OMNIVIEWS

IST–1999-29017

Omni-directional Visual System

Final Report

Covering period 1.9.2000-1.9.2001

Report Version: 1

Report Preparation Date: September 17, 2001

Classification: Public

Contract Start Date: September 1st, 2001 Duration: 12 Months

Project Co-ordinator: DIST - University of Genova (Giulio Sandini)

Partners: Czech Technical University in Prague (Tomás Pajdla)

Instituto Supérior Tecnico in Lisbon (José Santos-Victor)

Project funded by the European Community under the "Information Society Technologies" Programme (1998-2002)

Document Roadmap

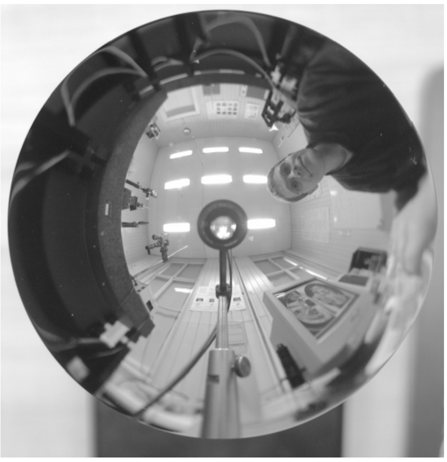
We tried to compile this document so that is should be self-contained and should not require the reading of the 8 annexes and the original Technical Annex of the proposal. The annexes are included because they contain the details of the different activities and for completeness. More documents and reports (plus copies of published paper) can also be found at the OMNIVIEWS web site: *http://cmp.felk.cvut.cz/cmp/omniviews/*. Furthermore at the review meeting a CDROM containing all these documents plus some software code and other material pertaining the demonstrations will be distributed. Note that the electronic version of this document has been produced with a low-compression and no-subsampling of the images. This results in a much bigger file but allows to enlarge the images when seen on the vide screen.

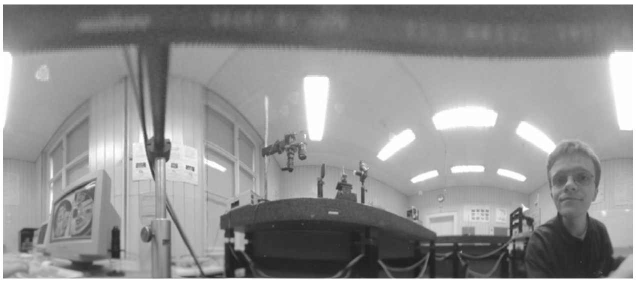
After the *Executive Summary* reported in **Section 1**, **Section 2** is a summary of the objectives and assessment criteria stated in the Technical Annex. **Section 3** reports how we proceeded with the mirror design and realization. **Section 4** reports our main results and achievements including a section reporting software that could be used for future applications. **Section 5** is the most relevant because reports a one-by-one discussion about the assessment criteria. **Section 6** lists the formal deliverables and how there have been compiled from the working documents of the project. A list of the software developed and publications is also included here. Finally in **Section 7** a draft outlook of future development in presented.

Contents

1 Executive Summary

The main objective of the 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 deg) images and performing a useful low-level processing on the incoming stream of images. In particular the goal was to integrate a space-variant CMOS visual sensor incorporating novel and unique technologies, with a mirror with a specially designed curvature. This matching, if demonstrated, would provide 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. During this assessment phase the consortium has realized the camera and the matching mirror, tested its performance in 3 different application areas, developed general purpose tools for the design of omnidirectional mirrors, published scientific papers. Furthermore ideas for new technologies and applications extending the original objectives have been proposed and could be exploited in a follow-up project. We believe that 7 assessment criteria described in the proposal have been successfully met. Among them the most important is related to the fact that the panoramic images obtained by our technology are not only equivalent to the ones obtained with conventional ones but these images can be obtained at no computational cost. For example with our current prototype a panoramic image composed of about 33,000 pixels is obtained by simply reading out the pixels (i.e. 33,000 operations) while with a conventional solution the same image would required more than 1.7 million operations (about 50 times more). 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. For this reason we believe the project was indeed very successful (see the figure below).
From the consortium perspective, all three partners contributed significantly to the project (the results obtained would have been impossible if any one of the partners would have been missing). In particular, the group at DIST – University of Genova, contributed in the realization of the space-variant camera and its software components. The group at the CMP –Czech Technical University contributed to the design and realization of the mirror and its use in different application areas, The group at IST in Lisbon contributed to the use of panoramic images and to the development of mirror design software tools.





Above: Image acquired by an "OMNIVIEWS" camera. **Left**: 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.

As to future plans, the fact that we believe we have successfully met all the assessment criteria, convinced us that it is indeed possible to prepare a follow-up proposal. This proposal will be aimed at increasing the overall resolution of the device by realizing a higher-resolution sensor (possible because of the on-going improvements of CMOS technology) while reducing the overall size with the adoption of new optical components. Furthermore we intend to investigate application areas where the use of panoramic images is not only preferable but indeed required. For example for the realization of endoscopes for the inspection of pipes and for the imaging of internal cavities of the human body (e.g. the esophagus, or other segments of the digestive system). In these application areas the need to miniaturize the device is a very strong requirement against which testing and improving the technology. In this follow-up phase we intend to extend the consortium by including industrial partners as well as potential users. Some of these potential partners already expressed a strong interest for the medium-long term objectives of the project.

2 Project Objectives and Success Criteria[1]

The main objective of the project is 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 is to design and realize a low-cost, miniaturized digital camera acquiring panoramic (360 deg) images and performing a useful low-level processing on the incoming stream of images.

The key technologies are two: panoramic mirrors and space-variant CMOS visual sensors. By designing an appropriate matching between the curvature of the mirror and the geometry of the photosite array of the visual sensor the microcamera will deliver panoramic images immediately usable by a human observer that can be processed without previous mapping. These two technologies are an ideal match to obtain a small-size and low-cost device for application areas such as surveillance, navigation control in robotics and telepresence.

One of the crucial aspects is to define the best match between the profile of the mirror and the distribution of the sensitive elements of the CMOS sensor. Another crucial aspect is the study and definition of the best match between image geometry resulting from the mirror/sensor coupling, and the visual measures required. For example, starting from a log-polar sensor, the optimal profile may be a conic mirror because, a proper match between the rate-of-increase of the log-polar geometry and the angle of the cone, will produce panoramic images so that: a) their elaboration will demand less effort and, therefore, simpler and faster hardware structures and b) their geometry will be directly usable by a human operator. On the other hand, conic mirrors seem not optimal for some classes of visual processing (e.g. stereo matching). Aspects of this kind need to be properly addressed taking into consideration technological, processing and application-derived constraints.
At the end of this first phase the aim is to obtain the following information:

- Definition of the optimal profile of the mirror
- Definition of the matching distribution of the sensing elements in the visual sensor
- Selection of processing primitives to be realized in hardware
- Estimate of the size of device and the speed of operation
- Preliminary demonstration of a potential application

2.1 Success Criteria

The success criteria listed in the original Technical Annex are summarized here.

2.1.1 Direct read-out of panoramic images

In OMNIVIEWS the most important criteria for success is the demonstration that it is possible to realize a camera capable of **acquiring panoramic images by simply reading out the pixels** as opposed to traditional implementations based on standard, uniform resolution, visual sensors, which require software remapping of the acquired images.

