Key Technologies of Seam Fusion for Multi-view Image Texture Mapping Based on 3D Point Cloud Data

Xiaoxue Yang*
Beijing Polytechnic College, Beijing 100042, China

*Corresponding author e-mail: yangxiaoxue@bgy.edu.cn

Abstract. With the rapid development of computer technology and measurement technology, three-dimensional point cloud data, as an important form of data in computer graphics, is used by light reactions in reverse engineering, surveying, robotics, virtual reality, stereo 3D imaging, Indoor scene reconstruction and many other fields. This paper aims to study the key technology of 3D point cloud data multi-view image texture mapping seam fusion, and propose a joint coding and compression scheme of multi-view image texture to replace the previous independent coding scheme of applying MVC standard compression to multi-view image texture. Experimental studies have shown that multi-view texture depth joint coding has different degrees of performance improvement compared with the other two current 3D MVD data coding schemes. Especially for Ballet and Dancer sequences with better depth video quality, the performance of JMVD is very obvious. Compared with the KS_IBP structure, the gain can reach as high as 1.34dB at the same bit rate.

Keywords: 3D Point Cloud Data, Multi-View Images, Texture Mapping, Texture Fusion

1. Introduction
The computer virtualization representation of cloud point data includes two aspects: cloud point data model reconstruction and texture mapping of the reconstructed model [1-2]. Among them, the reconstruction of data cloud points is mainly to solve the problem of cloud point data holes, so as to form a smooth connection surface between the points, thereby forming a virtual entity visually [3-4]. The texture mapping of the cloud point model is mainly to highlight the authenticity and good vibration of the model, and it is also considered to be a very important virtual reality technology [5-6]. The high-precision, low-quality spatial point data obtained by Didin in measurement technology, due to its large quantity, scattered distribution, lack of organization, etc., makes high-precision cloud point data reconstruction and point model cloud reconstruction become people's attention hot spot. In the field of graphics computing, spatial information processing, computer vision and computer simulation [7-8].

Considering the recording relationship between the geometric reconstruction model and the multi-projection image, there are two main forms of expression for texture mapping. One of them is to process cloud point data immediately. The color of a specific point on the cloud point is the weighted
average of the color values of the corresponding points in all input images. The weight depends on the corresponding point of the input image. The angle between the observation direction and the normal direction of the space point [9]; another method is to first convert the cloud point data into a polygonal mesh structure, and then obtain the corresponding texture map for each mesh surface, depending on how to deal with it textured seams [10].

This paper introduces the inter-layer motion image information prediction technology in the svc coding standard through the independent digital compression model system using mvc encoding for the existing multi-view texture image video and depth image-based video, and integrates the two. The ground is tightly integrated. Utilizing the complementary functions between the prediction of inter-layer motion information and the prediction of inter-layer disparity information, a network model with high compression ratio, low complexity and stable feature performance is constructed with multiple views and texture depths that are jointly encoded, which effectively improves The compression speed and accuracy of 3D images.

2. Research on the Key Technology of Seam Fusion of Multi-View Image Texture Mapping Based on 3D Point Cloud Data

2.1 Application Fields of 3D Point Cloud Data Technology

(1) Reverse engineering

Reverse engineering uses 3D point cloud data acquisition technology to survey the surface of the target object, reconstruct the 3D geometric model of the object surface, generate 3D digital design documents necessary for product manufacturing, and use CAD/CAM technology to adjust and optimize, and better improve and Create a better product. In the reverse engineering of 3D point cloud data, the preprocessing technology of massive point cloud data is of vital importance to the influence of 3D reconstruction, and it is a prerequisite and important link in reverse engineering.

