CENTER FOR
MACHINE PERCEPTION

CZECH TECHNICAL
UNIVERSITY

RESEARCH REPORT

ISSN 1213-2365

ANNEX 1

Panoramic Imaging with SVAVISCA Camera - Simulations and Reality

Tomáš Pajdla and Hadas Roth

Department of Electrical Engineering
Technion
Israel
www.technion.ac.il
shadasr@techst02.technion.ac.il

CTU–CMP–2000–16

October 20, 2000

Available at
ftp://cmp.felk.cvut.cz/pub/cmp/articles/pajdla/Pajdla-TR-2000-16.pdf

This research was supported by the EU Fifth Framework Programme project OMNIVIEWS No. 1999-29017, by the Grant Agency of the Czech Republic under the grant GACR 102/01/0971, and by the Research Programme J04/98:212300013 Decision and control for industry.

Research Reports of CMP, Czech Technical University in Prague, No. 16, 2000

Published by

Center for Machine Perception, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University
Technická 2, 166 27 Prague 6, Czech Republic
fax +420 2 2435 7385, phone +420 2 2435 7637, www: http://cmp.felk.cvut.cz

Abstract

An OMNIVIEWS catadioptric camera combines a curved mirror with a variant resolution imaging sensor. The goal is to obtain panoramic images with large view angle, suitable image formation geometry, and controlled resolution variation across the field of view. We present simulations of an OMNIVIEWS camera and compare them with real images taken by a combination of a hyperbolic mirror with a log-polar SVAVISCA imager. The experiments show that the simulations correspond to reality and can be used for further OMNIVIEWS camera design. Two auxiliary observations are made. Firstly, it is better to use pointed mirrors because they allow to exploit imager completely. Secondly, the hyperbolic mirror used altogether with the SVAVISCA imager provided a non-uniform resolution across the field of view.

1 Introduction

Conventional catadioptric panoramic cameras [7] combine curved mirrors with conventional CCD cameras to achieve large field of view. The price to pay for a large view angle is lower and (with a conventional CCD sensor) varying resolution across the field of view.

With the state of the art CCD imagers, the resolution of panoramic catadioptric cameras cannot be as high as the resolution achieved with conventional cameras having a narrow field of view. Assume a camera with a conventional $1/2''$ CCD element as used e.g. for the CCIR television format, i.e. having 576×768 pixels and a conventional 12 mm focal length lens. Such a camera would have 28 degrees wide field of view to cover the longer side of the sensor leaving us with 24.4 pixels per degree. In order to achieve the same resolution for a panoramic camera with 360×90 degrees angle of view, one needs 9800×2500 pixel sensor. It is difficult to get such a sensor for a reasonable price. However, the total number of available pixels per sensor still grows and therefore it will be available sometimes in future.

The change of the resolution across the field of view has two main reasons. Firstly, mirrors used for imaging rarely have constant angular gain, i.e. mirrors often map one pixel on the sensor to angular segments of varying size, depending on the distance of the pixel from the image center, see e.g. the mirror on Figure 1. Nevertheless, even if the mirror is designed to have a constant angular gain, like the mirror in the works [2, 3, 4], the still resolution changes. It is so because the number of pixels on a pixel ring corresponding to a set of rays reflected with a constant elevation gets smaller as the radius of the ring decreases, see Figure 2. While the circle with the

