3D Panoramic Mosaicking to Suppress the Ghost Effect at Far-Range Scene for Urban Area Visualization

Jaechoon CHON[1], Yongcheol SUH[2], Eihan SHIMIZU[3], and Ryosuke SHIBASAKI[1]

1: Center for Spatial Information Science, The University of Tokyo
4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, Japan
E-mail: jjc7151_at_hanmail.net
2: Pukyong National University, Korea
3: Dept. of Civil Engineering, The University of Tokyo, Japan

Abstract: 3D image mosaicking is useful for 3D visualization of the roadside scene of urban area by projecting 2D images to the 3D planes. When a sequence of images are filmed from a side-looking video camera passing far-range areas, the ghost effect in which same objects appear repeatedly occurs. To suppress such ghost effect, the far-range areas are detected by using the distance between the image frame and the 3D coordinate of tracked optical flows. The ghost effects are suppressed by projecting the part of image frames onto 3D multiple planes utilizing vectors passing the focal point of frames and a vertical slit. The vertical slit is calculated by utilizing the first and last frames of the far-range areas. We demonstrated algorithm that creates efficient 3D panoramic mosaics without the ghost effect at the far-range area.

Keywords: 3D Image Mosaicking, 3D Panoramic Mosaicking, Cross-Slit Projection, Ghost Effect.

1. Introduction

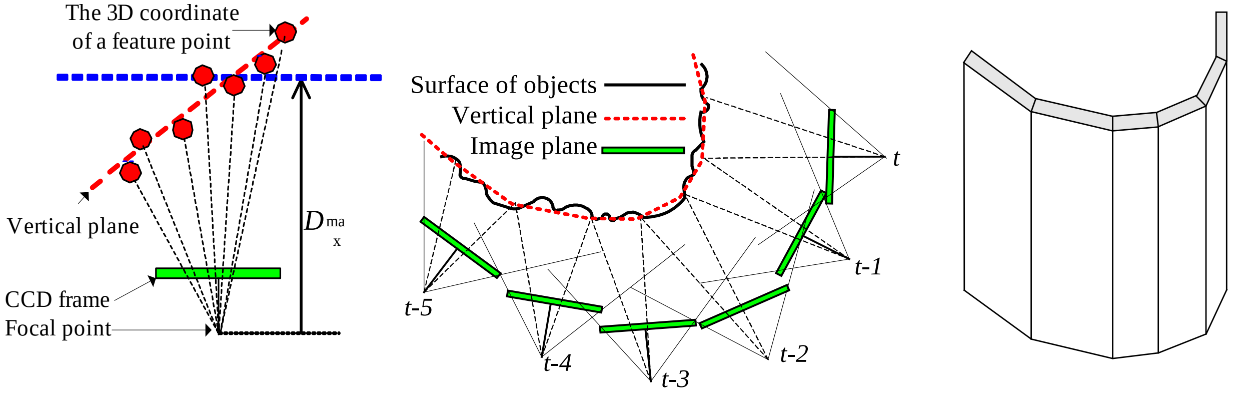
Environment visualization becomes popular gradually on Internet websites (or mobile systems) for tourism or virtual reality. So far, most of these websites provide still image-based visualization. Therefore, it is monotonous due to its fixed angle and viewpoint. Recently, panoramas have appeared to give more impressive visualization. Although the panorama enables the user to pan and zoom, the viewpoint of the panorama remains fixed. More impressive form of environment visualization is required to allow the user to view environment scenes from arbitrary viewpoints and angles. Three-dimensional Geographic Information System (3D GIS) data well meet the requirement.

A main approach for creating 3D GIS data is detailed 3D reconstruction of a scene [1-4]. Unfortunately, it is difficult to apply these detailed 3D surfaces to current Internet or mobile systems, because of their limited real-time transmission speeds. The image mosaicking technique is considered to be another efficient approach for environment visualization on websites. However, direct application of ordinary image mosaickings cannot create 3D sensation [5-8].

The goal of the 3D image mosaicking is to give 3D feeling for urban visualization on internet websites or mobile systems [9-11]. The created 3D image mosaics are textured 3D vector data generated from a side-looking camera along a road in a city or town area. The 3D feeling can be obtained through watching 3D vectors and textured image slits. The 3D vectors and textured image slits give imagination of the global abstract and detail part of objects, respectively. The 3D image mosaicking uses a roadside scene acquired by a side-looking video camera as a continuous set of textured vertical planar faces named "multiple projection planes". The scene geometry is approximated to multiple vertical planes using sparsely distributed optical flows. These vertical planes are concatenated to create an approximate model on which the images could be back-projected as textures and then blended together using seam-line or cut-line detection algorithm [11,12].