2.1.2 Frame rate

As a consequence of the direct read-out the frame-rate achievable will be higher than that obtainable with traditional cameras. In fact the frame-rate will be bounded by the read-out frequency of the sensor array that we expect to be at least 25 frames/s. This will be obtained at no extra computational cost by exploiting the duality between the sensor and mirror profile.

2.1.3 Resolution and layout of the sensor

An important criterion is the resolution achievable in terms of number of pixels in the panoramic image. It is clear that it is impossible to define an absolute optimum as, ideally, the better the resolution the better the image. However when images have to be processed in some way, the number of pixels has to be matched to other parameters such as: computational power and cost. These parameters are mostly defined by the application. As far as the prototype camera we intend to build in this assessment phase, the available CMOS sensor that has approximately 33000 pixels bounds the number of pixels. Therefore the overall number of pixels of the panoramic image will depend on how "efficient" will be the matching between the sensor, the

[1] Parts of this section are extracted from the Technical Annex of the project

mirror and the lens (part of the image may be covered by the reflection of the camera itself and, therefore, it is not useful).

2.1.4 Mirror profile and size

The profile of the mirror cannot be defined independently of the sensor's layout and resolution. The goal here is to define the profile so that it matches the space-variant and polar layout of the sensor and is of small size..

2.1.5 Lens characteristics

The main features of the lens are the width of the Field of View (FOV) and its focal length. Both depend on the size of the sensor and affect the overall size of the camera. Considering a sensor's size of 1 cm^2, the expected focal length is 8mm with a field of view 60deg with a distance between the lens and the mirror of approximately 50mm. Other parameters like the sensitivity of the CMOS element may however affect the optics characteristics.

2.1.6 Camera cost.

The cost of the camera is composed of the cost of:

1) the sensor,
2) the mirror,
3) the lens
4) the computational power required to produce the panoramic image.

The first criterion is that item 2 and 3 should be the same as that of standard panoramic cameras; secondly the cost of item 4 should be negligible. Therefore the cost of the camera, compared with traditional solutions, will be defined by the cost of the sensor. This cost is, independently of the sensor's layout, defined by the chip size. Our target here is to remain in the range of 1 cm^2, which is approximately the size of our current retina-like sensor as well as that of standard CMOS sensors.

2.1.7 Image quality

The definition of the "optimal" image quality, against which we will assess the success of the project, is driven by the targeted application (indirectly defining also the overall cost). The project will be successful if we demonstrate that it is possible to create virtual images by simple reading out the pixels from the proposed sensor and to use such images in the aimed applications in an equivalent way as images from a moving standard camera can be used.

3 Methodologies

As stated in the proposal during the assessment phase we concentrated on the design of mirrors matching the currently available log-polar camera with the SVAVISCA sensor and the demonstration of their utility in different application areas. The main features of the SVAVISCA layout superimposed to a image reflected by a spherical mirror are shown in Figure 1. In the table below the main features of the SVAVISCA layout is reported.

Exponential grow factor	Total pixels	Arrangement (periphery)	Arrangement (Fovea)	Min pixel size (µm)	Max pixel size (µm)	Sensor's size (mm)
1.02337	32,887	252x110 (27,720)	42 rings (5,167 pixels)	6.8x6.45	82x82	3.6 (radius)

Table 1: parameters of the SVAVISCA log-polar sensor used for the first phase of the project.

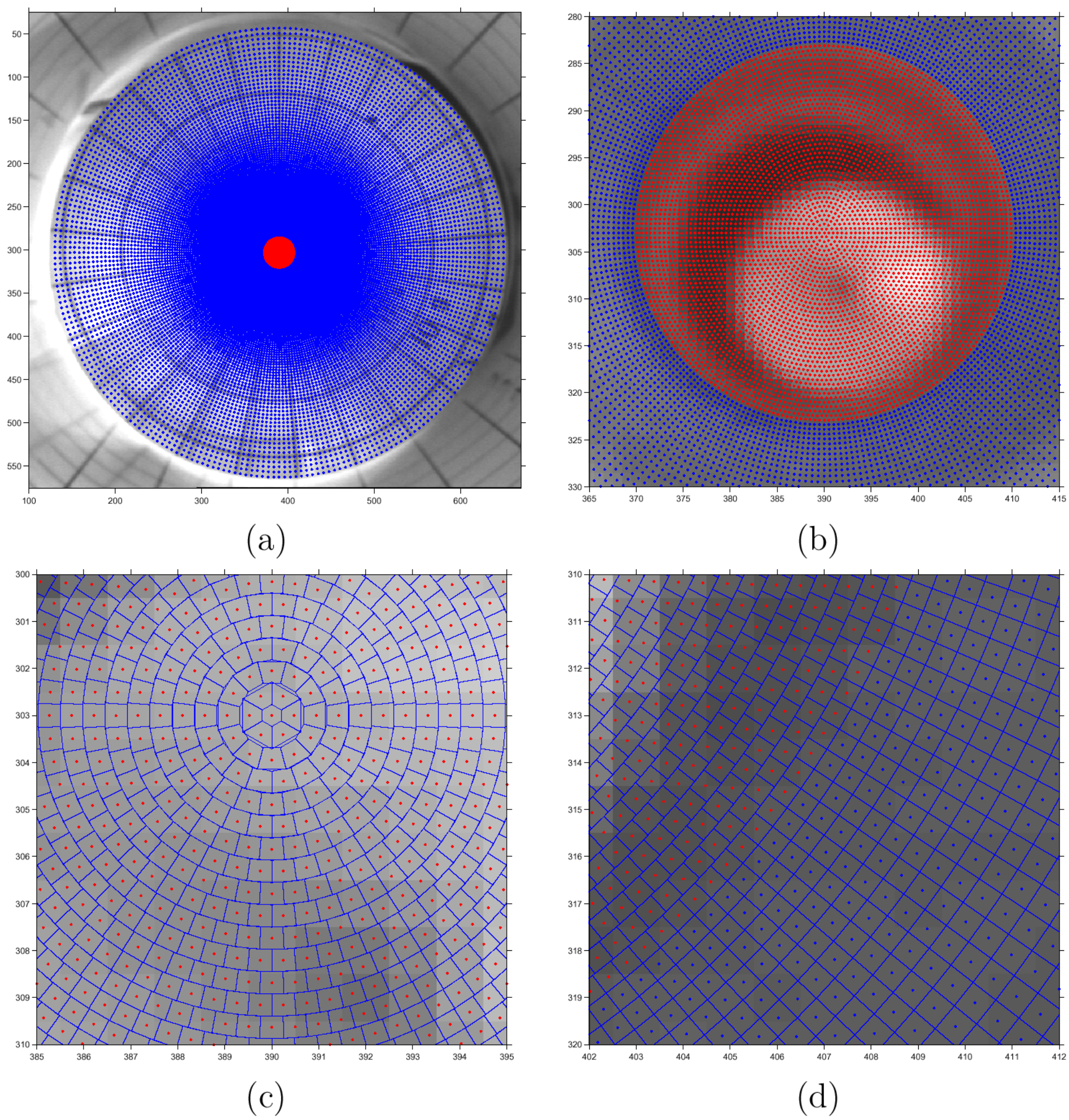


Figure 1: Arrangement of pixels in the SVAVISCA retina. (a) SVAVISCA consists of constant-resolution fovea (red) and a space-variant resolution periphery (blue). (b) The fovea; (c) detail of the fovea center; (d) Detail of the border between the fovea and the periphery.