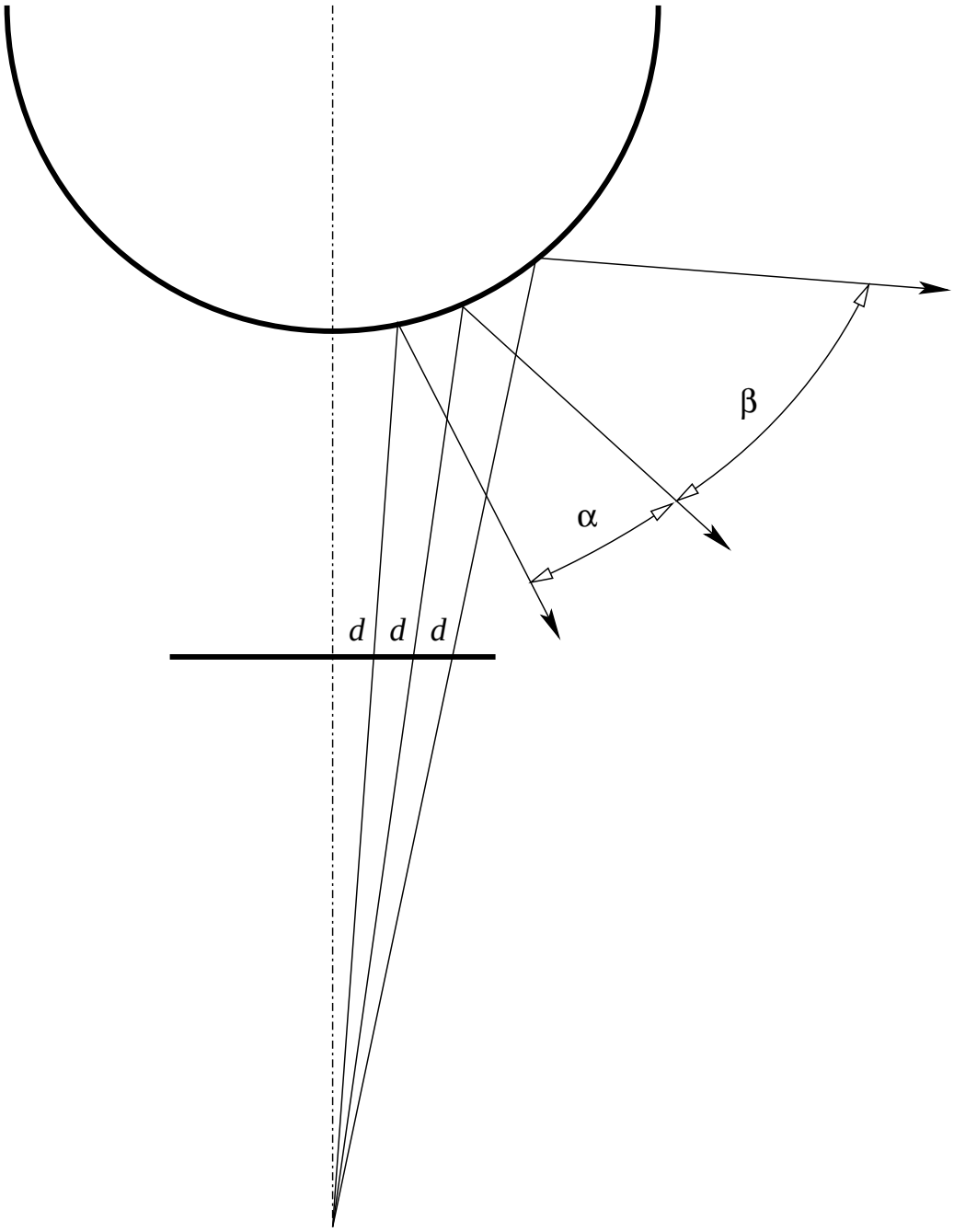


Figure 1: Often, mirrors are not designed to have a unit angular gain and thus the pixels cover varying spatial angles. The variation depends on the mirror shape as well as on the distribution of the pixels in the image plane.

(a)

(b)

Figure 2: (a) An original image of a mirror. (b) The warped image. Resolution in the upper part of the image (b) is lower because the upper part is transformed from the center of the image (a) where a small number of pixels covers a large view angle.

highest number of pixels, which is the one close to the border of the image, may have $2\pi 576 = 3619$ pixels, a small circle in the center is mapped to a single pixel on a CCD chip. The course of reduction is linear when neglecting the discrepancies caused by the anisotropy of two-dimensional sampling on a rectangular grid.

The key idea behind the OMNIVIEWS project [6] is to adapt the image acquisition sensor as well as the shape of the mirror such that the distribution of the pixels of the sensor combined with the angular gain of the mirror would lead to (more) constant angular resolution of the whole sensor. The SVAVISCA CMOS camera [5], originally inspired by anatomy of a human eye, was designed to have a logarithmically decreasing resolution from the center (fovea) to the periphery. In the fovea, pixels are arranged in a six-fold symmetry allowing for the most economic usage of space. In the periphery, a constant number of pixels per ring is implemented. Thus the arrangement in the periphery naturally leads to a constant angular resolution in azimuthal direction. An idea to use a non-uniform distribution of pixels also appeared independently in [1]. In this report we simulate image formation of a sensor built by combining a mirror with the CMOS SVAVISCA camera. We will

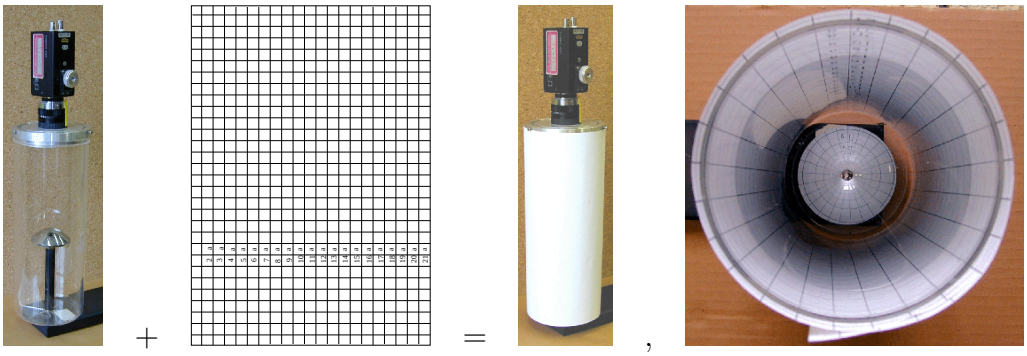


Figure 3: Acquisition of the artificial scene. (a) Catadioptric camera, (b) regular grid, (c) wrapped around a transparent cylinder, and (d) mirror inside the cylinder.

