Noise Reduction of 3D Human Body Surface Reconstruction Algorithm and Surface Continuous Optimization

Fengjun Hu

Institute of Information Technology, Zhejiang Shuren University, Hangzhou 310014, Zhejiang, China

Abstract:
Because of complexity and motility of human body surface, reconstruction of human body surface cannot be implemented perfectly through current existing algorithms compared to the ordinary objects. Also when we use Kinect camera to obtain the three-dimensional point cloud of human body, it is easy to be affected by all kinds of random noise due to the camera itself and the influence of environmental factors. Thus, to improve the current algorithms, this paper proposes a 3D point cloud registration algorithm based on depth image noise reduction and a 3D surface reconstruction method based on the improved SDF and Poisson optimization. The key points of the paper are as follows: (1) the depth image of the human body is obtained through Kinect camera, and the minimum energy is calculated by using the energy model and Gauss-Newton Algorithm to implement noise reduction for depth image. Then, the global registration is carried out by using the curve interpolation method to minimize the error in the point cloud registration. (2) By using the weight optimized directional distance function to combine the point clouds, the 3D human body surface is reconstructed through the estimation model of Poisson optimization to improve the continuity and completeness of the 3D curve.

Key words: Human body surface reconstruction, Depth image noise reduction, Energy model, Curve interpolation, Directional distance function, Poisson optimization

1. INTRODUCTION

The 3D rapid modeling technology is one of the most important research topics in the fields of the computer vision, virtual reality technology and graphics. By using the 3D rapid modeling technology to reconstruct the object model has become the development trend currently (Henry, Krainin, Herbs, et al, 2012). The 3D object reconstruction can be realized easily just like taking photos and videos. It can be applied to more and more fields, such as e-commerce platform, online shopping websites, gaming, etc. (Lomonosov, Chetverikov and Ekárt, 2013). Therefore, realizing a human body modeling quickly and cost-effectively has become the research focuses and an important goal in the fields of the computer visualization and graphics (Tong, Zhou, et al, 2012).

The 3D human body reconstruction technology has always been a hot research topic in the fields of computer graphics, visualization and industry. Using Kinect camera to realize 3D human body reconstruction is a common method currently. Zhang and others (Zhang and Li, 2016) proposed a new method of 3D human reconstruction, which uses Kinect to identify and obtain 3D human body data, and then, combined the point cloud library to reconstruct the human body model. Zhu and others (Zhou, Wu and Zhou, 2015) combined Kinect fusion technology to design the 3D human reconstruction system and realized it based on multi-Kinects. Han and others (Han, Pang and Wang, 2015) presented a high identification 3D human body model reconstruction method based Kinect. The method can combine advantages of multi-Kinects to get the high precision body surface point cloud data. Thus, the high precision 3D human body model can be obtained. Ma and others (Ma and Xue, 2014) also proposed a new reconstruction method which need one Kinect only, which combines with the feature point detection algorithm and error check processing to detect the feature points of a human in depth video frames.

There are also many researchers who adopted other methods to realize the reconstruction of 3D human body. Chen and others (Chen, Wu and Lu, 2009) used the surface reconstruction method to acquire the human surface model, and implemented parameterized deformation to the parameterization human body model. Chen and others (Chen and Deng, 2013) utilized the stereo vision measurement to implement 2D measurement of the model set feature points, and converted the 2D to 3D data according to the binocular principle of vision measurement. Then, through the interpolation algorithm for more data information, they made the surface data information of model smoother. Qi and others (Qi, Hou and Huang, 2015) proposed a virtual human modeling method combining 19 sets of anthropometry feature data of GB10000-88. The authors constructed a virtual human skin mesh by 3dsmax and a virtual bone joint movement model by hierarchy modeling, then combined the quaternion spherical vertex blending interpolation technology to establish the virtual human model. Zhao
and others (Zhao and Li, 2012) used Otsu algorithm based on the feature of shadow to complete shadow removal in the moving target direction, which can get human contour precisely. Then, the authors preprocessed the human contours by morphological refined method, used the new connectivity structure standard and the human joint location algorithm to establish the human skeleton model. However, under the condition of big data noise, the effect of the 3D human body reconstruction through above mentioned method is not very good. To deal with this problem, this paper proposes a 3D point cloud registration algorithm based on depth image noise reduction and a 3D surface reconstruction method based on the improved SDF and Poisson optimization.