3.1 Mirror design and realization

The main guiding design rule here was to obtain a mirror for an omnidirectional camera with a uniform vertical resolution when using the SVAVISCA sensor. In practice this means that a cylinder around the camera will be mapped, by the camera itself, into a rectangular image with equal pixel height irrespective of the elevation (see Figure 2). Another guiding principle was to reduce as much as possible the size of the camera self-reflection.

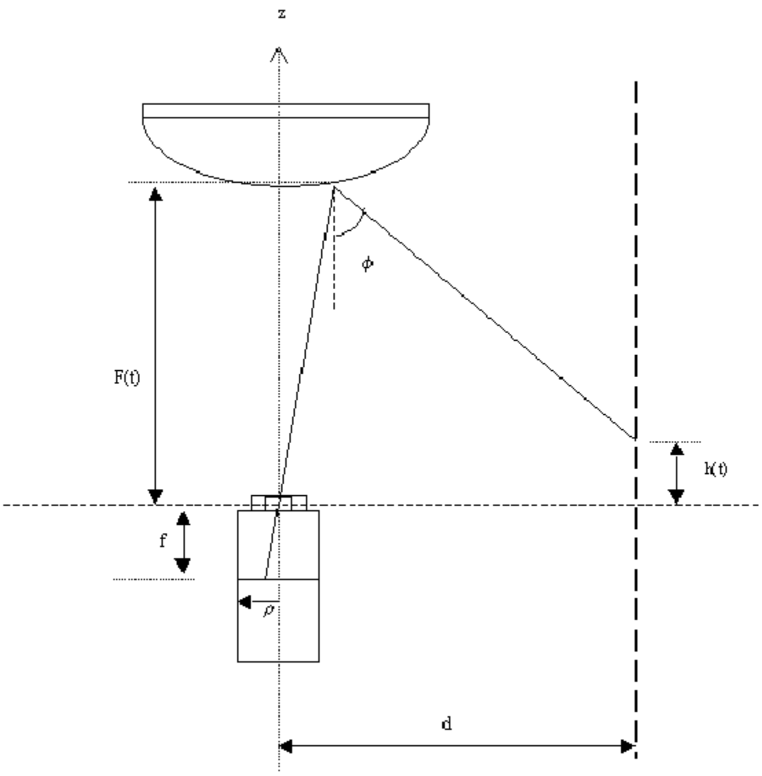


Figure 2: Projection of the rays of vertical resolution.

The roadmap of mirror design and fabrication was carried out, as reported in Deliverable 2 (see section 6 of this document), in the following way. First a simulation of the SVAVISCA imagining (see Figure 1) by re-sampling conventional images was implemented and tested (see Annex 1 of this report). A sample of such simulation is shown in Figure 3. A comparison between the spherical and the hyperbolic mirror shows that for the spherical one the reflection of the camera on the mirror is larger than for the hyperbolic mirror (approximately 70 rows of the 152 available corresponding to the dark part of Figure 3d). Based on the simulations, the mirror providing uniform cylindrical projection with the SVAVISCA sensor was designed using MATLAB routines (see Annex 2 of this report) and manufactured. The mirror was tested then as detailed in Annex 3 of this report.

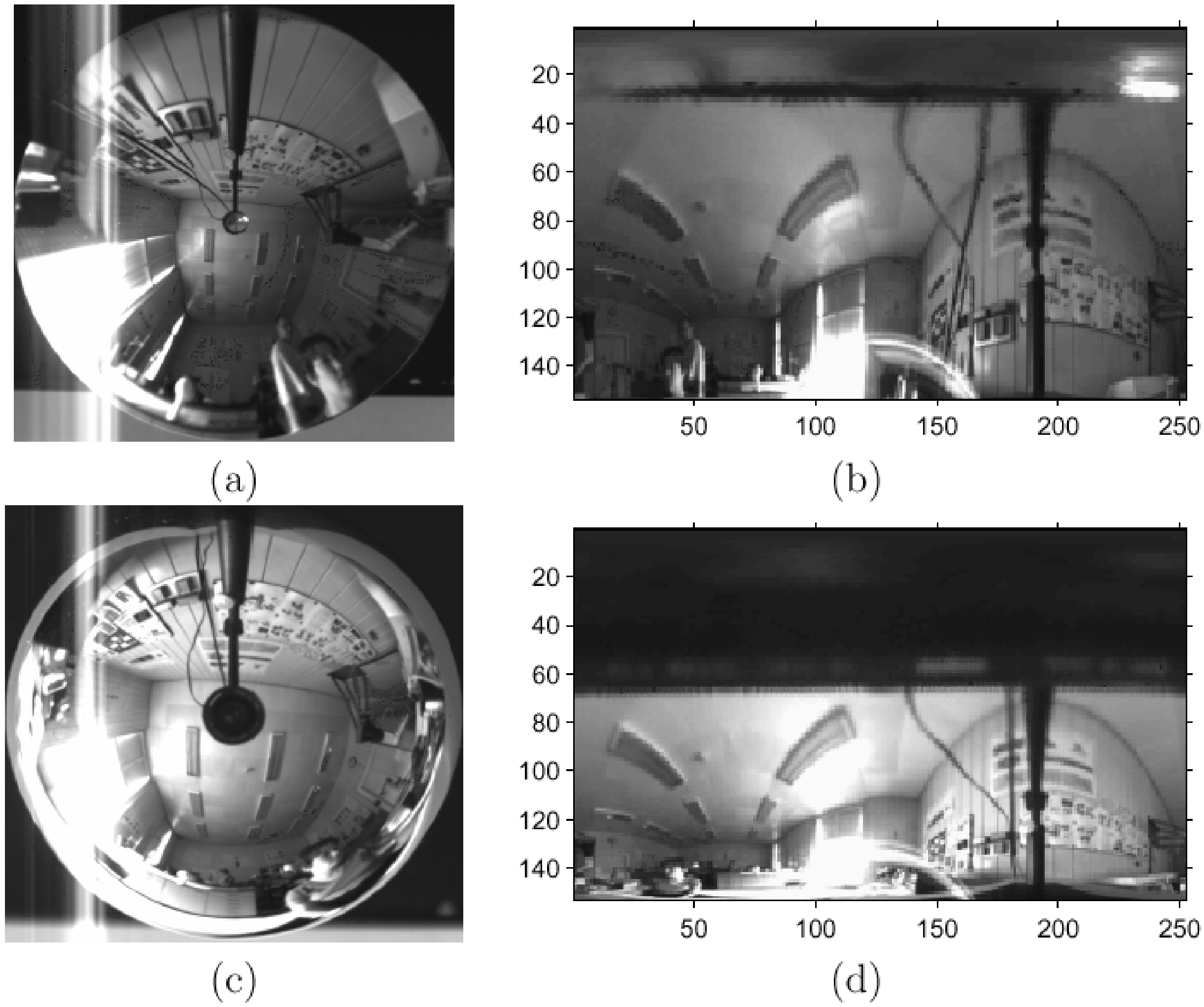


Figure 3: Simulated SVAVISCA images of a natural scene. (a) Image acquired with a conventional camera (576x768 pixels) and a hyperbolic mirror; (b) Image remapped by simulating the SVAVISCA sensor. (c) Image acquired with a conventional camera and a spherical mirror; (d) Image remapped by simulating the SVAVISCA sensor.

