It is important that the 3D information of periodic motion can be reconstructed based on the image. The methods based on multi-view techniques can acquire much in-
formation from each view. The correlation among these views is useful to acquire the 3D information.

The focus of this work is not on the species of tracking points of interest—instead, we choose to focus on the problem of reconstructing periodic motions trajectory in 3D using a single view. To extract gait features, we transform the multiple circles obtained from a single view into a single circle in multi-view, and then adopt the existing multi-vision theory to reconstruct 3D human walking trajectory at some potentially useful positions, such as the ankle. It ensures that under the global coordinates, the reconstructed 3D trajectory based on the 2D image sequence is no longer highly dependent on the view angle and also give a good individual gait features. Compared with the 3D reconstruction using binocular vision method, to the proposed method in this paper reduces the redevelopment cost, and yields a better adaptability.

The main development of this work is comprised of the following components: (1) to acquire the camera parameters through camera calibration; (2) identify the periodicity of the movement and calculate the length of the cycle base on the geometric restriction conditions; (3) to construct the virtual view of the object. This process carries out the transformation of the image from single view into stereo vision and would allow the identification of the object location in global coordinate using binocular vision theories.

This paper is organized as follows. Section II formulates the reconstruction problem and proposes a method to estimate the periodic motion parameters from a single camera. Experimental results are presented in Section III. Section IV concludes the paper and discusses future work.

II. The 3D Reconstructing Method of Periodic Movements

Previous works by Belongie et al.[8] have shown that constructing the virtual view is a key to 3D reconstruction of the trajectory using a single view. The stereoscopic vision theory is used to obtain the virtual view. Using the data from a single calibrated camera, the algorithm is divided into the following steps: Firstly, after knowing the geometric constraints, the movement cycles are divided ones and each cycle length is calculated. Secondly, a virtual view is built based on the previous calculated circle length; At last, the stereoscopic vision theory is used to complete 3D reconstruction of the periodic trajectory and extract the gait features of the monitoring object.

1. Periodic motion detection

The movement cycle detection is crucial for 3D reconstruction in single view. There are many ways to identify the movement cycles. Ben Xian-Ye[10] proposed a gait cycle detection method based on the regional characteristics, which is a change in each frame. The signals of periodic motions acquired from this method show a smoothing, strong noise robustness, and also in variance both in scale and translation. Another method is based on the characteristics of the motion curve to solve the cycle length. Ribnick[6] proposed a method based on Fourier analysis to divide the cycle by extracting the main components of the trajectory curve.

In Fig.1(a), the solid curve shows the trajectory of the movement in a single view. Supposing the object to move along the same direction, a virtual camera considered to be shifted along with the object to keep parallel to the object. The shifting distance is the same as the moving period, say ΔL. The periodicity of the movement yields that the imaging of the second period equal to the imaging of the virtual camera (Fig.1(b)) in the first period. In such a way, the movement of the object in two consecutive periods in a single view angle is equivalent to the movement in one period but from two different view angles. The shifting distance between two camera axes is ΔL. The methods of 3D reconstruction can be adopted after the translation is completed.

It is important to acquire the distance ΔL between the virtual camera and the original camera. More constraints need to be added since the distance between two points in the space is not unique in single image. Taking the gait as an example, during the walking process, the positions that heals in touch with the ground can be considered as one of the geometry constraints. The shifting distance ΔL is acquired by confirming the distance of the gait period in the global coordinate XwOwYw.

