NEW VIEW SYNTHESIS WITH NON-STATIONARY MOSAICING

Doron Feldman, Assaf Zomet, Daphna Weinshall, and Shmuel Peleg

peleg@cs.huji.ac.il
School of Computer Science and Engineering
The Hebrew University of Jerusalem, 91904 Jerusalem, ISRAEL

ABSTRACT

Practically all image rendering methods today use the perspective projection, or try to approximate it. We develop a new image-based rendering approach which is based on a new projection model – the Crossed-Slits *X-Slits* projection. The model is obtained by relaxing the requirement that all projection rays intersect at a single point (the focal-point). Instead, projection rays are required by definition to intersect two curves in space.

Image based rendering of X-Slits images is very simple, and is performed by non-stationary mosaicing from a sequence of images captured by a translating pinhole camera. The generated mosaic images are closer to perspective images than traditional pushbroom mosaics. Rendering of virtual walkthroughs can therefore be done merely by using non-stationary mosaicing, without using any 3D model.

1. INTRODUCTION

Perspective projection forms the foundation of imaging. Since our eyes, as well as most of our cameras, observe the world through a pinhole (via a lens whose effects we shall ignore), we are used to viewing images that are generated by perspective projections. Therefore, perspective projection is used in photography and in rendering from 3D models, and presents no special difficulty in these cases. The situation is different in image-based rendering (IBR), where rays are collected from a set of input images and resampled for the generation of new images [12, 4, 11, 8]. In order to create new perspective images of reasonable quality, the requirements become prohibitive: the number of stored rays becomes larger than available memory, and those rays are derived from a very large collection of carefully taken pictures.

In this paper we introduce and study a family of alternative projection models. Specifically, we propose a family of projections defined by two slits - the Crossed-Slits (X-Slits) projection. In the X-Slits model, the projection ray of every 3D point is defined by the line that passes through the point and intersects both slits. The image of

a point will be the intersection of the projection ray with the image surface.

Curiously, a physical X-Slits camera was designed in the 19th century by one of the pioneers of color photography, Ducos du Hauron [9], under the title "transformisme en photographie". Ducos du Hauron thought that X-Slits images would be used in the following (20th) century to "create visions of another world" [13].

Independent of the physical device, we argue that the X-Slits projection model is useful and worthy of our attention. This is because new X-Slits images can be easily generated by non-stationary mosaicing from a sequence of images captured by a translating pinhole camera.

In non-stationary mosaicing vertical strips are sampled from a sequence of frames and stitched together, just like in regular mosaicing. We will often refer to the vertical strips as "columns", with the understanding that they can be wider than a single pixel if necessary. What makes the process "non-stationary" is that the locations of sampled columns are different at different input frames. Non-stationary mosaicing can be used for rendering movies of virtual walkthroughs very efficiently, while providing a strong 3D sensation of parallax, reflections and occlusions.

1.1. The X-Slits Camera: Outline

Fig. 1a shows the basic design of the X-Slits camera as built by Ducos du Hauron in 1888 [9]. A more general design is shown in Fig. 1b. Specifically, the X-Slits camera has two slits l_1, l_2 which should be two different lines in $\mathcal{R}^3$, together with an image plane Π that does not contain any of the slits. For every 3D point not on either of the slits there is a single ray which connects the point with both slits simultaneously. The intersection of this ray with the image plane defines the projected image of the 3D point. The camera in Fig. 1a is a special case of the X-Slits camera, where the two slits are orthogonal to each other and parallel to the image plane. We call this special arrangement the Parallel-Orthogonal X-Slits camera (POX-Slits camera).

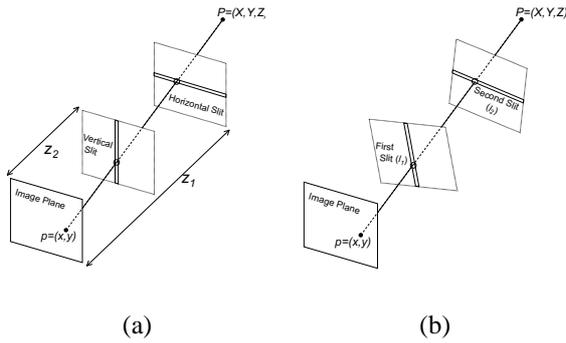The X-Slits model is a valid 3D to 2D projection, defining

Figure 1: *(a) A design of a X-Slits camera where the slits are orthogonal to each other and parallel to the image plane (POX-Slits camera). The projection ray of a 3D point $\mathbf{p} = (\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ is shown, with circles showing its intersection points with the 2 slits. (b) A general X-Slits design, with two arbitrary slits $\mathbf{l}_1, \mathbf{l}_2$. Note that the camera is define by the specific configuration of the two slits and the image plane; any of these three factors can change independently, giving rise to a different X-Slits camera.*

a many-to-one mapping from the 3D world to the 2D image plane. In Section 2 we develop the specific equations of the camera mapping as a function of the slits $\mathbf{l}_1, \mathbf{l}_2$ and the image plane Π.

