Stitching and Reconstruction of Linear-Pushbroom Panoramic Images for Planar Scenes

Chu-Song Chen[1], Yu-Ting Chen[2,1], and Fay Huang[1]

[1] Institute of Information Science, Academia Sinica,
Nankang, Taipei, Taiwan, R. O. C.,
`song@iis.sinica.edu.tw`
[2] Department of Computer Science and Information Engineering,
National Taiwan University, Taipei, Taiwan, R.O.C.

Abstract. This paper proposes a method to integrate multiple linear-pushbroom panoramic images. The integration can be performed in real time. The technique is feasible on planar scene such as large-scale paintings or aerial/satellite images that are considered to be planar. The image integration consists of two steps: stitching and Euclidean reconstruction. For the image stitching, a minimum of five pairs of non-collinear image corresponding points are required in general cases. In some special configurations when there is column-to-column image correspondence between two panoramas, the number of image corresponding points required can be reduced to three. As for the Euclidean reconstruction, five pairs of non-collinear image corresponding points on the image boundaries are sufficient.

1 Introduction

The image mosaicing techniques have been applied in photogrammetry back in 80's for constructing large aerial and satellite photographs [16]. However, only until 90's, intensive researches on automatic construction of panoramic image mosaics were carried out in the fields of computer vision and computer graphics [20]. Two main features of the image mosaicing concept are the abilities to increase the resolution and to enlarge the field of view of a camera.

In computer vision, panoramic image mosaics often serve as representations of visual scene for a wide diversity of applications [14, 21, 6, 12, 8]. In the case when multiple panoramic images are provided, the depth information or other geometric properties of the 3D scene can be recovered [11, 18, 22]. In computer graphics, panoramic image mosaics play an important role in the technique of image-based rendering [2, 15, 4, 13, 10, 19]. The key idea of this technique is to rapidly generate novel views from a set of existing images. Panoramic images are also used widely in virtual reality systems to provide an immersive and photorealistic environment [1].

The traditional way of constructing a panoramic image mosaic is to align a set of matrix images of a common view by performing image transformations. When

images are acquired by unknown camera poses, there is a need to solve camera calibration problem before the image transformations may take place. Up until more recently, the line-camera concept for creating panoramic image mosaics were introduced [9, 5, 17], in which a sequence of slit (or line) images are used as the base elements instead of the matrix images for composing a panoramic image. Such a panorama is generated by joining together a sequence of line images side by side, and is called line-based panoramic images. The main advantage of using line images is to ease or even to avoid the camera extrinsic calibration problem so that panoramic mosaics can be generated simultaneously during the image acquisition process. One major trade-off of the line-based panoramic images is that the vertical image field of view is constrained by the resolution of the line image. There is lack of research on stitching two line-based panoramic images vertically to increasing the panorama's field of view.

Linear-pushbroom camera model was first introduced by R. I. Hartley in 1997 [5], which belongs in the line-based panorama category. The main characteristic of linear-pushbroom panoramic images is that the line-camera moves along a straight line during image acquisition. We investigated the possibility of integrating two such panoramic images under some additional geometric constrains. It is found that two linear-pushbroom panoramas are geometrically related by an affine transformation if they capture a common planar scene. In this paper, the integration of two linear-pushbroom panoramic images of planar scene is established for the first time. We conclude that only few image corresponding points are needed to perform the integration. This integration technique can be used for digitizing the large-scale 2D artworks in the museums or documenting the huge historical paintings on the wall. It can also be used on aerial or satellite images that are considered to be planar.

The paper is organized as follows: linear-pushbroom camera model is summarized in Section 2, in which the projection matrix and the LP-fundamental matrix of this camera model are recapped. The image integration method is reported in Section 3, in which the image transformation equations for image stitching and Euclidean reconstruction are elaborated respectively. Section 4 illustrates the integration result of a large-scale famous Chinese painting. Finally, conclusions are drawn in Section 5.

2   Review of Linear-Pushbroom Camera Model

A linear-pushbroom camera can be considered as a perspective line-camera[1] moving in a linear orbit with a constant velocity and a fixed orientation. As the line-camera moves, the view plane[2] sweeps out a region of space and 1-D images are captured. Finally, the whole 1-D images constitute a 2-D image which lies on a plane called the image plane in 3-D space.

---

[1] An optical system projecting an image onto a 1-D array sensor, typically a CCD array, is called a line-camera.

[2] The plane defined by the optical center and the sensor array is called a view plane.