2. THE 3D POINT CLOUD REGISTRATION BASED ON KINECT

2.1. The Noise Reduction of Depth Image based on Energy Model

Since the precision of the reconstruction can be affected by the noise in depth image, noise reduction is a very important process for noise reduction. An energy model is introduced into this paper and also obtain the minimum energy through Gauss-Newton algorithm(Yang, 2014) to get the ideal depth image.

$$\min_{X} (E_{fill} + E_{ppi} + E_{smooth})$$

$$E_{ppi} = \sum_{k=1}^{N} \sum_{j \in N(k)} \|u(k) - u(j)\|^2$$

where $u(k), u(j) \in X$ represent the adjacent points in $X$, and $N(k)$ represents the index of adjacent points of $k$ index corresponding points. And $E_{ppi}$ makes the adjacent points as close as possible to improve the resolution.

$$E_{smooth} = \sum_{k=1}^{N} \sum_{j \in N(k)} \|X(u(k)) - X(u(j))\|^2$$

where $X(u(k))$ and $X(u(j))$ represent the depth values of the adjacent points in $X$. $E_{smooth}$ represents the degree of smoothing the depth map and burr reduction.

$$E_{fill} = \sum_{i=1}^{n} \| W_i \odot (D_i - X) \|^2$$

where $D_i$ represents the $i$-th frame of the original depth map. $\odot$ represents the multipication of the corresponding elements of the same size matrix, and the size of the matrix is still the same after calculation. $E_{fill}$ represents the integration of multiple depth map and empty filling

$$W_i(u(k)) = \begin{cases} 0, & D_i(u(j)) - X(u(j)) > Z_{threshold} \\ 1, & D_i(u(j)) - X(u(j)) \leq Z_{threshold} \end{cases}$$

$$\sigma_{i}(\theta, z) = 0.8 + 0.0035 \frac{\theta}{\pi - \theta} z \frac{p_i}{f_x}$$

where $p_i$ represents the pixel area in depth map, $f_x$ represents the focal length.

$$\sigma_{i}(\theta, z) = 0.0012 + 0.0019(z - 4)^2 + \frac{0.001}{\sqrt{2}} \frac{\theta}{\pi - \theta}$$

where $\theta$ represents the angle between the camera coordinate axis in Z direction and object surface normal vector. $W_i(u(k))$ represents the position weight of the $i$-th frame depth image $u(k)$. Based on this value, we can avoid and reduce the noise accumulation phenomena appeared in the process of depth map integration. The points cloud shown in figure 1 are the original ones, and the points cloud shown in Figure 2 are the processed ones.
2.2. The 3D Point Cloud Registration Based on the Curve Interpolation

Lots of frame point clouds can be produced during the human body scanning process. After the in-pairs local registration, we hope the result of the frame overlap of head and end can be obtained. Also, the rotation and translation errors of the local registration will be rejusted, as shown in Figure 3. However, the ideal result can not be acquired in the real situation, and the matching point of head and end of n frame matching point sequences can not be overlapped, which is called as loop closure problem.

For the above mentioned problem, the global registration is implemented based on the method of curve interpolation (Ye, Li and Zeng, 2013). The process is as follows:

(1) Resampling is carried out for the matching points and n frames of the match point sequences are constructed under the different heights of the human body models. The match point of head and end frame is corrected. Then, we set the center of the head and end frame match point as a closure match point, and use B spline interpolation curve method three times to construct the closure fitting curve. Parameter $t_i$ of
each point $P_i$ can be obtained through arc length parameterization. Thus, we can use these parameters to calculate point position $P'_i$ in the closure fitting curve. (2) For the match points under different heights in each frame, we recalculate the fixed rotation $R_i$ and translation matrix $T_i$. The total offset should be at minimum value, which can be calculated through the least-square method.

\[
\min \sum_{j=1}^{n} \left\| R_i \cdot P_j + T_i - P'_j \right\|
\]

(8)

(3) By using the corrected transformation matrix to adjust each frame, the global registration result can be obtained, which is shown as below in figure 4.

![Figure 4. 3D point cloud registration result](image)

3. 3D POINT CLOUD MODEL SURFACE RECONSTRUCTION

3.1. The Integration of Point Cloud base on Directional Distance Function

A complete point cloud model of human body is obtained through the global registration. However, in this model, the same surface can contain the multi-layer surface point. The point cloud registration can make the space between multiple frames smaller, yet they cannot be integrated into one frame cloud point. If we use these multiple-frame point clouds to reconstruct the surface directly, the effectiveness can be affected. This paper uses the directional distance function (Deng, 2016) to integrate the point cloud first, then the surface reconstruction is carried out.

