

Application research on 3D reconstruction based on Kinect

Xiao Zhiguo^{1, a}, Jiang Junhai^{2, b}, Han Jiawei^{3, c}, Li Nianfeng^{4, d}

^{1,3} *College of International Education, Changchun University, Changchun city of Jilin province satellite road No. 6543, 130022, China*

^{2,4} *Changchun University, Changchun city of Jilin province satellite road No. 6543, 130022, China*

^a398837379@qq.com, ^b108271843@qq.com, ^cJason.hjw@gmail.com, ^d287219611@qq.com

Abstract

3D reconstruction technology is widely used in computer teaching and rehabilitation of virtual reality, robot autonomous navigation, interactive, film and television animation, etc. The core problem is to achieve a scene with different Angle of view of information and effective registration of fusion, is essentially solve the space coordinate transformation relation in the multiple points of view. Using access relatively cheap sensor as input equipment, produced color data and depth of the data preprocessing, Coordinate transformation, scene, point cloud registration was aligned fusion, said polygon mesh and texture mapping technology is studied and the theoretical analysis, Proposed a 3D reconstruction scheme based on access depth sensors.

Keywords: 3D reconstruction; Kinect; Point cloud integration; Virtual reality index.

Introduction

3D reconstruction technique involves many subjects of computer graphics, computer vision, pattern recognition etc.. The technology of 3D reconstruction of existing, according to the way of obtaining the depth information, can be divided into passive technology and active technology. Passive technology using natural light reflection, generally through the camera image, the 3D coordinates of objects and information obtained through a series of calculation algorithm, such as motion image sequence method and stereo vision method. Active technology includes a light source, a direct measurement of the depth information of object, so it is very easy to do real-time effect such as Time of Flight [1] [2] technology and USES the structured light technology to access. Time of Flight technology USES sophisticated hardware to get high precision of 3D coordinate object, but the hardware equipment complex, expensive, greatly limits the use of the technology. Since the 2010 release of Kinect, According to the Kinect research of 3D reconstruction began to spread. In 2011, Shahram Izadi, Richard A, Newcombe et al. Construction of A system, for A small range of static scene and separate objects for 3Dreconstruction, get the reconstruction of the relatively fine effect, but the reconstruction efficiency remains to be further improved. At home, pattern recognition, institute of automation, Chinese academy of sciences state key laboratory of bass, etc. [4] implements a fast object reconstruction method based on the device sensor, and through the GPU acceleration, parallel implementation of rotating table rotation of 3Dreconstruction of the object, but the reconstruction process depends on the speed of turntable, delimits the scope of its application.

This paper proposes a generic object 3Dreconstruction method based on sensor device, using the device sensor to get the color and depth image of the object, of different coordinate system transformation, the depth image preprocessing to smooth noise and get the 3Dpoint cloud, image segmentation and characteristics, using collected point cloud image feature points, complete scene 3Dreconstruction.

The research content

Kinect USES is Israel Prime Sensor company's end-to-end solutions, to obtain depth imaging technology is based on Light Coding technology, Light Coding using infrared laser Light emitter technology theory, through camera grating or spread sheet will be the projection of the laser Light evenly to measure space, to encode measurement space, and then through the infrared camera to record each speckle in space, combined with the original speckle pattern, calculated with 3Ddepth of image.

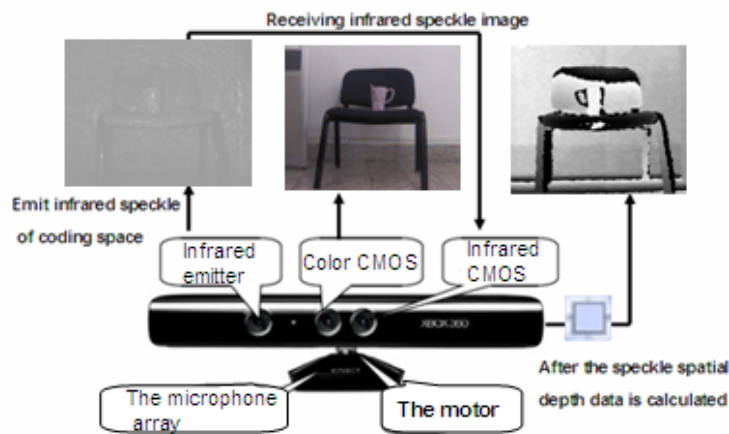


Fig.1. Structure, working principle diagram of the Kinect
Original coordinate data were collected by the Kinect are the coordinates of the depth image, where $X \in [0,640]$, $Y \in [0,480]$, with the help source library OpenNI depth image coordinate system X coordinate and Y coordinate transformation to the frame of reference for device camera coordinates X and Y coordinates of the world. Depth image coordinate system and Coordinate system to Kinect depth camera as reference system as shown.

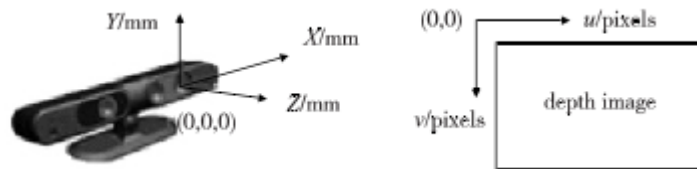


Fig.2. To access the depth camera as reference point coordinate system in the world
Set the resolution of the RGB image and depth are 640 by 480, according to

the principle of simple similar triangles, X, Y coordinates are:

$$X=(u-320) \times Z \times \frac{1}{f} \tag{1}$$

$$Y=(v-320) \times Z \times \frac{1}{f} \tag{2}$$

X, Y, respectively the abscissa and ordinate of the real world coordinate system; U, v, the abscissa and ordinate in depth image; The depth of information is Z for measurement. F is the focal length of the image. Finally the three-dimensional point cloud data by photo source database implementation [5].