An arbitrary point $\mathbf{x} = (x, y, z)^T$ in space is imaged and represented by two coordinates u and v. It has been shown in [5] that the linear-pushbroom camera model can be conducted as follows:

$$(u, wv, w)^T = \mathbf{M}(x, y, z, 1)^T \tag{1}$$

where w is a scale factor and $\mathbf{M}$ is a 3×4 projection matrix. The linear-pushbroom camera model, $(u, wv, w)^T = \mathbf{M}(x, y, z, 1)^T$, should be compared with the basic pin-hole camera model. An obvious difference is that the matrix of the pin-hold camera model is homogeneous; however, the linear-pushbroom camera matrix is not. That is, by multiplying linear-pushbroom camera matrix $\mathbf{M}$ with an arbitrary factor k, the v coordinate is unchanged while the u coordinate is scaled by k.

Consider a point $\mathbf{x} = (x, y, z)^T$ in space viewed by two linear-pushbroom cameras with projection matrices $\mathbf{M}$ and $\mathbf{M}'$. Let $\mathbf{u} = (u, v)^T$ and $\mathbf{u}' = (u', v')^T$ be the mappings of point $\mathbf{x}$ on these two panoramas respectively. A cubic equation $p(u, v, u', v') = 0$ called *fundamental polynomial* corresponding to these two cameras is introduced in [5], where the coefficients of p are determined by the entries of $\mathbf{M}$ and $\mathbf{M}'$. It concludes in the paper that there exists a 4×4 matrix F such that the equation $p(u, v, u', v') = 0$ may be rewritten as follows:

$$(u', u'v', v', 1)F_{4\times 4}(u, uv, v, 1)^T = 0.$$

The matrix F is called the *LP-fundamental matrix* corresponding to the linear-pushbroom camera pair $\{\mathbf{M}, \mathbf{M}'\}$. The matrix expresses the relationships between corresponding curves in these two linear-pushbroom panoramic images.

3   Integration of Linear-Pushbroom Panoramic Images

Consider two linear-pushbroom panoramic images (or LP-mosaic images) targeting at a planar scene (such as a large painting on the wall). In this section, we propose an image integration method to stitch these two panoramic images by image correspondence information. A key point to achieve this purpose is that, for an arbitrary point on the first image, can its corresponding point in the second image be determined and vice versa? In general, only curve-to-curve relationships can be established for two LP-mosaic images according to the theory of LP-fundamental matrix. Hence, the correspondence is ambiguous (to a curve) for a point specified in the first image and vice versa.

For perspective case, the point-to-point relationship can be established by imposing some scene constraints, such as co-planarity. The so-called planar homography [3] can be determined by four given pairs of image correspondences, and a complete point-to-point relationship can be exactly established from the planar homography thus determined. Planar homography has been widely adopted for applications such as image mosaicing [1, 21] and panorama construction [20].

As to the linear-pushbroom case, what we are interested is whether the point-to-point relationships can be determined as well when the scene is planar? If the

answer is yes, how many image correspondence pairs are required? These issues will be addressed in the following.

3.1   Image Stitching

Let $\mathbf{x}_i = (x_i, y_i, z_i)^T$ denote points in space that lie on a plane with planar equation $E : ax_i + by_i + cz_i + d = 0$ and are viewed by two linear-pushbroom cameras. Let $\mathbf{u}_i = (u_i, v_i)^T$ and $\mathbf{u}'_i = (u'_i, v'_i)^T$ be the mapping of point $\mathbf{x}_i$ on the source and the destination LP-mosaic images respectively. We intend to find transformation equations, which transform all the image points of the source panorama to the destination panorama, based on a set of corresponding points $\mathbf{u}_i$ and $\mathbf{u}'_i$.

According to linear-pushbroom camera model discussed in the last section (equation 1), we have

$$
\begin{aligned}
(u_i, w_i v_i, w_i)^T &= \mathbf{M}(x_i, y_i, z_i, 1)^T \\
(u'_i, w'_i v'_i, w'_i)^T &= \mathbf{M}'(x_i, y_i, z_i, 1)^T
\end{aligned}
\tag{2}
$$

where $\mathbf{M}$ and $\mathbf{M}'$ are 3×4 projection matrix associated to the source and the destination panoramic images, respectively. Let m_{jk} and m'_{jk}, where $1 \le j \le 3$ and $1 \le k \le 4$ denote the elements of $\mathbf{M}$ and $\mathbf{M}'$ respectively. Equation 2 plus the planar equation $E : ax_i + by_i + cz_i + d = 0$ can be rearranged into the following seven equations:

$$
\begin{cases}
u_i = m_{11}x_i + m_{12}y_i + m_{13}z_i + m_{14} \dots \text{(i)} \\
w_i v_i = m_{21}x_i + m_{22}y_i + m_{23}z_i + m_{24} \dots \text{(ii)} \\
w_i = m_{31}x_i + m_{32}y_i + m_{33}z_i + m_{34} \dots \text{(iii)} \\
u'_i = m'_{11}x_i + m'_{12}y_i + m'_{13}z_i + m'_{14} \dots \text{(iv)} \\
w'_i v'_i = m'_{21}x_i + m'_{22}y_i + m'_{23}z_i + m'_{24} \dots \text{(v)} \\
w'_i = m'_{31}x_i + m'_{32}y_i + m'_{33}z_i + m'_{34} \dots \text{(vi)} \\
ax_i + by_i + cz_i + d = 0 \dots \text{(vii)}
\end{cases}
\tag{3}
$$

Because w_i and w'_i are not necessary to be the same for each corresponding pair $\mathbf{u}_i$ and $\mathbf{u}'_i$, we may deal with these two variables separately. First, for a given $\mathbf{u}_i$, we use equations (i), (ii), (iii), (iv), and (vii) in equation 3 to find its corresponding u'_i to avoid the influence of w'_i, and the following equation holds:

$$
\begin{bmatrix}
m_{11} & m_{12} & m_{13} & m_{14} - u_i & 0 \\
m_{21} & m_{22} & m_{23} & m_{24} & v_i \\
m_{31} & m_{32} & m_{33} & m_{34} & 1 \\
m'_{11} & m'_{12} & m'_{13} & m'_{14} - u'_i & 0 \\
a & b & c & d & 0
\end{bmatrix}
\begin{bmatrix}
x_i \\
y_i \\
z_i \\
1 \\
-w_i
\end{bmatrix}
= 0,
\tag{4}
$$

where the left 5×5 matrix is denoted as W_1. This is a set of five homogeneous equations with five unknowns. Because one of the five unknowns is 1, this means that this equation has a non-zero solution. Note that only $\det(W_1) = 0$ can allow equation 4 to have a non-zero solution. Because the determinant of W_1 consists

of terms in u_i, v_i, u'_i, $u_i v_i$, and $u'_i v_i$, the following equation with six coefficients, $a_0 \sim a_5$, exists:

$$a_0 + a_1 u_i + a_2 v_i + a_3 u'_i + a_4 u_i v_i + a_5 u'_i v_i = 0. \tag{5}$$

Similarly, for a given $\mathbf{u}'_i$, we use equations (i), (iv), (v), (vi), and (vii) in equation 3 to find its corresponding u_i to avoid the influence of w_i. We again obtain a set of five homogeneous equations with five unknowns. By the same argument as above, we may conclude that the following equation with six coefficients, $b_0 \sim b_5$, exists:

$$b_0 + b_1 u_i + b_2 u'_i + b_3 v'_i + b_4 u_i v'_i + b_5 u'_i v'_i = 0.$$

Suppose a_i and b_i are known, given a point $\mathbf{u}_i = (u_i, v_i)^T$, its corresponding point $\mathbf{u}'_i = (u'_i, v'_i)^T$ can be calculated by the following equations:

$$\begin{cases} u'_i = \frac{-(a_0 + a_1 u_i + a_2 v_i + a_4 u_i v_i)}{a_3 + a_5 v_i} \\ v'_i = \frac{-(b_0 + b_1 u_i + b_2 u'_i)}{b_3 + b_4 u_i + b_5 u'_i}. \end{cases} \tag{6}$$

These equations can be applied to transform all the image points of one panorama to the other panorama. After the transformation, we obtain an panoramic image. Therefore, the problem left to be solved is determining the values of the twelve coefficients, $a_0 \sim a_5$ and $b_0 \sim b_5$, by image corresponding points provided.

