Mosaicing with Strips on Dynamic Manifolds [1]

Shmuel Peleg Benny Rousso Alex Rav-Acha Assaf Zomet
School of Computer Science and Engineering
The Hebrew University of Jerusalem
91904 Jerusalem, ISRAEL

Contact E-Mail: peleg@cs.huji.ac.il

[1]This research was partially funded by DARPA through ARL Contract DAAL01-97-0101 and by the European ACTS project AC074 "Vanguard".

ABSTRACT

Panoramic image mosaicing is commonly used to increase the visual field of view by pasting together many images or video frames. Existing mosaicing methods are based on projecting all images onto a pre-determined single manifold: a plane is commonly used for a camera translating sideways, a cylinder is used for a panning camera, and a sphere is used for a camera which is both panning and tilting. While different mosaicing methods should therefore be used for different types of camera motion, more general types of camera motion, such as forward motion, are practically impossible for traditional mosaicing.

A new methodology to allow image mosaicing in more general cases of camera motion is presented. Mosaicing is performed by projecting thin strips from the images onto manifolds which are dynamically determined by the camera motion. While the limitations of existing mosaicing techniques are a result of using predetermined manifolds, the use of dynamic manifolds overcomes these limitations. With manifold mosaicing it is now possible to generate high-quality mosaicing even for the very challenging cases of forward motion and of zoom.

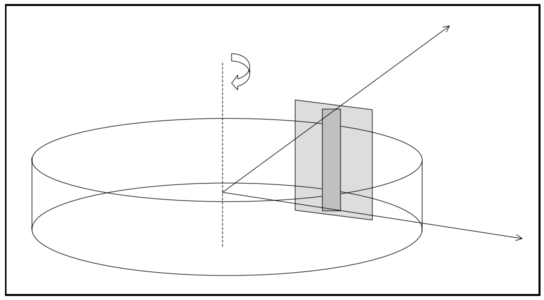

FIGURE 1. A panoramic image can be generated from a panning camera by combining the images on the surface of a cylinder.

1 Introduction

Creating pictures having larger field of view, by combining many smaller images, is common since the beginning of photography, as the camera's field of view is smaller than the human field of view. In addition, some large objects can not be captured in a single picture as is the case in aerial photography. Using omnidirectional cameras [24] can sometimes provide a partial solution, but the images obtained with such cameras have substantial distortions, and capturing a wide field of view with the limited resolution of a video camera compromises image resolution. A common solution is photo-mosaicing: aligning and pasting pictures, or frames in a video sequence, to create a wider view. Digital photography enabled new implementations for mosaicing [22, 23, 25, 10, 16, 33], which were first applied to aerial and satellite images, and later used for scene and object representation.

The simplest mosaics are created by panning the camera around its optical center using a special device, in which case the panoramic image can be created on a cylindrical or a spherical manifold [20, 12, 21, 32, 18, 33]. The original images, which are a perspective projection onto an image plane, are warped to be perspectively projected into an appropriate cylinder, where they can be combined to a full 360 degrees panorama as in Fig. 1. While the limitations to pure sideways camera rotation enables easy mosaicing without the problems of motion parallax, this approach can not be used with other camera motions.

Simple mosaicing is also possible from a set of images whose mutual displacements are pure image-plane translations. And for somewhat more general camera motions more general transformation for image alignment can be used, like a global affine transformation or a planar-projective transformation [11, 14, 17, 30, 16]. In most cases images are aligned pairwise using the global parametric transformation, a reference frame is selected, and all images are aligned to this reference frame and combined to create the panoramic mosaic. Such methods imply the perspective projection of

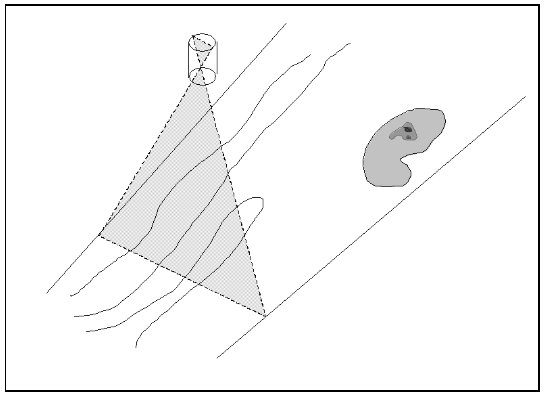

FIGURE 2. An Aerial Pushbroom Camera.

all the images onto the planar manifold corresponding to the image plane of the reference frame. Using a planar manifold, and aligning all frames to a single reference frame, is reasonable only when the camera is far from the scene and its motion is mainly a translation and a rotation around the optical axis. Significant distortions are created, for example, when the camera motion includes sideway rotation.

Most restrictions on the motion of the camera used for mosaicing can be eliminated by using a dynamic manifold whose shape is determined during the mosaicing process. To enable undistorted mosaicing the selected manifold should have the property that after projecting the images onto the manifold the *optical flow*[1] vectors become approximately uniform: parallel to each other and of equal magnitude. Typical cases for this optical flow is sideways image translation, where the manifold is a plane, and a panning camera, where the manifold is a vertical cylinder, and the optical flow in a central vertical strip of the image is approximately uniform. It will also be shown that in the general case of a translating camera the manifold should be a cylinder whose axis is the direction of motion

When a perspective camera moves in a general scene, the optical flow is not uniform and depends on the scene depth. A solution for mosaicing such scenes is the "slit camera", or the "pushbroom camera", used in aerial photography [15]. This camera can be modeled as a 1-D sensor array which collects strips by "sweeping" the scene, as described in figure 2.