In Section 3 we discuss how to generate new X-Slits images from images taken by a regular pinhole camera translating along a line in 3D space. Specifically, we show that virtual X-Slits images can be generated by non-stationary mosaicing of perspective images, or equivalently by slicing the space-time volume[1] of images. A readily available tool for slicing this volume is the "video cube" [10].

Applications are discussed in Section 4. The first application includes the generation of an arbitrary virtual walkthrough from a single sequence of images without using any 3D model. The results show that in many practical cases the discrepancies between X-Slits images and perspective images are hardly noticeable. The second application is 3D object visualization. The X-Slits model allows us to virtually place one slit behind an object, keeping the other slit in front of the object. Such visualization can be useful when one wants to see simultaneously different sides of an object. We note that both applications are done by plain mosaicing, without recovering the scene or object geometrical model.

1.2. Relation to Previous Work

Much work was done on non-perspective projection models [14] and the generation of non-perspective images from

video sequences [15, 16, 20, 7]. In most cases these images are used as a visual summary of the video, or for 3D visualization. The present work uses a similar mosaicing technique with one important difference: the mosaiced strips are sampled from varying positions in the input images. This makes the generation of virtual walkthroughs possible.

In many image-based rendering (IBR) techniques rays from a set of input images are collected and a new image is rendered by resampling the stored rays [12, 4, 11]. In order to create new perspective images of reasonable quality, the requirements become prohibitive: the number of stored rays becomes larger than available memory, and those rays are derived from a very large collection of carefully taken pictures. There are attempts to make IBR more efficient and more general [3, 1], or to use such approximations as moving the camera in a lower dimensional space [17, 18]. The present work is mostly related to [17, 18, 3] with several differences: First, rather than trying to approximate the perspective projection, we accurately define the projection geometry of the resulting images, and analyze the model limitations. Second, the rendering tool is very simple - slicing of the space-time volume obtained by a simple motion of a perspective camera. Consequently the most important feature of our technique is the fact that ray-sampling for the generation of new views does not require detailed accounting of the parameters of the generating images. As we show below, if the camera's motion is sideways and constant, every vertical planar slice of the space-time volume gives *some* valid X-Slits image.

The idea of using linear non-stationary mosaicing for IBR and the generation of a virtual walkthrough was originally reported in [19]. But while the bi-centric camera model analyzed in [19] is mathematically equivalent to the POX-Slits analyzed below, the general X-Slits model and the realization with two slits are new. This realization allows for a wider range of applications, as discussed in Section 4.

2. THE X-SLITS PROJECTION

In Section 2.1 we derive the projection model of the X-Slits camera. Properties of this model are discussed in Section 2.2. Special configurations of slits and planes are discussed in Section 2.3. All derivations are done in the projective spaces $\mathcal{P}^2$ and $\mathcal{P}^3$, where points in $\mathcal{R}^2$ and $\mathcal{R}^3$ respectively are represented by homogeneous coordinates.

2.1. Projection Geometry

Consider the camera configuration as shown in Fig. 1b. The *projection ray* of a point $\mathbf{p} \in \mathcal{P}^3$ should intersect the two camera slits $\mathbf{l}_1, \mathbf{l}_2$. The point $\mathbf{p}$ defines a plane with each of the slits, since a plane is defined by a point and a line. These two planes intersect in a line, the projection

---

[1]The space-time volume is simply a stacking of the input images, a.k.a. the epipolar volume.
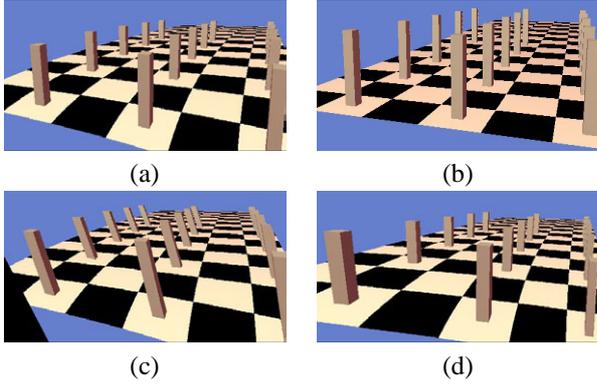
Figure 2: *Simulated X-Slits images of an artificial scene; (a) POX-Slits image using the projection equation from Eq. (2). (b) A regular pinhole image. (c) Same as (a), but with the vertical slit rotated about the Z axis. (d) Same as (a), but with the vertical slit rotated about the X axis.*

ray of $\mathbf{p}$. The image of $\mathbf{p}$ is the intersection of this projection ray with the image plane. Furthermore, all points on a projection ray will project to the same image point (unless the ray lies on the image plane).

