3D Digitization of Human Foot Based on Computer Stereo Vision Combined with KINECT Sensor

Hai-Qing YANG\textsuperscript{a,*}, Li HE\textsuperscript{b}, Geng-Xin GUO\textsuperscript{c} and Yong-Jun XU\textsuperscript{d}

College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China
\textsuperscript{a}yanghq@zjut.edu.cn, \textsuperscript{b}646498438@qq.com, \textsuperscript{c}guogengxin@zjut.edu.cn, \textsuperscript{d}xuyongjun@zjut.edu.cn; *Corresponding author

Keywords: 3D Digitization of Human Foot, Computer Stereo Vision, KINECT Sensor, Feature Extraction and Matching, Harris-SIFT Algorithm.

Abstract. It is significant for shoe-making industry to develop three-dimension digital models of human foot. In this paper, a system combining a computer stereo vision with a KINECT sensor was proposed. Based on the system, an algorithm of feature extraction and matching was examined. Experiments show that the system and relevant algorithms are successful for 3D digitization of human foot. The proposed method could be further developed to obtain characteristic parameters of human feet, which can be used directly by shoe-making machines.

Introduction

In traditional shoe-making procedures, shoe lasts are essential for shoe lofting. Currently, shoe lasts are mainly hand-made, which often takes long time and skill needs [1]. It is necessary for human foot digitization to develop efficient technology of foot modeling. One method of three dimension digitization is using Laser scanner, whose main problem lies in high price and long scanning time[2,3].Another method is using structured light technology to form feature points or lines on the surface of an object. Although this technology is useful to build a 3D model of human foot, it is sensible to light interference and makes low precision to 3D reconstruction as well [4,5].

In this paper, a new system combining computer stereo vision with a KINECT sensor was proposed. Based on the system, feature points of human foot were extracted and used for 3D reconstruction of foot model.

Fundamentals of Computer Stereo Vision

Figure 1 shows the basics of computer stereo imaging, where \( b \) is the distance between two cameras, \( D \) is the distance between the object to be measured and the cameras, \( f \) is focal length of camera lens, \( L \) and \( R \) represent left camera and right camera.

Let world coordinates of the object \( P \) be \((X, Y, Z)\), imaging coordinates in the left camera \( P_l(x_l, y_l) \), and imaging coordinates in the right camera \( P_r(x_r, y_r) \). According to similar triangle rules, the world coordinates of the object \( P \) can be obtained through the Eq.1, Eq.2, and Eq.3. It is clear that the world coordinates depend on pixel point in the left and right cameras and the \( d \), which represents the parallax formed on the left and right image planes.

![Figure 1. Schematic for Binocular Vision System.](image-url)
3D Digitization System of Human Foot

The system device mainly includes two parts, binocular camera system and KINECT sensor, as shown in Figure 2.

KINECT sensor and binocular camera system structure is designed for the upper and lower structure, fix them together, it can be regarded as rigid structure. Then KINECT and binocular camera are calibrated, KINECT coordinate center and binocular system coordinate center exists below conversion relationship:

\[
\begin{bmatrix}
X_k \\
Y_k \\
Z_k
\end{bmatrix} = R_{kb} \begin{bmatrix} x_b \\ y_b \\ z_b \end{bmatrix} + T_{kb}
\]  

(5)

In the above formula, \((X_k, Y_k, Z_k)\) is the center of the KINECT coordinate system, \((x_b, y_b, z_b)\) is the center of the binocular camera coordinate system, and \(R_{kb}, T_{kb}\) is the rotation matrix and the translation matrix of the coordinate transformation.

3D Reconstruction of Foot System

We use the computer stereo vision and geometric affine 3D image reconstruction principle for the study theoretical basis, use binocular systems and KINECT sensor to obtain binocular images and depth information images. It is important to choose the coordinate system reasonably in the 3D reconstruction of foot which will reduce the reconstruction time and the complexity of the subsequent algorithms. 3D reconstruction system program of foot is shown in Figure 3:

![Figure 2. 3D Digitization System.](image-url)
1. Image acquisition

As shown in Fig.3, firstly fix the KINECT and binocular camera system, adjust the camera's perspective to make it fixed, the KINECT sensor optical axis and binocular camera optical axis is parallel in ideal mode, the device is seen as a rigid model, and collect binocular images and depth information images through computer.

2. Camera calibration

Firstly, the binocular camera system is calibrated by the Zhang Z calibration method, and then the binocular camera system and the KINECT sensor are calibrated to obtain \( (R_{kb}, T_{kb}) \).