The imaging process of the pushbroom camera can be modeled by a multi-perspective projection: For each strip the projection is perspective, while different strips may be acquired from different centers of projections. Thus in the direction of the strips, the projection is perspective, while in the

[1] Image motion is represented by the optical flow: the displacement vectors associated with each image point, which specify the location of the image point in the next frame relative to its location in the current frame.

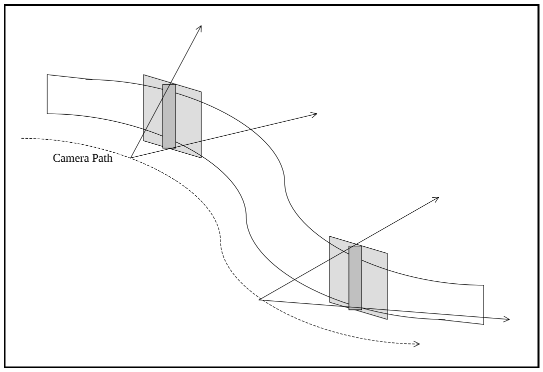

FIGURE 3. A panoramic image generated from a 'vertical "slit" moving on a smooth path on a horizontal plane.

direction of advance the projection is parallel. Under parallel projection, there is no parallax, so the optical flow in the result image is uniform.

The mosaicing techniques described in this paper process video sequences acquired by a standard perspective camera moving on a smooth route. They approximate the mosaic image which would have been acquired by a pushbroom camera moving on the same route. This is done by reprojecting thin strips from the images onto a dynamic manifold, such that the optical flow becomes approximately uniform: parallel and of equal magnitude. Both the dynamic manifold and the reprojection transformation are computed implicitly.

Each region in the mosaic is taken from that image where it is captured at highest resolution. While this could have been neglected in the traditional mosaicing which do not allow any scale changes, it is critical for general camera motions where, for example, a region is seen at higher resolution when closer.

A mosaicing approach which constructs for the first time multi-perspective panoramic views on general manifolds has been described in [35]. A video camera is continuously scanning the scene through a vertical "slit", and the one-dimensional vertical slits are then combined into a panoramic image. In the case of a purely panning camera the panoramic images generated by this approach are similar to the cylindrical case of Fig. 1, and when the slits are narrow and continuous there is no need to warp the images from a plane to a cylinder. A more general case is also presented: a camera moving on a smooth path on a horizontal plane, as described in Fig. 3. The mosaic is generated in this case on a more general manifold. It was assumed in [35] that the motion of the camera is measured by external devices rather than being computed from the video itself.

Practical implementations of general manifold mosaicing will be presented, based on the computed motion between the images.

Unlike other methods for multi-perspective mosaics [27, 34] the mosaics are constructed without knowing or recovering the structure of the scene, and without knowing explicitly the full motion and calibration of the camera.

We assume that the camera motion is a pure rotation or a pure translation. In case the camera motion comprises of both rotation and translation it is assumed that the rotation can be cancelled [28].

The method is based on building a mosaic images by collecting strips from the images, satisfying the following conditions:

- The width of the strips should be proportional to the motion.
- The collected strips should be warped and pasted into the panoramic image such that after warping, their optical flow becomes parallel to the direction in which the panoramic image is constructed, and of equal magnitude.
- In order to avoid global resizing, each image strip includes a feature (the *anchor*) which does not change under the warping. This anchor determines the form of the “broom”.
- It is recommended to have *the anchor* perpendicular to the optical flow. This maximizes the information collected by the virtual 1-D sensor array.

Examples of manifold mosaicing using strips will be given for cases of almost uniform image translations caused by a panning camera [26], for 2D planar projective transformation caused from a forward moving camera [29], and for 2D planar projective transformation caused by a tilted panning camera, or a tilted camera translating in a planar scene. Mosaics generated in this manner can be considered as similar to the vertical “slits” [35] or the“linear push-broom cameras” [15]. However, unlike the straight “slit” or “broom”, the broom in manifold mosaicing will change its shape dynamically from a straight line to a circular arc, to become mostly perpendicular to the optical flow.

In cases where strips are wide, it is possible to reduce the parallax and simulate the parallel projection by generating intermediate views [31, 13]. The introduction of intermediate views simulates a denser image sequence, where the strips are narrower, with smaller discontinuities due to motion parallax.

2 Mosaicing with Strips

Most existing mosaicing systems align and combine full images or video frames [33, 16, 30, 19]. The combination of full frames into mosaics introduces some difficulties:

- It is almost impossible to align accurately complete frames due to lens distortion, motion parallax, moving objects, etc. This results in "ghosting" or blurring when the mosaic is constructed.
- It is difficult to determine the mosaicing manifold. E.g., If all images are aligned to one reference image, different reference images will give different mosaics. In the case of a projection onto a cylinder it is important that the camera motion is a pure sideways rotation.

In order to overcome the above difficulties, we propose to use mosaicing with strips. A preliminary system was proposed in [35], where the strip was a vertical "slit", and the camera motion was limited to sideways camera translations and rotations measured by external devices. In this simple case vertical sections were taken from each image and pasted side by side (see Fig. 3. The same vertical slit is useless with vertical image motion, as the optical flow is parallel to the scanning slit (Fig. 5.b). No mosaic will be created as no image area will pass through the slit. Optimal mosaics are achieved only with a slit which is perpendicular to the optical flow.