We can show that the projection equations from $\mathcal{P}^3$ to $\mathcal{P}^2$ of the X-Slits projection model can be described as follows:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} \propto \begin{pmatrix} \mathbf{p}^T \mathbf{S}_1 \mathbf{Q}_1 \mathbf{S}_2 \mathbf{p} \\ \mathbf{p}^T \mathbf{S}_1 \mathbf{Q}_2 \mathbf{S}_2 \mathbf{p} \\ \mathbf{p}^T \mathbf{S}_1 \mathbf{Q}_3 \mathbf{S}_2 \mathbf{p} \end{pmatrix} \qquad (1)$$

where the 3D point $\mathbf{p}$ is projected to the 2D image point $(\frac{x}{w}, \frac{y}{w})$. All 4×4 matrices $\mathbf{S}_1, \mathbf{S}_2, \mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3$ are antisymmetric; moreover, each $\mathbf{S}_i$ depends only on optical slit i, while $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3$ depend only on the plane of projection Π. Matrices $\mathbf{Q}_1$, $\mathbf{Q}_2$, and $\mathbf{Q}_3$ are the Plücker matrix representations of the image X axis, the image Y axis and the image line at infinity, respectively (all in 3D).

2.2. Properties of the X-Slits Camera

To get some intuition for X-Slits images, Fig. 2 shows examples of X-Slits images as compared with pinhole images of the same scene. We also show the effects of varying the relative geometry of the slits and image plane.

2.2.1. Projection of Straight Lines

It can be shown that in X-Slits images, the image of a 3D line is a conic. The distortion of straight lines is illustrated in Fig. 2. In practice this distortion is not very disturbing, as can be seen in the examples in Section 4. In the various scenes we have experimented with, this distortion was

rather minor. Many scenes, particularly natural scenes, do not have very dominant straight lines, in which case this distortion is hardly noticed. Furthermore, people are accustomed to similar effects caused by lens distortions.

2.3. Special Case: the POX-Slits Projection

A configuration of special interest is when the two slits are orthogonal to each other and parallel to the image plane. We will refer to this configuration as POX-Slits (see Fig. 1a).

We first fix the slits by assigning $\mathbf{u}_1 = (1,0,0,0)^T$, $\mathbf{v}_1 = (0,0,1,-Z_1)^T$, $\mathbf{u}_2 = (0,1,0,0)^T$, and $\mathbf{v}_2 = (0,0,1,-Z_2)$. This defines a vertical slit at $X = 0$, $Z = Z_1$ and a horizontal slit at $Y = 0$, $Z = Z_2$. We also denote $\Delta = Z_2 - Z_1$. Plane Π is the $X - Y$ plane at $Z = 0$, and therefore we can assign $\mathbf{m} = (0,0,0,1)^T$, $\mathbf{j} = (1,0,0,0)^T$, $\mathbf{k} = (0,1,0,0)^T$. From (1) it follows that

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -Z_1 \frac{X}{Z-Z_1} \\ -Z_2 \frac{Y}{Z-Z_2} \end{pmatrix} \qquad (2)$$

These projection equations are identical to the model analyzed in [19], where it was called bi-centric projection.

2.3.1. Projection of Straight Lines

Recall that under the perspective projection straight lines are projected to straight lines, and under the X-Slits projection straight lines are projected to conic sections. Let us look more closely at the images of lines in the POX-Slits case. If the 3D line is perpendicular to the Z axis, i.e., of the form $(a + c\lambda, b + d\lambda, e)$, then its image is the curve $(x,y) = \left(-Z_1 \frac{a+c\lambda}{e-Z_1}, -Z_2 \frac{b+d\lambda}{e-Z_2} \right)$, which is a line. If the 3D line is *not* perpendicular to the Z axis, i.e., of the form $(a + c\lambda, b + d\lambda, \lambda)$, then its image is the curve

$$(x,y) = \left(-Z_1 \frac{a+c\lambda}{\lambda - Z_1}, -Z_2 \frac{b+d\lambda}{\lambda - Z_2} \right) \qquad (3)$$

Solving for λ, we can find that the straight line is projected to an hyperbola.

2.3.2. Aspect Ratio Distortions

The most apparent aspect of the distortion in POX-Slits images is the variation of aspect-ratio (especially in pushbroom images [5]). The apparent aspect-ratio of objects in the image depends on their depth. This is unlike the perspective model, in which the distortion in aspect-ratio is constant for all objects.

From (2) it follows that an object at depth Z with aspect-ratio of 1 would appear on the image plane to have an

aspect-ratio of

$$\frac{\Delta y}{\Delta x} = \frac{Z_2}{Z_1} \cdot \frac{Z - Z_1}{Z - Z_2} \tag{4}$$

In practice, we found this distortion to be typically rather insignificant. If the range of depths of scene objects is not too large, we can normalize the image to compensate for this distortion by scaling, as discussed in Sect. 3.2.2. Specifically, if we cancel the aspect-ratio distortion for some intermediate depth value Z_0, the distortion at depth Z would be

$$\frac{Z - Z_1}{Z - Z_2} \cdot \frac{Z_0 - Z_2}{Z_0 - Z_1} \tag{5}$$

To demonstrate the magnitude of the distortion, consider the following example: Suppose the depth range of objects in the scene is $3 - 5$ meters (measured from the horizontal slit at Z_1, i.e., $3m < Z - Z_1 < 5m$), and assume that the images are normalized so that objects at the depth of $3.84m$ appears undistorted (i.e., $Z_0 - Z_1 = 3.84m$). If the vertical slit is behind the horizontal slit at $\Delta = -2.5m$, the aspect-ratio distortion would not exceed 10%.