At a later stage a more general tool for the design of mirrors was developed at IST Lisbon. The MATLAB code is distributed in a CDROM (see section 6 of this document). With this tool another interesting design possibility was investigated for applications in the area of robot navigation. Specifically the possibility of preserving ratios of distances measured on the ground plane. In such a case, one can directly use image measurements to obtain ratios of distances or angles on the ground floor (which can greatly facilitate navigation problems or visual tracking). Such images are also termed as “bird ’s eye views”. Figure 4 shows how the ground plane is projected onto the image plane, where the camera-to-ground distance is represented by h and d(t) represents radial distances taken on the ground plane. In section 7.3 the possibility of designing mirrors with a mixed constant (i.e. horizontal and vertical) resolution is presented as a possibility for future work.

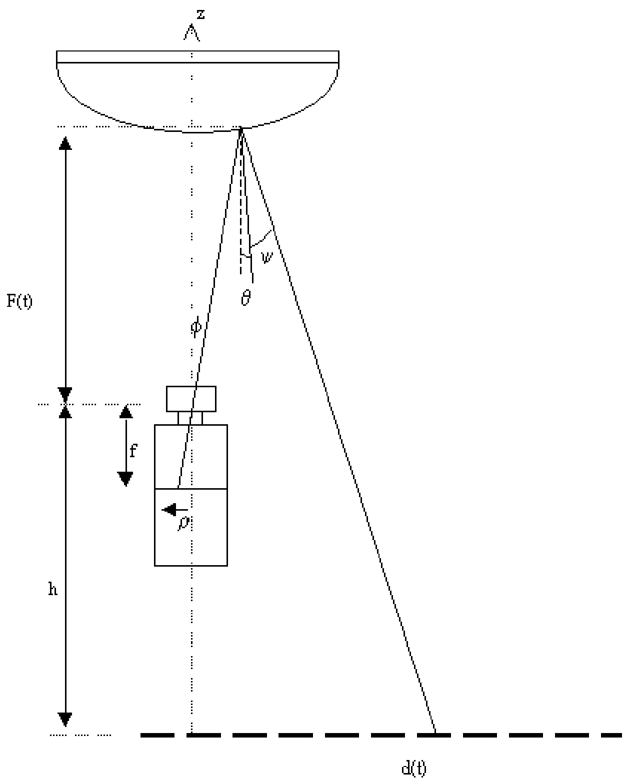


Figure 4: Projection of the rays of horizontal resolution.

Finally the work on demonstrations was started along with the initial discussion about the features to be added/modified during the follow-up project (provided we produce convincing results at the end of the assessment phase).

3.2 Definition of Demonstrations

To test the performance of the omnidirectional log-polar camera (OLPC), three main demonstrations were planned:

1) Surveillance demo (CMP-CTU): this demo will show the possibility of detecting and tracking people in an indoor environment acquired with a OLPC. The demo will run in real-time using software developed in MATLAB.
2) Robot navigation demo (IST): for this demo the OLPC will be mounted on board a mobile base and the images will be processed in real time to guide and self-locate the robot in an indoor environment.
3) Remote Transmission (DIST); for this demo the real-time transmission of panoramic images over telephone and GSM lines will be demonstrated.

4 Results and Achievements

The main results of the project will be presented here subdivided into three headings: 1) Hardware set-ups, 2) Software Design Tools and 3) Demonstrations.

4.1 Hardware set-ups

In Figure 5 the mirror as well as the hardware set-up realized and used to test the performance of the prototype is shown.

The set-up is composed of the camera, the mirror, plus the holder. The camera has a standard C-mount lens and is connected to a PC through the parallel (printer) port. Considering the bandwidth limitation of the interface up to about 10 frames/s. can be acquired.
The actual profile of the mirror is shown in Figure 5 along with a picture of the mirror itself. More details about the curvature and its design principle are included in Deliverable Item 2. The main features of the mirror can be qualitatively seen in Figure 6 where a picture obtained with the set-up of Figure 5 is shown and compared

with one obtained with a hyperbolic mirror. In particular the implicit normalization of size of objects lying on a cylinder around the mirror can be seen: the size of the black dots is the same when acquired with the OMNIVIEWS set-up while they are different when acquired with an hyperbolic mirror.

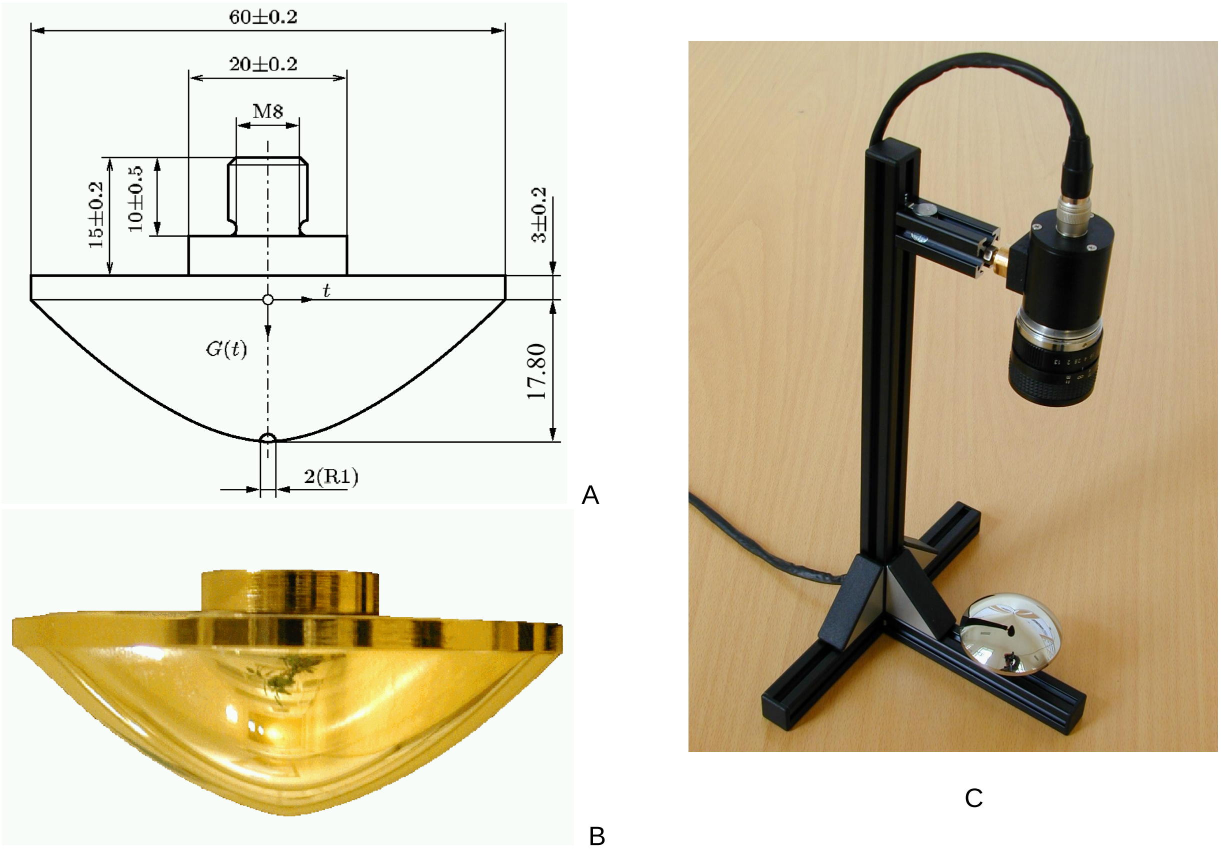


Figure 5: A) Profile of the OMNIVIEWS mirror; B) mirror realized (before coating); C) Set up used including the mirror and the log-polar camera.