An example for the determination of the shape of the slit is given for image motion generated by a pure translation of the camera, as shown in Fig. 4. In this case the image motion can be described by a radial optical flow emanating from the focus of expansion (FOE), and the field of view of the camera (FOV) can be described as a circle on the image plane. The optimal slit will be the longest circular section having its center at the FOE and passing through the FOV. This is the longest curve in the FOV that is perpendicular to the optical flow.

The definition of the scanning slit as perpendicular to the optical flow is very simple for some cases.

- In sideways image motion the optimal slit is vertical (Fig. 5.a).
- In image scaling (zoom), and in forward motion, the optimal slit is a circle (Fig. 5.c).
- In image motion generated by camera translation, the optimal slit is a circular arc (Fig. 5.d).

Image motion is usually more general than these simple special cases. However, in most cases circular or elliptic curves are sufficient for mosaicing.

The shape of the slit determines the shape of the manifold on which the mosaic is created. The circular slit, for example (Fig. 5.c), can be combined on a cylindrical manifold.

3 Cutting and Pasting of Strips

The mosaic is constructed by pasting together strips taken from the original images. The shape of the strip, and its width, depend on the image

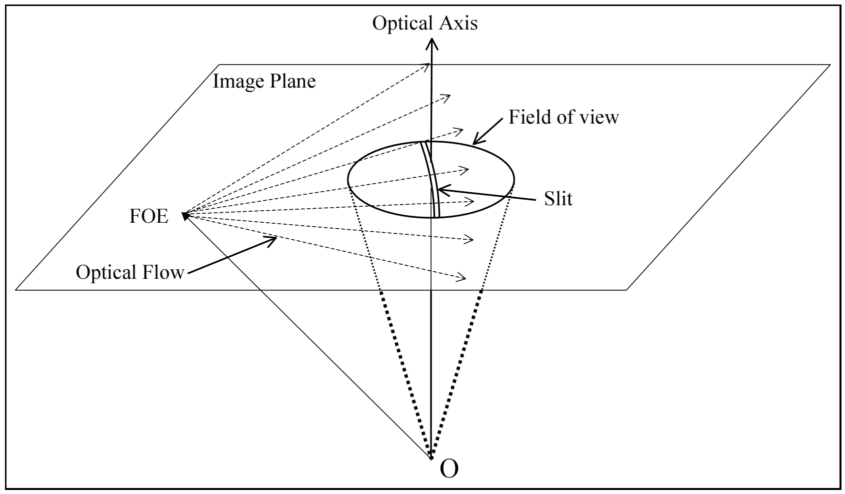

FIGURE 4. Determining the shape of the slit with camera translation. In this case the optimal slit will be the longest circular section having its center at the FOE and passing through the field of view. This is the longest curve in the FOV that is perpendicular to the optical flow.

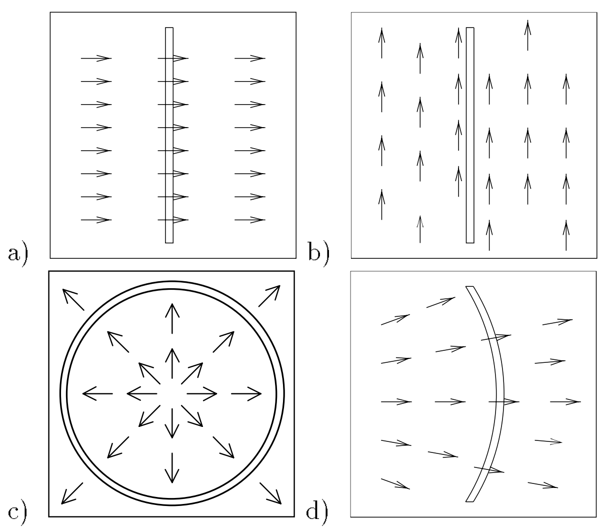

FIGURE 5. The mosaicing process and the direction of the optical flow.
(a) A vertical slit is optimal when the optical flow is horizontal. (b) A vertical slit is useless when the optical flow is vertical. (c) A circular slit is optimal when the optical flow is radial. (d) For general motion optimal slits should be perpendicular to the optical flow, and bent accordingly.

motion. After choosing the strip, it should be warped such that the optical flow becomes parallel and of equal magnitude to allow for mosaicing. This section describes how to select and to warp these strips.

3.1 Selecting Strips

In order to determine the strip to be taken from Image I_n, the preceding frame, I_{n-1}, and the succeeding frame, I_{n+1}, should be considered.

Let $\mathcal{A}_n$ be the transformation relating points $p_n = (x_n, y_n)$ in Image I_n to the corresponding points $p_{n-1} = (x_{n-1}, y_{n-1})$ in Image I_{n-1}, and let $\mathcal{A}_{n+1}$ be the transformation relating points p_{n+1} in Image I_{n+1} to the corresponding points p_n in Image I_n.

Given the transformations $\mathcal{A}_n$ and $\mathcal{A}_{n+1}$, the lines $\mathcal{F}_n(x_n, y_n) = 0$ and $\mathcal{F}_{n+1}(x_{n+1}, y_{n+1}) = 0$ are selected respectively (see Fig. 6.a-c). The line $\mathcal{F}_n(x_n, y_n) = 0$ in I_n corresponds to the line $\mathcal{F}'_n(x_{n-1}, y_{n-1}) = 0$ in I_{n-1} using the transformation $\mathcal{A}_n$. In the same way, the line $\mathcal{F}_{n+1}(x_{n+1}, y_{n+1}) = 0$ in I_{n+1} corresponds to the line $\mathcal{F}'_{n+1}(x_n, y_n) = 0$ in I_n using the transformation $\mathcal{A}_{n+1}$.

