

8.5: OMNIVIEWS: Direct Omnidirectional Imaging Based on a Retina-like Sensor

Giulio Sandini

LIRA-Lab
DIST - Università di Genova
Viale F. Causa, 13
16145 Genova – Italy
sandini@dist.unige.it

Tomas Pajdla

Center for Machine Perception
Dept. Of Cybernetics, Faculty of EE
Czech Technical University in Prague
pajdla@cmp.felk.cvut.cz

José Santos-Victor

Instituto Superior Tecnico
ISR
Lisboa – Portugal
jasv@isr.ist.utl.pt

Fabio Berton

Aitek S.r.l.
Via della Crocetta, 15
16122 Genova – Italy
fberton@aitek.it

Abstract

Traditionally the acquisition of real time panoramic images has been performed by the usage of lenses or mirrors coupled with standard image sensors, which give distorted images. In this paper we will present a system composed of a mirror and a log polar sensor, which is able to provide directly understandable images.

Keywords

Log-polar, panoramic image, surveillance, navigation.

INTRODUCTION

The main objective of the OMNIVIEWS project was to integrate optical, hardware, and software technology for the realization of a smart visual sensor, and to demonstrate its utility in key application areas. In particular our intention was to design and realize a low-cost, miniaturized digital camera acquiring panoramic (360°) images and performing a useful low-level processing on the incoming stream of images.

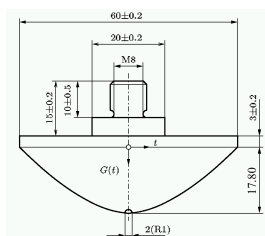


Figure 1. Left: Detail of the mirror profile. Right: the system camera-mirror.

The classical acquisition of panoramic images is based on the use of mechanical or optical devices. The solution proposed in OMNIVIEWS was to integrate a retina-like CMOS visual sensor incorporating novel and unique technologies, with a mirror with a specially designed, matching curvature. This matching, if feasible, provides panoramic images without the need of computationally intensive processing and/or hardware remapper as required by conventional omnidirectional cameras. Therefore reducing overall cost, size, energy consumption and computational power with respect to the currently used devices. In the paper we will present the results of our assessment showing, for 3 different application areas, that the panoramic images obtained by our technology are not only equivalent to the ones obtained with conventional ones but also these images can be obtained at no computational cost. For example with our current prototype a panoramic image composed of about 27,000 pixels is obtained by simply reading out the pixels (i.e. 27,000 operations) while with a conventional solution the same image would required more than 1.7 million operations (about 50 times more). Besides that, unlike with a warped traditional image, we get the interesting side effect of a uniform resolution along the entire panoramic image.

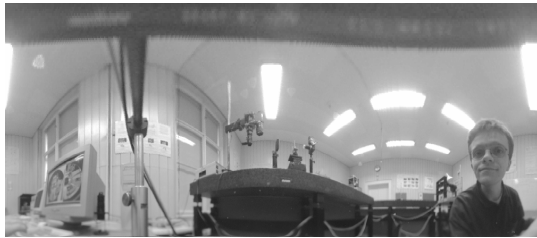
The results of demonstrations in areas such as surveillance, robot navigation and image transmission support the fact that, in spite of this enormous saving, the use of OMNIVIEWS images is the same as for a conventional approach. No extra cost is required for the components.