2.3.3. X-Slits vs. Other Projections

Many projection models are special cases of the POX-Slits camera:

- Perspective Projection – the two slits intersect; the intersection point is the optical center of the perspective projection.

- Parallel Projection – both slits are at infinity.

- Linear Pushbroom [5] – the vertical slit resides on the plane at infinity. It was shown in [5] that a line in 3D is projected by this projection model to a hyperbola in the image. This is a special case of the result shown in Section 2.3.1.

3. NON-STATIONARY MOSAICING

Our goal is to synthesize new X-Slits views from "regular" perspective images. The input sequence is assumed to be captured by a pinhole camera translating along a horizontal line in 3D space in a roughly constant speed, and without changing its orientation or internal calibration. As we show below, we can generate a new X-Slits image where the two slits of the underlying virtual X-Slits camera are defined as follows:

1. A horizontal slit lies on the path of the optical center of the moving pinhole camera.

2. A vertical slit that is parallel to the image's vertical axis, and its location is determined by the parameters of the mosaicing process.

New view synthesis is performed by non-stationary mosaicing defined as follows:

- From each frame t, sample the vertical column (strip) centered on the horizontal coordinate $s(t)$.

- Paste the strips together into a single image, as in "regular" mosaicing.

In the general case we may sample slanted strips rather than vertical columns (strips), and the orientation may also change as a function of t. In this case the vertical slit of the underlying virtual camera will not be the image's vertical axis, but some other direction parallel to the image plane. The parameters of the strip sampling function $s(t)$ determine the exact location of the vertical slit of the virtual camera. Using this observation, a virtual walkthrough is the result of generating a sequence of X-Slits images via non-stationary mosaicing, while moving the vertical slit along a planar path. Adjusting the image plane orientation is done by warping the mosaiced image.
In Section 3.1 we show how to sample vertical strips from the input images in the sequence in order to generate a valid X-Slits image. We also discuss the relation between the sampling function $s(t)$ and the parameters of the virtual X-Slits camera. In Section 3.2 we discuss implementation issues, including the deviation from constant speed and aspect-ratio normalization.

3.1. Non-Stationary Strip Sampling

We start our analysis in with the simplest case where the input image sequence is generated by a camera moving sideways in a direction parallel to the X-axis of the image. The camera is also assumed to be internally calibrated. In this simple case the new synthesized image is in fact a POX-Slits image (see Section 2.3), and the non-stationary strip sampling is a linear function. We show below the exact relation between the parameters of the linear sampling function and the parameters of the virtual POX-Slits camera. It can also be shown that a linear strip sampling function, even when the camera is not internally calibrated, always results in an X-Slits image (but not necessarily POX-Slits). When the motion of the camera is not parallel to the image plane, the sampling function is not linear anymore. When the basic assumptions of the analysis are violated, namely, the camera changes its orientation and internal calibration arbitrarily along the input sequence, we need to preprocess the sequence. One solution involves registering all the images with each other using the homography of the plane at infinity. This computation requires,
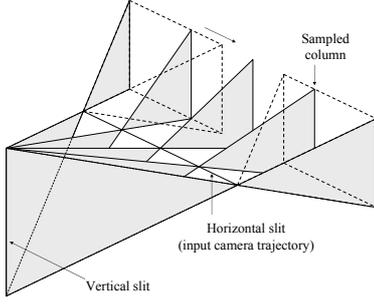
Figure 3: *The non-stationary column sampling routine which is used to synthesize new images.*

a)

b)

Figure 4: *New image formation with two possible positions of the vertical slit (see text).*

however, either internal camera calibration or some domain knowledge (such as parallel lines in the scene).

Let our input be a sequence of images captured by a pinhole camera translating in constant speed along the X axis from left to right. We generate a new panoramic image by pasting columns from the input images, as illustrated in Fig. 3. We start by sampling the left column of the first (leftmost) image, and conclude by sampling the right column of the last (rightmost) image. In between, intermediate columns are sampled from successive images using a linear sampling function.

A schematic illustration of this setup is given in Fig. 4a, in a top-down view. A sequence of positions of the real pinhole camera is shown, together with the corresponding field of view. The moving input camera, whose optical centers are located at positions $\mathbf{c}(t) = (X_t, 0, 0)$, generates images according to the following mapping:

$$\mathbf{p} = (X, Y, Z) \implies p = (x, y) = (f\frac{X - X_t}{Z}, f\frac{Y}{Z}) \quad (6)$$

We denote the range of columns (x) in each pinhole image as $[-r, r]$, and the range of camera pinhole positions (X_t) as $[-l, l]$ (see Fig. 4a). The new synthesized image is constructed by pasting columns from the input images. The range of columns in the synthesized image is $[-(r + l), r + l]$. For each $t \in [-1, 1]$, we assign to the $(l + r)t$ column of the new image the image values at the rt column of the pinhole camera positioned at $(lt, 0, 0)$ (i.e., $X_t = lt$, see Fig. 4a). It now follows from Eq. (6) that $rt = f\frac{X - lt}{Z}$. In addition, for each column $x \in [-(r + l), (r + l)]$ in the new image, $t = \frac{x}{l+r}$ and therefore