Define $\phi(\vec{x})$ as an implicit function. If $\phi(\vec{x}) \leq 0$, $\vec{x}$ is located in the inner region $\Omega^-$, if $\phi(\vec{x}) > 0$, $\vec{x}$ is located in the outer region $\Omega^+$, and if $\phi(\vec{x}) = 0$, $\vec{x}$ is located in boundary area $\partial \Omega$. Then, we can define a distance function to show the distance to the boundary.

\[
d(\vec{x}) = \min_{\vec{y} \in \partial \Omega} |\vec{x} - \vec{y}|, \forall \vec{x} \in \Omega
\]

(9)

From the above distance function, we can find that, if $d(\vec{x}) = 0$, $\vec{x} \in \partial \Omega$. The directional distance function can be defined as follows: for all $\vec{x}$, the implicit function $f(\vec{x})$ can meet $f(\vec{x}) = d(\vec{x})$, where

\[
f(\vec{x}) = \begin{cases} 
-d(\vec{x}), & \vec{x} \in \Omega^- \\
0, & \vec{x} \in \partial \Omega \\
d(\vec{x}), & \vec{x} \in \Omega^+
\end{cases}
\]

(10)

The common method is to discretize the directional distance function into the space voxel grid $m \times m \times m$ (usually set $m$ as 512). The whole space is in a cube. Each voxel grid contains two sets of data $\{d, w\}$, where $d$ represents the distance between the space grid and surface grid, $w$ represents the distance merge weight of multiple grids in the same direction. If $d < 0$, $v$ is located in the internal surface, if $d > 0$, $v$ is located in the external surface, and only if $d = 0$, $v$ is on the surface. The method uses the interception of bandwidth, which means that only the voxel grid within the range of $-T < d < T$ (usually set $T$ as 0.03m) can be calculated, and other data in the grid are set as constants.

The directional distance function can also be denoted as below:
Each grid weight is set as one through the method. The actual merge operation is to use $d$ and $w$, which are taken from the point cloud computing voxel grid of the previous frame, and add the point cloud of the next frame, then recalculate the distance in the voxel grid through formula (12) and (13).

$$D_{k+1} = \frac{D_k + w_{k+1}d_{k+1}}{w_k + w_{k+1}}$$  \hspace{1cm} (12)

$$w_k = w_{k+1} = 1$$  \hspace{1cm} (13)

Cumulative error can be easily produced in the point cloud merging based on the directional distance function (SDF), which can affect the final construction result. The error mainly includes the following two aspects: first, due to the limitations of the scanner itself, the obtained point cloud data has errors. In order to solve this problem, this paper improves the method of weight values of the original SDF method.

$$w_k = 1 - \text{error}_{n_k} - \text{error}_{n_i}$$  \hspace{1cm} (14)

$$\text{error}_{n_k} = \frac{1}{k} \sum_{i=1}^{k} \cos(n_i \cdot n_{k+1})$$  \hspace{1cm} (15)

$$\text{error}_{n_i} = \frac{1}{k} \sum_{i=1}^{k} \frac{||P_i - P_{k+1}||}{||P_i||}$$  \hspace{1cm} (16)

where $n_{k+1}$ represents the weighted average normal vector of the selected three blocks of point clouds in the $k+1$-th frame, and $n_i$ represents the weighted average normal vector of the selected three blocks of point clouds in the $i$-th frame. The cosine of the angle between $n_i$ and $n_{k+1}$ shows the rotating error. The translation error is calculated by using ratio of the distance between the matching point of the $i$-th and $i+1$-th frames and the distance from the matching block of the $i+1$-th frame to the original point. When the estimation error of transformation matrix is zero, $w_k = 1$, and the weight value method in the standard SDF is a special case here. Figure 5 below shows the effect of the merging.

![Figure 5. Point cloud merge effect map](image)

### 3.2. The Surface Reconstruction Based on Poisson Optimization

Based on the Poisson reconstruction method (Zhao and Wang, 2011), this paper propose a method to combine the advantages of the global and local fitting methods. Since it is global, forming adjacent areas, selecting type of surface and adjusting the weight are not related to the heuristic decision making. However, because the basis function is related to the surrounding space, not the data points, supported locally and has a hierarchical structure, a sparse good state system can be generated. We can estimate the indicating function of the model and extract the isosurface through Poisson optimization, reconstruct a seamless triangle approach, which can achieve the effect of voids patching to obtain the surface of the reconstruction model. Figure 6 shows the process of Poisson optimization.
In order to optimize the point cloud data, an octree structure relation is established. The point cloud data with the void information is added to the octree. Function is defined in each node of the tree, and expanded at node \( p \) with its size:

\[
F_p = F\left( \frac{p - p_c}{p_w} \right) \frac{1}{p_w^3}
\]  

\[
p_c = \frac{1}{k} \sum_{j \in k} p_j
\]

where \( p_c \) represents the center of the node \( p \), \( p_j \) represents the \( j \)-th node, and \( p_w \) represents the width of the node. The surface of the directed point set is reconstructed and transferred to the Poisson problem with one space.