A

B

C

Figure 6: A) Set-up with a test pattern composed of black dots at the same distance but at different elevation. B) Image acquired with an hyperbolic mirror; C) Image acquired with the OMNIVIEWS mirror. The size of the dots are the same in C while in B they are different. In both cases the images are shown as they are acquired (i.e. no remapping and warping is required).

Furthermore the mirror was designed so that the self-reflection of the camera is made minimal. This was obtained by shaping the tip of the mirror so that even centermost part of the mirror reflects points of surrounding scene. Therefore SVAVISCA mirrors have pointed tips instead of the flat tips of hyperbolic mirrors.

4.2 Software

The software realized in OMNIVIEWS can be classified in two groups: 1) software tools for mirror simulation and design; 2) Processing of omnidirectional images.

4.2.1 Software tools for mirror design

Software tools to design the various types of mirrors described in this report have been developed. This software has been developed in MATLAB and allows for designing a rich set of profiles simply by modifying the input parameters. In particular, we have defined a set of specifications in tune with the particular applications envisaged within the project. The Graphical User Interface (GUI) of MATLAB is shown in Figure 7. The mirror shapes for both conventional and log-polar cameras can be visualized as well as the curves that allow us to analyze the curvature of the mirrors. These mirrors will be projected by considering a linear relation between the position of the world point and the radial position of an image point.

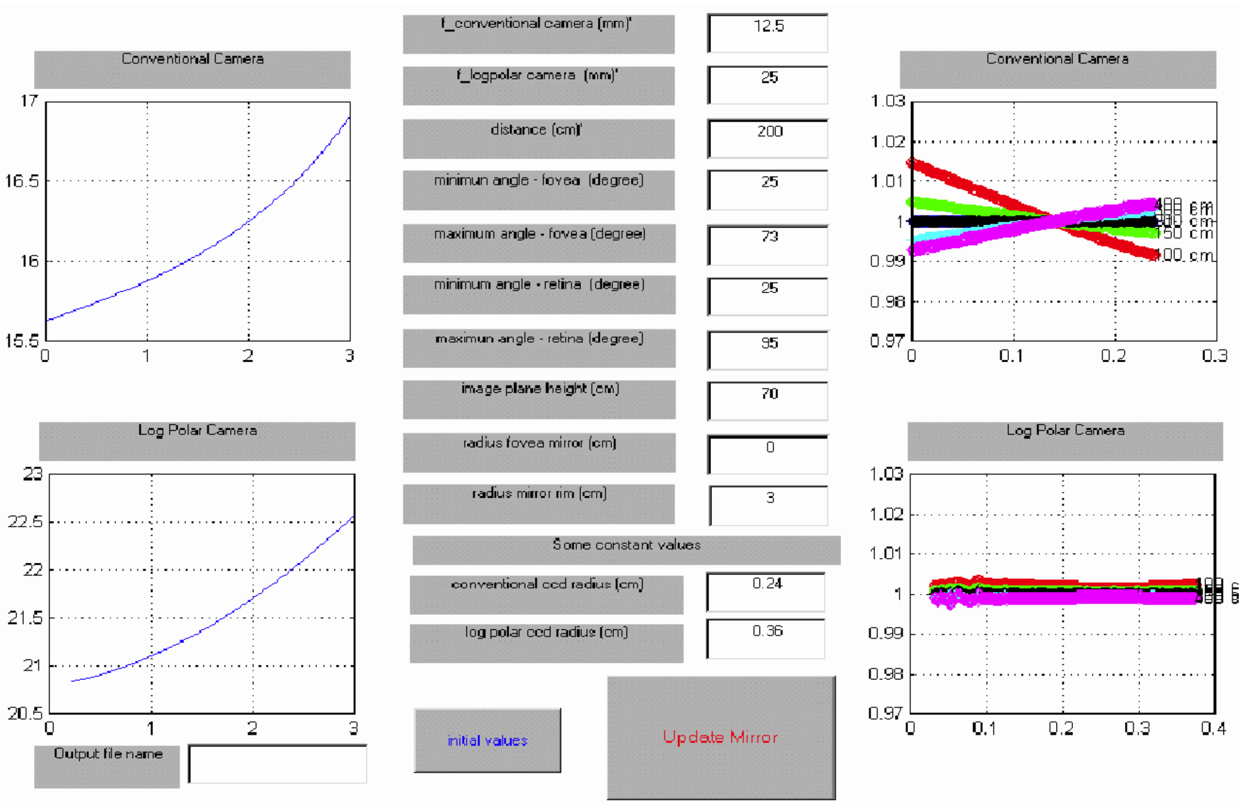


Figure 7: Graphical User Interface of the MATLAB tool developed for mirror's design.

Furthermore, software tools to acquire and process OMNIVIEWS images have been realized both in C++ and MATLAB.

4.2.2 Software for omnidirectional image processing

In this category of software we describe the work done not devoted to specific demonstrations at this stage but that could be exploited in a later stage of development. Within this category we include: 1) Teleoperation and 3D reconstruction; 2) Visual Tracking, and 3) Self-location on the basis of AGAMI fiducials.

Teleoperation and 3D reconstruction. In order to be able to teleoperate the mobile robot, an intuitive user interface is being developed. The main idea behind this approach is to present the user with a simple 3D model of the environment. The user may then pick locations, for example, a certain office, from within the model thus defining the robot 's destination. Along this line at IST an algorithm that can reconstruct environmental scene from a single omnidirectional image has been developed. A result of the algorithm is shown in Figure 8 (see Annex 8 for more details).

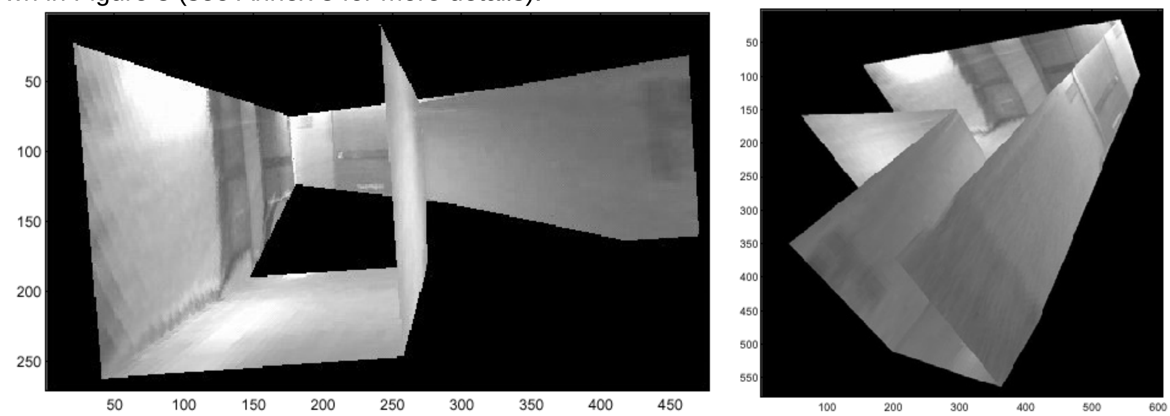


Figure 8: Views of a reconstructed 3D model of a corridor obtained from a single panoramic image.