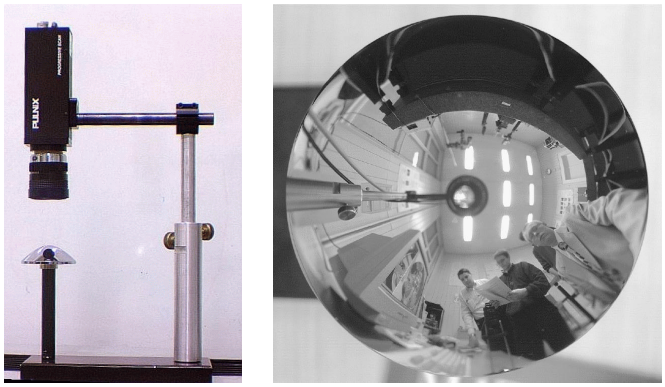
further call such a sensor the *OMNIVIEWS sensor* and the image taken by an OMNIVIEWS sensor the *OMNIVIEWS image*.

Section 2 briefly describes the sensor arrangements used in this work. A simulation technique for creating OMNIVIEWS images is described in Section 3. Simulations produced from conventional images of artificial (Section 4) and natural (Section 5) scenes are presented. Images obtained by a real OMNIVIEWS camera are presented in Section 6. Section 7 compares the simulations with the real data and the conclusions are drawn in Section 7.

2 OMNIVIEWS sensor

In order to study the mapping of scenes into OMNIVIEWS images, an experimental setup depicted on Figure 3 was proposed. A catadioptric camera consists of a mirror and a conventional 576×768 CCD camera held by a transparent cylinder of the radius equal 5.3 cm, see Figure 3(a). A grid with constant spacing equal 1 cm, Figure 3(b), was wrapped around the cylinder, Figure 3(c). Images obtained by a catadioptric camera then show the mapping of the height coordinate of the cylinder to the image induced by the combination of a mirror shape, optics, and a camera sensor used, see e.g. Figure 7(b,d) for examples of images.

Figure 4 shows a conventional panoramic catadioptric camera consisting of a hyperbolic mirror [8] and a CCD camera used to capture real panoramic images in this work.



(a) (b)

Figure 4: Acquisition of natural scenes. (a) a hyperbolic mirror and a conventional CCD camera, (b) a panoramic image.

3 Simulating OMNIVIEWS sensor from images taken by a CCD camera

OMNIVIEWS images were simulated by resampling conventional images since there was no camera available at the time of doing the experiments described here. The geometric arrangement of pixels on the SVAVISCA sensor was modeled according to the SVAVISCA specifications [5] and mapped into the conventional image such that the image of the mirror was completely covered by pixels of the SVAVISCA sensor, as if the mirror was optimally adjusted with a SVAVISCA camera, see Figure 5(a). All pixels of SVAVISCA sensor are arranged in rings that are symmetrical w.r.t. the center of the sensor. The width as well as the number of pixels per ring vary across the sensor as a function of the radius of the rings.

There are two regions with different geometry of pixel arrangements on the rings. In the fovea, which is in the center of the sensor shown on Figure 5(a,b), pixels are arranged in a fix-fold symmetry as shown in Figure 5(c). There are 41 rings of fovea pixels. In the periphery, which is around the fovea as shown in Figure 5(a), there is a constant number of 251 pixels in 110 rings, see Figure 5(d).

An OMNIVIEWS image is obtained by averaging intensities from a conventional image inside regions covered by SVAVISCA pixels. As the size of SVAVISCA pixels increases with increasing radius more conventional pixels cover one SVAVISCA pixel. Similarly, as the radius decreases, less conventional pixels cover one SVAVISCA pixel until more SVAVISCA pixels are

