Rendering Almost Perspective Views from a Sparse Set of Omnidirectional Images*

Hynek Bakstein, Tomáš Pajdla, and Daniel Večerka
Center for Machine Perception, Dept. of Cybernetics
Faculty of Electrical Eng., Czech Technical University
121 35 Prague, Czech Republic
{bakstein,pajdla,vecerka}@cmp.felk.cvut.cz

Abstract

Non-central (X-slits) images can be used in image based rendering (IBR) for new view synthesis for a viewer moving in a restricted region. Novel, i.e. previously unseen, views are rendered from virtual camera positions from a set of images captured at some predefined camera positions. The novel views correspond to camera positions not contained in the input image set. For instance, the input sequence consists of images captured on a circular path, but the viewer can move in a disk defined by that path. We investigate a class of non-central images, the omnidirectional X-slits images, which capture environment around the viewer completely. The virtual viewer can rotate his head freely in each position without the need of creating another image for different viewing direction. New images have to be generated only in case of a change in position of the viewer. This approach allows efficient representation of high resolution (and thus space demanding) images for IBR in memory.

1 Introduction

Image based rendering (IBR) is an approach to representation of scenes by a set of light rays without explicit 3D model of the scene, which can be hard or impossible to create for a complex real scene from 2D images. A straightforward way to a representation of a scene is to capture a complete Plenoptic function [1]. But even if we ignore the time and the wavelength, we have to capture a five parameter set of rays, which results in unfeasibly huge amount of data.

There are numerous approaches to reduction of the dimensionality of the plenoptic function [5, 6, 7, 4, 3, 11, 10]. In general, they result in a restriction of the movement of the viewer. For example, when just a panoramic image is used [5], the viewer is only allowed to rotate his head. The amount of data can also be reduced by incorporating some information about the scene, for example assumption about depth of the objects in the scene. For an overview of IBR methods, see [8].

*This work was supported by the following grants: GACR 102/01/0971, BeNoGo IST-2001-39184, MSMT Kontakt 22-2003-04, and MSM 212300013.

BMVC 2003 doi:10.5244/C.17.25

We focus on IBR using X-slits camera model [11], where the X-slits camera consist of two so called slits to which all the light rays are incident. Therefore, the X-slits camera is an example of a non-central camera. X-slits images are composed from perspective images captured by a pinhole camera moving on some path, for example line or circle. One of the slits is the path on which a pinhole camera capturing an input sequence was moving. The second slit is, e.g., a line, which can be moved in some limited region, see Figure 1(a). In our case, the pinhole camera was following a circular path.

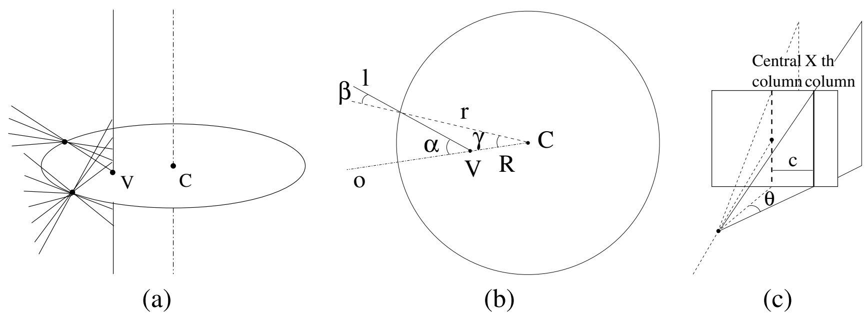

Figure 1: (a) X-slits camera with one circular slit, the circle of rotation, and one vertical slit, the line perpendicular to the plane containing the circle of rotation and intersecting this plane at a point V. (b) Column α of an X-slits image is composed from light rays containing angle β with the central column in an image at position γ. (c) Angle θ corresponds to the column c in the image.

The main contribution of this paper is aspect ratio compensation of omnidirectional X-slits images. Aspect ratio of some perceptionally dominant planar object in the scene, e.g. a computer screen in an office, is important for the amount of presence, the viewer perceives while looking at the images. If the aspect ratio of such object is right, minor deviations from correct perspective in the scene is not noticed. If the aspect ratio is not right, the viewer immediately notices that something is wrong.

The remainder of the paper is organized as follows. In Section 2 we briefly introduce omnidirectional X-slits images and a concept of a viewing window used to present almost perspective views with a limited field of view to the viewer. Section 3 describes aspect ration compensation of omnidirectional X-slits images. Finally, we demonstrate the theory in practical experiments in Section 4.