In Fig.2, the ankle is marked as the object and then the trajectory of ankle is acquired from a single view camera. The circle in each frame represents the position of ankle. The distribution of the circles is not uniform, and the areas with many circles overlaps in correspond to no movement of the left ankle while the right ankle moves.
Consider a position $i$, we can calculate the distance between position $i$ and the center of its adjacent window $C(i+1, i+n)$:

$$s(i) = \sqrt{(x_i - C_x)^2 + (y_i - C_y)^2}$$  \hspace{1cm} (2)

Threshold $t$ is decided by the camera view and sampling interval and $t$ is very small. When $s_i > t$, the sliding window will move backward; but when $s_i < t$, the center of all points in the sliding window will be recorded as $G_j = C(i, i+n-1)$, $(j = 1, 2, \cdots, m)$ and then the sliding window will move backward continued and then $i = i + 1$; this process will not complete until the last point in the sliding window is calculated. As a result, a series of center points $G_1, G_2, \cdots, G_m$ is obtained. We can determine the projection on the image of $\Delta L$ by connecting $G_1, G_2, \cdots, G_m$ in order.

The distance between two points in the 3D space using one 2D image measurement doesn’t normally have a sole solution. However, if we impose additional constraints, the true distance $\Delta L$ in the global coordinate system can be determined. Therefore, consider the positions in $X_w, O_w, Y_w$-plane where the heel touches the ground during walking as the geometric constraints, we can calculate the length of the walking gait circle which is the translation amount $\Delta L$, by the projection on the image of $\Delta L$, coupled with the $X_w, O_w, Y_w$-plane constraints.

2. The camera model and calibration

In this paper, a pin-hole model\cite{11} is used to model the geometry of camera imaging principles, as shown in Fig.3.

![Fig. 3. The geometry of the camera imaging described using pin-hole model](image)

The origin of image coordinates $O_i$ is the point of intersection between principle axis and image plane. The horizontal direction is considered as parallel to $x$-axis and the vertical direction parallel to $y$-axis. The camera coordinates $O_w - X_w, Y_w, Z_w$ is established as follows: the center of camera $O_w$ is considered as the origin of the camera coordinate, $Z_w$ axis is considered as the principle axis, the line parallel to $x$ axis while passing $O_w$ is named as $X_w$ axis and the line parallel to $y$ while passing $O_w$ is named as $Y_w$ axis. The global coordinate $O_w - X_w, Y_w, Z_w$ is used to describe the position of the point and camera in the space.

The coordinates of a point $P$ is noted as $(X_w, Y_w, Z_w)$ in the camera coordinates. The point which is corresponding to the point $P$ in the 2D image plane coordinate is denoted as $p(x, y)$. The relationship of these two coordinates system is given by the following equation:

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x X_c + c_x Z_c \\ f_y Y_c + c_y Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$  \hspace{1cm} (3)

where $(x_c, y_c)$ is the coordinates of image center $O_i$, $f_x$ and $f_y$ are the local lengths along the $x$ and $y$ axis.

The coordinates of point $P$ is noted as $(X_w, Y_w, Z_w)$ in the global coordinate. The relation between $(X_c, Y_c, Z_c)$ and $(X_w, Y_w, Z_w)$ is described in Eq.(4), where $R$ is the rotated matrix and $T$ is the shifted vector.

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = R \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + T$$  \hspace{1cm} (4)

Combining Eq.(3) and Eq.(4) gives:

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = M \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$  \hspace{1cm} (5)

where $M$ is the camera matrix. The camera inner parameters include $f_x$, $f_y$, $c_x$, and $c_y$, while the camera outside parameters include $R$ and $T$.

Eq.(5) combines two classical calibration methods to acquire the parameters of camera. Firstly, we adopt Zhang’s calibration method\cite{12} to calculate the camera inner parameters. Then, we use a stereo calibrated board and adopt the method proposed by Tsai\cite{13} to calculate the camera outside parameters.

3. 3D reconstruction of the trajectory

Using the camera parameters obtained from Section II.2, we can now use the binocular perspective principle to reconstruct the 3D trajectory of the motion.

As shown in Fig.1, using camera matrix of the first calibrated camera $M_1$ in combination with the translation amount $\Delta L$, we can calculate the camera matrix of the second camera $M_2$.