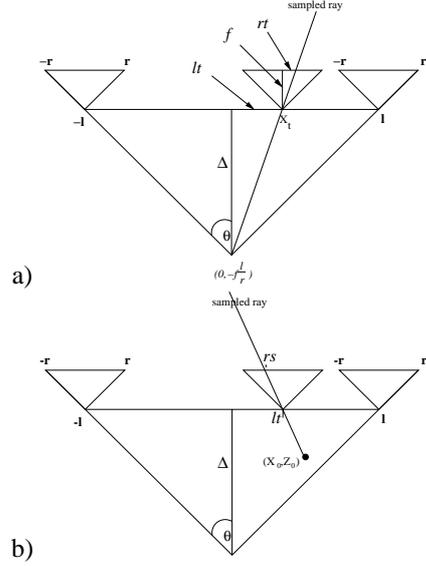$$X = \frac{rt}{f}Z + lt = x \left(\frac{r}{l+r} \cdot \frac{Z}{f} + \frac{l}{l+r} \right)$$

or

$$x = \frac{l + r}{r}f \cdot \frac{X}{Z + f\frac{l}{r}}$$

Observe that this defines a vertical slit at $Z = -f\frac{l}{r}$ (see Fig. 4a). The horizontal slit is at $Z = 0$, same as for each pinhole camera. Eq. (6) therefore becomes the following projection

$$\mathbf{p} = (X, Y, Z) \implies p = (x, y) = (f_x\frac{X}{Z + \Delta}, f_y\frac{Y}{Z}) \quad (7)$$

where $f_x = \frac{l+r}{r}f$ is the horizontal focal length, $f_y = f$ is the vertical focal length, and $\Delta = f\frac{l}{r}$ is the distance between the two slits.

Suppose next that instead of taking the rt column from the camera at $(lt, 0, 0)$, we choose an arbitrary linear column sampling function. More specifically, for $t = \alpha s + \beta$, we take the rs column of the lt camera, see Fig. 4b. (Recall that r, l are fixed, while t, s are free parameters which determine the rate of column sampling). Let the field of view of the original pinhole camera be 2θ. It can be shown that such a choice of columns defines the mapping

$$(x, y) = ((f + \frac{\alpha l}{\tan \theta})\frac{X - \beta l}{Z + \frac{\alpha l}{\tan \theta}}, f\frac{Y}{Z}) \quad (8)$$

This can be written simply as

$$(x, y) = (f_x\frac{X - X_0}{Z + \Delta}, f_y\frac{Y}{Z}) \quad (9)$$

where $X_0 = \beta l$, $\Delta = \frac{\alpha l}{\tan \theta}$, $f_y = f$ and $f_x = f + \Delta$.

The method described so far produces images which do not follow the perspective projection model, since the focal lengths are not the same vertically and horizontally. They do, however, follow the X-Slits projection model defined above. To see this, we observe that all rays producing the image must intersect the following two lines:

1. The line of camera motion; this is because each projection ray must be collected by some camera whose optical center is on this line.

2. The vertical line located at (X_0, Z_0) (as in Eq. (9), where $Z_0 = Z + \Delta$); this is shown above to be the vertical slit.

The projection model is therefore defined by a family of rays intersecting a pair of lines (*"slits"*), projecting 3D points onto a plane. Moreover, the model is POX-Slits (compare Eq. (9) with Eq. (2)).

In the derivation leading to (9) we effectively showed that any linear sampling function yields a valid new POX-Slits image. Furthermore, we can set the location of the vertical slit to (X_0, Z_0) by fixing $\alpha = -\frac{Z_0}{l} \tan \theta$ and $\beta = \frac{X_0}{l}$. This result enables us to synthesize new views of the scene with any vertical slit of our choice, by sampling the columns of the original input sequence according to $t = \alpha s + \beta$, with α and β assigned the appropriate values.

3.2. Implementation Issues

In this section we discuss what to do when the motion of the camera deviates from constant speed (Section 3.2.1), and how the aspect ratio of the resulting mosaic is determined (Section 3.2.2). We also present an alternative implementation of mosaicing, namely the slicing of space-time volume (Section 3.2.3).

3.2.1. Variable Camera Speed

When the camera moves in a linear trajectory but varying orientation and speed, we compensate for this variability by estimating camera motion (see [6]) and by derotating the image planes. We found that when the changes in camera orientation are small, a simple approximation is sufficient. Specifically, we compute the 2D rotation and translation between consecutive input frames using the method described in [2], and warp the images to cancel any 2D rotation and vertical translation. The residual 2D translation is used to estimate the 3D velocity of the translating camera, and which determines the thickness of the vertical strip. This approach is similar to the pushbroom mosaicing technique described in [15].

3.2.2. Aspect-Ratio Normalization

The most apparent aspect of the distortion in X-Slits images is the variation of aspect-ratio, as analyzed in Section 2.3.2. To reduce this distortion, we vertically scale the new images. This normalization is essential for achieving compelling results.