Given n pairs of corresponding points, namely $\{u_i, v_i, u'_i.v'_i\}$, $i \in [1..n]$, where not all image points $(u_i, v_i)^T$ lie on the same image row or image column. Equation 5 can be restated as follows:

$$a_0 \mathbf{1} + a_1 \mathbf{U} + a_2 \mathbf{V} + a_3 \mathbf{U}' + a_4 \mathbf{W} + a_5 \mathbf{W}' = \mathbf{0}, \tag{7}$$

where $\mathbf{1}$, $\mathbf{U}$, $\mathbf{V}$, $\mathbf{U}'$, $\mathbf{W}$, $\mathbf{W}'$ and $\mathbf{0}$ are all n-vectors. Note that vector $\mathbf{1}$ has all its elements equal to one and vector $\mathbf{0}$ has all elements equal to zero. In order to have a non-trivial solution of $a_0 \sim a_5$, the rank of matrix $[\mathbf{1}, \mathbf{U}, \mathbf{V}, \mathbf{U}', \mathbf{W}, \mathbf{W}']$ must equal to five. Moreover, if $\{a_0, a_1, a_2, a_3, a_4, a_5\}$ is a solution of equation 7, then $\{ka_0, ka_1, ka_2, ka_3, ka_4, ka_5\}$ will also be solutions for all $k \in \mathbb{R}$. However, all solutions lead to the same value of u'_i as shown in equation 6. Hence, we aim to find any set of $a_0 \sim a_5$ that satisfies equation 7.

Since there are six unknowns and the six-dimensional solution vector is up to a common scale factor, equation 7 can be solved with at least five pairs of image correspondences. In our work, we assume $a_0^2 + a_1^2 + \ldots + a_5^2 = 1$. When $n \geq 5$, a least-squared-error solution can be obtained by solving the eigenvalue problem in association with the scatter matrix of the linear equation system. Similar arguments also apply to $b_0 \sim b_5$.

Singular case occurs when the rank of matrix $[\mathbf{U}, \mathbf{V}, \mathbf{U}', \mathbf{W}, \mathbf{W}']$ is less than five. A common situation which causes the singular case is when vectors $\mathbf{U}$ and $\mathbf{U}'$ are linearly dependent, that is when we have $\mathbf{U} = A\mathbf{U}' + B$ for some $A, B \in \mathbb{R}$. This situation happens when the two line sensors used for grabbing the two LP-mosaic images are parallel to each other. (Note: we explain this situation in Appendix.) We despite the case when there are only few image

corresponding points provided, so two or more vectors of $\mathbf{U}$, $\mathbf{V}$, $\mathbf{U}'$, $\mathbf{W}$, and $\mathbf{W}'$ happen to be linear dependent because of poor sampling.

When vectors $\mathbf{U}$ and $\mathbf{U}'$ are linearly dependent, instead of using equation 6, we derive another set of transformation equations to transform the image. First, since we know $\mathbf{U} = A\mathbf{U}' + B$, the values of A and B can be obtained straightforwardly by solving a system of linear equations with at least two pairs of image correspondences. Secondly, by substituting equations (i) and (iv) in equation 3 into $u_i = Au'_i + B$, we get

$$\begin{cases} (m_{11} - Am'_{11})x_i + (m_{12} - Am'_{12})y_i + (m_{13} - Am'_{13})z_i + (m_{14} - Am'_{14} - B) = 0 \\ m_{21}x_i + m_{22}y_i + m_{23}z_i - v_iw_i + m_{24} = 0 \\ m_{31}x_i + m_{32}y_i + m_{33}z_i - w_i + m_{34} = 0 \\ m'_{21}x_i + m'_{22}y_i + m'_{23}z_i - v'_iw'_i + m'_{24} = 0 \\ m'_{31}x_i + m'_{32}y_i + m'_{33}z_i - w'_i + m'_{34} = 0 \\ ax_i + by_i + cz_i + d = 0. \end{cases}$$

It implies

$$\begin{bmatrix} m_{11} - Am'_{11} & m_{12} - Am'_{12} & m_{13} - Am'_{13} & m_{14} - Am'_{14} - B & 0 & 0 \\ m_{21} & m_{22} & m_{23} & m_{24} & -v_i & 0 \\ m_{31} & m_{32} & m_{33} & m_{34} & -1 & 0 \\ m'_{21} & m'_{22} & m'_{23} & m'_{24} & 0 & -v'_i \\ m'_{31} & m'_{32} & m'_{33} & m'_{34} & 0 & -1 \\ a & b & c & d & 0 & 0 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \\ w_i \\ w'_i \end{bmatrix} = 0,$$

where the left $6{\times}6$ matrix is denoted as W_2. It has non-trivial solution if $\det(W_2) = 0$, which means there exist coefficients $c_0 \sim c_3$ such that the following equation holds:

$$c_0 v_i + c_1 v'_i + c_2 v_i v'_i + c_3 = 0.$$

Given at least three pairs of image corresponding points, we are able to determine solutions of $c_0 \sim c_3$. The resulting transformation equations in this case are as follows:

$$\begin{cases} u'_i = \frac{u_i - B}{A} \\ v'_i = \frac{-c_0 v_i - c_3}{c_1 + c_2 v_i}. \end{cases}$$