If the multiple projection planes are created around the far-range area in the same way around the close-range area, then the ghost effect will occur [13]. To suppress ghost effect, we proposed 3D panoramic mosaicking as on an expanded concept of the crossed-slits projection [7]. The far-range areas are detected by using the distance between the image frame and the 3D coordinate of the detected optical flows. The ghost effects are suppressed by projecting the part of image frames onto 3D multiple planes utilizing vectors passing the focal point of frames and a vertical slit. The vertical slit is calculated by utilizing the first and last frames of the far-range areas.

2. 3D Image Mosaicking Using Multiple Projection Planes



(a) (b) (c)

Fig. 1. Concatenation of vertical planes formed with the feature points of each frame; (a) The vertical plane of each frame, (b) multiple projection planes in 2D space, (c) multiple projection planes in 3D space

The multiple projection planes around the close-range area involved a continuous set of vertical planes. The vertical plane for each frame was calculated from the sparsely distributed feature points belonging to a frame. The least median of squares (LMedS) was applied to estimate the projection plane, which was calculated from the regression line, $Z = aX + b$, in (x,z) coordinates. During the estimation, only the feature points closer than $D_{\max}$ were used (Fig 1(a)). The distance $D_{\max}$ was defined as being 1.5 times the average distance between the focal point of the first frame and the feature points with 3D coordinates that appeared in the first and second frames, and was specified by trial and error. Assume that the outer covering of urban objects and the position of a side-looking and moving video camera are built like the thick curve and rectangles shown in Fig. 1(b). In this case, the multiple planes will be designed as a dotted curve in Fig. 1(b) and will be presented as linked planes in 3D space, as in Fig. 1(c).

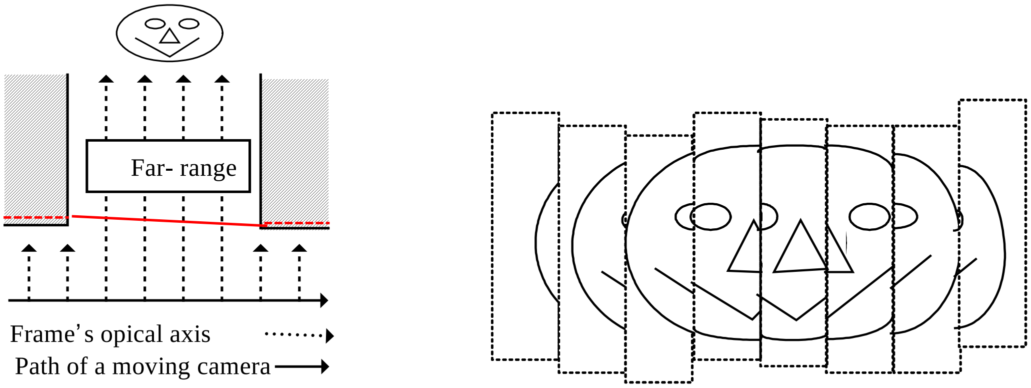
3. 3D Panoramic Mosaicking Using a Expended Crossed-Slit Projection

In the case where the distance between one frame and a building exceeded the distance $D_{\max}$, the frame was determined to be in the far-range area frame. For the far-range area frame, a virtual vertical plane was needed to create a multiple projection plane. The virtual vertical plane was calculated using the path of the camera and the two neighboring vertical planes in the close-range area. The terms P_t and P_{t+n} depicted in Fig. 2(a) represent the right-hand and the left-hand endpoints of the two neighboring vertical planes in the close-range area, respectively. The virtual vertical plane was designed in the same way as the camera path. The virtual vertical plane of Frame t+k was set as the line passing through Positions P_{t+k-1} and P_{t+k+1}. Position P_{t+k} was computed as follows:

$$P_{t+k} = P_t + \frac{l(k)(P_{t+n} - P_t)}{L}, \tag{1}$$

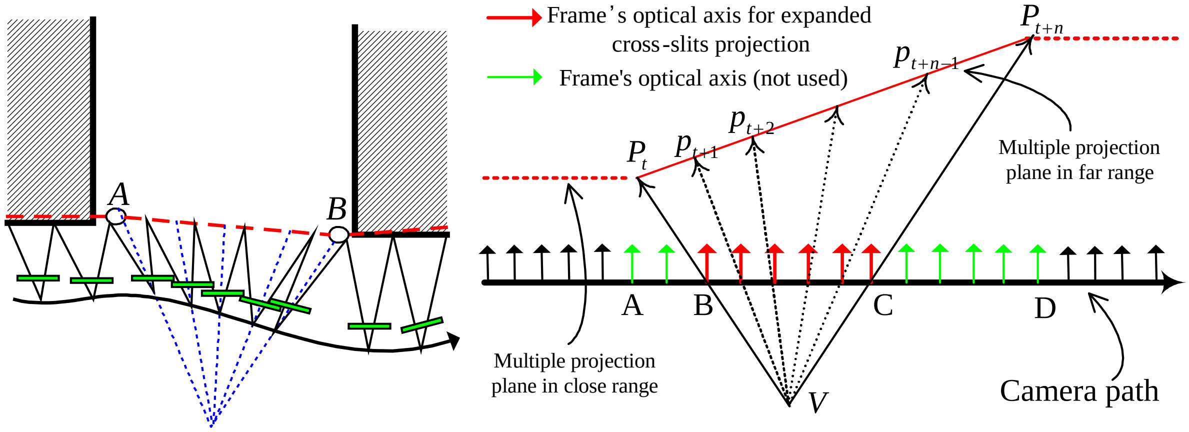
where $L = \sum_{i=t}^{t+n} | F_i - F_{i+1} |$ and $l(k) = \sum_{i=t}^{t+k} | F_i - F_{i+1} |$ are the lengths along the path of the transverse view camera from Frame t to Frame t+n and to frame t+k, respectively. The term F_{t+k} denotes the focal point of Frame t+k, where k = 1…*n*.

If the same way used around close-range areas is directly applied to the virtual vertical planes to project images onto the virtual vertical planes as textures, then certain urban objects will appear repeatedly in a continuous set of textured planar faces (see Fig. 2) [13]. This is known as the "ghost effect". To reduce the ghost effect, we proposed the use of an expanded crossed-slits projection technique [7] (see Fig. 3(a)).



(a) A camera passing an intersection area (b) Ghost artifacts

Fig. 2. Ghost effect by applying the same way used in close-range areas to virtual vertical planes directly.



(a) (b)

Fig. 3. Concept of expanded crossed-slits projection technique in a far-range area.

The technique determined Point p, as shown in Fig. 3(b). Point p is the point at the intersection of the virtual vertical plane with a ray passing through the focal point and a vertical slit, V. Vertical slit V was calculated using Frames B and C. Frame B is the final frame that includes Point P_t, and Frame C is the first frame that includes Point P_{t+n}.

Vertical slit V was calculated by using two vectors, $\overrightarrow{P_t B}$ and $\overrightarrow{P_{t+n} C}$, as below

$$\begin{bmatrix} X_V \\ Z_V \end{bmatrix} = \begin{bmatrix} X_{P_t} \\ Z_{P_t} \end{bmatrix} + K_t\left(\begin{bmatrix} X_B \\ Z_B \end{bmatrix} - \begin{bmatrix} X_{P_t} \\ Z_{Pt} \end{bmatrix}\right) = \begin{bmatrix} X_{P_{t+n}} \\ Z_{P_{t+n}} \end{bmatrix} + K_{t+n}\left(\begin{bmatrix} X_C \\ Z_C \end{bmatrix} - \begin{bmatrix} X_{P_{t+n}} \\ Z_{P_{t+n}} \end{bmatrix}\right), \tag{2}$$

where $K_{t+n} = \dfrac{X_{P_t} + K_t\left(X_B - X_{P_t}\right) - X_{P_{t+n}}}{X_C - X_{P_{t+n}}}$ and $K_t = \left(Z_{P_{t+n}} + \left(\dfrac{X_{P_t} - X_{P_{t+n}}}{X_C - X_{P_{t+n}}}\right)\left(Z_C - Z_{P_{t+n}}\right) - Z_{P_t}\right) / \left(Z_B - Z_{P_t} - \dfrac{\left(X_B - X_{P_t}\right)\left(Z_C - Z_{P_{t+n}}\right)}{X_C - X_{P_{t+n}}}\right)$ are the proportionality coefficients. The point p on each virtual vertical plane was determined as an intersection point

$$\begin{bmatrix} X_p \\ Z_p \end{bmatrix} = k\left(\begin{bmatrix} X_F \\ Z_F \end{bmatrix} - \begin{bmatrix} X_V \\ Z_V \end{bmatrix}\right) + \begin{bmatrix} X_V \\ Z_V \end{bmatrix} \tag{3}$$

where F is the focal point of each frame, $k = \dfrac{aX_V + b - X_V}{Z_{F_t} - Z_V - a\left(X_{F_t} - X_V\right)}$ is a proportionality coefficient, and (a,b) represents the set of coefficients of the vertical plane of the frame.