A function space is defined, in which distance translation and scaling can be realized. When the function space is near the surface of the model, it has relatively high model resolution. On the other hand, when it is far away from the surface of the model, it has low model resolution. The basis function of function space is defined as \( F \). In this paper, we use the \( n \)-th order of box filtering to define the basis function \( F \):

\[
F(x, y, z) = (B(x)B(y)B(z))^n
\]

where \( x, y, z \) represents the coordinate of any point in point cloud data, \( n \) represents the order of the filter, and \( B(t) \) represents filter function, which is defined as follows:

\[
B(t) = \begin{cases} 
1, & |t| < 0.5 \\
0, & \text{else}
\end{cases}
\]

This basis function \( F \) can be more closer to Gaussian filter with the increase of number \( n \), and its support range also is enlarged. In this paper, we use three sectional second order approximation, that is, \( n = 3 \).

Under the condition of uniform sampling, the model surface can be estimated approximately based on gradient of vector field \( V \) approaching the index function. In order to avoid the sampling point being fixed in the center of sampling point leaf node, the sampling point is distributed to eight nearest neighbor nodes by using the linear interpolation method to achieve the accuracy of the child notes. According to function \( F_p \) of each node, the gradient approximation of the indicator function is defined as:

\[
V(p) = \sum_{s \in S} \sum_{p \in Nb_D(S)} \partial F_p N_p
\]

where \( S \) represents the point set of the point cloud data, \( s \) represents the nearest neighbor region of point \( p \), \( Nb_D(S) \) represents the eight nodes with the depth \( D \) in the nearest neighbor region of point \( p \), \( \partial \) represents the linear coefficient, and \( N_p \) represent the fixed point normal vector of point \( p \).

After getting the calculation result of vector field \( V \), we can use Laplacian iteration to solve Poisson equation.

\[
\Delta \vec{\xi} = \nabla \cdot \vec{V}
\]
For the reconstruction of surface, a proper threshold is selected, and iso-surface is extracted through MC (mobile cube) algorithm. Finally, we match the extracted triangular patches to get the reconstructed 3D surface model, which is shown in figure 7 as follows.

**Figure 7.** Surface reconstruction effect

4. THE EXPERIMENT RESULTS AND ANALYSIS

In order to verify the effectiveness of the improved algorithm proposed in this paper, simulations are implemented. The system programme is based on C/C++ language and Microsoft Visual Studio 2010. The raw data of Kinect is obtained through Official Microsoft Kinect SDK 1.7. And the hardware is Inter(R) Core(TM) i3-2100CPU @ 3.10Hz, graphics card is GeForce GTS 450, and memory is 1 g with Windows 7 system. The 3D object surface reconstruction is carried out based on teeth, which is shown in figures 8-10.

**Figure 8.** Noise reduction and point cloud image registration

**Figure 9.** Point cloud merge results
Then, 3D object surface reconstruction is carried out based on human body, which is shown in figure 11.

From the above results, we can see that, the improved method proposed in this paper has the good effect on 3D human body surface reconstruction. The reconstructed surface not only has the smooth surface, but also has the ability of noise resistance in extracting point cloud.

5. CONCLUSION

3D human body reconstruction is a very important research topic in the fields of computer graphics, virtual reality, computerized vision and 3D animation, etc. Since human body has the special features of non-static characteristic and the complex surface, human body reconstruction is more difficult compared with the surface reconstruction of other ordinary objects or scenes. For the defects of the current 3D human body reconstruction algorithm, this paper proposes the 3D point cloud registration algorithm based on depth image noise reduction and the 3D surface reconstruction method based on the improved SDF and Poisson optimization. Simulation results have shown that the improved method has good effect on 3D human body reconstruction. The constructed surface not only has the high degree of smoothness and continuity, but also has the ability of noise resistance in extracting point cloud.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China(Grant No. 51675490).
REFERENCES:


