Video Mosaics in 3D Space

Jaechoon CHON, Takashi FUSE, Eihan SHIMIZU

SHIMIZU Lab., Department of Civil Engineering, The University of Tokyo
Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-8656, Japan
jjc7151@planner.t.u-tokyo.ac.jp, fuse@planner.t.u-tokyo.ac.jp, shimizu@planner.t.u-tokyo.ac.jp

Abstract: Video mosaicing techniques have been widely used in virtual reality environments. Especially in GIS field, video mosaics are becoming more and more common in representing urban environments. Such applications mainly use spherical or panoramic mosaics that are based on images taken from a rotating camera around its nodal point. The viewpoint, however, is limited to location within a small area. On the other hand, 2D-mosaics, which are based on images taken from a translating camera, can acquire data in wide area. The 2D-mosaics still have some problems : it can't be applied to images taken from a rotational camera in large angle.

To compensate those problems , we proposed a novel method for creating video mosaics in 3D space. The proposed algorithm consists of 4 steps: feature-based optical flow detection, camera orientation, 2D-image projection, and image registration in 3D space. All of the processes are fully automatic and successfully implemented and tested with real images.

Keywords: Video mosaics, 3D space, Camera orientation, feature-based optical flow.

1. Introduction

Applications for video mosaics techniques are used in virtual reality environments, computer-game settings, and movie special effects. Especially in GIS field, video mosaics are becoming more and more common in civil engineering and administration that are representing urban environments, construction sites management, building permission and inspection, sign management, and facilities management. The mosaics algorithms can be mainly divided into three categories: a 360-degree panorama [1], a full sphere omni-directional image [2], and general video mosaics projected 2D space [3]. The spherical mosaics generated from the rotations relating images taken from a common optical center allow any number of images to be merged into a view (see fig.1) [2]. The panorama can be captured or generated in several ways. 1) By using a single camera and a special lens or some kind of mirror (conical, spherical, hyperboloidal, parabolpidal, etc.). Through the mirror a single image can view the entire scene, or at least a very large field of view. 2) By using multiple cameras, or one rotating camera, and stitching their images into a single panorama (see fig.2) [1]. The camera mounted with fixed optical center on an indexed pan tilt head, itself attached to a tripod base. In those kinds of mosaics, the viewpoint is limited to locations within a very small area, usually along a circle of a meter in diameter. Under more general motion other than rotation, the method of video mosaics for visual navigation uses images taken from a translating camera mounted on an airborne or hand- hold and can generate mosaics projected 2D space (see fig.3) [3]. Based-on 2D space projection, its result is not applied to images taken from a camera in large angle.

Fig. 1.Spherical mosaics

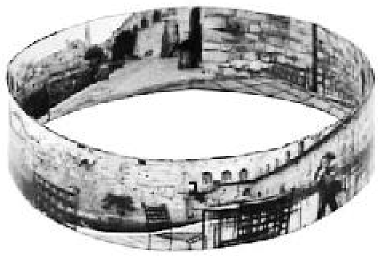


Fig. 2. Cylinder mosaics (panorama)

Fig. 3. General video mosaics in 2D space

To compensate those problems, we propose a novel method for creating video mosaics in 3D space (see fig.4). The proposed algorithm consists of 4 steps: feature-based optical flow detection, camera orientation, 2D-image projection, and image registration into 3D space. The overall flow of processing in our mosaicing system is illustrated in Fig. 5.