3.2 Euclidean Reconstruction

Through the stitching method introduced above, two LP-mosaic images can be integrated into a single one. In our work, one of the LP-mosaic images is selected as the reference image, and the other image is transformed (or wrapped) so that the two images can be registered. By doing so, the integrated panoramic image can be treated as being obtained by enlarging the viewing range of the line sensor used for grabbing the reference image and then employing this range-enlarged sensor to scan the same planar scene. However, such a scanning process cannot preserve some desired properties such as right angle and parallelism.

From the reconstruction point of view, only an affine reconstruction can be obtained by using only image correspondence information as introduced in [5]. Without imposing other constraints on the scene or the camera, Euclidean reconstruction is impossible. Availability of constraints is application dependent. In

the following, we focus on the application of reconstructing a large-scale painting. The priori knowledge that the painting is of rectangular shape provides scene constraints for Euclidean reconstruction.

To upgrade the original reconstruction to a Euclidean one, let us image a virtual line sensor that scans the painting in the way that this sensor is aligned with the painting plane and is moved on this plane (that is, the painting is scanned by a 'virtual' Xerox machine). Furthermore, we assume that the line sensor is placed parallel to one of the four borderlines of the paintings, and the moving path is perpendicular to the line sensor. Then, the painting scanned with this virtual camera shall be of rectangular shape as well, and we call this rectangular frame the destination image. We aim to reconstruct the rectangular frame by performing some transformation from the integrated panoramic image to the destination image. However, there is no image pixel information for the virtual rectangular frame. Thus, the only correspondence knowledge can be applied are the boundaries of the frame and the image. By assuming that the ratio of the width to the height of the rectangular frame is known, we will show that a Euclidean reconstruction can be achieved by employing the four pairs of borderline-to-borderline correspondences of the painting.

Let (u_i, v_i) be the image coordinates of the source panoramic image (the integrated mosaic), and (u'_i, v'_i) be the image coordinates of the destination rectangular frame. Based on the linear-pushbroom camera model as discussed in Section 2, we have $(u'_i, w'_i v'_i, w'_i)^T = \mathbf{M}'(x_i, y_i, z_i, 1)^T$. Since the destination image is obtained by a 'virtual' Xerox machine, w'_i becomes constant for all i and the value w'_i can be absorbed by the second and the third row of matrix $\mathbf{M}'$. Hence, we have

$$(u_i, w_i v_i, w_i)^T = \mathbf{M}(x_i, y_i, z_i, 1)^T$$
$$(u'_i, v'_i, 1)^T = \mathbf{M}'(x_i, y_i, z_i, 1)^T$$

and these two equations can be expanded as follows:

$$\begin{cases} u_i = m_{11}x_i + m_{12}y_i + m_{13}z_i + m_{14} \dots \text{(i)} \\ w_i v_i = m_{21}x_i + m_{22}y_i + m_{23}z_i + m_{24} \dots \text{(ii)} \\ w_i = m_{31}x_i + m_{32}y_i + m_{33}z_i + m_{34} \dots \text{(iii)} \\ u'_i = m'_{11}x_i + m'_{12}y_i + m'_{13}z_i + m'_{14} \dots \text{(iv)} \\ v'_i = m'_{21}x_i + m'_{22}y_i + m'_{23}z_i + m'_{24} \dots \text{(v)} \\ 1 = m'_{31}x_i + m'_{32}y_i + m'_{33}z_i + m'_{34} \dots \text{(vi)} \end{cases} \quad (8)$$

From (i), (iv), (v) and (vi) in equation 8, we obtain the following:

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} - u_i \\ m'_{11} & m'_{12} & m'_{13} & m'_{14} - u'_i \\ m'_{21} & m'_{22} & m'_{23} & m'_{24} - v'_i \\ m'_{31} & m'_{32} & m'_{33} & m'_{34} - 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} = 0,$$