2 Rendering from omnidirectional X-slits images

The theory of X-slits cameras was recently presented in [12]. However, [12] supposes that the simulated camera has a certain, limited, field of view. We have proposed generalization of concentric mosaics to omnidirectional concentric mosaics [2] and the same approach can be used also for X-slits images. Since we moved an omnidirectional camera (with 180° field of view) on a circular path, we can compose omnidirectional X-slits images covering the whole environment around the viewer, see Figure 8(b) for an example image. Note significant distortion of objects especially at the top and bottom part of the image. This is due to the fact that omnidirectional X-slits images cannot be mapped on a plane. They can be, however, mapped on a torus.

We can compose the X-slits images from omnidirectional images, provided that we can select pixels corresponding to columns in classical images. Each column should correspond to pixels lying in a plane containing an angle θ with the plane containing the optical center of the camera and the central column, as it is illustrated in Figure 1(c). That means that the omnidirectional sensor has to be calibrated.

The omnidirectional X-slits image is composed in the same way as it was proposed in [12] for images, where one of the slits was circular

$$\beta = \arcsin\left(\frac{R}{r}\sin(\alpha)\right) \ , \qquad (1)$$

$$\gamma = \beta - \alpha \ . \qquad (2)$$

Note that the position of columns in the omnidirectional and pixels in the original images corresponds to angles α and β respectively. The index of image in the sequence represents the angle γ, see Figure 1(b). To create omnidirectional X-slits images, we let the value of α go from 0 to 2π. The fraction $\frac{R}{r}$, where R denotes distance of the vertical slit from the center of the circle of rotation and r stands for the radius of this circle, determines the position of the vertical slit. The value of γ determines the image index. It's zero value specifies the direction in which the viewer moves. The direction in which the viewer looks does not affect the mosaic construction, since it covers all 360° in horizontal direction.

One omnidirectional X-slits image allows the viewer to look around in arbitrary direction. When the user moves, we have to select different omnidirectional X-slits images. However, these images will be close to the current image, supposing that the user does not jump but moves smoothly. Therefore, we can keep more than one omnidirectional X-slits image in the memory and while the user moves, we just load predictively more images into the memory while discarding the old ones. This assures, that the memory requirements are reasonable. When the user stops, he can start looking around without any need to load other images into the memory. The density of X-slits images, and thus their number, determines whether the viewer will perceive continuous or discontinuous movement. The necessary number of images can be estimated in a similar was as in [9].

Omnidirectional X-slits mosaics are not intended for direct displaying to the viewer. Usually, only a small part of the whole mosaic is mapped into an image presented to the viewer. This applies also to omnidirectional mosaic images, where strong distortions have to be compensated prior to mapping the part of the image. The whole mosaic image cannot be mapped to a plane, but part of it can. We realize this by a so called *viewing window*, which is sliding on the mosaic images as the viewer changes his viewing direction, see Figure 3. Note that the images represented by the viewing window are not perspective, there is no single center of projection. For a human brain, however, they are 'almost' perspective, if aspect ratio of some perceptionally dominant object is correct.

3 Aspect ratio compensation

We simulate movement of the viewer by moving the vertical slit. The resulting omnidirectional X-slits image should correspond to that movement. Not only the occlusions should change, but also image of part of the scene towards which the viewer moves shall

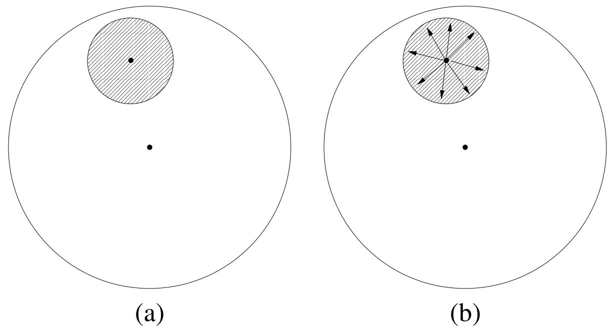

Figure 2: (a) We keep several omnidirectional X-slits images in memory representing a region around the viewer. When the viewer moves, we do not have to load the images into the memory. (b) Each omnidirectional X-slits image allows the viewer to freely look around with the need to change the image, it captures light rays in every direction from the viewers position.