4. Experimental Results

We used Visual C++ 6.0 and OpenGL to realize our proposed method and to display the results in 3D space. We applied our method to a real sequential image taken from a moving train equipped with a side-looking video camera. The extracted and tracked feature points, and the calculated camera positions and orientations are shown in Fig. 4(a). In Fig. 4, the red and green points denote feature points belonging to the close- and far-range areas, respectively. The computed vertical planes in the close-range area and the virtual vertical planes in the far-range area are denoted by the white and green lines in Fig. 4(b), respectively. The multiple projection planes obtained using the panoramic mosaicking algorithm to obtain the virtual vertical planes (green line) are shown in Fig. 4(c).

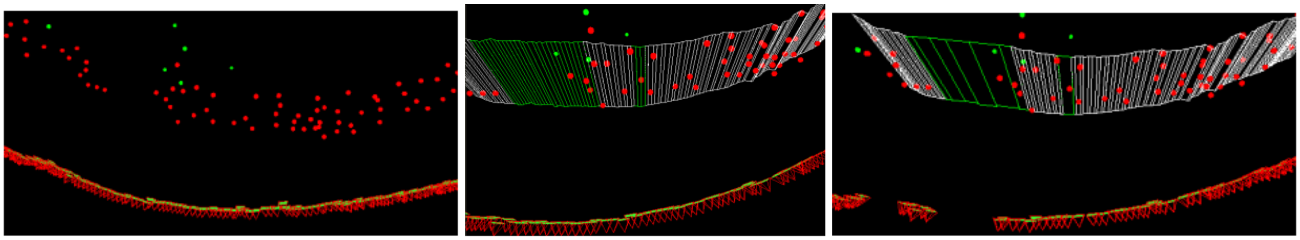
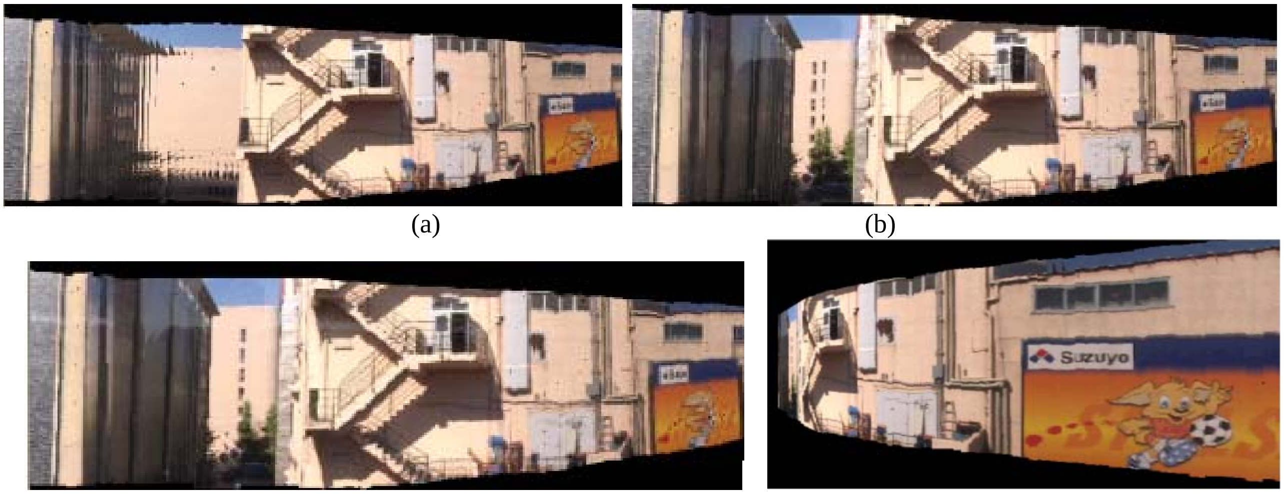

(a) (b) (c)

Fig. 4. The multiple projection planes around the far-range area.