The determinate of the left 4×4 matrix must equal to zero, hence there exist coefficients $d_0 \sim d_3$, such that the following equation holds:

$$d_0 u_i + d_1 u'_i + d_2 v'_i + d_3 = 0. \quad (9)$$

Moreover, from (ii), (iii), (iv), (v) and (vi) in equation 8, we have

$$\begin{bmatrix} m_{21} & m_{22} & m_{23} & m_{24} & -v_i \\ m_{31} & m_{32} & m_{33} & m_{34} & -1 \\ m'_{11} & m'_{12} & m'_{13} & m'_{14} - u'_i & 0 \\ m'_{21} & m'_{22} & m'_{23} & m'_{24} - v'_i & 0 \\ m'_{31} & m'_{32} & m'_{33} & m'_{34} - 1 & 0 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \\ w_i \end{bmatrix} = 0,$$

and by the same reason as above, we obtain the following equation:

$$e_0 v_i + e_1 u'_i + e_2 v'_i + e_3 v_i u'_i + e_4 v_i v'_i + e_5 = 0. \tag{10}$$

Let the four corner points of the virtual rectangular frame be $(0,0)$, $(W,0)$, $(0,H)$, and (W,H), respectively, where W and H are the width and height of the frame. Use these corner points as four inputs (u'_i, v'_i) together with the corresponding corners of the integrated panoramic image (u_i, v_i), we are able to determine the values of $d_0 \sim d_3$ in equation 9.

Consider a boundary point $(u_i, v_i)^T$ lying on one of the border lines of the painting in the integrated panoramic image. One of its corresponding values u'_i and v'_i is known, (which is equal to either W or H), because the painting's borderline-to-borderline correspondences are available. The unknown one can be obtained by equation 9 based on the determined values of $d_0 \sim d_3$. Thus, by given at least five image correspondences on the boundaries, we are able to determine the values of $e_0 \sim e_5$ in equation 10. Once all the values of the coefficients are known, the transformation equations can be derived. This transformation enables us to refine the integrated panoramic image to a Euclidean reconstruction.

4   Experimental Results

We conducted a synthetic experiment to demonstrate how the image correspondence error (i.e. the input noise) affects the image stitching result. The experiment was designed as follows. There are 250 coplanar points randomly distributed in a bounded space. For each trial, two linear-pushbroom panoramic images with image resolutions of 170×550 are captured by two virtual line-cameras, whose intrinsic parameters are identical and set to be constant during the image acquisition. The starting positions and the moving velocities of these two cameras vary in each trial. The values of the position and the velocity vectors are randomly chosen within practical ranges. The image correspondence error is introduced by corrupting the ideal image projections by some random noise up to two image pixels. The image stitching error is measured as the average square-norm distance of all pairs of image corresponding points after merging. The average stitching error of 1000 trials is calculated for each noise scale.

Figure 1 is an illustration of our synthetic experiment. The image corresponding points with labels are shown in two linear-pushbroom panoramic images as well as in resulting image after stitching process. Table 1 summarizes our experiment results, which suggests that the stitching error increases linearly as the input noise increases linearly from zero to two pixels.
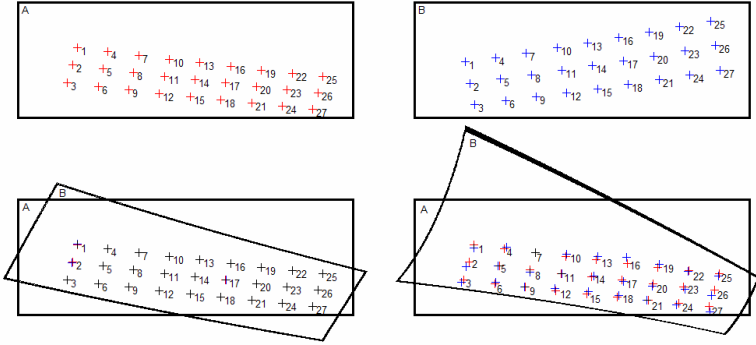
Fig. 1. Image A and B represent two linear-pushbroom panoramic images. Only 27 image corresponding points (instead of 250) are shown here for clarity. The bottom figures illustrate the stitching results of two cases: input noise free (left) and noise up to two pixels (right).

Table 1. The image correspondence errors (in different noise scale) vs. the image stitching errors (the average square norm).

Noise (pixel)	0	0.25	0.5	0.75	1	1.25	1.5	1.75	2
Error (pixel)	0.05	0.69	0.82	1.37	1.77	2.34	2.49	2.66	3.48

Moreover, a real image example is given in figure 2. The painting is named "Lang Shih-Ling One Hundred Stallions". Sony DCR-VX2000 camera was used and only the central image column of each shot was employed for generating panoramic image. Two linear-pushbroom panoramic images of a small portion of the painting were acquired with certain overlapping, which are shown in figure 2 (A) and (B). The resolutions of these two panoramas are both 450×2000 pixels. Figure 2 (C) shows the image stitching result based on 53 identified image corresponding points. The resulting image after Euclidean reconstruction is shown in figure 2 (D), which has resolutions 600×1700 pixels. This width/height ratio has been adjusted to meet the true ratio of the selected portion.