The strip that is taken from the image I_n is bounded between the two lines $\mathcal{F}_n(x_n, y_n) = 0$ and $\mathcal{F}'_{n+1}(x_n, y_n) = 0$ in I_n (see Fig. 6.a-c).

Line $\mathcal{F}_n$ will be the first boundary of the strip, and will be orthogonal to the optical flow with regard to the previous image. Line $\mathcal{F}'_{n+1}$ will be the second boundary of the strip, which is the projection of line $\mathcal{F}_{n+1}$ onto image I_n.

This selection of the boundaries of the strip ensures that no information is missed nor duplicated along the strip collection, as the orthogonality to the optical flow is kept.

3.2 Pasting Strips

Consider the common approach to mosaicing where one of the frames is used as a reference frame, and all other frames are aligned to the reference frame before pasting. In term of strips, the first strip is put in the panoramic image as is. The second strip is warped in order to match the boundaries of the first strip. The third strip is now warped to match the boundaries of the *already warped* second strip, etc. As as result, the mosaic image is continuous. However, major distortions may be caused by the accumulated warps and distortions. Large sideways rotations can not be handled, and cases such as forward motion or zoom usually cause unreasonable expansion (or shrinking) of the image.

To create continuous mosaic images while avoiding accumulated distortions, the warping of the strips should depend only on an adjacent original frame, independent of the history of previous warpings.

for example, we may choose the anchor as the back side of each strip. This is the side of the strip which corresponds to the boundary between

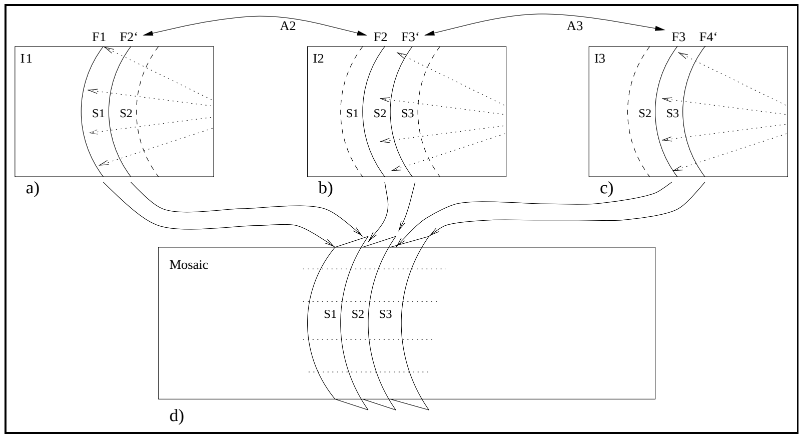

FIGURE 6. Cutting and pasting strips.
(a)-(c) Strips are perpendicular to the optical flow. (d) Strips are warped and pasted so that their back side is fixed and their front side is warped to match the back side of the next strip.

image I_{n-1} and image I_n and is defined by $\mathcal{F}_n$. In this case, the front side of the strip is warped to match the back side of the next strip defined by $\mathcal{F}'_{n+1}$.

In the example described in Fig. 6.d, we warp the strip from image I_1 such that its left side does not change, while its right side is warped to match the left side of the strip coming from image I_2. In the second strip, the left side does not change, while the right side is warped to match the left side of the third strip, etc.

The warping done when strips are pasted together, which is necessary in order to get a continuous mosaic, is actually the projection of the strips onto the mosaicing manifold. This is done without the explicit computation of that manifold. After that warping, the original optical flow becomes parallel to the direction in which the panoramic mosaic is constructed. No accumulative distortions are encountered, as each strip is warped to match just another original strips, avoiding accumulative warps. Having an anchor in the strip prevents change of scale in the warping. The anchors are placed parallel along the mosaic, completing a parallel projection in the direction of the motion of the camera. Assuming the motion between successive frames is small, canceling the parallax by some smooth interpolation of the coordinates is a satisfying approximation for the narrow gaps between the anchors. In case the strips are wide, and the gaps between the anchors are big, view interpolation can be used, as described in 6.

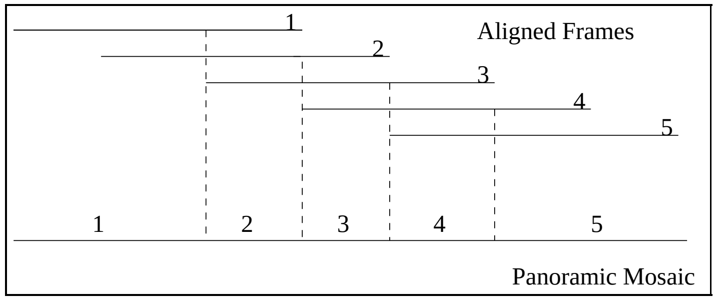

FIGURE 7. Pixel values in the panoramic mosaic are taken from a single image whose center, after alignment, is closest to the corresponding pixel. All pixels in Region 2 of the mosaic are therefore taken from the strip at the center of Image 2, etc. This constructions corresponds in 2D to the Voronoi tessellation.

4 Examples of Mosaicing Implementations

In this section two implementations of manifold mosaicing are described. The simplest implementation uses only straight slits, and the other implementation uses curved slits. Implementation issues, like strip cut and paste, and color merging across seams, are also described.

