3d Motion Capture Method Based on Kinect Sensor

Lingda Meng[a], Lingdi Kong[b]

School of software, zhengzhou university, Zhengzhou 450000, China

[a]913896134@qq.com, [b]764452379@qq.com

Keywords: Kinect, motion capture, spatical attitude transformation

Abstract: It is of great significance to analyze human posture movement in behavior monitoring, movement analysis, medical rehabilitation and other aspects. The traditional attitude sensor needs to be worn on the body, so the movement is easy to be hindered, and it is difficult to restore the real action state. In this paper, a real-time motion capture system is designed based on Kinect camera to capture the motion data of human body. The method and principle of Kinect camera coordinate system conversion to world coordinate system are given in detail. The data can drive the human body bone model drawn by OpenGL in real time after the three-dimensional attitude transformation, which solves the problem that only the human body model established by Kinect with the default bone level can be driven, and accurately reflects the attitude change process of human in the process of motion.

1. Introduction

In recent years, Kinect camera equipment is widely used and motion capture technology, the equipment is less affected by the light, the performers are not limited to dress, do not need to wear sensor equipment. Kinect V2 achieves the acquisition of Depth image through TOF technology (Time of Flight), and calculates the Depth information by measuring the Time that the infrared ray projected by the infrared transmitter returns to the infrared receiver (Depth sensor). Human pose recognition mainly uses Kinect's real-time bone tracking technology and logic regression classification algorithm. Human skeleton is mainly composed of joints, so it can analyze human pose well through joint features, Literature [1] using Kinect bone tracking technology for medical rehabilitation training. [2]Zengxiao used the human body image obtained by Kinect's depth sensor to recognize 3D human body posture. [3]Mengwei Han captured the coordinates of human joints based on the Kinect camera coordinate system, calculated the Angle by using the two-point method and defined the pose library of human body. The above research is based on Kinect's own coordinate system for data processing. In this paper, the kinect coordinate system is converted into the world coordinate system by using the principle of three-dimensional attitude transformation, and the captured joint data is transferred into the bone model, so as to drive the human skeleton model drawn in real time, which improves the reusability and accuracy of the capture system.

2. Capature human body coordinates

Kinect obtains the depth image information through the depth sensor, and separates the human body target (T-shaped object, which will be recognized as human body by Kinect) from the environment background through edge detection, noise threshold processing and other technologies, to obtain a depth image of human body [4]. Then, BPC algorithm (body part classification) is used to locate the joints. BPC algorithm uses the forest classification method of depth random decision to segment the human body parts from the depth image information, and marks the pixels in the center of each part. These markers are very close to the actual position of the bone joint, so as to locate the joint point of the bone. These joint points are mapped into coordinate system and connected into human skeleton map.

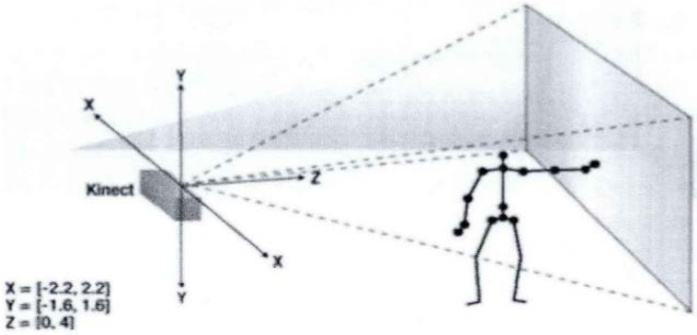

Figure 1. 3D coordinate system of Kinect

Using the scene depth information obtained by Kinect, the actual distance between human and camera can be calculated. Kinect obtains the actual distance d from the measured object to Kinect sensor by acquiring the depth value $d = Ktan(Hdraw + L) - O$. Where, draw is the 3d depth value of the joint node read by Kinect, where K, H and L are fixed values provided by Kinect, K=12.36cm, L=1.18rad, O=3.7cm. Then the pixel coordinates of the depth map are transformed to the actual coordinates relative to Kinect. The spatial coordinates of any joint point in the world coordinate system are obtained, and the transformation formula [5] is as follows:

$$\begin{cases} x_{kinect} = \left(x_{image} - \dfrac{w}{2}\right)(z_{kinect} + D)F\dfrac{w}{h} \\ y_{kinect} = \left(y_{image} - \dfrac{h}{2}\right)(z_{kinect} + D)F \\ z_{kinect} = d \end{cases}$$

Where d = - 10, f = 0.0021, w * h indicates Kinect resolution: 640 * 680.

3. Spatial coordinate transformation

The coordinates obtained by Kinect in capturing human body are relative to those obtained by Kinect's own coordinate system. To drive the OpenGL human skeleton model, we need to convert the coordinates relative to Kinect coordinate system to the coordinates relative to human posture.