Specifically, the distortion on the *image plane* of objects at depth Z given in Eq. (4) can be written as $\frac{Z}{Z+\Delta} \cdot \frac{f_y}{f_x}$ in the notation of Eq. (9). In order to keep the horizontal field-of-view angle constant in the walk-through animation, we sample all the columns from left to right (from the appropriate frames, according to the column sampling function). Without any scaling, this process generates an image in which only the plane at infinity ($Z = \infty$) appears undistorted. Therefore, in order to cancel the distortion at depth Z_0, we must scale the image vertically by the factor:

$$1 + \frac{\Delta}{Z_0} \qquad (10)$$

3.2.3. The Space-Time Volume

In Section 3.1 we described how to synthesize a X-Slits image by sampling columns from the input images. Specifically, we discussed the following linear sampling formula:

$$t = \alpha s + \beta \qquad (11)$$

where t denotes the camera translation. Recall that α, β are free parameters which control the location of the vertical slit.

A useful representation for the visualization of this process is the Space-Time Volume (or the *epipolar volume*), which is constructed by stacking all input images into a single volume. In case of constant sideways camera motion, any vertical planar slice in the volume according to (11) is a X-Slits image. This process is illustrated in Fig. 5; it assumes that the input sequence has high frame-rate and negligible spatial aliasing, so that simple interpolation (such as bilinear or bicubic) of the volume is sufficient. Thus rendering new X-Slits images is as simple as slicing a plane in the space-time volume.

4. EXPERIMENTAL RESULTS

We use a camera moving in an horizontal plane. As discussed above, new view generation in this case is done by sampling vertical strips from successive images and pasting them together into individual X-Slits images. The parameters of the atrip sampling function determine the location of the vertical slit of the X-slit camera. In our experiments below, we manipulated the parameters of the sampling function so that the location of the vertical slit
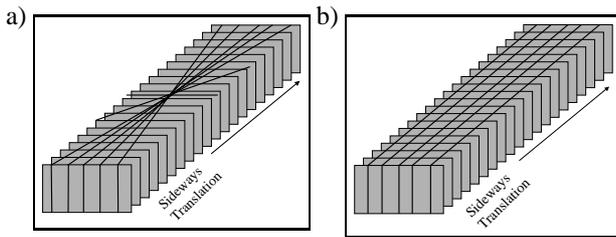
Figure 5: *A schematic description of images generated as slices in the space-time volume. a) Changing the orientation of the slice moves the vertical slit inside and outside the scene. b) The central slice gives a pushbroom image (the "traditional" mosaic). Sliding parallel slices in the space-time volume results in different viewing directions of skewed pushbroom images.*

moves according to the desired ego-motion. A very compelling impression of camera motion is obtained, even though the horizontal slit of the X-slits camera, which is the trajectory of the input camera, remains fixed.

Next we discuss two applications: the generation of a virtual walkthrough from a sequence of perspective images (Sections 4.1 and 4.2), and 3D object visualization (Section 4.3). Another application, which is not demonstrated here, is the generation of oblique stereo panoramas. Specifically, we can generate two X-Slits images using the sampling function from (11) with the same α but two different β's. This gives us a stereo pair, where the two images have an identical horizontal slit but a shifted vertical slit.

The input video sequences used in our examples, as well as the synthesized walkthrough movies, are currently available on the web at
http://www.cs.huji.ac.il/~daphna/demos.html.

4.1. Virtual Walkthrough

In the first experiments (Figs. 6-8) we synthesized new sequences which correspond to a camera motion that has forward motion component, with visible parallax and lighting effects. In addition, the direction of the image plane was changed quite a bit.

Another example used a real sequence taken by a helicopter flying along a rocky coast in an unknown path and viewing direction (Fig. 9). Here we synthesized a new sequence which corresponds to a forward moving camera. This sequence was more challenging since the input sequence was taken in free motion with random disturbances (e.g., the effect of wind), and thus motion compensation was required.
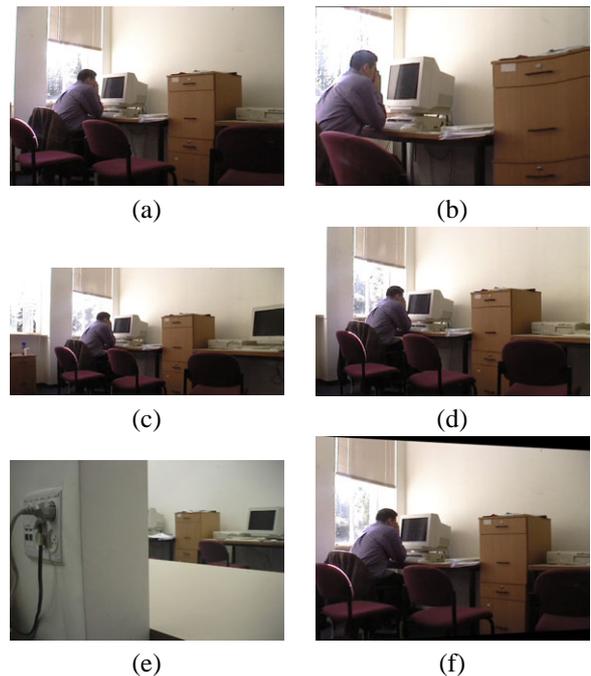

Figure 6: *This scene is located in a small room where moving backward to capture the whole room is impossible. The scene was filmed by a sideways moving camera, total of 591 frames; one of the original frames is shown in (a). We show three new images: one where the vertical slit is located in front of the original track (b), and two where the vertical slit is located behind the original track (c-d). For comparison, we took a normal (pinhole) picture from the same location as (c), where part of the scene is obscured by the wall; this picture is shown in (e), and it demonstrates our ability to make images from impossible camera positions. Finally, (f) shows a simulated image where the camera was translated and rotated.*

4.2. New Views of Extended X-Slits Images

In this example we show how to generate new views from a sequence of a rotating object, where the new sequence demonstrates forward motion with parallax (Fig. 10). The projection model of the new images corresponds to to the case where one slit is a circular slit.