4.1 Strip Cut and Paste

Combination of the sequence of aligned image frames into a single panoramic mosaic can be done in several ways. In those cases where image alignment is close to perfect, it is possible to use all overlapping images to produce the mosaic.

The most common approach to combine the overlapping parts of the images is averaging. Averaging, however, may result in blurring when the alignment is not perfect. In this case it is preferred to select only one of the input images to represent a region in the mosaic. Such a selection should be done to minimize effects of misalignment. The most logical selection is to select from each image the strip closest to its center. There are two reasons for that selection:

- Alignment is usually better at the center than at the edges of the pictures.
- Image distortion is minimal at the center of the images

This selection corresponds to the Voronoi tessellation [8], and is shown in Figure 7. Using the Voronoi tessellation for image cut-and-paste also serves to minimize visible misalignment due to lens distortions. Voronoi tessellation causes every seam to be at the same distance from the two corresponding image centers. As lens distortions is a radial effect, features that are perpendicular to the seam will be distorted equally on the seam, and therefore will remain aligned regardless of lens distortion.

4.2 Color Merging in Seams

Changes in image brightness, usually caused by the mechanism of automatic gain control (AGC), cause visible brightness seams in the mosaic between regions covered by different images. These seams should be eliminated in order to get a seamless panorama.

The process of blending the different images into a seamless panorama must smooth all these illumination discontinuities, while preserving image sharpness. A method that fulfills this requirement is described in [10]. In this approach, the images are decomposed into band-pass pyramid levels, and then combined at each band-pass pyramid level . Final reconstruction of the images from the combined band-pass levels give the desired panorama.

4.3 Mosaicing with Straight Strips

Manifold mosaicing can be implemented very efficiently when the optical flow is approximately parallel, as in camera translations or sideways rotation. In this case a simple 2D rigid image alignment (only image translations and rotations) can be used, and the strips can be straight. Construction is very fast, and has been demonstrated live on a PC [26]. Results are impressive in most cases, and have the desired feature of manifold mosaicing: each object in the mosaic appears in the same shape and size as it appears in the video frames, avoiding any scaling, and therefore avoiding distortions and loss of resolution. In this system, the manifold is defined to follow the center strip of the images as seen in Fig. 1 and Fig. 3. Mosaicing was done without the explicit assumption of pure rotation, and without the need to project the images onto a cylinder before mosaicing.

Fig. 9 and Fig. 8 show panoramic mosaic images created with an implementation of the manifold mosaicing on the PC [26].

4.4 Mosaicing with Curved Strips: Forward Motion

Mosaicing of images from forward moving cameras can be done using the more general manifold mosaicing. In Fig. 10 the camera was moving and looking forward inside a small canyon. The computed image motion was an homography, and the slit was an elliptic shape.

FIGURE 8. Manifold mosaicing with vertical scanning. The curved boundary is created by the unstabilized motion of the hand-held camera.

FIGURE 9. An example of panoramic imaging using manifold mosaicing with straight strips. The curved boundary is created by the unstabilized motion of the hand-held camera.

a) b)

c)

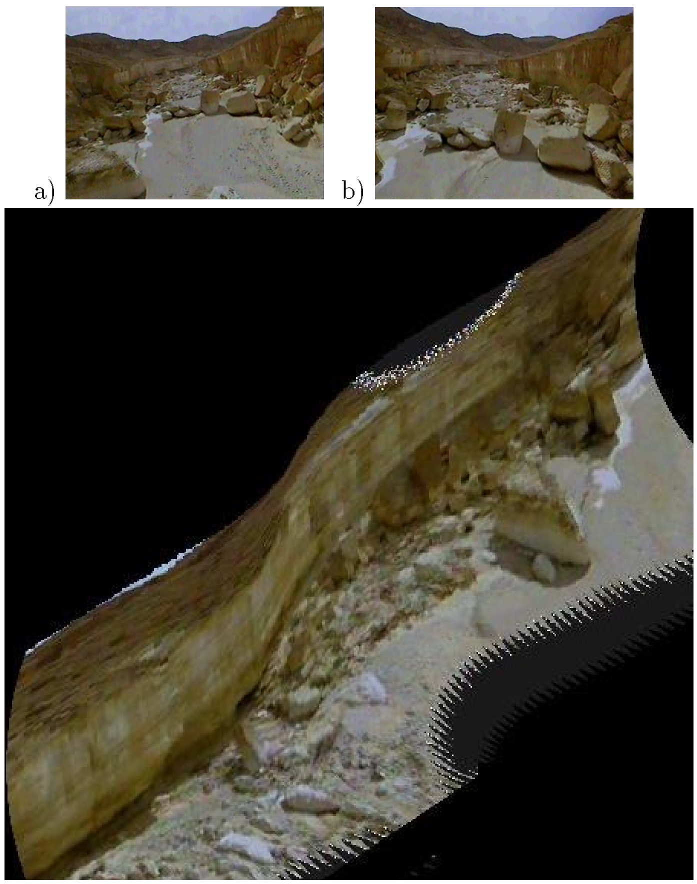

FIGURE 10. Forward motion in a small canyon.
(a-b) Two original frames. (c) Mosaic generated from curved strips.

5 Rectified Mosaicing: A Tilted Camera

The mosaicing algorithm described in Sect. 4 handles flawlessly the following two cases:

- A panning camera, when the optical axis is perpendicular to the rotation axis. E.g. - When a camera is panning from left to right with a vertical rotation axis, its optical axis must be horizontal.

- A translating camera scanning a planar scene, when the viewing direction is in the plane defined by the direction of motion and the normal to the plane. E.g. - The camera may look normal to the plane or have a forward view.