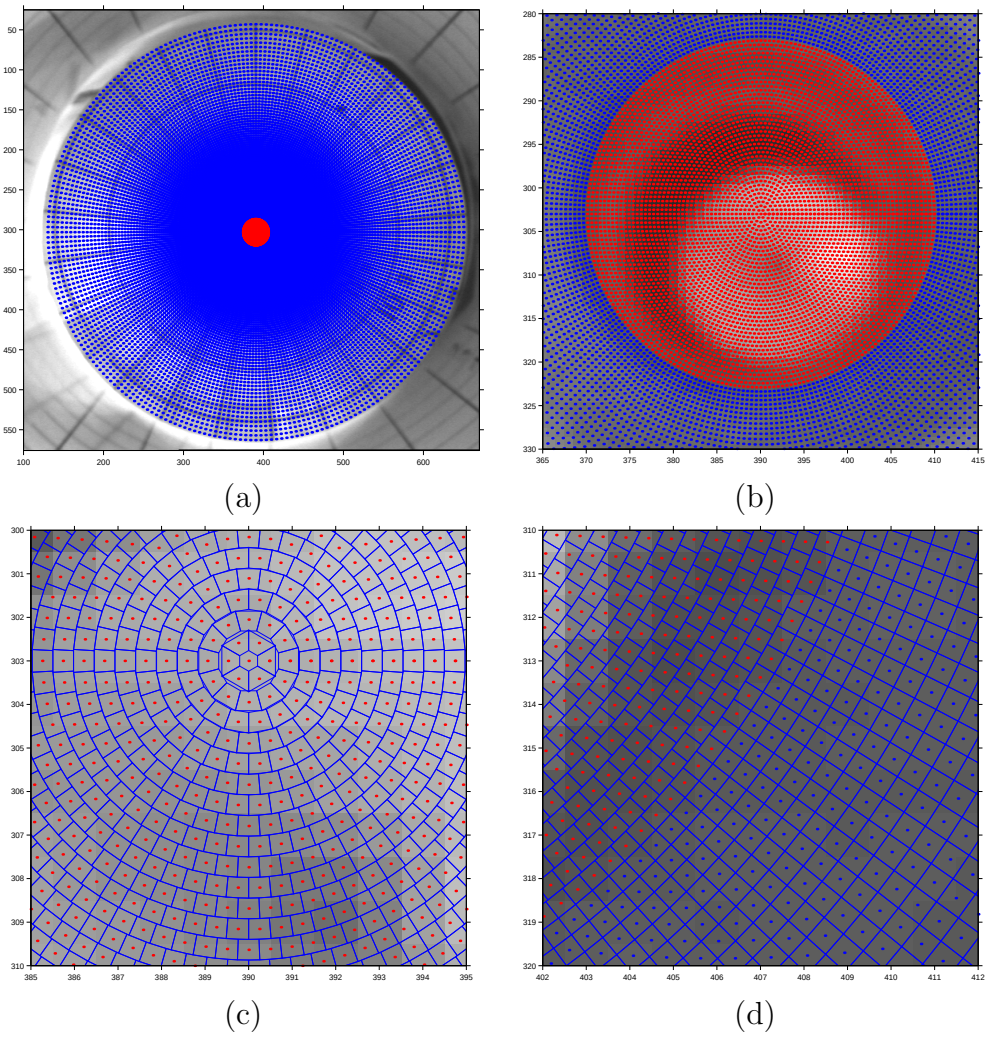


Figure 5: Arrangement of pixels in SVAVISCA retina. (a) SVAVISCA consists of a fovea and a periphery, (b) the fovea, (c) a detail of the fovea center, (d) a detail of the border between the fovea and the periphery.

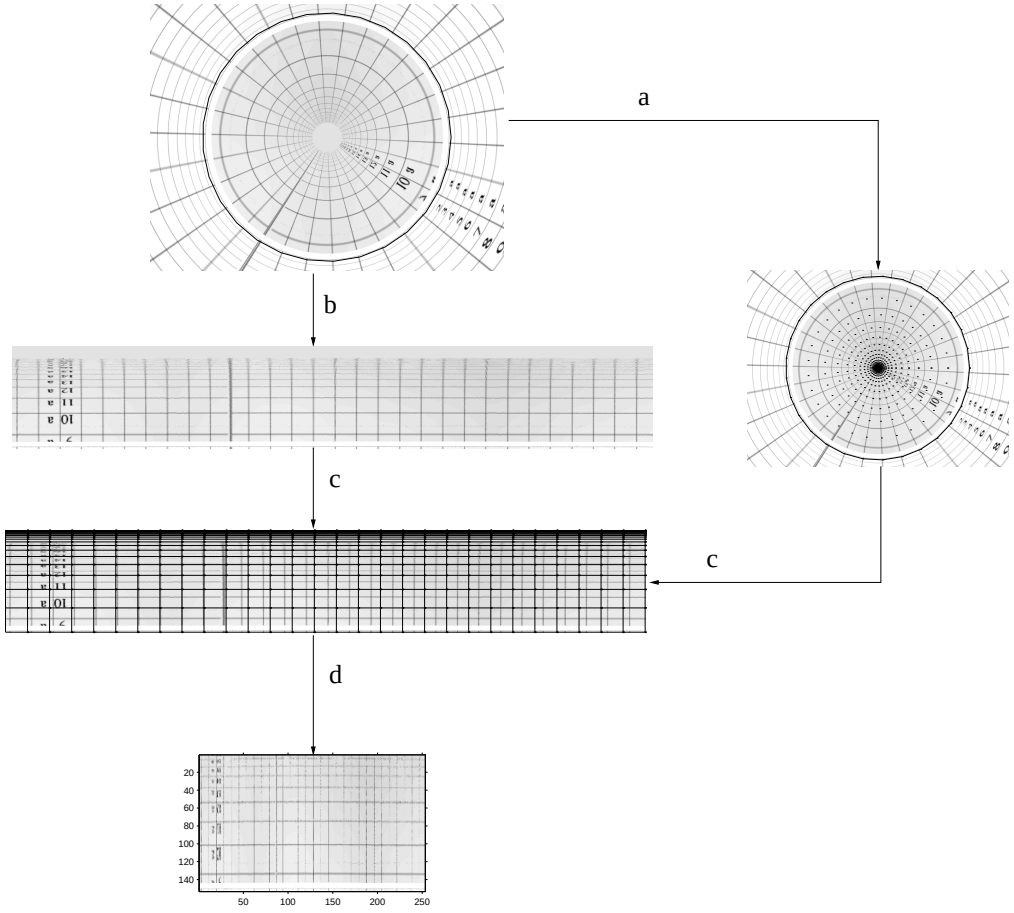


Figure 6: Conventional image resampling for simulating OMNIVIEWS image formation.

mapped to one conventional pixel in the fovea.

Figure 6 shows how the simulated images are actually computed. Original image is first, following the arrow b, transformed into polar coordinates. The image is interpolated using the nearest neighbors and the resolution of the polar image is chosen to be approximately equal to the resolution of the original image, i.e. one pixel in polar image roughly corresponds to one pixel in the original image.

Following the arrow a from the original image we arrive at the image showing the centers of the pixels of OMNIVIEWS sensor mapped into the original image. The SVAVISCA pixels are transformed to polar coordinates, arrow c, to for regions fro averaging. The regions are rectangular with sides parallel to the polar image borders. It is easier to compute average values in such regions compared to computing average values in non-rectangular regions corresponding to SVAVISCA pixels in the original image.

Continuing along the arrow d we arrive at the final OMNIVIEWS image where each pixel is obtained as the average intensity value inside the corresponding region in the polar image. Again, the nearest neighbor interpolation is used to interpolate intensity values.

The simulated images obtained by the previous procedure are certainly not the best images one could obtain by a simulation. The main loss of quality results from the double resampling of images with the nearest neighbor interpolation. The image transformation could certainly be improved by using more advanced interpolation as well as by performing it only once. However, even though the quality of images might be worse, the geometry of projection is modeled with high fidelity. The loss on image quality is well balanced by the simplicity of the implementation of the simulation algorithm.

4 Simulated images of an artificial scene

Figures 7(a,c) show images of the artificial scene with the grid from Figure 3(b) taken with a hyperbolic resp. spherical mirror. Figures 7(b,g) show respective simulated SVAVISCA images. We can clearly see how the first 65 rows in the image (b) and 110 rows in the image (d) are covered by the self-reflection of the camera. Both images exhibit a nonlinear distortion of the grid. Originally uniform grid on the cylinder in space is mapped to the grid in the OMNIVIEWS image with increasing spaces between the grid lines along the row coordinate of the image. For the image (b) (taken with a hyperbolic mirror), the height of grid squares at the row 140 is roughly twice the height of grid squares at the row 80. A similar behavior can be observed in the image (d), which was taken with a spherical mirror.

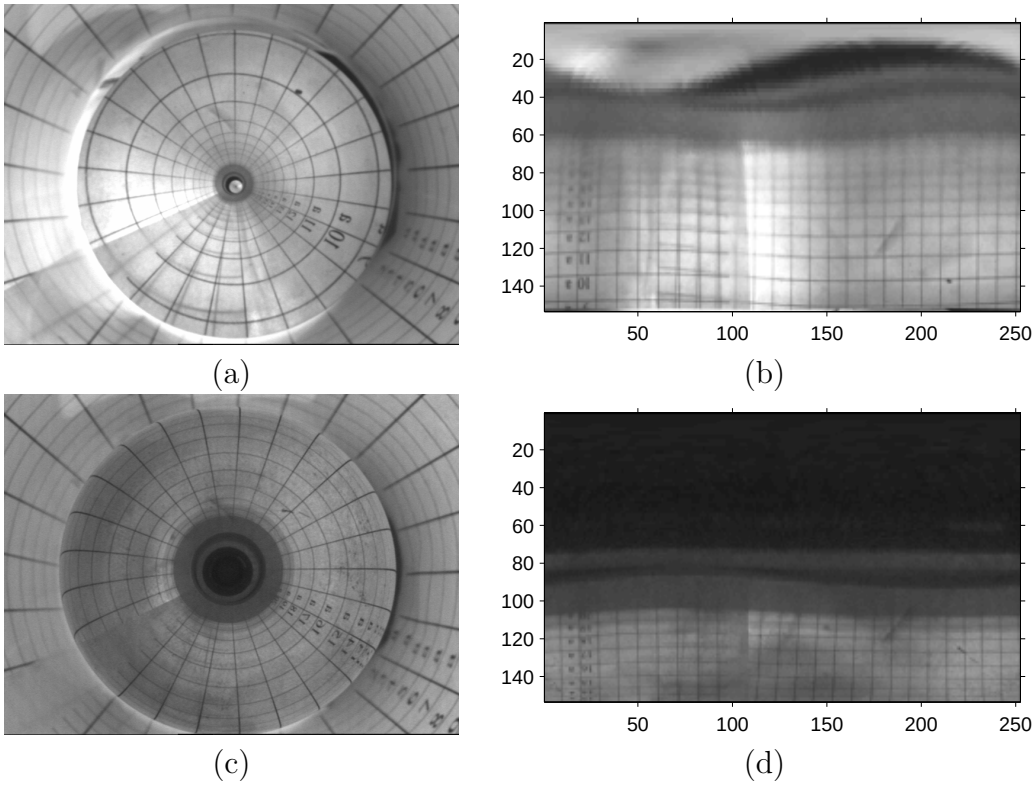


Figure 7: Simulated OMNIVIEWS images of the artificial scene. See text.

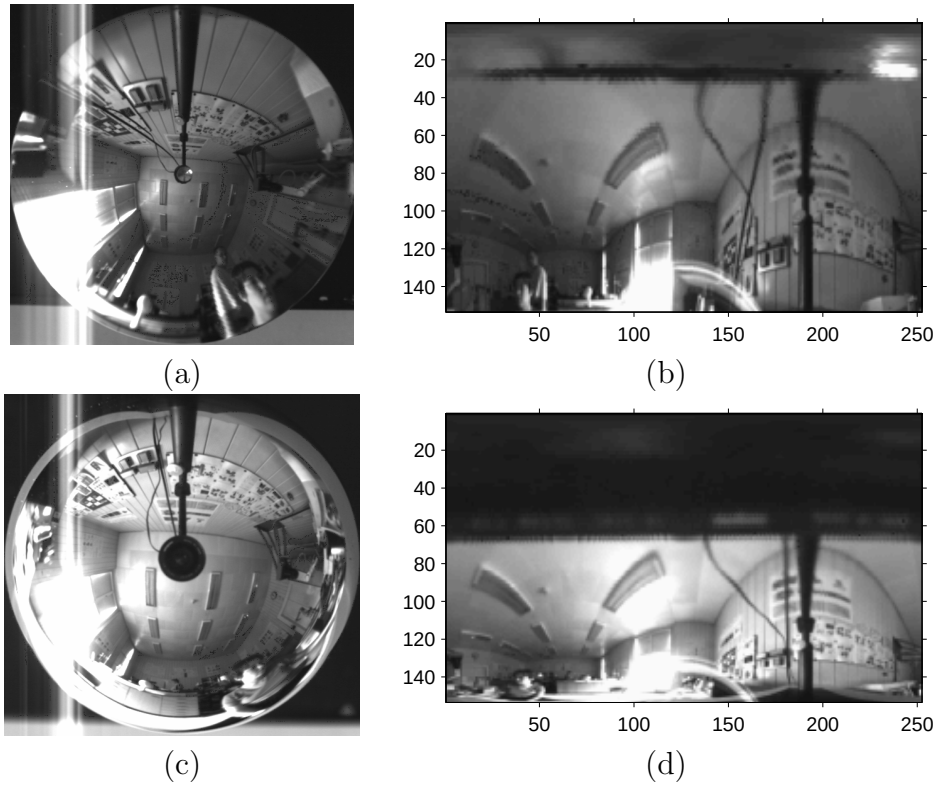


Figure 8: Simulated SVAVISCA images of a natural scene. See text.

5 Simulated images of a natural scene

Figures 8(a,c) show two images of a natural scene taken by a camera in arrangement as shown in Figure 4 with 576×768 CCD sensor and a hyperbolic resp. spherical mirror. Figure 8(c) shows the simulated OMNIVIEWS image for the hyperbolic mirror. The self-reflection of the CCD camera occupies only a small upper part (the first 30 rows) of the OMNIVIEWS image. On the other hand, Figure 8(c) shows that the area covered by the self-reflected camera on the image taken with the spherical mirror is quite large (the first 70 rows). The spherical mirror is more flat around the point where camera is reflected and thus its self-reflection occupies a larger area.

Figure 9(a,c) show two images of a natural scenes capturing a group of people at distance one to two meters from the camera arranged as shown in Figure 4 with 1000×1000 CCD sensor. Figures 9(b,d) show respective simulated SVAVISCA images. We can see that the first 60 rows, i.e. the whole retina that is covered by the first 41 rows, are occupied by the self-reflection of the camera. Second, even though the SVAVISCA images are of very low

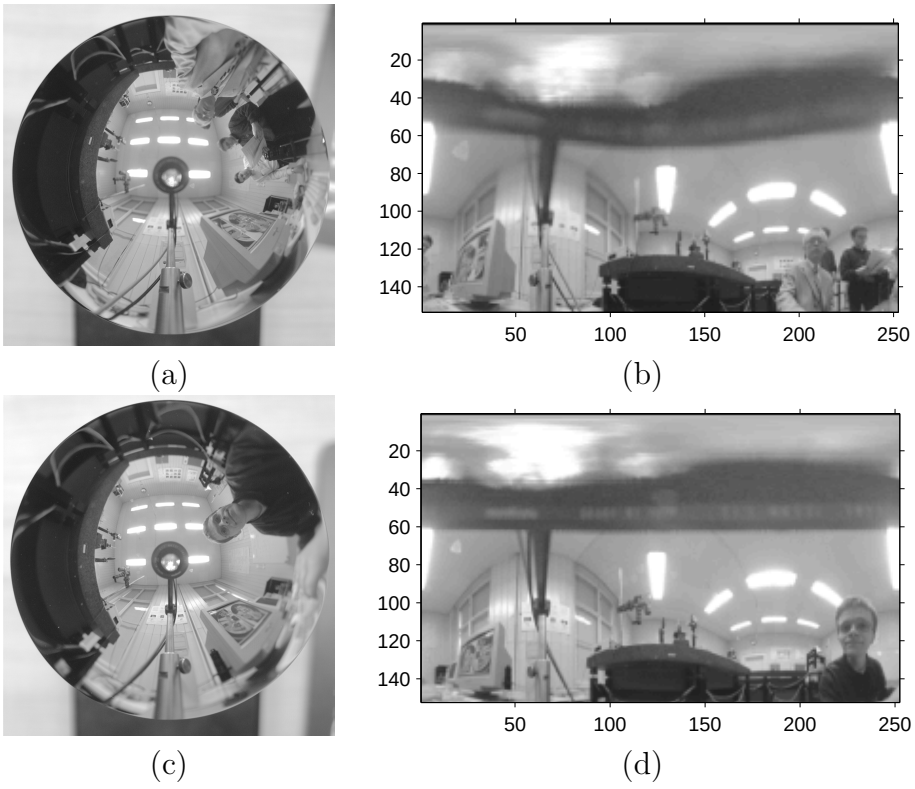
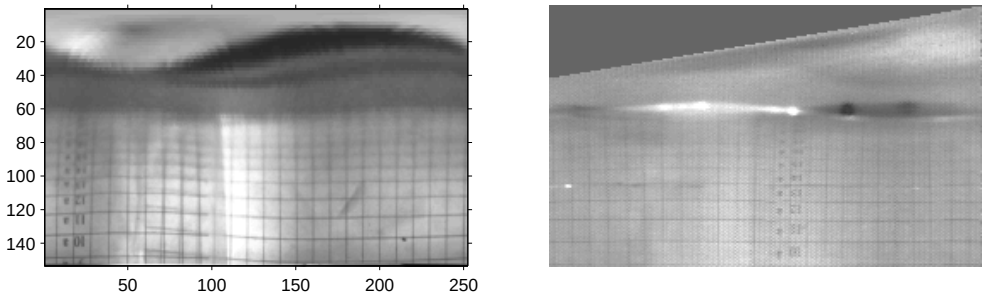


Figure 9: Simulated SVAVISCA images of faces. See text.



(a) Simulated image (b) Real image

Figure 10: Real and simulated OMNIVIEWS images. See text.

resolution, it is still possible to recognize persons from the background as well as some objects at distance four to five meters on the walls of the room.

6 Real OMNIVIEWS images

A set of real images with a color SVAVISCA 33000 pixel chip were taken to compare the simulations with reality. The artificial scene described above as well as a natural scene were captured.

6.1 Artificial scene

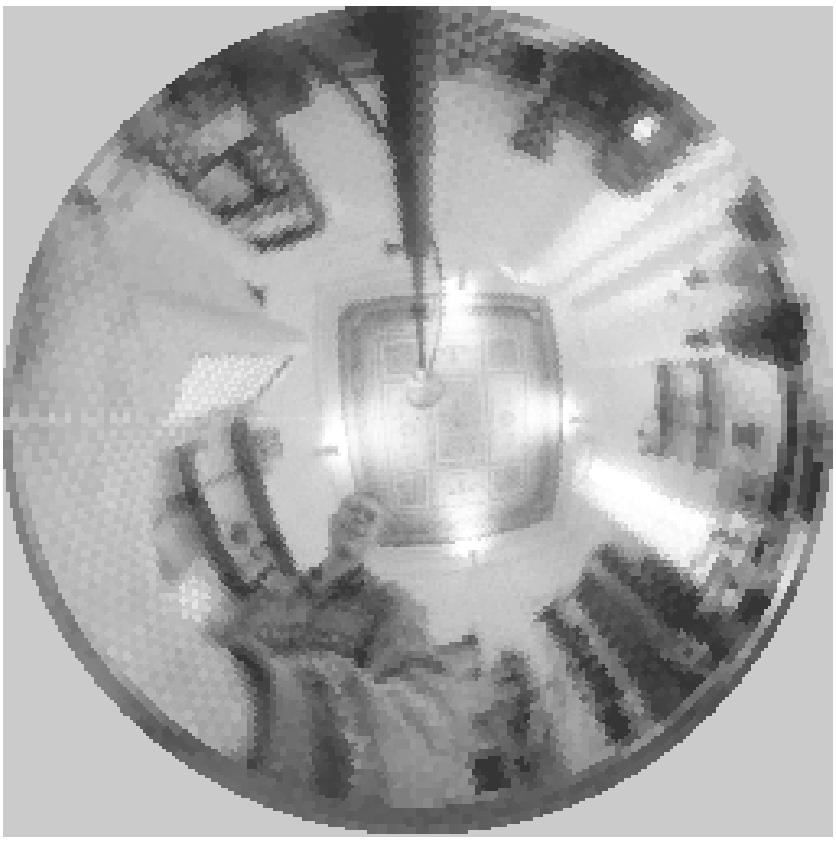
Figure 10 shows a set of images of a grid wrapped around the cylinder as described above. The images were taken by a SVAVISCA camera held by a transparent cylinder above a hyperbolic mirror ATC-03 and Computar 25 mm, 1:1.13 lens, with zoom set to 0.7 m, and iris set to 2.8. The distance between the mirror tip and the lens C-Mount was equal to 19 cm. All images were converted to gray values before saving.

The foveal part (the first 41 rows) of the real SVAVISCA image in Figure 10(b) is shown as it was read from the sensor. The number of pixels per row increases linearly with the increasing row coordinate. In order to obtain perceptually meaningful image, the pixels on each row have to be remapped and interpolated to fill the whole row.

6.2 Natural scene

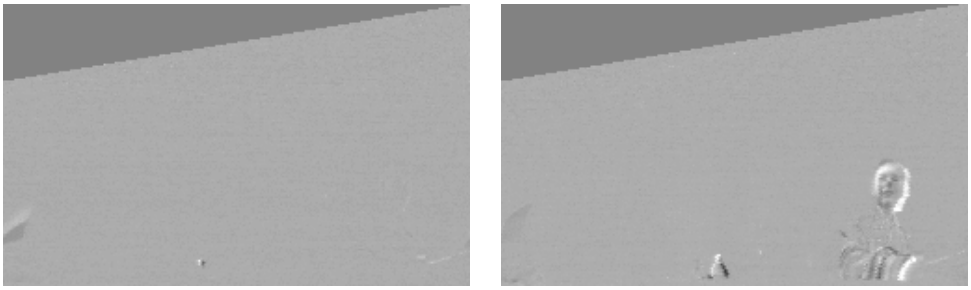
Figure 11 shows an image of a room of size roughly 8×8 meters taken by the OMNIVIEWS sensor using Computar 25 mm, 1:1.3 zoom lens with iris set to 5.6, zoom set to 1 m, and the hyperbolic mirror ATC-03. The distance

(a) Output for the SVAVISCA sensor.



(b) Panoramic image after re-mapping back to the SVAVISCA retina plane.

Figure 11: Images of a natural scene taken by the OMNIVIEWS camera. See text.



(a) (b)

Figure 12: The SVAVISCA camera is photometrically calibrated by subtracting a fixed pattern image. If image of the scene is taken as the fixed pattern image (a) the camera outputs a constant if the scene remains the same and (b) the difference image if the scene changes.

between the tip of the mirror and the C-Mount of the lens was equal to 21.5 cm. All images were converted to gray values before saving.

7 Conclusions

The experiments show that the simulation of SVAVISCA resampling correctly models a real OMNIVIEWS sensor. Not only that images are visually very similar, see Figure 11, but also the simulated geometrical mapping corresponds to the real mapping, see Figure 10.

Images shown in Figure 8 further suggest that pointed (e.g. hyperbolic) mirrors are suitable for capturing large view angle as the camera self-reflection occupies a small area in the resulting OMNIVIEWS image. Ideally, the mirror should be designed such that the camera self-reflection occupied only the fovea or it was not seen at all.

The hyperbolic mirror ATC-03 imaged by the SVAVISCA 33000 pixel log-polar sensor does not map a regularly spaced grid on a cylinder into a regularly spaced grid in the image, see Figure 10. The size of imaged objects depends on the distance of the objects from the camera as well as on the position where they project into the image. Thus an object would be digitized with varying resolution even if moving around the OMNIVIEWS sensor at a constant distance. It is a matter of future research to design a mirror shape and a pixel distribution on the imager in order to arrive at a constant resolution OMNIVIEWS sensor.

References

[1] A. Bruckstein and T. Richardson. Omniview cameras with curved surface mirrors. In *IEEE Workshop on Omnidirectional Vision*, June 2000.

[2] T. Conroy and J. Moore. Resolution invariant surfaces for panoramic vision systems. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 1, pages 392–7, 1999.

[3] A. Hicks and R. Bajcsy. Reflective surfaces as computational sensors. In *CVPR-Workshop on Perception for Mobile Agents*, June 1999.

[4] A. Hicks and R. Bajcsy. Catadioptric sensor that approximate wide-angle perspective projections. In *IEEE Workshop on Omnidirectional Vision*, June 2000.

[5] A. Mannucci, P. Questa, and D. Scheffer. D1-Document on Specification, Esprit Project N. 31951 - SVAVISCA, Version 1.0. 1999.

[6] T. Pajdla, G. Sandini, and J. Santos-Victor. OMNIVIEWS - omnidirectional visual system. http://cmp.felk.cvut.cz/cmp/omniviews/, Sep 2000. RTD - Future and Emerging Technologies Project No: IST-1999-29017.

[7] T. Pajdla, T. Svoboda, and V. Hlaváč. Epipolar geometry of central panoramic cameras. In Ryad Benosman and Sing Bing Kang, editors, *Panoramic Vision : Sensors, Theory, and Applications*. Springer Verlag, Berlin, Germany, 1 edition, 2000.

[8] T. Svoboda. *Central Panoramic Cameras Design, Geometry, Egomotion*. PhD Thesis, Center for Machine Perception, Czech Technical University, Prague, Czech Republic, April 2000.