Visual Tracking. The interaction with dynamic environments requires a constant monitoring of the changing position of moving objects and the perception of self-motion is greatly enhanced by visual perception. The use of wide field-of-view visual sensors (e.g. panoramic), further improve the reaction time of the system, since an extended part of the environment is monitored without the need for mechanical control of the sensor. Also the navigation capabilities of the system greatly benefit from the translation-rotation ambiguity reduction of omnidirectional systems. With conventional cameras, the 3D rigid motion of a plane can be described by a 2D transformation (homography - 8 degrees of freedom) in the image plane. Usual omnidirectional systems distort the viewing geometry in such a way that the motion of planar surfaces can no longer be described by such simple transformations. However with a panoramic representation (such as the one obtained with the OMNIVIEWS approach), the imaging geometry recovers a quasi-Cartesian description, allowing approximating the image observed deformations with 2D homographies, with the additional advantage of the observed surfaces “rarely ”leaving the field of view. Along this line the algorithm implemented is based on a succession of processing steps involving local image warping (acting as a sort of virtual pan-tilt unit) and template matching. Some frames of a simulated tracking sequence illustrating translations and rotations is presented in Figure 9. The initial region is represented by solid lines. The small circles represent the output of the algorithm and are supposed to track as close as possible the corners and the center of the selected region as it changes along time. Notice that the algorithm copes with the typical angular wrapping of panoramic sensors and the high precision of the tracker even with significantly blurred images. For more details abut the algorithm refer to Annex 8.

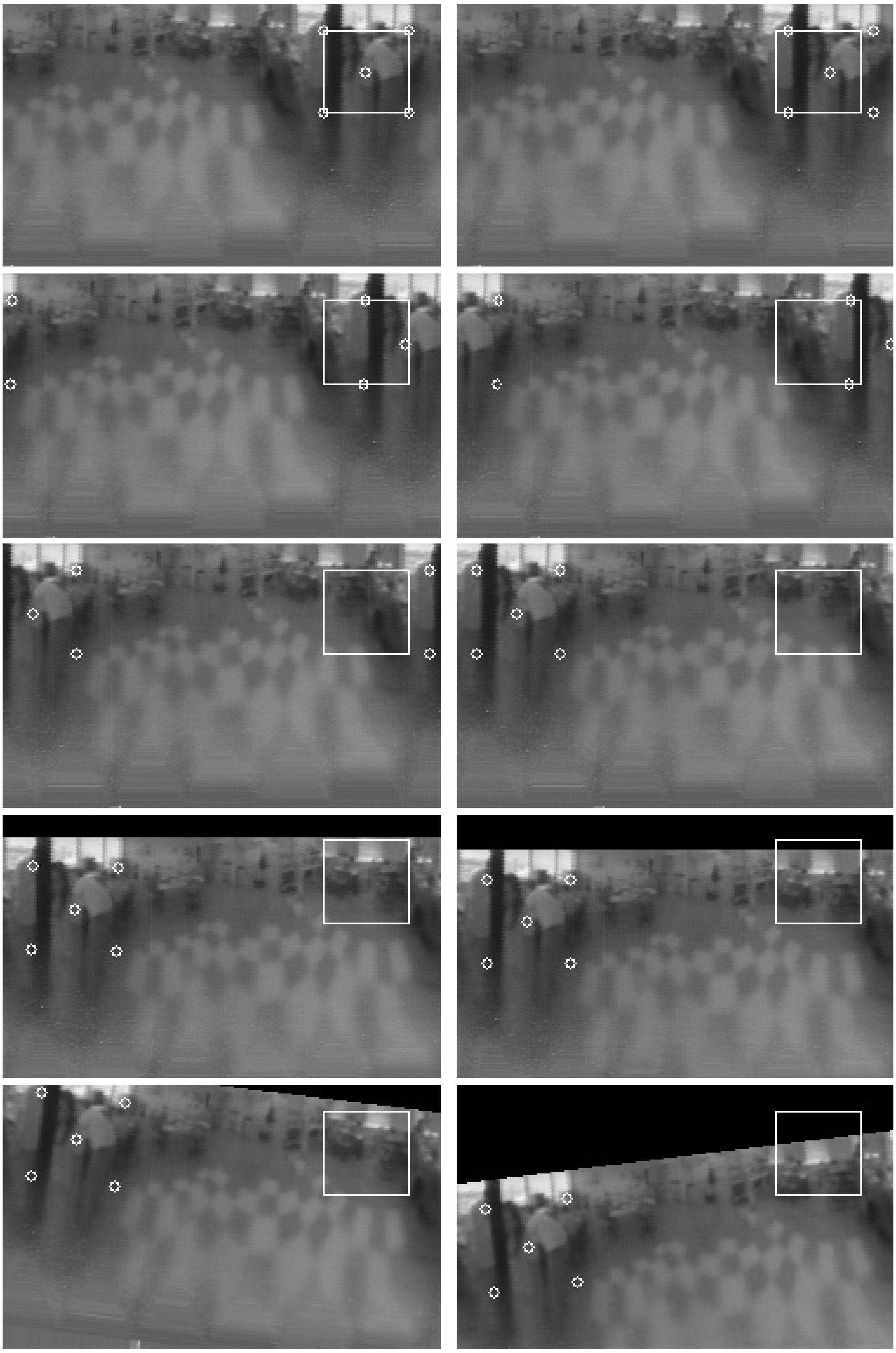


Figure 9: Tracking sequence. The solid square represents the original location of the selected region to track. The small circles correspond to the center and the corners of the estimated location of the tracked area.

Self-location on the basis of AGAM fiducials. AGAM fiducials are particular visual patterns that can be used to estimate their position relative to a camera on the basis of gray level distribution (ref: A.M. Bruckstein, R.J. Holt, Huang T.S., and Netravali A.N. New devices for 3D pose estimation: Mantis eyes, agam paintings, sundials, and other space fiducials. IJCV , 39(2), 2000). The aim of this activity was to study how AGAMI fiducials are imaged by the OMNIVIEWS set-up and to test their use in a self-location task (see Annex 5). The main feature of AGAM fiducials is that their observed intensity depends on the viewing angle as shown in

Figure 10.

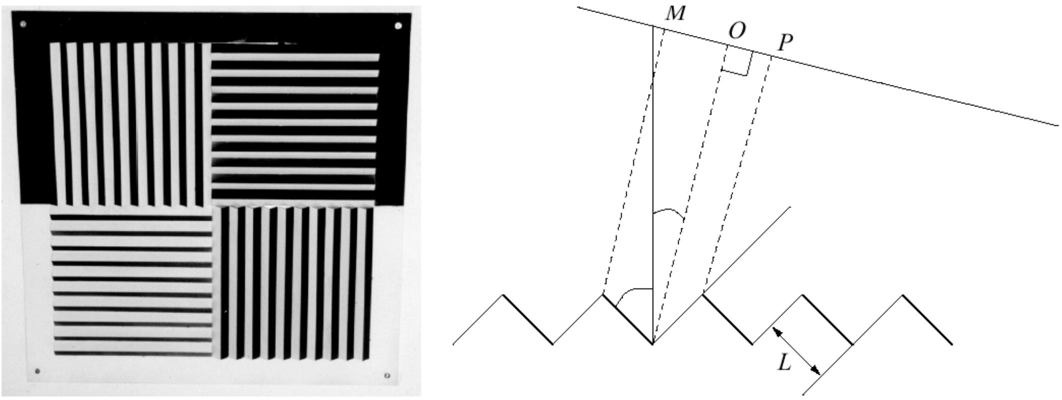


Figure 10: Left: image of an AGAM fiducial; right: Height profile of an AGAM fiducial. The difference in intensity as a function of the viewing angle is caused by "grooving" of the horizontal and vertical gratings.