We use center points $G_1, G_2, \cdots, G_m$ to divide the gait trajectory into $n - 1$ cycles; the trajectory of each cycle corresponds to the image in each camera plane. We assume that $p \leftrightarrow p'$ is a pair of corresponding points from two camera views. Using the camera matrices $M_1$ and $M_2$, the reverse ray $l_p$ of $p$ and $l_{p'}$ of $p'$ can be calculated. If the coordinates on the image are error free gives no error, $l_p$ and $l_{p'}$ can be used to determine a plane $\pi_t$ through the two cameras’ optical center. So $l_p$ and $l_{p'}$ have to intercept at a point in space.

Here the reverse ray $l_p$ of $p$ and $l_{p'}$ of $p'$ respectively refer to the collection of all the space points including image point $p$ and $p'$ under the action of the camera $M_1$ and $M_2$. As an example, $p$ intersects at a point in space.

$$l_p = \{ P | p = M_1 P \}$$  \hspace{1cm} (6)
Geometrically, it can be seen that the reverse projection of the image point \( p \) is a ray which is starting from the camera center and travelling through the image point \( p \). If we can identify the space coordinates of the camera center \( O_c \) and another point in the ray \( l_p \), the equation of the ray in space can be obtained.

If the camera matrix \( M_1 \) is known, the coordinate value of the camera center in the global coordinate system can be obtained by solving the equation \( M_1P = 0 \).

Let \( M_1 = (H, m_4) \), where \( H \) is the first three columns of the matrix \( M_1 \), \( m_4 \) is a column vector taking from the fourth column of \( M_1 \), the homogeneous coordinates of the camera center in the global coordinate can be obtained by solving the equation \( M_1P = 0 \), where the center

\[
O_c = \begin{pmatrix}
-H^{-1}p_4 \\
1
\end{pmatrix}
\]

Another point needed can be chosen as at the infinity on the ray \( l_p \) as \( P_\infty \). Let \( M_1P_\infty = p \). Thus, after getting the camera center \( O_c \) and the point \( P_\infty \), the ray parameter equation can be written as:

\[
P(u) = u\begin{pmatrix}
H^{-1}p \\
0 \\
1
\end{pmatrix} + \begin{pmatrix}
-H^{-1}m_4 \\
1
\end{pmatrix}
= \begin{pmatrix}
H^{-1}(up - m_4) \\
1
\end{pmatrix}
\tag{7}
\]

This 3D reconstruction algorithm is base on the assumption of zero error for the image coordinate. However in real world, the error from the measurement data is inevitable, and then their reverse-projection rays are unlikely to intersect at one point in space.

Therefore, we should find the second point \( P \) from the measurement data which corresponds to \( p \leftrightarrow p' \) by minimizing the geometric distance between the measurement point and its two projection point \( \hat{p} = MP \) and \( \hat{p'} = M'P \) in the two images, respectively.

Here we apply an approximate method called Sampson estimate \[^{14}\] \( \phi(P) = p^TFp \). It works for the algorithm of 3D space point construction when their reverse-projection rays generally do not intersect in the image.

Let \( P = (u, v, u', v')^T \), \( \hat{P} = (\hat{u}, \hat{v}, \hat{u}', \hat{v}')^T \), \( \phi(P) = p^TFp \). Sampson estimates of \( \Delta P = (\hat{P} - P) \) is

\[
\Delta P = -\left(\begin{matrix}
\frac{\partial \phi}{\partial P}
\end{matrix}\right)^T \left(\begin{matrix}
\frac{\partial \phi}{\partial P}
\end{matrix}\right)^{-1} \phi(P)
\tag{8}
\]

where

\[
\frac{\partial \phi}{\partial P} = ((Fp)_1, (Fp)_2, (F^Tp)_1, (F^Tp')_2)
\tag{9}
\]

As a results,

\[
\hat{P} = P - \frac{p'Tp}{(Fp)^2 + (Fp')^2 + (F^Tp)^2 + (F^Tp')^2} \left(\begin{matrix}
\frac{\partial \phi}{\partial P}
\end{matrix}\right)^T
\tag{10}
\]

It is worth pointing out that the pair of points \( \hat{p} \leftrightarrow \hat{p}' \) obtained using this estimation does not strictly satisfy the geometric constraints. However, if the noise of measurement point is within the range of a single pixel, the reconstruction algorithm can provide a satisfied reconstruction result.