Figure 5(a) shows the textured surface of multiple projection planes corresponding to the data shown in Fig. 5(b). It can be seen that the result around the far-range area includes the ghost effect. Most of the ghost effect disappeared in the results obtained when using the expanded crossed-slits projection algorithm, as shown in Fig. 5(b). Although the expanded crossed-slits projection algorithm was applied in this case, the result still included a partial ghost effect. To eliminate the ghost effect, we applied our proposed seam-line detection algorithm to the above results [11,12]. Figure 13(c) shows the results obtained using our proposed algorithm, and Fig. 5(d) shows the results obtained from another viewpoint.



(a) (b)

(c) (d)

Fig. 5. Back-projection of images using multiple projection planes around the far-range area.

5. Conclusions

When applying the 3D image mosaicking into frames capturing far-range areas, the ghost effect occurred in the created 3D image mosaics. To suppress the ghost effect, we proposed 3D panoramic mosaicking based on the crossed-slit projection. The ghost effects are suppressed by projecting the part of image frames onto 3D multiple planes utilizing vectors passing the focal point of frames and a vertical slit. The vertical slit is calculated by utilizing the first and last frames of the far-range areas. We demonstrated algorithm that creates efficient 3D panoramic mosaics without the ghost effect at the far-range area.

Since the textured projection plane of each frame consists of four 3D coordinates and a part of the image, the results obtained by using the proposed method can be the form of MPEG-4 data. One of the requirements of MPEG-4 composition for streaming of 3D worlds will be the Virtual Reality Modeling Language (VRML) that has made viewing 3D content on Internet websites possible. Therefore, these results as next generation navigation data can be widely applied to 3D virtual visualization and games on websites, cellular phones, and PDAs.

References

[1] Zhao, H. and R. Shibasaki, "A vehicle-borne urban 3D acquisition system using single-row laser range scanners," *IEEE Transaction onSMC Part B: Cybernetics*, vol.33, no. 4, 2003.

[2] Fruh, C. and A.Zakhor, "Data processing algorithms for generating textured 3D building facade meshes from laser scans and camera images," *Proc. 3D Data Processing*, Visualization and Transmission 2002, Padua, Italy, 2002.

[3] Pollefeys, M., R. Koch, M. Vergauwen, L. Van Gool, "Automated reconstruction of 3D scenes from sequences of images," *ISPRS Journal Of Photogrammetry And Remote Sensing*, vol. 4, no. 55, pp. 251-267, 2000.

[4] Sato, T., M. kanbara and N. yokoya, " Dense 3D reconstruction of an outdoor scene by hundreds-baseline stereo using a hand-held video camera. *In IJCV*, vol. 47, no. 1/2/3, pp. 119-129, 2002.

[5] Zomet, A., S. Pleeg and C. Arora, “Rectified Mosaicking: Mosaics without the Curl,” *International Conference on Computer Vision and Pattern Recognition*, II: pp. 459–465, 2000.

[6] Zhu, Z., A. R. Hanson and E. M. Riseman, “Generalized parallel-perspective stereo mosaics from airborne video,” *IEEE Transactions on PAMI*, vol. 26, no. 2, pp. 226–237, 2004.

[7] Zomet, A., D. Feldman, S. Peleg and D. Weinshall, “Mosaicking new views: the crossed-slits projection,” *IEEE Transactions on PAMI*, vol. 25, no. 6, pp. 741–754, 2003.

[8] Rom, A., G. Garg and M. Levoy, “Interactive Design of Multi-Perspective Image for Visualizing Urban Landscapes,” *Proc. Visualization 2004*.

[9] Jeachoon CHON, Takashi FUSE, and Eihan SHIMIZU, “Urban Visualization through Video Mosaics Based on 3D Multibaselines”, *20th Congress of the International Society for Photogrammetry and Remote Sensing*, Istanbul, Turkey, 2004.

[10] Jaechoon CHON, Takashi FUSE, and Eihan SHIMIZU, “Video Mosaics in 3D Space”, *The 24th Asian Conference on Remote Sensing & 2003 International Symposium on Remote Sensing*, Busan, Korea, 2003.

[11] Jaechoon CHON, “3D image mosaicking using multiple projection planes”, *The University of Tokyo*, Ph.D thesis, 2005.

[12] Jaechoon CHON, Hyongsuk Kim, Takashi Fuse and Eihan Shimizu, “Visually Optimal Seam-Line Determination in Image Mosaicking”, *The 25th Asian Conference on Remote Sensing*, Chiang Mai, Thailand, 2004.

[13] Uyttendaele, M., A. Eden and R. Skeliski, “Eliminating ghosting and exposure artifacts in image mosaics,” *Proc. 2001 IEEE Computer Society Conference on CVPR 2001*, Vol. 2 , pp. 8–14, 2001.