A comparison between AGAM fiducials as seen from a conventional and from the OMNIVIEWS set-up is shown in Figure 11. The aim of the experiments performed was to determine how far an observer can be from a fiducial before the localization breaks down. The distance is primarily determined by the size of fiducial projection into panoramic images. If the projection of a fiducial becomes too small, the localization information, which is the average brightness in fiducial squares, is lost. The main conclusion of this experiment (which is reported in details in Annex 6) was that, with the current log-polar sensor the method can be used successfully inside a 4m x 4m room.

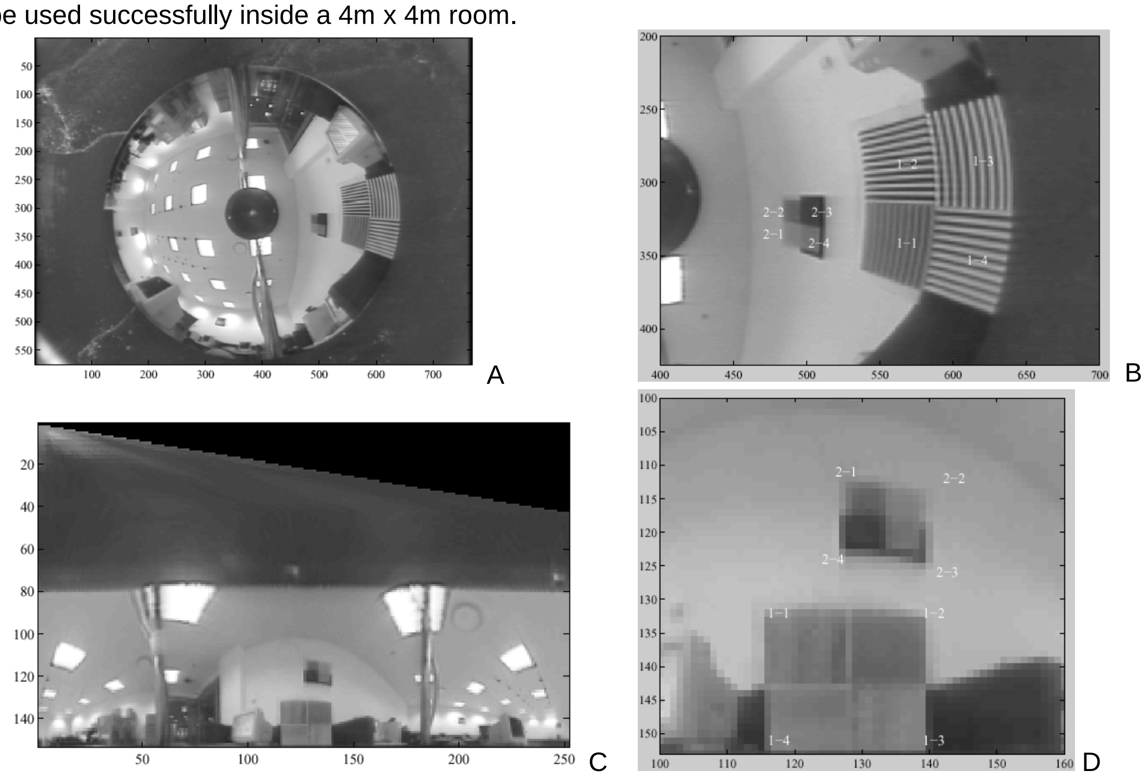


Figure 11: A) omnidirectional view of an office scene taken from a conventional camera containing AGAM fiducials. B) The part of the image containing the fiducials is shown enlarged. C) Same as A) but acquired by an OMNIVIEWS set-up. D) Part of the panoramic image containing the fiducials.

This is lower than with a conventional camera (acquiring a much larger number of pixels) but the limitation is due to the resolution of the sensor and can be improved, without loosing the computational advantage of the OMNIVIEWS approach, by increasing the number of pixels in the log-polar sensor.

4.3 Demonstrations

4.3.1 Robot Navigation

We propose two main navigation modalities: 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. The navigation process combines appearance based methods and visual servoing upon some environmental features in a bird 's-eye view of the ground plane. Visual Path Following is required for local, very precise navigation, for e.g. door traversal, docking. The robot is controlled to follow a pre-specified path accurately, by tracking visual landmarks in the scene.

4.3.2 Surveillance

The surveillance demonstration presents a method for motion detection based on background subtraction (see Annex 7 of this report). The main point here is to show that the detection and tracking of moving objects (such as people entering a room) is possible using the panoramic image sequence directly obtained from the OMNIVIEWS set-up. The moving object detection algorithm is based on image subtraction using a model of the background constantly updated on the basis of spatio-temporal information (stationary or slowly moving objects becomes part of the background). Further for fast background changes a reset method for the algorithm is derived. The thresholds for the temporal and background change detection are determined depending on the noise in the images, where the noise is characterized by the standard deviation estimated with the least median of squares. The algorithm was implemented in MATLAB and runs on a portable computer in real-time (the code is listed in Annex 7 of this report).

4.3.3 Omnidirectional Image transmission

In this demonstration the transmission of omnidirectional images will be presented through telephone lines. Once more the key point is to demonstrate that useful information can be derived from images acquired and transmitted directly from the OMNIVIEWS set-up. The low number of pixels allows a more efficient transmission compared to conventional solutions. The demonstration will show transmission of live images from a remote server performing acquisition, compression and transmission, to a PC performing decoding and visualization.

5 Assessment of results vs. proposed success criteria[2]

The project has demonstrated the feasibility of the approach proposed in the sense that it has been shown that it is possible to design a sensor-mirror pair producing panoramic images without software/hardware remapper needed by using standard uniform resolution cameras. The demos will show that the images obtained, even if constrained by the currently available sensor, are indeed usable in different application areas. Before entering into the details of the success criteria mentioned in the original proposal we would like to stress that besides the original criteria during the development of the project the idea of a new kind of mixed-mirror have arisen and general-purpose mirror design tools have been developed. Furthermore a new, cheaper, technology for the fabrication of the mirror has been tested (not for the actual OMNIVIEWS mirror).

5.1.1 Direct read-out of panoramic images

> ***…the project will be successful if we demonstrate that: 1) it is possible to design a mirror-sensor match producing desired resolution panoramic images; 2) the mirror and the sensor can be coupled by a realizable lens (meaning a lens with a suitable viewing angle, depth of field and focusing surface); 3) the image can be acquired without any extra processing.***

Examples of an image acquired with a direct read-out from a retina-like sensor coupled to the mirror designed in the project are shown in Figure 12. An image acquired from a similar environment with a standard camera is shown for comparison.