### III. Results and Analyses

#### 1. Reconstruction result of the simulation data

The simulated trajectory curves of several known equation groups are built; and they are projected onto the camera imaging planes. The proposed method is then used to reconstruct the curve in 3D space and then the results are compared with the original curve in terms of the reconstruction error.

Firstly, three groups of simulation data given by the known equation groups are used to rebuild the trajectories which are shown in the dashed lines in Fig.4.(a). The reconstruction results are shown by the solid lines with scatters (shown as Fig.4(a)). Then, 100 sampling points, on the real trajectories and the reconstruction trajectories are used to calculate their errors compared with the two values shown as Fig.4(b)). Finally, the simulation results show that this method can reconstruct the trajectory curve of the periodic motion.

![Figure 4. Simulation results for the trajectory reconstruction.](image)

(a) show the simulation trajectory and reconstruction trajectory in each figure; (b) show the errors at the sampling points between real trajectory and reconstructed trajectory.

#### 2. Reconstruction of measurement data

The proposed method is used to reconstruct the motion trajectories and compared with the actual measurements. In order to validate the reconstructed data, Experiments are carried out to measure the static and dynamic periodic trajectories respectively. For the static periodic trajectory, a curve of simulated trajectory for simple motions is applied. For the dynamic periodic trajectory, the curve of human joint points is extracted by using the Microsoft Kinect camera \[^{15}\] system.

1. **Reconstruction of static periodic trajectory**

In order to obtain the real coordinate of trajectory curves more conveniently, the research builds a model to simulate
periodic spatial trajectory curves and marks twenty-five corresponding points on each periodic curve and then extracts the curve of trajectory at 2D camera plane. The curves of the trajectories of 2D coordinate of the marked points are extracted from the video camera’s 2D image plane. The simulated trajectory curve is reconstructed in 3D coordinate. We calculate the error between real data extracted from experiments and reconstructed data of the fifty marked points in two periods, as shown in Table 1.

Table 1. Errors between the real data from measurements and the reconstructed data

<table>
<thead>
<tr>
<th>Point ID</th>
<th>X (mm)</th>
<th>Y (mm)</th>
<th>Z (mm)</th>
<th>Point ID</th>
<th>X (mm)</th>
<th>Y (mm)</th>
<th>Z (mm)</th>
<th>Error (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>699</td>
<td>456</td>
<td>211</td>
<td>1</td>
<td>699.27</td>
<td>456.15</td>
<td>209.87</td>
<td>1.1736</td>
</tr>
<tr>
<td>2</td>
<td>686</td>
<td>480</td>
<td>227</td>
<td>2</td>
<td>674.06</td>
<td>482.27</td>
<td>220.65</td>
<td>13.7148</td>
</tr>
<tr>
<td>3</td>
<td>672</td>
<td>504</td>
<td>213</td>
<td>3</td>
<td>665.07</td>
<td>506.61</td>
<td>205.97</td>
<td>10.2151</td>
</tr>
<tr>
<td>4</td>
<td>672</td>
<td>506</td>
<td>183</td>
<td>4</td>
<td>674.58</td>
<td>508.26</td>
<td>181.69</td>
<td>3.6809</td>
</tr>
<tr>
<td>5</td>
<td>686</td>
<td>480</td>
<td>165</td>
<td>5</td>
<td>679.09</td>
<td>479.88</td>
<td>161.98</td>
<td>7.5377</td>
</tr>
</tbody>
</table>

Total points: 50, Mean error: 6.8181mm, Max error: 13.7148mm

(Note: we only show a part of our acquired data from the experiment.)

Table 1 shows that both the mean error and maximum error between real and reconstructed data are small. It proves that we achieve good reconstruction results.