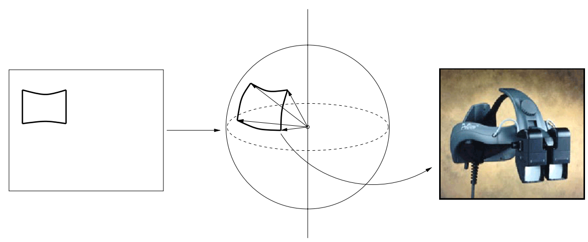

Figure 3: Viewing window simulates the limited field of view of the eyes of the viewer. Image of the HMD is from Kaiser Electro-Optics, Inc.

be enlarged, while the part of the scene in the opposite direction should be displayed smaller. The occlusion do indeed change and due to the fact that closer parts of the scene are sampled with higher density that distant parts, the objects in the direction of movement do appear larger in the horizontal direction. Higher density of sampling results in growth of object in the horizontal direction. But since we create a mosaic image by stacking columns from original images with a constant height, the objects are not enlarged in the vertical direction. The same applies to shrinkage in the opposite direction. Therefore, parts of the omnidirectional X-slits image do not have to have the same aspect ratio, see Figure 10, the top row, where the computer screen is more far away in the left image and closer in the right image than in the middle image, which has correct aspect ratio. Note that the screen is smaller in the horizontal direction, but in the vertical direction, its size is the same, thus resulting in incorrect aspect ratio.

The mosaic images have to be rescaled by some function operating on image columns in order to create images with correct aspect ratio. The rescaling function depends on depth in the scene on which objects should have correct aspect ratio. If we assume constant scene depth D, as it is depicted in Figure 4, then, using basic trigonometric identities,

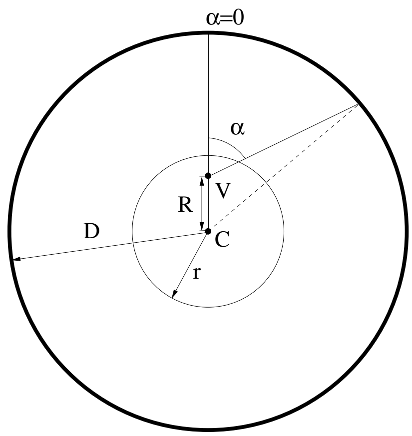

Figure 4: Constant depth assumption is used in aspect ratio correction. The thick circle illustrates the constant scene depth D. Parameter r is the radius of the circle of rotation.

we can write the following rescaling function

$$s = \sqrt{1 - \left(\frac{R}{r}\right)^2 \sin^2(\alpha) d^2} + \frac{R}{r} cos(\alpha) d \quad . \tag{3}$$

R is the position of the vertical slit V, r is the radius of the circle of rotation, and α determines the mosaic column. It is the same α as in (1) and (2). Note that $\alpha = 0$ in the direction of movement of the viewer. The factor $0 \leq d < 1$ depends on the inverse of the depth of the scene and has to be determined empirically. Closer objects will be scaled more than the distant ones, objects at infinity will not be scaled at all.

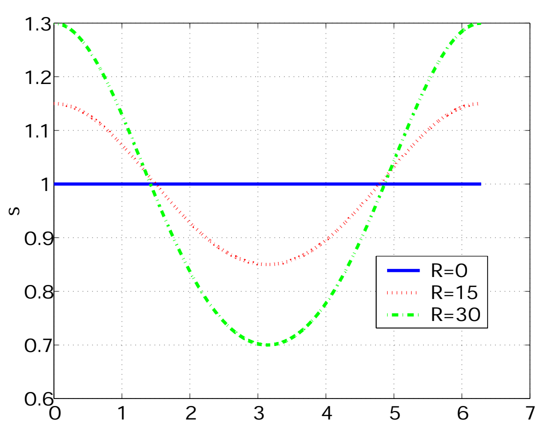

Figure 5: Values of the rescaling function (3) for different positions of the vertical slit R, constant scene depth, and radius of the circle of rotation $r = 30cm$.