4.3. Object Visualization

Here we demonstrate the use of the X-Slits projection for object visualization. Specifically, we show how an object can be "flattened", revealing several of its sides simultaneously. This is done by positioning the vertical slit behind the object (Fig. 11). Since the image is a valid X-Slits image that can be characterized and analyzed, we need not worry about such issues as duplicate images, which usu-
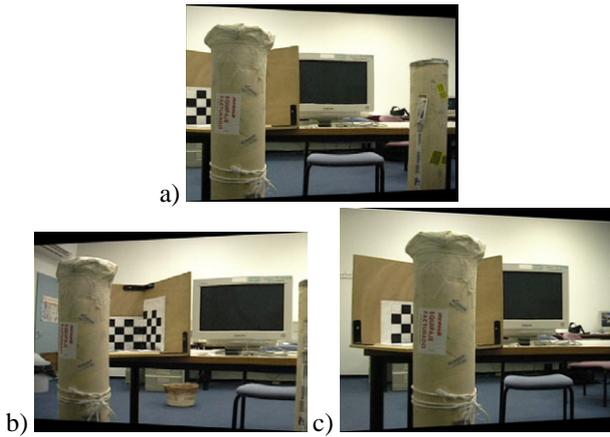
Figure 7: *This scene was filmed by a sideways moving camera in our lab, total of 567 frames. We generated a X-Slits movie where the virtual camera rotates about an object in the scene (a-b), and then translates ahead in a diagonal (c).*

ally require hand-crafted stitching.

5. CONCLUDING REMARKS

We presented a new non-perspective projection model, which is defined by two slits and a projection surface. The main application we pursued in this paper is new view generation, or image based rendering. New view generation with the X-Slits camera is greatly simplified as compared with perspective new view generation, as it is performed by non-stationary mosaicing, or by slicing the space-time volume. The X-Slits theory helps the user to "drive" the slicing process in order to get the desired effect. When compared to traditional mosaicing, X-Slits images can be shown to be closer to perspective images than linear push-broom images.

Using our method we can also generate new images taken from "impossible" positions, like behind the back wall of a room or in front of a glass barrier. Movies with new ego motion can also be generated, such as forward-moving movies from a side-moving input sequence. Although not perspective, the movies generated in this way appear compelling and realistic.
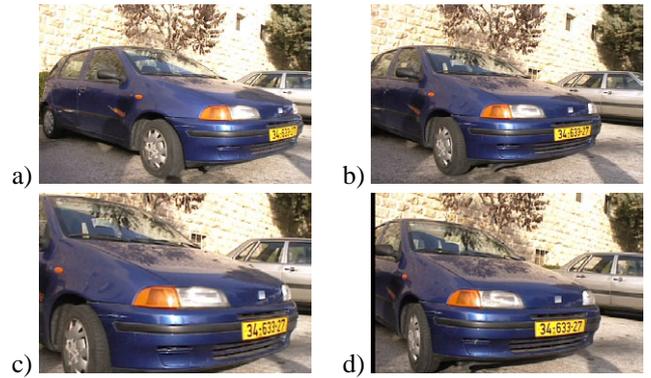
Acknowledgments

Figure 8: *Virtual walkthrough from a translating camera. a, b) Two frames from the input sequence. c, d) Two images rendered in forward motion. Note the apparently realistic changes in parallax and reflection.*
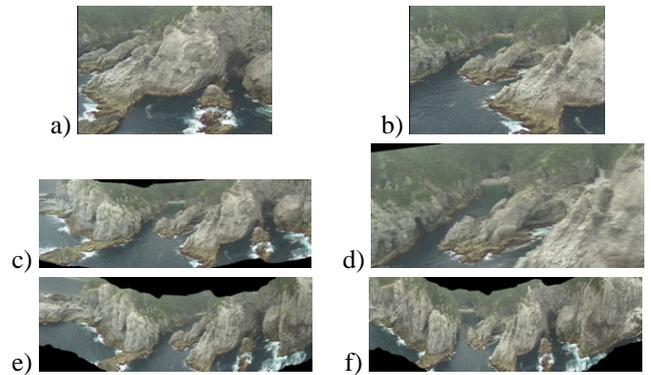