DESCRIPTION

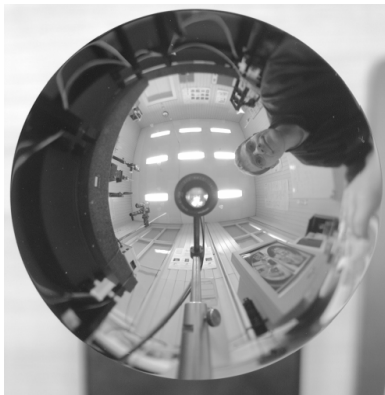
The classical acquisition of panoramic images is based on the use of mechanical or optical devices. The mechanical solutions are based on motorized linear or array-based cameras, usually with a 360° rotation, scanning the visual world. The main advantage of the mechanical solution is the possibility of acquiring very high-resolution images and

its major drawback the time required to mechanically scan the scene to obtain a single image. Optical solutions provide lower resolution images but they are the most appropriate for real-time applications and this is, therefore, the solution adopted. Two optical alternatives have been proposed in the scientific literature, namely the use of mirrors and the use of special purpose lenses (such as fish eye lenses). While the lenses are optimal for a forward looking cameras, e.g. when providing images in order to steer a vehicle, mirrors are optimal for sideways and panoramic views, e.g. when inspecting pipes, veins, or acquiring panoramic views for surveillance.

When this solution is used with traditional cameras the resulting image is not immediately understandable because of the geometric distortion introduced by the mirror to obtain a panoramic (view Figure 2b). To transform Figure 2b into Figure 2a requires, in a traditional camera, to map the original (i.e. distorted) panoramic image onto a cylinder and then the cylinder into a plane, thus, providing a complete panoramic image.



a



b

Figure 2. a): Image acquired by an “OMNIVIEWS” camera. b): Image acquired by a conventional omnidirectional camera Note that the image from OMNIVIEWS camera is immediately understandable while the image from a conventional camera requires more that 1.5 million operations to be transformed into something similar with no added advantage.

In the following Figure 3 the principle that guided the design of the mirror during the assessment phase of the project is presented. Basically the guiding principle is to design the profile of the mirror so that if the camera is inserted inside a cylinder, the direct camera output provides an un-

distorted, constant resolution image of the internal surface of the cylinder.

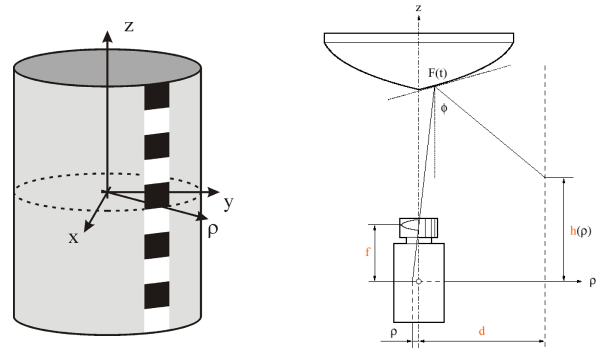


Figure 3. The mirror is designed so that vertical resolution of a cylindrical surface is mapped into constant radial resolution in the image plane

The advantage of such an approach lays in providing the observer with a complete view of its surrounding in one image which can be refreshed at video rate.

In order to achieve this we needed to design not only the mirror but also a matching visual sensor implementing a space-variant polar mapping of the image reflected by the surface of the mirror.

The sensor geometry we used is, therefore, space-variant (i.e. with variable resolution), and with a log-polar structure (retina like).

The matching between the optical part and a log-polar visual sensor can be better understood considering two aspects. First, the size of the photosites increases linearly with eccentricity and secondly the photosites are arranged over concentric rings with each ring composed by an equal number of pixels (Figure 4).

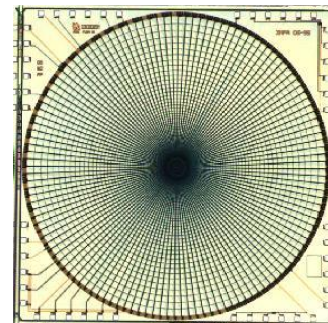


Figure 4. Structure of the sensor

The read-out of the sensor array is performed so that a point of the polar structure at coordinates (ρ, θ) is mapped into Cartesian plane at coordinates $(\log(\rho), \theta)$. It should be intuitively clear that, if the structure of the sensor is matched to the curvature of the mirror, simply reading out the pixels of the retina-like sensor would provide images directly in a panoramic form. The increase of resolution in the sensor array can be designed so that the resulting panoramic image will have a constant vertical resolution eliminating the lack

of homogeneity observed in a traditional mirror-based panoramic image.