(2) Reconstruction of dynamic periodic trajectory

In this work, we track and extract the 2D curves and 3D real coordinate of ankle joint points during the human gait using the Microsoft Kinect camera and OpenNI SDK. The 2D curve is used as input and trajectory curve of human gait is 3D reconstructed by using the method proposed in this paper. The extracted 3D real coordinate of ankle joint points is used as the reference data to evaluate the accuracy of the reconstruction results, as shown in Fig.5.

As for the dynamic periodic trajectory, the coordinate of human gait for chosen positions such as joints and ankles needs to be extracted. In this experiment, three persons’ ankle joint positions are extracted in nine groups for both the 2D curves and the 3D real coordinates in the global coordinate. The human gait can then be reconstructed in 3D using the method proposed in this paper and the results from one group are presented in Fig.6.

During two motion periods, twenty positions of each nine groups of reconstructed trajectory curves and Kinect-camera-extracted curves are sampled and the errors between the two results are calculated, as shown in Table 2.

Table 2 shows that the maximum error is 6.89 cm and the mean error is 3.79cm which proves that this method achieves effective reconstruction results.

Table 2. Error between extracted dynamic trajectory data and reconstructed data (mm)

<table>
<thead>
<tr>
<th>Point ID</th>
<th>X (mm)</th>
<th>Y (mm)</th>
<th>Z (mm)</th>
<th>Error (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−39.4122</td>
<td>−644.5428</td>
<td>96.5386</td>
<td>3.7596</td>
</tr>
<tr>
<td>2</td>
<td>2.8787</td>
<td>−543.8296</td>
<td>108.3418</td>
<td>5.6345</td>
</tr>
<tr>
<td>3</td>
<td>28.0723</td>
<td>−415.2761</td>
<td>126.3348</td>
<td>3.1298</td>
</tr>
<tr>
<td>4</td>
<td>47.2568</td>
<td>−300.5571</td>
<td>143.8267</td>
<td>3.1294</td>
</tr>
<tr>
<td>5</td>
<td>69.1693</td>
<td>−206.1069</td>
<td>160.7194</td>
<td>2.5353</td>
</tr>
</tbody>
</table>

Total points: 180, Mean error: 37.8756mm, Max error: 68.8793mm

(Note: we only show a part of our acquired data from the experiments.)

The errors of reconstruction results shown in Table 2 are larger than those in Table 1 due to two reasons: (1) The data of human joint points extracted by Microsoft Kinect camera and OpenNI SDK could be one of the sources of errors, including tracking errors in the 2D curves and data errors in the 3D real coordinate of human joint points. The errors from data extraction are the main reasons resulting in the errors for the 3D reconstruction. (2) The 3D reconstruction theory is established in the ideal situation where the curve of human gait trajectory in each period is completely coherent with that from other periods. However, the consistent human gait is hardly to be predicted in real world.
IV. Conclusion

Based on the nature of the periodicity of human gait, this paper proposes a method to calculate the periodic length of the human gait and builds a virtual perspective by the use of scene calibration and the geometric constraints of the trajectories. As a result, the 3D reconstruction of the periodic motion trajectory is implemented through a stereo vision method. The proposed method overcomes the problems of the camera angle-dependence for 3D periodic motion extraction using only a single camera. For further work, it is proposed to apply the presented method to solve more realistic practical problems for human gait motioning, in addition it may become necessary to develop additional algorithmic tools to handle the specific complexities of natural human motion.

References


ZHANG Jun was born in 1975. Currently she is an associate professor in Department of Information System and Management, National University of Defense Technology. Her main research interests include image processing, image coding, compressive sensing, image fusion, target detection and data mining. (Email: zhangjun1975@nudt.edu.cn)

HUANG Weiqiang was born in 1985. He received B.S. degree and M.S. degree in control science and engineering from National University of Defense Technology, Changsha, China, in 2009 and 2011, respectively. His research interests include image processing, virtual reality and multimedia information system.

LUO Tingjin was born in 1989. He is currently a M.S. candidate in control science and engineering from National University of Defense Technology. His research interests include image processing, virtual reality, target detection and multimedia information system.