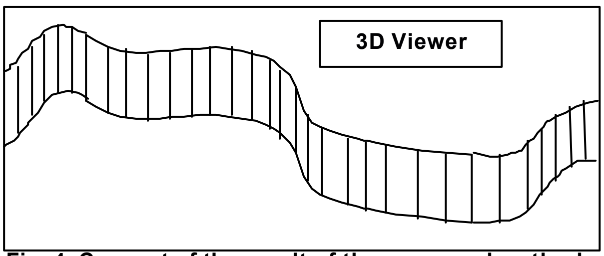


Fig. 4. Concept of the result of the proposed method

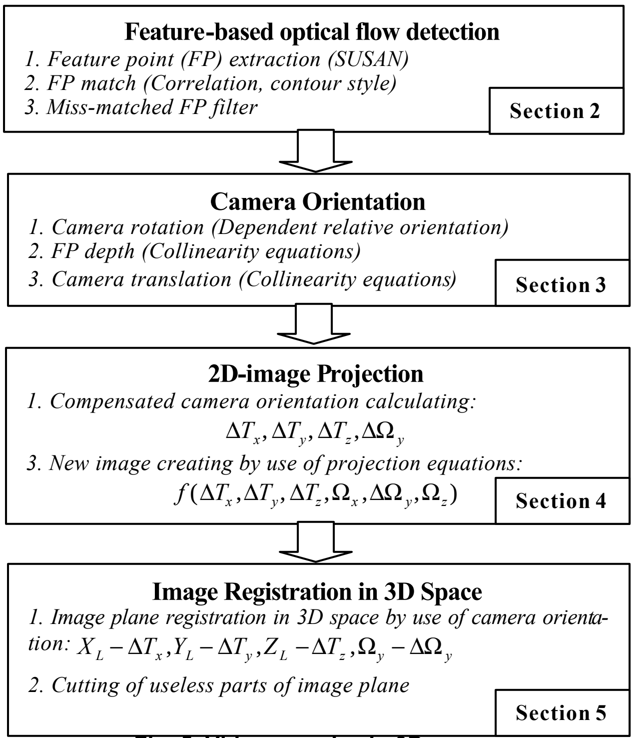


Fig. 5. Video mosaics in 3D space

2. Feature-based Optical flow detection

Camera orientation is computed utilizing optical flows which are obtained from sparsely located feature points that are detected using SUSAN algorithm [4]. Based on such feature points, the correlation and the contour style are computed and utilized to determine the best matching pair of feature points. The false optical flows, which are significantly different from others, are removed in the procedure of the repeated conversion using a median filter.

3. Camera Orientation

In our research, we assume that the interior orientation of the camera is already established and discuss about the exterior orientation. Let *P(X,Y,Z)* denote the Cartesian coordinates of a scene point with respect to the camera (see Fig. 6), and let *(x,y)* denote the corresponding coordinates in the image plane. The image plane is located at the focal length *f* from the focal point $o(X_L, Y_L, Z_L)$ of a camera. The perspective projection of a scene point *P(X,Y,Z)* on the image plane at a point *p(x,y)* is expressed by:

$$F_1 = x - fU/W \qquad (1)$$
$$F_2 = y - fV/W$$

Where $U = m_{11}(X - X_L) + m_{12}(X - Y_L) + m_{13}(X - Z_L)$,
$V = m_{21}(X - X_L) + m_{22}(X - Y_L) + m_{23}(X - Z_L)$,
$W = m_{31}(X - X_L) + m_{32}(X - Y_L) + m_{33}(X - Z_L)$,
X_L, Y_L, Z_L is a camera station, $M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$ is a 3×3 rotation matrix. To solve the relative orientation with the collinearity equations (Eq.1), we can transfer from the non-linearity of the equations to linearized form, Eq.1, using Taylor series approximations.

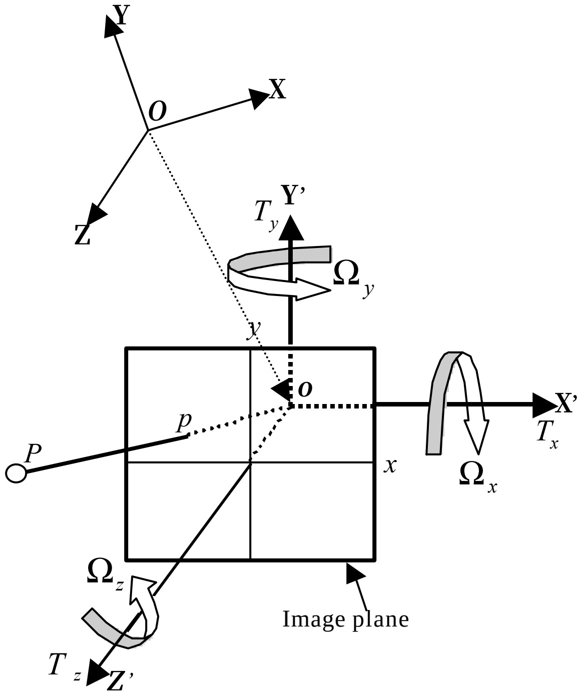


Fig. 6. The coordinate system. The camera orientation from the global coordinate system XYZ to the camera coordinate system X'Y'Z' is expressed by the classical form of collinearity equations. T_x, T_y, T_z are camera translations and $\Omega_x, \Omega_y, \Omega_z$ are camera rotations.