The principle of space attitude coordinate conversion is as follows: In any spatial coordinate system, the point P can be expressed as $^{A}P = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}$

The spatial coordinate system I is translated to obtain the spatial coordinate system j, the coordinates of the origin of the coordinate system j in the coordinate system I are represented by $^{i}P_{oj}$, and the relation of the same point P in the two coordinate systems can be expressed as: $^{i}P = {}^{j}P + {}^{i}P_{j}$

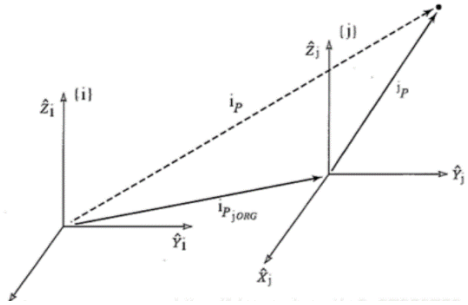

Figure 2. Coordinate translation

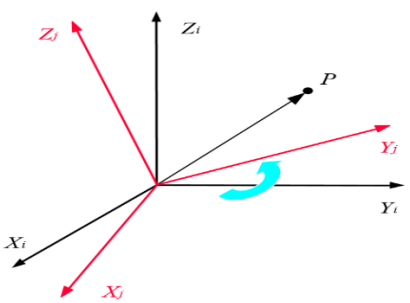

Figure 3. Coordinate rotation

When coordinate system I and origin of coordinate system J coincide, coordinate system J rotates a certain angle theta around any direction axis of coordinate system I. The coordinate of point P in J coordinate system is expressed as ${}^{j}P=[x_j,y_j,z_j]^T$, The coordinate of point Pin I coordinate system is expressed as ${}^{i}P=[x_i,y_i,z_i]^T$

According to the mathematical relation ${}^{i}P={}_{i}^{j}R{}^{j}P$, ${}_{i}^{j}R$is the rotation matrix rotating around the current direction axis. $\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix}=\begin{bmatrix} \cos(\angle x_i,x_j) & \cos(\angle x_i,y_j) & \cos(\angle x_i,z_j) \\ \cos(\angle y_i,x_j) & \cos(\angle y_i,y_j) & \cos(\angle y_i,z_j) \\ \cos(\angle z_i,x_j) & \cos(\angle z_i,y_j) & \cos(\angle z_i,z_j) \end{bmatrix}\begin{bmatrix} x_j \\ y_j \\ z_j \end{bmatrix}$ The rotation matrix around the x-axis, Y-axis and z-axis of the original coordinate system I is as follows:

$${}_{i}^{j}R(x_i,\theta)=\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}\quad {}_{i}^{j}R(y_i,\alpha)=\begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix}\quad {}_{i}^{j}R(z_i,\beta)=\begin{bmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

When the object has an Angle of rotation in both the X and the Y and the Z axes. The rotation matrix $R(\theta,\alpha,\beta)$ can be represented as $R(\theta,\alpha,\beta)={}_{i}^{j}R(x_i,\theta){}_{i}^{j}R(y_i,\alpha){}_{i}^{j}R(z_i,\beta)$

$$R(\theta,\alpha,\beta)=\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix}\begin{bmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Calculate joint angle:

In the process of coordinate transformation, it is necessary to calculate the rotation angle of human body relative to the coordinate axis in Kinect camera coordinate system. According to the joint coordinates in the Kinect camera coordinate system, the vector between the joints is calculated, and the motion Angle of the joints is calculated by the included Angle relationship between the vectors.

For example, when the coordinate of the hand joint in the plane is T(x,y,z), the hand joint vector is $\vec{a}=(x,y,z)$, The unit vector in the X-axis can be expressed as$\vec{\lambda}=(1,0,0)$, Then the angular relation of hand joint rotation relative to X-axis can be expressed as: $\cos\theta=\dfrac{\vec{a}\,\vec{\lambda}}{\left|\vec{a}\right|\left|\vec{\lambda}\right|}=t$. So we can figure out the Angle at which the knuckle moves in the X-axis: $\theta=\arccos t$. The unit vector in the Y-axis can be expressed as$\vec{\mu}=(0,1,0)$, Then the angular relation of hand joint rotation relative to Y-axis can be expressed as: $\cos\alpha=\dfrac{\vec{a}\,\vec{\mu}}{\left|\vec{a}\right|\left|\vec{\mu}\right|}=h$. So we can figure out the Angle at which the knuckle moves in the Y-axis: $\alpha=\arccosh$. The unit vector in the Z-axis can be expressed as$\vec{\nu}=(0,0,1)$, Then the angular relation of hand joint rotation relative to Z-axis can be expressed as: $\cos\beta=\dfrac{\vec{a}\,\vec{\nu}}{\left|\vec{a}\right|\left|\vec{\nu}\right|}=g$. So we can figure out the Angle at which the knuckle moves in the Z-axis: $\beta=\arccos g$.

Other joint in the coordinate direction of the axis of the solution method is also the use of vector included Angle relations.