TECHNICAL DETAILS

Geometry of the sensor's array

The space variant sensor has a log polar displacement of the pixels. This means that the pixels are arranged in concentric rings (*polar*), with a logarithmic increasing of the size with eccentricity (*log*). One of the peculiarities of log-polar arrays is that the size of the pixels increases linearly with eccentricity so that the number of pixels per ring can be kept constant. This produces rectangular images at least for the space variant part of the array. Besides the proportions of the pixels are constant along the whole image.

The centermost area of the array is covered with a matrix of constant resolution pixels whose size is constrained by technology. The adopted solution seems the most advantageous because it allows: i) polar arrangement; ii) maximum resolution; iii) pixels with aspect ratio very close to 1; iv) smooth blending of constant resolution fovea and space-variant periphery. It should be noted, however, that the extreme solution of leaving the centermost part empty because it may cover the part of the mirror where a self-reflection of the camera is seen, does not really provide a cost-advantage. Cost is mostly related to the surface of the silicon area used and it would be a waste to leave such area empty. However for some applications it may be more efficient to limit the readout to the peripheral part of the sensor.

Image Resolution

The sensor is composed of approximately 33,000 pixels arranged on 152 rings. The 110 most external rings have 252 pixels each ring, while the inner 42 have a constant width and a variable number of pixels. The first release of this sensor was completed at the end of 1999, using a 0.35 μm CMOS technology. This allows the size of the smallest pixel to be approximately of $7 \times 7 \mu\text{m}$. In Table 1 the physical parameters of the sensor are listed. It is interesting to note the Q parameter, which is a measure of the size of the constant resolution array that would be necessary to "simulate" a retina-like sensor by mapping a constant resolution image. This parameter makes evident the advantage of a silicon implementation of a space variant sensor with respect to its software simulation. To obtain 33,000 pixels with the "retina-like" geometry it would be necessary to acquire a constant resolution array of $1,303 \times 1,303$ pixels (i.e. 1.7 Million Pixels). This may be technologically possible but with a much higher cost, lower acquisition rate, higher power consumption and huge processing power to downsample the image.

Table 1: Physical parameters of the sensor. Min size and Max size refers to the size of the pixels. Fovea Radius is the radius of the central constant resolution part of the sensor. Q is the ratio between the diameter of the chip and the size of the smallest pixel. Max/Min is a

figure of "space-variancy" of the array expressed as the ratio between the maximum and the minimum pixel's size. Note that both the number of pixels and the number of rings refer to the pixels actually used.

Pixels/Ring	252
Rings	110
Min Size	6.8 μm
Max Size	100 μm
Fovea Radius	270 μm
Chip \emptyset	8,800 μm
Q	1,303
Total Pixels	27,720
Max/Min	15
Increase Factor	1.02337

Another key parameter especially for clinical applications is the quality of the chromatic information obtained and the trade-off between spatial resolution and color information. In the sensor the color information is obtained by a micro-deposition of colored filters on the pixels. As shown in Figure 5, the density of the red, green and blue filters is uniform.

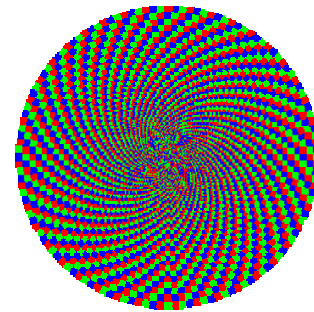


Figure 5. Color distribution on the sensor

APPLICATIONS

Typically, the main fields of application of panoramic images are in medical imaging, remote surveillance, video conferencing and robot navigation.