The two first images of the sequence are used to determine a reference frame. The world frame is typically aligned with the first frame. The camera orientation of 2nd frame can be calculate using the dependent relative orientation with the 5 optical flows and the $\tilde{X}_{L,2} = lX_{L,2}$ of current camera station, which is input by manual or average length of optical flows. We will show the method of l calculation in the geo-correction of camera orientation to get real camera translations as our future works. All of camera translations from 3rd frame to the last frame must be proportional to l. The camera rotation angles, $\Omega_X, \Omega_Y, \Omega_Z$, which are independent parameters to l can be calculated by dependent relative orientation per frame. From 3rd frame to the last image frame, the camera translation $X_{L,t}, Y_{L,t}, Z_{L,t}$ can be calculated the collinearity equations with the calculated camera rotations and the 2 optical flows tracked during 3 frames. The calculated camera orientation uses in the step of video mosaics in 3D space.

4. 2D-image Projection

The difficult point of creating mosaics in 3D space is from T_Z, Ω_X, Ω_Z excepted for the main camera motion, T_X, T_Y, Ω_Y. To remove the T_Z, Ω_X, Ω_Z effect to the creating mosaics in 3D space, we generate new image using projection function $f(\Delta T_x, \Delta T_y, \Delta T_z, \Omega_x, \Delta\Omega_y, \Omega_z)$. The compensated camera orientation, $\Delta T_x, \Delta T_y, \Delta T_z, \Delta\Omega_y$, are calculated as:

$$\Delta T_x = X_{L,t} - (X_{L,t+1} + X_{L,t-1})/2, \quad \Delta T_y = Y_{L,t} - (Y_{L,t+1} + Y_{L,t-1})/2 \tag{2}$$
$$\Delta T_z = Z_{L,t} - (Z_{L,t+1} + Z_{L,t-1})/2, \quad \Delta\Omega_y = \Omega_{y,t} - (\Omega_{y,t+1} + \Omega_{y,t-1})/2$$

Where t is image frame number. Fig 7 shows the process of generating a new image using the compensated camera orientation.

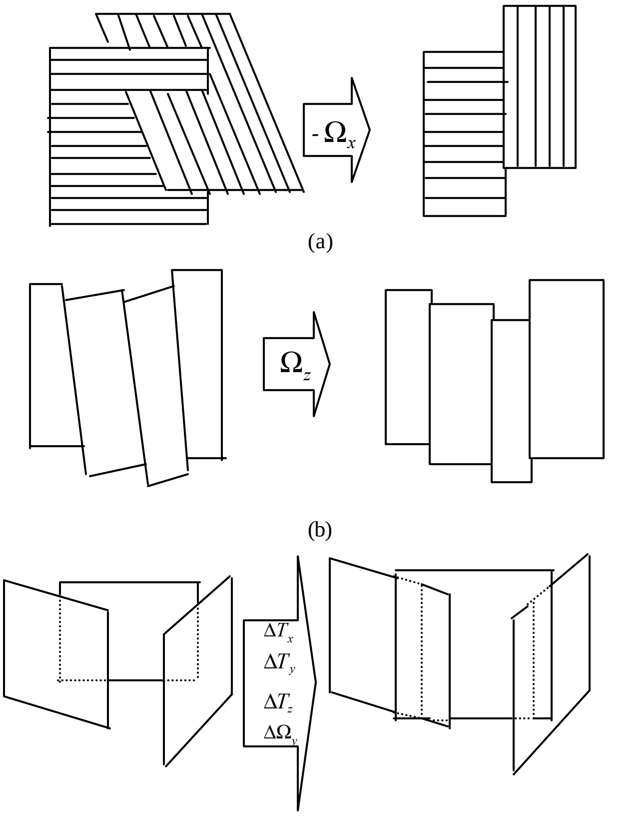
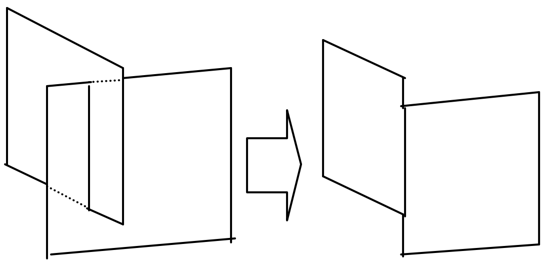


Fig.7 Creating new image by the compensated camera orentation. (a), (b), and (c) are Ω_x , Ω_z , *and* $\Delta T_x, \Delta T_y, \Delta T_z, \Delta\Omega_y$ compensations respectively.