According to coordinate translation and rotation, the relationship between the original coordinate system and the transformed coordinate system can be obtained: $P_i = {}^i_j RP_j + {}^i P_{oj}$

The coordinate data of kinect camera is transformed into the coordinate system of human skeleton model by using the transformation of three-dimensional coordinate and attitude, so as to drive the movement of human skeleton model drawn by OpenGL and obtain the real-time data of human motion.

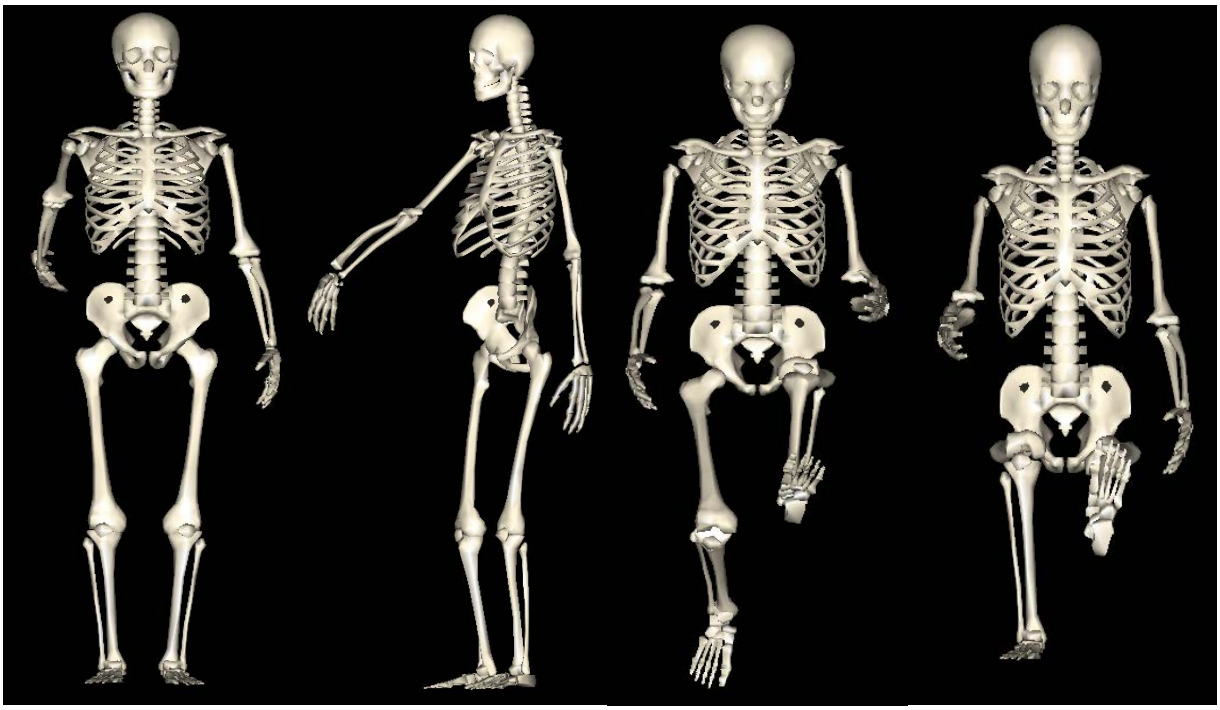

Figure 4. Human skeletal movement

4. Summary

In this paper, Kinect camera was used to capture the body posture and motion. Instead of using the human body model established at the default bone level of Kinect, the captured coordinate data was processed to establish the world coordinate system. The motion of the human skeleton model drawn by OpenGL was driven by the three-dimensional attitude transformation to improve the reusing of the motion capture system. Although it sensors to obtain depth information, the human body to calculate human body space, but for the joint point coincidence, on the recognition is not accurate enough, so, should be in focus at the same time, the development of the human behavior analysis field study hard skeleton correction problems, further improve the joint skeleton accuracy, capture the human movement data needs to be further analysis.

References

[1] Da Gana A, Chaves T, Figueiredo L. Poster: Improving motor rehabilitation process through a natural interaction based system using Kinect sensor [C]//2012 IEEE Symposium on 3D User Interfaces (3DUI),

[2] Zheng Xiao, Fu Mengyin, Yang Yi, ET al. 3D Human postures recognition using Kinect [C]//2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHM SC), 2012: 344-347

[3] Zhan yinwei Pose recognition algorithm based on Kinect angle measurement, [J]. Sensors and Microsystems (2014) 07—0129—04

[4] Meng Ming, Yang Fangbo, she Qingshan, sun Yao, Luo zhizeng. Human motion detection based on Kinect depth image information [J]. Journal of instrumentation, 2015, (02): 386-393

[5] Satyavolu S, Bruder G, Willemsen P Analysis of IR-based virtual reality tracking using multiple Kinects [C]//IEEE Virtual Reality Conference, 2012: 149—150.