(2) Machine vision

Accurate perception, correct cognition and recognition of target scenes are the primary prerequisites for mobile robots to perform tasks in specific environments. The current mobile robot's recognition of the scene and the target object in the scene is mainly based on visual image processing technology, but the acquisition of visual images is too much interfered by environmental factors, especially the quality of the image obtained in harsh environments such as dim light. With the development of acquisition equipment, 3D point cloud data processing technology has entered the field of machine vision research. Robots can better realize the cognition and identification of scenes and target objects through the acquired depth information. Therefore, 3D point cloud data processing Technology will become a strong support for applications and research in the field of robotics.

(3) Virtual reality

By simulating the functions of human sensory organs, the human body can be immersed in a computer-generated virtual scene in terms of audiovisual and touch, and the human body can use natural language and physical behavior to interact with the virtual scene in real time, creating a super-realistic virtual reality information space is a computer man-machine interface for visual and entity behavior simulation of three-dimensional dynamic scenes with the characteristics of multi-source information fusion, immersion, conception and interaction. Virtual reality technology currently has many mature application cases, such as virtual surgery training devices in medical laboratories, 3D virtual reality games, emergency deductions, etc., which have realized the existence of virtual reality and realistic human-computer interaction experience.

2.2 Methods of Texture Mapping

(1) Forward texture mapping

By analyzing the existing algorithms, it can be known that forward texture mapping is a sequential processing of the input image sequence, and each image undergoes a transformation from two-dimensional to three-dimensional and back to two-dimensional. Therefore, some points on the screen
space will appear holes, and some will overlap. Relatively speaking, the model obtained by this method is rather confusing. But forward texture mapping also has its own advantages: the processing of texture images is performed sequentially; it does not need to take up a lot of storage space.

(2) Reverse texture mapping
Reverse texture mapping is just the opposite of forward texture mapping, which refers to the mapping from screen space to texture space. Although this method requires a large storage space, the quality of the obtained model is high, the occurrence of confusion and other situations is avoided well, and the application range is wide.

(3) Two-step texture mapping
The first step: the mapping of the texture image to the intermediate surface. The corresponding relationship between the texture image and the intermediate surface is established. The intermediate surface is generally a simple representable surface, such as a plane, a spherical surface, a cylindrical surface, and so on. The corresponding mapping formula is \( T(u,v) \rightarrow T'(x',y',z') \), and this step is also called S-mapping. The second step: the mapping of the intermediate surface to the three-dimensional model. Correspondingly The mapping formula is \( T'(x',y',z') \rightarrow 0(x,y,z) \), and this step is also called O-mapping.

(4) Spherical texture mapping
Spherical texture mapping refers to a process of mapping the texture image and the surface of the sphere. The specific method can be described as follows: We define a coordinate system based on texture space. The point on the horizontal axis is the longitude of a sphere. Each point above represents the constant vertical line \( u \). After being mapped to the sphere, the longitude coordinate \( \theta \) and the corresponding point is a meridian. The vertical axis is the latitude coordinate line of the sphere, and each point above represents the horizontal line \( v \) of normal value. After the latitude coordinate \( \phi \) is mapped to the sphere, the latitude coordinate \( \phi \) corresponds to the latitude and longitude line. In this way, the problem of solving the texture mapping of the spherical surface can be equivalent to the solving of the texture coordinates \((u,v)\) and the longitude and latitude coordinates \((\theta,\phi)\) of the spherical surface.

2.3 Multi-View Image Fusion Quality Evaluation Standards
(1) Qualitative evaluation
Qualitative evaluation, that is, subjective evaluation, is the perceptual perception of the fusion quality by the evaluator, and a score evaluation of the fusion quality based on the human perception results and certain evaluation criteria, such as judging whether the image detail features after multi-view image fusion keep it well, whether the image is blurred, etc. Since qualitative evaluation is affected by observers, image types, applications, etc., a statistical evaluation process is required, that is, a large number of people are required to participate in the evaluation process of multi-view image fusion quality, otherwise subjective evaluation is meaningless.