The scaling function has three parameters, d, r, and R, and one variable, α. Instead of

Figure 6: The X-slits mosaic camera has a blind spot (no light rays captured) depicted by the solid circle centered at O. When the viewer moves (simulated by moving the vertical slit V), we have to rescale the resulting X-slits mosaic image in order to get correct aspect ratios. Black regions appear in the image at places we have no data. The white lines in the mosaic depict the original mosaic size.

supplying both r and R we can use just their ratio $\frac{R}{r}$ thus the radius of the viewing circle does not have to be known. Variable α corresponds to columns in the omnidirectional X-slits mosaic image and the rescaling function operates on these columns so that the height of the column α_i is multiplied by the factor $s(\alpha_i)$. In [12], only narrow angle X-slits images were generated, therefore it was sufficient to rescale the whole image by the same factor. Omnidirectional X-slits images require non-uniform scaling by (3). If the vertical slit is placed into the center of the circle of rotation C, then $R = 0$ and we get uniform mosaic everywhere. In general, we get maximal expansion in the direction of movement $\alpha = 0$, maximal shrinking in the opposite direction $\alpha = \pi$, and no scaling in the directions $\alpha = \frac{\pi}{2}$ and $\alpha = \frac{3\pi}{2}$, see Figure 5.

Since we capture the input image sequence by a camera with a FOV 180° moving on a circle and looking outwards, we get a set of light rays which intersect a circle defined by the optical center of the camera moving on a circular path. We do not get any light rays looking inside a cylinder defined by the rotating camera. It is a "blind spot" of our mosaic camera. This cylinder is depicted by the solid circle, centered at O, in Figure 6.

When we move the vertical slit to a position V, we should capture light rays intersecting the dashed circle in Figure 6. Instead, we sample them from light rays intersecting the solid circle. Note that we simulate uniform sampling of the dashed circle by non-uniform sampling of the solid circle. This is one of the reasons for different aspect ratios of objects in the mosaic image because we sample objects closer to the vertical slit by more rays than the more distant objects. The change in a distance to an object results also in change of elevation of rays by which we observe the object. Moving closer results in higher elevation, moving away in lower one. In the mosaics, the rows correspond to these elevations therefore, after rescaling the mosaics by (3), we get black regions in some parts of the mosaic, which should be filled by the light rays from the blind spot (elevation higher than the FOV of the camera), which are not available. Fortunately, the resulting black regions are negligible compared to the whole image and can be filled by some augmented texture or we can use slightly smaller field of view in the vertical direction. See Figure 6, where

(a) (b)

Figure 7: (a) Original image captured by the Canon D1s camera. (b) Resampled images used for mosaic composition.

(a) (b)

Figure 8: (a) Detail of the mosaic (b) illustrating the level of details contained in the high resolution image. The detail is marked by a black rectangle in the mosaic images.

the white lines in the mosaic depict the size of the original mosaic (before rescaling). Before mapping the rescaled mosaic on a sphere, we have to crop it to the original size throwing out the regions above and bellow the white lines in Figure 6.

4 Experimental results

Our experimental setup consisted of a Canon D1s digital SLR camera equipped with a Sigma AF 8/4 EX circular fish eye lens providing a 180° field of view mounted on a motorized turntable. We captured 265 images with resolution 4064×2704 pixels on a circular path with a $30cm$ radius. We resampled the circular images into rectangular images where columns corresponded to angles like it is depicted in Figure 1(c). These resampled images have resolution 2500×2500 pixels, see Figure 7. The omnidirectional X-slits images composed from the resampled images have resolution 5000×2500 pixels. The level of detail the resolution provides can be seen in Figure 8(a) where an enlarged region from the mosaic Figure 8(b) is shown.

The first experiment, depicted in Figure 9 illustrates the simulated motion of the

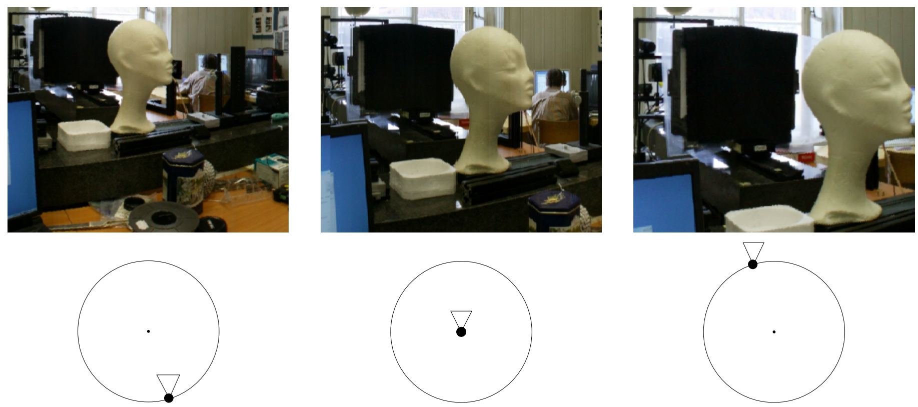

Figure 9: Simulated movement of the viewer in a disc with radius 30*cm*. Note the occlusion changes.