3. Reconstruction of 3D model

The 3D reconstruction module consists two parts: point cloud stitching and carved surface reconstruction of different viewpoints. Firstly, the coordinates of feature points from different viewpoint are calculated by iteration to obtain the optimal solution through ICP algorithm. The optimal parameters obtained are substituted into the 3D points obtained by KINECT to stitch points, until get a complete 3D point cloud map of foot, finally Delaunay algorithm is used for triangular meshing, carved surface reconstruction [6,7].

4. Rendering of 3D model

Render the data that have been rebuilt through GEOMAGIC software, to get 3D effect, and measure various parameters information of foot.

**Experimental Environment and Equipment**

Experimental computer CPU is the Intel (R) Core (TM) i5, clocked at 2.5Ghz, memory 2G. The binocular camera used in this study is composed of two digital camera that its model is OLYMPUS-C500-ZOOM, the camera image resolution of 1280 pixels *960 pixels, the image file is JPG format. KINECT sensor used is KINECT2.0. This article uses the development tools: MATLAB R2013a, VS2012, KINECT Studio, programming language: C/C++, M language, OPENCV. The computer experiment system in this paper is: Windows7.
Results and Discussion

Human Foot Reconstruction

The reconstruction of the human foot need people natural sitting in a chair and feet in a natural floating state, then move the KINECT sensor to obtain a full range of foot depth images. For non-rigid objects, there will be a certain error, this error has little effect on the subsequent production of the shoe last, and it is within the acceptable range. The object measured for reconstruction experiment of the human foot is the author of this article and the laboratory member. Firstly, get different angle 3D point cloud of foot as shown in Fig. 4. Then stitch the obtained 3D point cloud of foot to get the full of 3D point cloud, and the Delaunay algorithm is used for triangular meshing to reconstruct the curved surface, as shown in the left figure in the Fig. 5. The right figure shows the reconstruction model of the foot after rendering.

![Figure 4. Part of Point Cloud of Human Foot Obtained From Different Angles.](image1)

![Figure 5. 3D Reconstruction Model of Human Foot.](image2)

3D Measurement of Human Foot Parameters

We combine several key parts when making shoes, measure the following parameters: foot length $L_f$, foot width $w = y_1 + y_2$, metatarsal circumference $C_1$, tarsal circumference $C_2$, scaphoid circumference $C_3$, metatarsal height $H_1$, tarsal height $H_2$ and scaphoid height $H_3$ [8].

As shown in Fig. 6, the foot length $L_f$ refers to the distance between the front end of the forefoot and the rear end of the foot, the metatarsal circumference $C_1$ refers to the circumference of the protruding point around the metatarsal joint, tarsal circumference $C_2$ refers to the circumference of the protruding point of front tarsal, scaphoid circumference $C_3$ refers to the circumference surrounded by the curved point of scaphoid and the heel, metatarsal height $H_1$ is the vertical distance between the highest point of the metatarsal and the measurement plane, tarsal height $H_2$ is the vertical distance between the protruding point of front tarsal and the measurement plane, scaphoid height $H_3$ is the vertical distance between the curved point of scaphoid and the measurement plane.
In this study, feet of adult students were measured 3 times with averages shown in Table 1, the measured staff 1 is the author himself. It can be seen from Table 1, the average deviation measurement of measured personal is +1.656mm, we can see the date difference between this paper system measured and manual measured is $\Delta$, a positive value, we can use $\Delta$ to correct the deviation, make the measurement closest to the value.

<table>
<thead>
<tr>
<th>Characteristic Parameters</th>
<th>Left foot</th>
<th>Right foot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Manual measurement</td>
<td>Automatic measurement</td>
</tr>
<tr>
<td>Foot length</td>
<td>266.0</td>
<td>267.9</td>
</tr>
<tr>
<td>Foot breadth</td>
<td>102.0</td>
<td>103.6</td>
</tr>
<tr>
<td>Metatarsal circumference</td>
<td>223.0</td>
<td>225.1</td>
</tr>
<tr>
<td>Tarsal circumference</td>
<td>255.0</td>
<td>257.4</td>
</tr>
<tr>
<td>Scaphoid circumference</td>
<td>333.4</td>
<td>336.4</td>
</tr>
<tr>
<td>Metatarsal height</td>
<td>43.0</td>
<td>43.7</td>
</tr>
<tr>
<td>Tarsal height</td>
<td>65.0</td>
<td>66.8</td>
</tr>
<tr>
<td>Scaphoid height</td>
<td>88.0</td>
<td>89.1</td>
</tr>
</tbody>
</table>

**Conclusion**

In this paper, a scheme of 3D reconstruction of human foot was proposed based on computer stereo vision combined with a KINECT sensor. The system runs successfully and costs inexpensive. The system could be further developed to obtain characteristic parameters of human feet, which can be used directly by shoe-making machines.

**References**