Figure 9: *Virtual walkthrough generated from a sequence taken by a freely flying helicopter. a, b) Two frames from the input sequence. c, d) Two images rendered in forward motion (diagonal slices). e, f) Two images rendered in different viewing angles (parallel slices).*
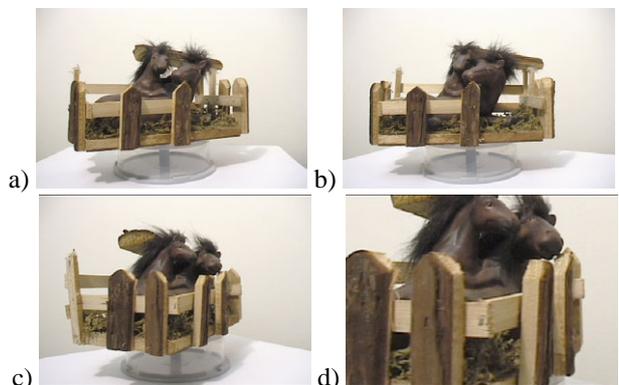

Figure 10: *A rotating object: a, b) Two images from the original sequence of a rotating object. c, d) Two synthesized images from a forward moving viewpoint.*
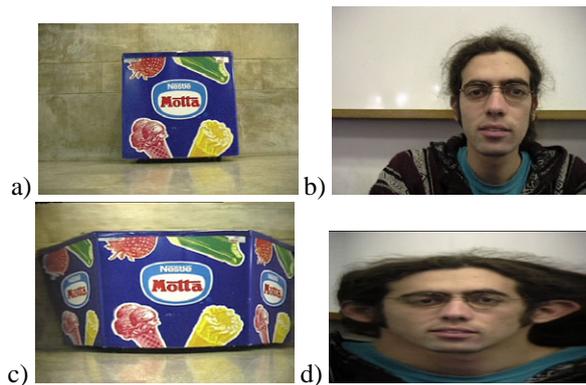
Figure 11: *Object representation. a, b) The original input images. c, d) Visualization with the vertical slit located behind the object. The object is seen as if "opened" inside-out, giving a cubist effect: multiple sides are seen in a single picture.*

us to the relevant events in the history of photography, and Tomas Pajdla for many useful comments.

6. REFERENCES

[1] D. Aliaga and I. Carlbom. Plenoptic stitching: A scalable method for reconstructing interactive walkthroughs. In *Proc. of ACM SIGGRAPH'01*, pages 443–450, 2001.

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

[3] C. Buehler, M. Bosse, S. Gortler, M. Cohen, and L. McMillan. Unstructured lumigraph rendering. In *Proc. of ACM SIGGRAPH'01*, pages 425–432, 2001.

[4] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen. The lumigraph. In *Proc. of ACM SIGGRAPH'96*, pages 43–54, 1996.

[5] R. Gupta and R. I. Hartley *Linear Pushbroom Cameras IEEE PAMI*, 9(19):963-975, 1997.

[6] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[7] H. Ishiguro and M. Yamamoto and S. Tsuji Omni-Directional Stereo *IEEE PAMI*, 14(2):257–262, 1992.

[8] S. Kang. A survey of image-based rendering techniques. In *Videometric VI*, 3641:2–16, Jan 1999.

[9] R. Kingslake. *Optics in Photography*. SPIE Optical Engineering Press, 1992.

[10] A. Klein, P. Sloan, A. Colburn, A. Finkelstein, and M. Cohen. Video cubism. Technical Report MSR-TR-2001-45, Microsoft Research Technical Report, 2001.

[11] M. Levoy and P. Hanrahan. Light field rendering. In *Proc. of ACM SIGGRAPH'96*, pages 31–42, 1996.

[12] L. McMillan and G. Bishop. Plenoptic modeling: An image-based rendering system. In *Proc. of ACM SIGGRAPH'95*, pages 39–46, Los Angeles, California, August 1995.

[13] B. Newhall. *The History of Photography, from 1839 to the present day*, p. 162. The Museum of Modern Art, 1964.

[14] T. Pajdla. Stereo with Oblique Cameras. *International Journal of Computer Vision*, 47(1/2/3):161-170, 2002.

[15] S. Peleg, B. Rousso, A. Rav-Acha and A. Zomet. Mosaicing on Adaptive Manifolds *IEEE PAMI*, 10(22):1144–1154, 2000.

[16] S.M. Seitz. The Space of All Stereo Images In *Proc. of ICCV'01*, pages I: 26-33, Vancouver, Canada, July 2001.

[17] H. Shum and L. He. Rendering with concentric mosaics. In *Proc. of ACM SIGGRAPH'99*, pages 299–306, Los Angeles, California, August 1999.

[18] T. Takahashi, H. Kawasaki, K. Ikeuchi, and M. Sakauchi. Arbitrary view position and direction rendering for large-scale scenes. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 296–303, Hilton Head, SC, June 2000.

[19] D. Weinshall, M. Lee, T. Brodsky, M. Trajkovic, and D. Feldman. New view generation with a bi-centric camera. In *Proc. of Seventh European Conference on Computer Vision*, 614–628, Copenhagen, Denmark, May 2002.

[20] J.Y. Zheng and S. Tsuji. Panoramic Representation for Route Recognition by a Mobile Robot *International Journal of Computer Vision*, 9(1): 55-76, October 1992.