viewer on a line from the most distant position (left), through the center of the circle of rotation (middle), to the closest position (right). The images in the bottom row depict the position of the viewer (big black dots) in the disc defined by the circle of rotation. The triangles show the viewing direction of the images. The movement was towards a point left from the objects in the images. Note occlusion changes as the viewer approaches the object in the scene. The aspect ratio of the objects was corrected using function (3).

In the next experiment, we demonstrate the rescaling of the images using (3). The result is shown in Figure 10. The positions of the camera are the same as in the previous experiment. The first row contains unscaled images, the second row shows the result of the aspect compensation. Note that the corrected aspect ratio of the objects in the scene is almost constant.

The last experiment demonstrates the stereo pair generation. Recall that we create the stereo views by selecting appropriate viewing windows from two X-slit mosaic images. Unlike in case of a singe image, we have to change the mosaic images when the viewer rotates his head. If we use only one omnidirectional X-slit image then just like in case of a traditional stereo, a disparity can be observed in the viewing direction, depicted in Figure 11, the first and the third thirds of the image, and no stereo disparity is observable along the baseline joining the two eyes of the viewer, see Figure 11, the middle part (the second third) of the image. The effect is clearly observable on the lights and on the table in the middle of the image.

5 Conclusion

We have presented an approach to representation of a real scene by a set of omnidirectional X-slits mosaics together with aspect ratio compensation of these images. We have shown that in a monocular case, when the viewer rotates his head we need just one such mosaic image in order to represent the scene. By preloading several mosaics corresponding to the neighboring positions of the viewer into the memory, we can keep the memory requirement reasonably low while still allowing the use of high resolution images. In the stereo case, we have to select different mosaics when the viewer rotates his head, but the

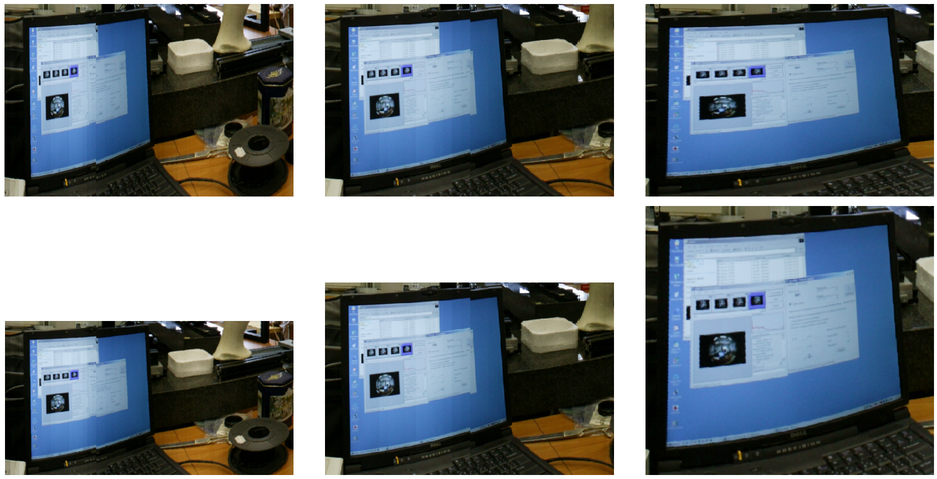

Figure 10: Aspect ratio compensation for the same movement as in Figure 9. The top row is without aspect ratio compensation, note shrunk resp. enlarged screen in the right and the left image, while the middle image has the correct aspect ratio. The bottom row shows compensated images.