In the medical field the possible applications are mainly in endoscopy. In this case we can use even the inner 42 rings to get a frontal view, which is useful to the operator for an easier navigation, while the remaining 110 rings will acquire the panoramic view. To enhance the usability of the device, we are planning to develop a specific video mosaic algorithm.

Regarding the surveillance/video conferencing application, the advantage of the panoramic view can be enhanced by coupling a fixed 360° camera with a standard pan-tilt camera, which will be controlled by the content of the panoramic view, so it can focus on the significant details of the scene. Furthermore, we could even implement an algorithm, which is able to perform a virtual pan-tilt, in order to get perspective-correct sub images from the original 360° one.

Omnidirectional images have been successfully used in both robot topological navigation and visual path following. Topological

Navigation is used for traveling long distances and does not require knowledge of the exact position of the robot but rather, a qualitative position on a topological map. Visual Path Following is required for local, very precise navigation, for e.g. door traversal, docking. In the first case the acquired images are compared to a previously stored set of images (landmarks), in order to estimate the position and the orientation of the robot. In the second case the moving images allow the robot to estimate the relative position respect to the desired target. The main advantage of panoramic image is that the robot can have a description of the complete environment in a single image, rather than having a partial view or needing multiple cameras.

ACKNOWLEDGMENTS

The research reported here was supported by EU FET-OPEN grant OMNIVIEWS (IST-1999-29017).

REFERENCES

- [1] T.Pajdla, "Epipolar geometry of some non-classical cameras", in Proceedings of Computer Vision Winter Workshop, pages 223-233, Ljubljana, Slovenian Pattern Recognition Society, Slovenia, February 2001.
- [2] T.Pajdla, V.Hlavac, "Image-based self-localization by means of zero phase representation in panoramic images", in Proceedings of the 2nd International Conference on Advanced Pattern Recognition, volume 2013 of Lecture Notes in Computer Science, pages 24-33, Heidelberg, Germany, March 2001. IAPR, Springer-Verlag.
- [3] T.Werner, T.Pajdla, "Cheirality in epipolar geometry", in Proc. Eight Intl. Conf. Computer Vision. IEEE Computer Society Press, July 2001.
- [4] J. Gaspar, E. Grossmann and J. Santos-Victor, "Interactive Reconstruction from an Omnidirectional Image" VisLab-TR 02/2001 - 9th International Symposium on Intelligent Robotic Systems - SIRS2001 - Toulouse, France, July 2001.
- [5] N. Winters and J. Santos-Victor, "Information Sampling for Optimal Image Data Selection", VisLab-TR 01/2001 - 9th International Symposium on Intelligent Robotic Systems - SIRS2001 - Toulouse, France, July 2001.
- [6] S.Gaechter, T.Pajdla and B.Micusik, "Mirror Design for an Omnidirectional Camera with a Space Variant Imager", Omnidirectional Vision Applied to Robotic Orientation and Nondestructive Testing Workshop. Budapest, Hungary, August 2001.
- [7] N.Winters, J.Gaspar, E.Grossmann and J.Santos-Victor, "Experiments in Visual-based Navigation with an Omnidirectional Camera", Proceedings of the IEEE ICAR 2001 Workshop: Omnidirectional Vision Applied to Robotic Orientation and Nondestructive Testing, Budapest, Hungary, August 2001. Invited Talk.
- [8] T.Svoboda, T.Pajdla, "Matching in Catadioptric Images with Appropriate Windows and Outliers Removal", Proc. of the 9th International Conference on Computer Analysis of Images and Patterns. Springer Verlag. Warsaw, Poland, 5-7 September, 2001.
- [9] N. Winters and J. Santos-Victor, "Visual Attention-based Robot Navigation using Information Sampling", Proceedings of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'01), Hawaii, USA, October 2001.
- [10] H. Bakstein and T. Pajdla, "3D Reconstruction from 360 x 360 mosaics", Proceedings of the CVPR'01 conference, Hawaii, USA. December 2001.