5   Conclusion

Planar homography, which can help determine complete point-to-point image relationships for a pair of images taken with perspective cameras, serves as a critical property for realizing many image mosaicing applications. Nevertheless, whether similar useful properties exist for other imaging model, such as the linear-pushbroom model, has not been well studied yet. In this paper, we demonstrate that with additional planarity constraint to the scene geometry, complete point-to-point image relationships can also be established between two linear-pushbroom panoramic images by employing at least five pairs of corresponding

Fig. 2. Portion of painting "Lang Shih-Ling One Hundred Stallions".

points. By the existence of such property, an image stitching method is developed for integrating two LP-panoramas to enlarge the panorama's field of view. Moreover, a Euclidean reconstruction method is presented to restore the properties of 2D Euclidean geometry for reconstructing a rectangular frame. Both methods required only few pairs of image corresponding points as the input. The image integration algorithm as a whole can be performed in real-time.

References

1. Chen, S. E.: QuickTime VR - An Image-Based Approach to Virtual Environment Navigation. Computer Graphics (SIGGRAPH'95) (1995) 29-38
2. Chen, S., Williams, L.: View Interpolation for Image Synthesis. Computer Graphics (SIGGRAPH'93) (1993) 279-288
3. Faugeras, O., Luong, Q.-T.: The Geometry of Multiple Images. The MIT Press, London, England (2001)
4. Gortler, S. J., Grzeszczuk, R., Szeliski, R., Cohen, M. F.: The Lumigraph. In Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH (1996) 43-54
5. Gupta, R., Hartley, R. I.: Linear Pushbroom Cameras. IEEE PAMI, **19**(9), (1997) 963-975
6. Hansen, M., Anadan, P., Dana, K., van de Wal, G., Burt, P.: Realtime Scene Stabilization and Mosaic Construction. In Proc. of IEEE CVPR (1994) 54-62
7. Hartley, R. I.: Multiple View Geometry in Computer Vision, Cambridge University Press (2000)
8. Irani, M., Anandan, P., Hsu, S.: Mosaic Based Representation of Video Sequences and Their Applications. IEEE Proc. Int'l Conf. Computer Vision (1995) 605-611
9. Ishiguro, H., Yamamoto, M., Tsuji, S.: Omni-Directional Stereo. IEEE PAMI, **14**(2) (1992) 257-262
10. Kang, S.: A Survey of Image-Based Rendering Techniques. Videometric VI, **3641** (1999) 2-16
11. Kumar, R., Anandan, P., Hanna, K.: Shape Recovery from Multiple Views: A Parallax Based Approach. In Image Understanding Workshop, Monterey, CA. Morgan Kaufmann Publishers (1994) 947-955

12. Kumar, R., Anandan, P., Irani, M., Bergen, J., Hanna, K.: Representation of Scenes from Collections of Images. In Proc. IEEE Workshop on Representations of Visual Scenes (1995) 10-17
13. Levoy, M., Hanrahan, P.: Light Field Rendering. In Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH (1996) 31-42
14. Mann, S., Picard, R. W.: Virtual Bellows: Constructing High Quality Stills from Video. In ICIP (1994)
15. McMillan, L., Bishop, G.: Plenoptic Modeling: An Image-Based Rendering System. Computer Graphics (SIGGRAPH'95) (1995) 39-46
16. Moffitt, F. H., Mikhail, E. M.: Photogrammetry. Harper & Row, New York, 3rd Edition (1980)
17. Peleg, S., Herman, J.: Panoramic Mosaics by Manifold Projection. IEEE CVPR Proceedings (1997)
18. Sawhney, H. S.: Simplifying Motion and Structure Analysis Using Planar Parallax and Image Warping. In 12th International Conference on Pattern Recognition (ICPR'94), Vol. A. Jerusalem, Israel. IEEE Computer Society Press (1994) 403-408
19. Shum, H. Y., He, L. W.: Rendering with Concentric Mosaic. SIGGRAPH'99 (1999) 299-306
20. Shum, H. Y., Szeliski, R.: Panoramic Image Mosaics. Tech. Rep. MSR-TR-97-23, Microsoft Research (1997)
21. Szeliski, R.: Image Mosaicing for Tele-Reality Applications. Technical Report 94/2, Digital Equipment Corporation, Cambridge Research Lab (1994)
22. Szeliski, R., Kang, S. B.: Direct Methods for Visual Scene Reconstruction. In IEEE Workshop on Representations of Visual Scenes, Cambridge, Massachusetts (1995) 26-33