(2) Quantitative evaluation
When people evaluate the fusion quality of multi-view images, it is difficult to grasp the evaluation criteria and scales. On the one hand, it is not clear to the general observer what the fusion image should be, and there is no good reference for comparison; on the other hand, Image fusion has different application scenarios and purposes, such as:

① In order to obtain clearer and more detailed images in visual effects;
② In order to get a more comprehensive description of the scene;
③ In order to realize the precise identification of the target;
④ To judge the status and location of the target;
⑤ In order to judge the authenticity of the target, etc.

Therefore, for those who perform subjective evaluation of multi-view image fusion, when faced with such a complex application background and purpose, even if they have considerable professional
knowledge background, they may not be able to make a reference evaluation. Therefore, expand the quality of multi-view image fusion. The qualitative and objective evaluations are very meaningful.

3. Experimental Research

3.1 Research Purpose
In order to test the key technology of multi-view image texture mapping seam fusion, we first built a codec system for actual compression of MVD data. The system is built based on the SVC standard reference software JSVM, and as described in the previous chapter, the inter-layer intra-prediction mechanism, the inter-layer residual prediction mechanism, and the limited intra-prediction mode mechanism for single-loop decoding are closed. Deep coding-independent modules improve the coding and decoding speed, and have many advantages such as high coding efficiency, low complexity, good forward compatibility, and strong scalability.

3.2 Experimental Procedure
(1) JMVDC: Based on the key technology of multi-view image texture mapping seam fusion implemented by the SVC standard reference software JSVM, the inter-view prediction uses a simple KS_IBP prediction structure. Therefore, the inter-layer disparity information prediction is only enabled in the anchor frame. Inter-layer motion information prediction is turned on for all non-anchor frames.

(2) MVC1: Multi-view texture video and multi-view depth video are respectively coded using the MVC standard reference software JMVC, and the inter-view prediction structure adopts a simple KS_IBP structure.

(3) MVC2: Multi-view texture video and multi-view depth video are respectively coded using the MVC standard reference software JMVC, and the inter-view prediction structure adopts a more complex AS_IBP structure.

3.3 Research on Image Fusion Algorithm
Image and fusion technology is an image processing technology that combines various image sensors working at different wavelengths and different imaging principles to form image information in the same scene through a specific calculation method. The fused image can have richer and more comprehensive image details, making the overall image more credible.

The number of column vectors in Y obtained after a certain orthogonal transformation T is far from that of the original matrix X, and the sum X restored by Y has the smallest mean square error, then this orthogonality becomes the PCA transformation, and the formula is

$$\epsilon(\Delta X) = \min E[\|X - \hat{X}\|^2]$$

Among them: is the normalized vector.

Recovering X by Y is shown in formula (2). For the convenience of expression, suppose the dimension of the Y vector is M and the dimension of the X vector is N.

$$X = T^T Y = [\phi_1 \phi_2 \ldots \phi_N]^T = \sum_{i=1}^{N} \phi_i y_i$$

Since Y is a vector containing only M components, can be written as:

$$\hat{x} = \sum_{i=1}^{M} \phi_i y_i + \sum_{i=1}^{N} \phi_i b_i$$

4. Experimental Analysis
The actual coding performance comparison with the current MVC independent compression scheme is the most direct way to verify the compression performance of multi-view texture depth joint coding. The coding scheme tested in this article is shown in Table 1:
Table 1 shows the performance gain of multi-view texture depth joint coding and two independent MVC coding methods. From the data in the table, it can be seen that the multi-view texture depth joint coding has different degrees of performance improvement compared with the other two current 3D MVD data coding schemes, especially for the Ballet and Dancer sequences with better depth video quality, the performance of JMVDC obviously, the highest gain can be as high as 1.34dB compared with the KS_IBP structure under the same bit rate. Because Balloons' depth sequence is of poor quality and low correlation with texture sequence, the application of this coding scheme to the sequence is not obvious, which also shows that the quality of depth sequence will greatly affect the coding of depth video effectiveness.

The figure below shows the corresponding RD performance curve.