To access the infrared transmitter and infrared CMOS camera can get around the scene of the combination of 3Ddata. As a result of the limitation of Kincet sensor equipment itself, measurement environment and surface characteristics under test, the influence of such factors as the depth of the image often has a lot of noise, based on literature [6]. In this paper, a fast bilateral filtering algorithm for image smoothing denoising depth.General image filtering can be represented as:

$$I^0(x, y) = \frac{1}{w_p} \sum_{i, j \in \Omega} w(i, j) \times I(i, j) \tag{3}$$

Among them: I (x, y) as the input image; Filter in (x, y) has a weight w (I, j); Pixel (x, y) as the scope of areas; W_p as the normalized parameter; I_0 (x, y) as the filter after image. Bilateral filtering coefficient by weights in the spatial domain and image domain weight value, That is:

$$w = w_s + w_r$$

$$w_s = \exp \left(- \frac{(i-x)^2 + (j-y)^2}{2\delta_s^2} \right)$$

(4)

$$w_r = \exp \left(- \frac{(I(i, j) - I(x, y))^2}{2\delta_r^2} \right)$$

(5)

(4) and (5), and the said standard deviation based on gaussian function [7].



(a) Depth map



(b) The depth of the image after bilateral filtering

Fig.3. Depth of the original image and the depth of the image after bilateral filtering

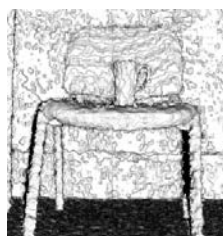
SIFT algorithm is a kind of image local feature description algorithm based on scale space, by Lowe DG in the existing technology based on invariant feature detection method is proposed on the basis of. SIFT feature description is only partial feature of the image, the characteristics of scale keep invariance, scaling, rotation, brightness change of noise, the change of perspective, affine transformation and maintain stability.

When two image feature point descriptor is generated, using feature vector of the Euclidean distance as the feature points in two images similarity measure. Take some feature points in an image, in another image Euclidean distance with the first two recent feature points, if recent distance and time close ratio is less than the threshold value of a setting, received the pair of matching points. Lower threshold, the number of matching points will reduce, but more stable.

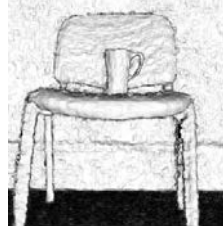
Test results

By measuring the chair to verify the validity of the algorithm in this paper,

by the local image and reconstruction of device to get the chair to get point cloud data; Local image and get the chair again with the previous measurement image overlap areas, using the algorithm of SIFT feature matching, using SIFT algorithm for image matching feature points between not precise enough, after purification of consistency, can eliminate false matching points, well known only from coarse registration will cause the registration precision is not enough, can't meet the requirements, with the improved SIFT algorithm after registration, can get a good registration accuracy, the result is shown in figure.



(a) Before optimization,3D



(b) After the optimization,3D

Fig.4. The generated point cloud and three-dimensional figure

Conclusion

This paper proposes a data based on Kinect depth and color data of 3D modeling method, this method can create according to the requirements of the simulation scene color with different levels of fine 3D model, can be seen from the above several types of the object reconstruction results, this method is convenient and practical, the reconstruction effect is better, to establish the three-dimensional model can meet the basic requirements of simulation scenario. Methods also have some limitations, this paper mainly includes: affected by calculation, require computer configuration is too high, to get relatively fine

image 3D reconstruction of time affected; To access acquisition precision co., LTD., the object surface finely are easy to be ignored; Kinect of depth measuring range is limited, effective minimum distance is about 0.35 m, which limits the size of the modeling objects, when the object through the hours it will not be able to obtain dense point cloud.

Acknowledgement

In this paper, the research was sponsored by The Ministry of education Chunhui project (Project No. Z2014137) and Youth Fund Project of Luoyang Institute of Science and Technology (Project No. 2010QZ16).

References

- [1] P. A. Karkantzakos. Time of flight and range of the motion of a projectile in a constant gravitational field[J]. *Journal of Engineering Science and Technology Review*, 2009, 2(1).
- [2] Ming-Qiang Huang, Zhen-Ya Wang, Li-Qing Hao et al. Laser aerosol time-of-flight mass spectrometry analysis of individual aerosol particles from photooxidation of toluene[J]. *Optica Applicata*, 2006, 36(1).
- [3] Newcombe, Richard A. Izadi, Shahram, Hilliges, Otmar et al. KinectFusion: Real-time dense surface mapping and tracking[C] // 2011 10th IEEE International Symposium on Mixed and Augmented Reality. [S.l.]: [s.n.], 2011:127-136.
- [4] Rebekah, Xu hua rong, Zhan-yi hu, Rapid object reconstruction based on GPU and access it [J]. *Journal of automation*, 2012, 38.
- [5] Alexiadis D, Zarpalas D, Daras P. Real-time, full 3-D reconstruction of moving foreground objects from multiple consumer depth cameras[J]. 2013.
- [6] Newcombe R A, Dabison A J. Izadi S, et al. Kinect Fusion: Real-time dense surface mapping and tracking[C] // Proceedings of 2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR). IEEE, 2011:127-136.

[7] Lowe D G. Distinctive image features from scale-invariant keypoints [J].
International Journal of Computer Vision, 2004,60(2):91-110.