Appendix

Let E and E' be two LP-mosaic images of a common planar scene. We explain and illustrate when the two line sensors used to grab the two LP-mosaic images are parallel to each other, the following statement holds: for any pair of image corresponding points $(u_i, v_i)^T$ and $(u_i', v_i')^T$, the values u_i and u_i' are related by the equation: $u_i = Au_i' + B$, where A and B are constants.

First, consider a plane in 3D and a line camera which moves along a straight line (set it to be the x-axis of the camera coordinate system) with constant velocity and taking line images at each position $\mathbf{C}_0$, $\mathbf{C}_1$, $\mathbf{C}_2$, and so on, as shown in the top-left of figure 3. The y-axis is defined to be parallel to the line-sensors and is perpendicular to the x-axis. The z-axis is defined following the right-hand-rule. The geometric relationship between the plane and the camera coordinate system is unknown.

The bottom-left of figure 3 shows the resulting LP-panoramic image E. The parallel lines $L_0 \sim L_4$ on the plane are projected to image columns $u = 0 \sim 4$ respectively. Since the distance between any pair of points $\mathbf{C}_i$ and $\mathbf{C}_{i+1}$ is constant (as defined in Section 2), lines L_i and L_{i+1} is a set of parallel lines with equal distance.

Then, consider another LP-panoramic image E', whose associated camera's moving path is rotated with respect to the y-axis, as shown in the right-hand-side of figure 3. The equal-distance parallel lines $L_0' \sim L_4'$ on the plane are projected to image columns $u' = 0 \sim 4$ respectively.

In fact, lines $L_0 \sim L_4$ are parallel to lines $L_0' \sim L_4'$ as they both are parallel to the y-axis of the camera coordinate system. So, it is possible that lines $L_0 \sim L_4$ also appear in
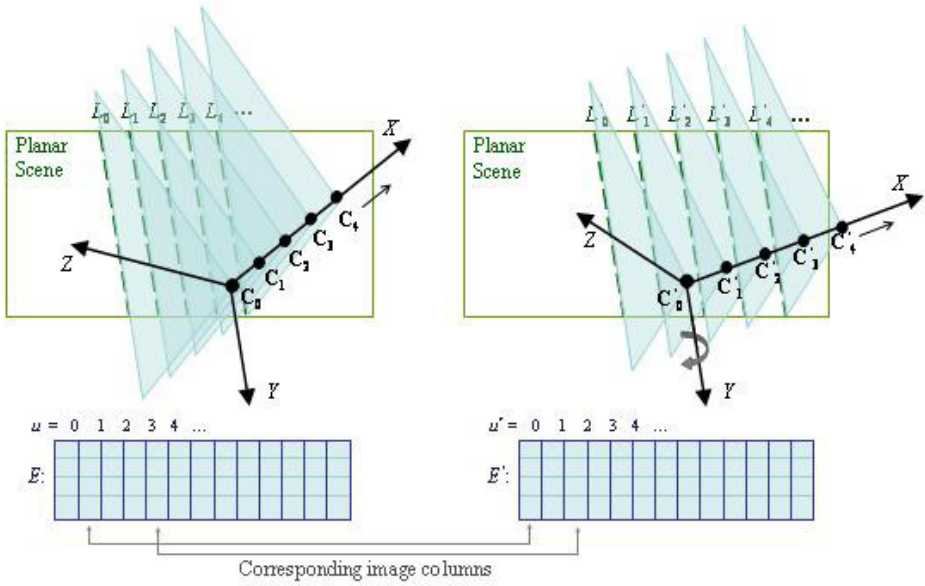
Fig. 3. Geometric configuration that illustrates the existence of image column-to-column correspondence.

the image E' and vice versa. Hence, we have column-to-column correspondence between two LP-panoramic images. According to the basic geometrical property, when there is a column-to-column correspondence between two images as described, the relationship between those corresponding columns can be expressed by equation $u_i = Au_i' + B$, where A and B are constants.

Finally, we may conclude that as long as the y-axes of two camera coordinate systems, which are associated to the different LP-panoramic images, are parallel, we have relation $u_i = Au_i' + B$ for all corresponding u_i and u_i'.