When the camera motion and the viewing directions are different, e.g. when the camera is tilted, this mosaicing algorithm constructs a curled mosaic. In this section we describe an algorithm for the mosaicing of sequences when the camera is tilted, and the two conditions above are not satisfied. The algorithm comes in two variants, one uses asymmetric strips, and the other uses symmetric strips. The image motion model used is a homography, which is assumed to be computed by one of many methods (e.g. [9]).

For the simplicity of explanation, we assume that the optical flow is close to parallel and close to horizontal, and therefore we simulate a pushbroom camera having a vertical straight slit. Therefore the "anchor", which is a feature in the strip that does not change with the warping of the strip, will also be a vertical straight line. The method can be easily adapted to more general motions using the methodologies described earlier in this paper.

In the first variant of the algorithm one side of the strip is taken as the anchor. In the second variant the anchor will be a vertical straight line at the center of the strip.

Examples for rectified mosaicing in the cases of a translating camera is shown in Fig. figure:translation, and in the case of a panning camera in Fig. figure:panning.

In the first algorithm, one of the borders of the strip is used as the anchor. It is simpler than the second algorithm, and useful when the borders of the strips are close to the borders of the image, which is recommended when the motion induces significant changes of scale in the image. In the second algorithm, a vertical line in the middle of the strip is used as the anchor. When the anchor is the central vertical line of the image, this algorithm is less sensitive to lens distortion, and less dependent on the direction of motion. For methodological reasons, we assume for both algorithms that the camera is translating to the right in front of a planar scene. In order to follow the technical details, we recommend the reader to use figures , .

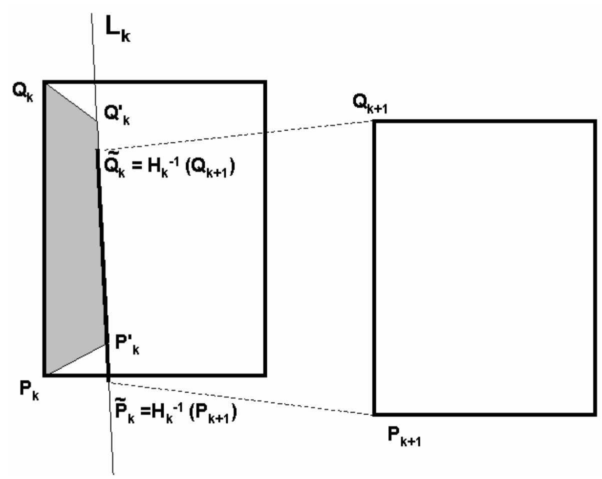

FIGURE 11. Non symmetric strip. The anchor is the left border of the strip.

5.1 Asymmetrical Strips

Assuming the camera motion is to the right, we use the left border of the strip as the anchor (Fig. 11). We mark the intersection of the anchor with the top and bottom image borders by P_k and Q_k. Given the homography H_k between Image I_k and Image I_{k+1}, let $\tilde{Q}_k = H_k^{-1}(Q_{k+1})$ and $\tilde{P}_k = H_k^{-1}(P_{k+1})$. $\tilde{Q}_k$ and $\tilde{P}_k$ are the mapping onto Image I_k of the anchor edges in Image I_{k+1}.

Let L_k be the line passing through $\tilde{Q}_k$ and $\tilde{P}_k$. We find on the line L_k two points Q'_k and P'_k such that their distance is like the distance between $\tilde{Q}_k$ and $\tilde{P}_k$, and their centroid is on the middle row of the image. The region in the image to be warped to a strip in the mosaic is defined by the quadrangle $\tilde{Q}_k \tilde{P}_k P_k Q_k$. The warping is done by smooth (e.g. bilinear) interpolation of the coordinates of $\tilde{Q}_k, \tilde{P}_k, P_k, Q_k$. The use of an interpolation is needed for strip alignment, and this is an approximation to the real transformation which is unknown. As the strips are very narrow, this approximation is satisfying.

The next strip in the mosaic is placed with vertical offset of $\| \tilde{Q}_k - Q'_k \|_2 * \frac{h}{\|Q'_k - P'_k\|_2}$ from the current strip, where h is the image height.

5.2 Symmetrical Strips

We assume similar imaging conditions as in 5.1.

We mark the vertical line at the center of the image as C_k, and its intersection with the top and bottom image borders by P_k and Q_k.

We would like to choose a region which is approximately symmetrical

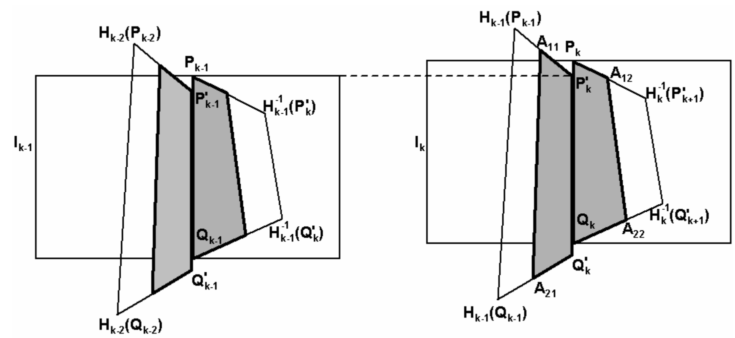

FIGURE 12. Mosaicing with symmetrical strips. A rectangular strip in the mosaic is mapped to the grey polygonal region in the image.

around C_k, to reduce lens distortion. (This is the reason for choosing C_k as the anchor, in general any other line an be used). This region is illustrated in Fig. 12

Given the homography H_{k-1} between Image I_k and Image I_{k-1}, Let O_{k-1} be the center of image I_{k-1}, and let d be the vertical offset between O_{k-1} and $H_{k-1}(O_{k-1})$. Let P'_k be a point shifted from P_k by d, and Let Q'_k be a point vertically shifted from Q_k by d. Based on the homography H_k between Image I_{k+1} and Image I_k, we apply a similar process between images I_k and I_{k+1}.