![Figure 1. Ballet](image-url)
As can be seen from the RD curves in Figure 1, Figure 2, and Figure 3, unlike the scheme that uses AS_IBP to improve coding performance, the compression performance of multi-view texture depth joint coding is always higher than the other two in the normal bit rate range. The coding scheme, especially the compression performance gain at the low bit rate end, is particularly prominent. The above test is performed under the condition of ensuring that the depth video has 4 different
quantization parameters, so that the cubic spline is used to calculate the performance difference between the three different encoding methods.

5. Conclusions
With the rapid development of science and technology such as computer graphics and computer vision in the field of three-dimensional media applications, a series of key technologies such as stereo vision, three-dimensional reconstruction, and virtual reality have gradually entered people's work and daily life. Three-dimensional media video, a dynamic form of media that can record three-dimensional situations, is no longer unfamiliar. Multi-view texture has been widely regarded as the most flexible, and it can also be said to be the most promising way of expressing large amounts of data in 3D video. However, it has the advantages and disadvantages of a large amount of data, which brings them expensive data storage/transmission costs. High-efficiency compression algorithms must be adopted to effectively reduce the amount of data and ensure that they can be used smoothly in various practical uses. This article first introduces some related fields in which cloud data technology is widely used in 3D point cloud information technology, and then introduces the methods and approaches for texture mapping in multi-view images and the evaluation standards for image fusion quality. Finally, an experiment is constructed. The joint coding model framework of texture depth in multi-view images with high compression ratio, low complexity and stable reliability, effectively improves and enhances the data compression quality and accuracy of multi-view images.

References
[1] Rui Z, Li G, Li M, et al. Fusion of images and point clouds for the semantic segmentation of large-scale 3D scenes based on deep learning. Isprs Journal of Photogrammetry & Remote Sensing, 2018, 143(SEP.):85-96.
[2] H Shen, Sun W, Fu J. Multi-view online vision detection based on robot fused deposit modeling 3D printing technology. Rapid Prototyping Journal, 2019, 25(2):343-355.
[3] Hong H. Effect of 3D depth on the resolution and the observed image for auto-stereoscopic multi-view 3D display. Journal of the Society for Information Display, 2018, 26(7-9):401-406.
[4] A H Y D, B Q M B, C Y H D, et al. Generation and evaluation of 3D digital casts of maxillary defects based on multisource data registration: A pilot clinical study. The Journal of Prosthetic Dentistry, 2017, 118(6):790-795.
[5] Wen C, Wang W. The Research on Integrated Production of Large Scale 3D Products Based on Mass Multi-angular Images. Journal of Physics: Conference Series, 2020, 1624(5):052010 (6pp).
[6] Liu Y, Peng M, Swash M R, et al. Holoscopic 3D Microgesture Recognition by Deep Neural Network Model Based on Viewpoint Images and Decision Fusion. IEEE Transactions on Human-Machine Systems, 2021, PP(99):1-10.
[7] Su M, Zhang X. 95.3D Data Model Acquisition of Binocular Vision Medical Instrument Based on Multi Data Fusion. Boletin Tecnico/Technical Bulletin, 2017, 55(20):667-672.
[8] Yan T, Chen B, Liu F, et al. Multi-focus Image Fusion Model for Micro 3D Reconstruction. Jisuanji Fuzhu Sheji Yu Tuxingxue Xuebao/Journal of Computer-Aided Design and Computer Graphics, 2017, 29(9):1613-1623.
[9] Dittrich A, Weimann M, Hinz S. Analytical and numerical investigations on the accuracy and robustness of geometric features extracted from 3D point cloud data. Isprs Journal of Photogrammetry & Remote Sensing, 2017, 126(APR.):195-208.
[10] Maimaitimin M, Watanabe K, Maeyama S. Stacked convolutional auto-encoders for surface recognition based on 3d point cloud data. Artificial Life & Robotics, 2017, 22(2):259-264.