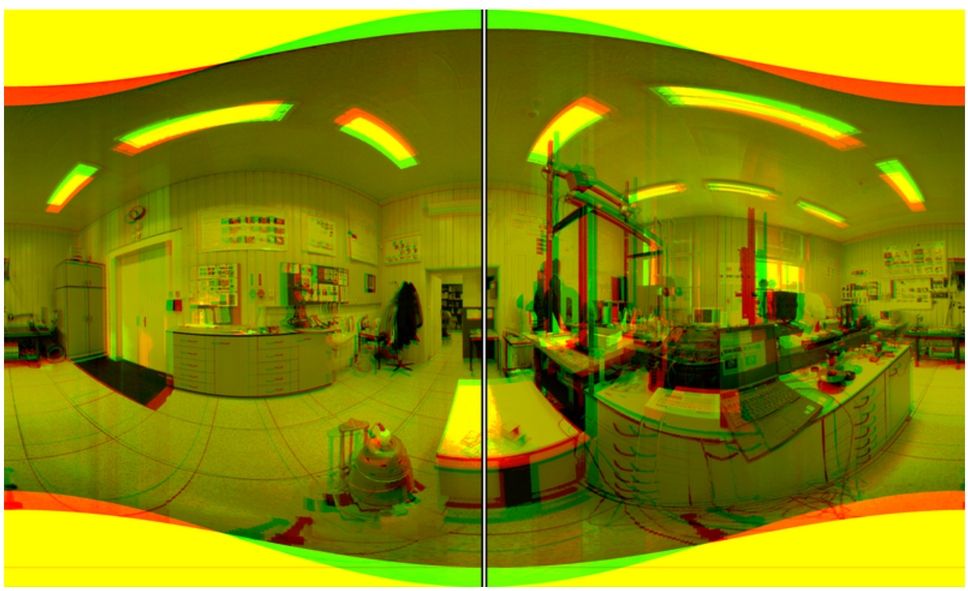

Figure 11: Stereo pair of omnidirectional X-slits images. The images are rectified so that the baseline of the two eyes intersects the image in the middle column (marked by the solid line). There is no disparity observable, just a scaling, see the table or the door in the back. There is disparity in both directions away from that line.

concept of preloading the mosaics into the memory can be again used. We have demonstrated the movement of the viewer in a circular region together with a stereo image pair generation from high resolution images.

References

[1] E. H. Adelson and J. Bergen. *The Plenoptic Function and the Elements of Early Vision*, pages 3–20. MIT Press, 1991.

[2] H. Bakstein and T. Pajdla. Rendering novel views from a set of omnidirectional mosaic images. In *Proceedings of Omnivis 2003: Workshop on Omnidirectional Vision and Camera Networks*, Los Alamitos, CA, June 2003. IEEE Press.

[3] J. Chai, S. B. Kang, and H.-Y. Shum. Rendering with non-uniform approximate concentric mosaics. In *Second Workshop on Structure from Multiple Images of Large Scale Environments, SMILE*, July 2000.

[4] J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum. Plenoptic sampling. In *Procceding of the SIGGRAPH 2000 conference, New Orleans, Louisiana*, volume 1, July 2000.

[5] S. E. Chen. QuickTime VR — an image-based approach to virtual environment navigation. *Computer Graphics*, 29(Annual Conference Series):29–38, 1995.

[6] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *Computer Graphics Proceedings (SIGGRAPH'96), New Orleans*, pages 43–54, August 1996.

[7] M. Levoy and P. Hanrahan. Light field rendering. In *Computer Graphics Proceedings (SIGGRAPH'96), New Orleans*, pages 31–42, August 1996.

[8] H.-Y. Shum and S. B. Kang. A review of image-based rendering techniques. In *IEEE/SPIE Visual Communications and Image Processing (VCIP) 2000*, pages 2–13, June 2000.

[9] H.-Y. Shum, L. Wang, J.-X. Chai, and X. Tong. Rendering by manifold hopping. *International Journal of Computer Vision*, 50(2):185–202, 2002.

[10] L. Wang, H.-Y. Shum, and S. B. Kang. Object representation and rendering using inverse concentric mosaics. In D. Suter and A. Bab-Hadiashar, editors, *Proceedings of the Fifth Asian Conference on Computer Vision : ACCV 2000*, volume 1, pages 150–155. Asian Federation of Computer Vision Societies, January 2002.

[11] D. Weinshall, M.-S. Lee, T. Brodsky, M. Trajkovic, and D. Feldman. New view generation with a bi-centric camera. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors, *ECCV'02: Proc. Eighth European Conference on Computer Vision*, volume 1, pages 614–628, Berlin, Germany, May 2002. Springer Verlag.

[12] A. Zomet, D. Feldman, S. Peleg, and D. Weinshall. Non-perspective imaging and rendering with the crossed-slits projection. Technical Report 2002-41, Leibnitz Center, Hebrew University of Jerusalem, 2002.