We now use the homographies to map points P'_{k+1} and Q'_{k+1} from Image I_{k+1}, and points P_{k-1} and Q_{k-1} from Image I_{k-1}, to Image I_k. We then find the middle points: Let F_L be the homography mapping an arbitrary rectangle UVWX to the points $H_{k-1}(P_{k-1}), P'_k, Q'_k, H_{k-1}(Q_{k-1})$ respectively, and Let F_R be the homography mapping UVWX to the points $P_k, H_k^{-1}(P'_{k+1}), H_k^{-1}(Q'_{k+1}), Q_k$ respectively. The region borders are defined by:

$$A_{11} = F_L(\frac{U+V}{2}), A_{21} = F_L(\frac{W+X}{2}),$$

$$A_{12} = F_R(\frac{U+V}{2}), A_{22} = F_R(\frac{W+X}{2})$$

The polygonal region in the image is comprised of two quadrangles: the left quadrangle, with the corners at P'_k, Q'_k, A_{11}, and A_{21}, and the right quadrangle, with the corners at P_k, Q_k, A_{12}, and A_{22}. Each of these quadrangles is mapped to a rectangle in the mosaic. We warp the left quadrangle to a rectangle in the mosaic by some smooth (e.g. bilinear) interpolation of the coordinates of the corners like in the asymmetric case. We apply a similar process to the right part of the strip (rectangle) and the right part of the region.

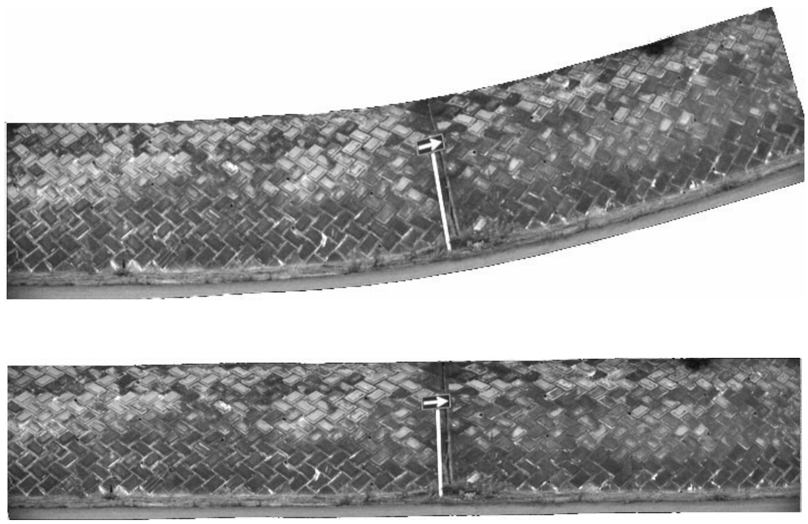

FIGURE 13. A translating camera mosaicing a slanted wall. Regular mosaicing results in a curled image, while rectified mosaicing results in a straight mosaic.

We place the left part of the strip at the same vertical offset as the right part of the previous strip, and the right side of the strip with vertical offset of d from the left part.

6 View Interpolation for Motion Parallax

Taking strips from different images when the width of the strips is more than one pixel would work fine only without parallax. When motion parallax is involved, no single transformation can be found to represent the optical flow in the entire scene. As a result, a transformation that will align a close object will duplicate far objects, and on the other hand, a transformation that will align a far object will truncate closer objects.

In order to overcome the problems of motion parallax in general scenes, instead of taking a strip with a width of N pixels, we can synthetically generate intermediate images, and use narrower strips. For example, we can take a collection of N strips, each with a width of one pixel, from interpolated camera views in between the original camera positions. In order to synthesize new views we can use various methods, such as optical flow interpolation [13, 31], trilinear tensor methods [28], and others. In most cases approximate methods will give good results. The creation of the intermediate views can involve only view interpolation, as in this application view extrapolation is not needed.

FIGURE 14. Mosaicing from a panning camera which is slightly tilted upward. Regular mosaicing results in a curled image, while rectified mosaicing results in a straight mosaic.

The use of intermediate views for strips collection gives the effect of orthographic projection, which avoids discontinuities due to motion parallax. This strategy can be combined with the methods that were described in the previous sections as a preliminary stage, such that a complete solution is given for general motion in general scenes.

7 Concluding Remarks

Mosaicing on a surface of a manifold, which is determined dynamically based on the motion of the camera, has been introduced. Strips from the images are reprojected onto the manifold using multi-perspective projection.

Manifold mosaicing can be performed by computing the manifold explicitly from the ego motion of the camera, and projecting the images onto that manifold. Alternatively, this projection can be done implicitly by the process of cutting and warping strips, and without explicit computation of the manifold.

Manifold mosaics represent the entire environment of a video shot in a single, static, image. This single image can be used as a summary of the video clip for video browsing, or as a compressed representation of the shot which can be approximately re-generated from the mosaic given the stored motion parameters.