5. Image Registration in 3D space

Image registration is composed of registering images into 3D space and cutting of useless parts of image planes. In the step of registering image into 3D space, the equation of image plane is calculated using camera orientation, ($X_L - \Delta T_x$, $Y_L - \Delta T_y$, $Z_L - \Delta T_z$, $\Omega_y - \Delta\Omega_y$) (see Fig. 8.a). A crossing line between two planes determines useless parts of image planes (see Fig. 8.b).



(a) (b)

Fig. 8. Cutting of useless parts of image plane.

6. Result

The propose method has been applied to real images taken from a digital camera for creating mosaics in 3D space. For realizing the proposed method, we use ***Visual C++ 6.0*** and ***OpenGL***. Fig. 9.a is about registering of images into 3D space. Fig. 9.b is about cutting of useless part of image planes and shows left side of a house. Fig 9.c shows the front side of the house and Fig. 9.d shows the result at view-port with downward angle.

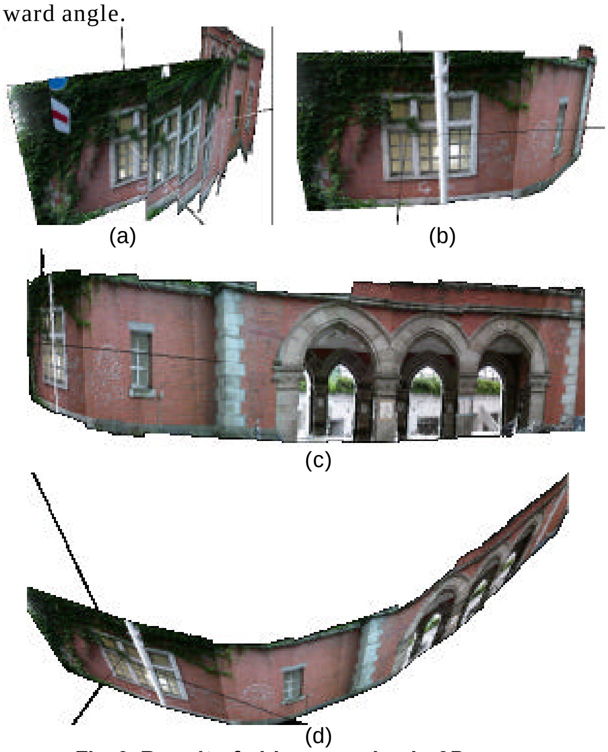


Fig. 9. Result of video mosaics in 3D space.

7. Conclusions

The novel method, creating video mosaics in 3D space, is proposed in this paper. The result created by the proposed method is more powerful environment visualization than that by the previous research such as panorama, spherical mosaics, and general mosaics in 2D space, because it compensates the defects of the previous research.

The results created by the proposed method will be of great use in GIS field and commercial software. We can also obtain inexpensive 3D data and textures from stereo video mosaics applied the proposed method.

References

[1] S. Coorg, S. Teller, " Spherical mosaics with quaternions and dense correlation, " *In IJCV,* vol. 37, No.3, pp. 259-273, June 2000.

[2] H.-Y. Shum, R. Szeliski, " Systems and experiment paper construction of panoramic image mosaics with global and local alignment, " *In IJCV*, vol. 36, No.2, pp. 101-130, 2000.

[3] Z. Zhu, A. R. Hanson, H. Schultz, E. M. Riseman, " Error Characterics of Parallel-Perspective Stereo Mosaics, " *IEEE Workshop on Video Registration (with ICCV01)*, Vancouver, Canada, July 13, 2001.

[4] S.M. Smith and J.M. Brady,"SUSAN- New Approach to Low Level Image Processing," *Technical Report TR95SMS 1c*, 1995.