8 References

[1] *ARPA Image Understanding Workshop*, Monterey, California, November 1994. Morgan Kaufmann.

[2] *Fifth International Conference on Computer Vision*, Cambridge, MA, June 1995. IEEE-CS.

[3] *Proc. IEEE Workshop on Representation of Visual Scenes*, Cambridge, MA, June 1995. IEEE-CS.

[4] *SIGGRAPH'95*, Los Angeles, California, August 1995. ACM.

[5] *IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, California, June 1996.

[6] *IEEE Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, June 1997.

[7] *SIGGRAPH'97*, Los Angeles, California, August 1997. ACM.

[8] F. Aurenhammer. Voronoi diagrams: A survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23(3):345–405, September 1991.

[9] J.R. Bergen, P. Anandan, K.J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *European Conference on Computer Vision*, pages 237–252, 1992.

[10] P.J. Burt and E.H. Adelson. A multiresolution spline with application to image mosaics. *ACM Trans. on Graphics*, 2(4):217–236, October 1983.

[11] P.J. Burt and P. Anandan. Image stabilization by registration to a reference mosaic. In *DARPA IUW'94* [1], pages 457–465.

[12] S.E. Chen. Quicktime VR - an image-based approach to virtual environment navigation. In *SIGGRAPH'95* [4], pages 29–38.

[13] S.E. Chen and L. Williams. View interpolation for image synthesis. In *SIGGRAPH'93*, pages 279–288, Anaheim, California, August 1993. ACM.

[14] M. Hansen, P. Anandan, K. Dana, G. van der Wal, and P.J. Burt. Real-time scene stabilization and mosaic construction. In *DARPA IUW'94* [1], pages 457–465.

[15] R. Hartley and R. Gupta. Linear pushbroom cameras. In J.O. Eklundh, editor, *Third European Conference on Computer Vision*, pages 555–566, Stockholm, Sweden, May 1994. Springer.

[16] M. Irani, P. Anandan, and S. Hsu. Mosaic based representations of video sequences and their applications. In *Fifth International Conference on Computer Vision* [2], pages 605–611.

[17] P. Jaillon and A. Montanvert. Image mosaicking applied to three-dimensional surfaces. In *12th International Conference on Pattern Recognition*, pages 253–257, Jerusalem, Israel, October 1994. IEEE-CS.

[18] A. Krishnan and N. Ahuja. Panoramic image acquisition. In *IEEE Conference on Computer Vision and Pattern Recognition* [5], pages 379–384.

[19] R. Kumar, P. Anandan, M. Irani, J. Bergen, and K. Hanna. Representation of scenes from collections of images. In *Proc. IEEE Workshop on Representation of Visual Scenes* [3], pages 10–17.

[20] S. Mann and R. Picard. Virtual bellows: Constructing high quality stills from video. In *First IEEE International Conference on Image Processing*, volume I, pages 363–367, Austin, Texas, November 1994.

[21] L. McMillan and G. Bishop. Plenoptic modeling: An image-based rendering system. In *SIGGRAPH'95* [4], pages 39–46.

[22] D.L. Milgram. Computer methods for creating photomosaics. *IEEE Trans. on Computers*, C-24:1113–1119, 1975.

[23] D.L. Milgram. Adaptive techniques for photomosaicking. *IEEE Trans. on Computers*, C-26:1175–1180, 1977.

[24] S.K. Nayar. Catadioptric omnidirectional cameras. In *IEEE Conference on Computer Vision and Pattern Recognition* [6], pages 482–488.

[25] S. Peleg. Elimination of seams from photomosaics. *Computer Graphics and Image Processing*, 16:90–94, May 1981.

[26] S. Peleg and J. Herman. Panoramic mosaics by manifold projection. In *IEEE Conference on Computer Vision and Pattern Recognition* [6], pages 338–343.

[27] Paul Rademacher and Gary Bishop. Multiple-center-of-projection images. In *SIGGRAPH'98*, pages 199–206, Orlando, Florida, July 1998. ACM.

[28] B. Rousso, S. Avidan, A. Shashua, and S. Peleg. Robust recovery of camera rotation from three frames. In *IEEE Conference on Computer Vision and Pattern Recognition* [5], pages 796–802.

[29] B. Rousso, S. Peleg, I. Finci, and A. Rav-Acha. Universal mosaicing using pipe projection. In *Sixth International Conference on Computer Vision*, pages 945–952, Bombay, India, January 1998. IEEE-CS.

[30] H.S. Sawhney, S. Ayer, and M. Gorkani. Model-based 2D & 3D dominant motion estimation for mosaicing and video representation. In *Fifth International Conference on Computer Vision* [2], pages 583–590.

[31] S. Seitz and C. Dyer. Physically valid view synthesis by image interpolation. In *Proc. IEEE Workshop on Representation of Visual Scenes* [3].

[32] R. Szeliski. Video mosaics for virtual environments. *IEEE Computer Graphics and Applications*, pages 22–30, March 1996.

[33] S. Szeliski and H.-Y. Shum. Creating full view panoramic image mosaics and environment maps. In *SIGGRAPH'97* [7], pages 251–258.

[34] D.N. Wood, A. Finkelstein, J.F. Hughes, C.E. Thayer, and D.H. Salesin. Multiperspective panoramas for cel animation. In *SIGGRAPH'97* [7], pages 243–250.

[35] J.Y. Zheng and S. Tsuji. Panoramic representation for route recognition by a mobile robot. *International Journal of Computer Vision*, 9:55–76, 1992.