[2] The parts in quote are extracted from the Technical Annex of the project

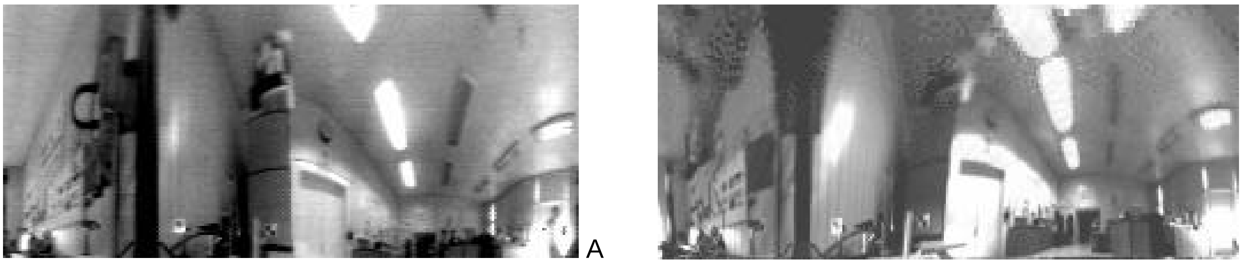


Figure 12: (A): Image of an office scene acquired with the OMNIVIEWS set-up; (B) Image of the same environment obtained by simulating a log-polar remapper on a standard constant resolution image of size 576x768. For technical reasons the two images were not acquired simultaneously as it is evident on the average illumination. The two images are not qualitatively different

In relation to this particular criterion let's try to quantify the computational load required by our solution in comparison to the load needed to generate panoramic images using a conventional rectangular pixel array. Let us have a conventional imager with a rectangular pixel grid. Figure 13 shows that pixels are read from the input imager (left) and according to a lookup table transformed into the output panoramic image (right).

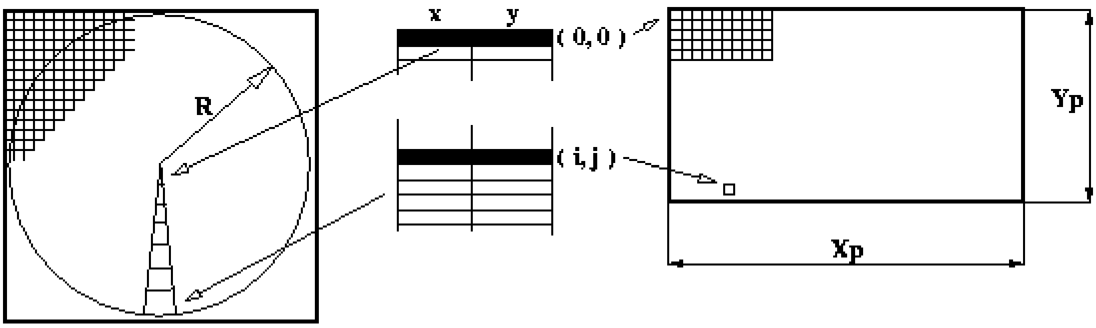


Figure 13: Graphical representation of the operations required to obtain a panoramic image from a conventional camera.

The mirror is projected into the circle with the radius equal R. The final panoramic image has size X_p x Y_p. One pixel in the panoramic image is mapped into a region in the original image. The size of the region grows from the center to the periphery. In order to transform the original image into the panoramic image, πR^2 pixels have to be read from the imager. If we assume that the output pixels are created from the input pixels by averaging, πR^2 additions are needed and X_p x Y_p divisions have to be done. It is worth noting, because it will become important in a later section, that, in order to simulate such a sensor (i.e. in order to obtain, from a constant resolution image, a image with the same "look" as, for example, the image directly obtained by the Giotto camera), we need to start from a constant resolution image composed of, at least, 3,600/6.8 (size of sensor/minimum pixel size – see also Table 1). In this way the smallest possible pixel in the log-polar array exactly matches the smallest possible pixel in the constant resolution array AND the field of view will be the same. From this it derives that the value of R required to simulate the SVAVISCA layout turns out to be about 530 pixels. The following table shows the difference between a conventional and SVAVISCA imager for R = 530 pixels, X_p = 252 pixels, and Y_p = 110 pixels.

Operations Required	Readout	Additions	Divisions	Total
CONVENTIONAL	882,400	882,400	27,720	1,792,520
SVAVISCA	27720	0	0	27720

The saving obtained is 64 times.

5.1.2 Frame Rate

… we expect to be at least 25 frames/s.

The frame rate at present is bounded by the acquisition rate of the retina-like sensor which is, in turn, limited by the parallel port of the retina-like camera (about 12 frames per second). In the future we plan to use a faster connection (e.g. USB or PCMICIA) and we will easily reach a frame rate of 100 frames/s. Even with a higher resolution sensor (e.g. with four times more pixels), the objective frame rate will be obtained.

5.1.3 Resolution and lay-out of the sensor

> ***The project will be successful if we: i) design and demonstrate a sensor with even resolution taking into account the constraints given by existing CMOS sensor; ii) propose and demonstrate in simulations a sensor with enough high and enough even resolution to be used in targeted applications.***

In order to assess this criterion the set-up shown in Figure 14 was used.

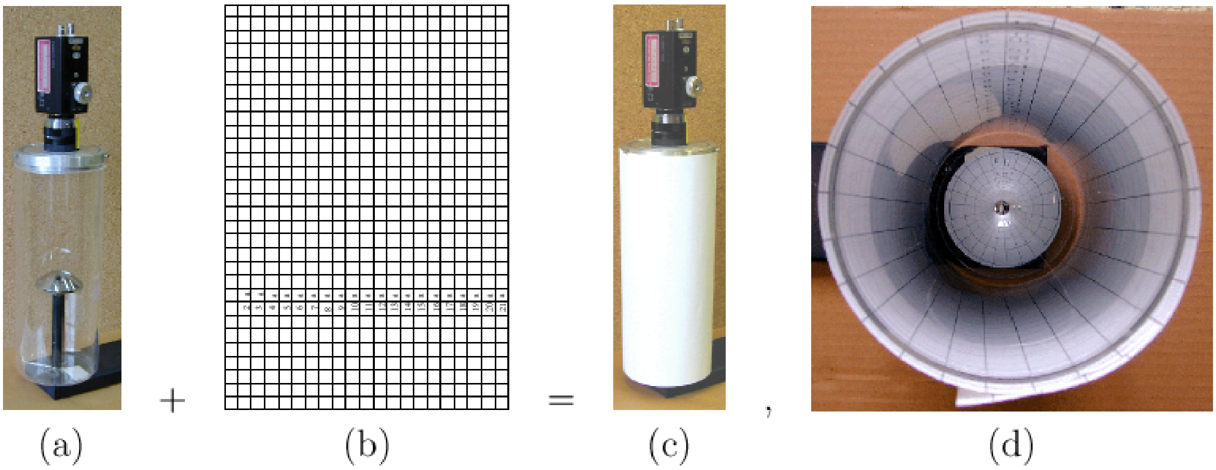


Figure 14: Set up used to test the uniformity of the images. (a) Transparent cylindrical holder of the camera with the mirror. (b) Regular grid that is wrapped around the cylinder as shown in (c). (d) Image seen by the camera. In this picture a standard camera is used to show the "look" of the image acquired. The grid as seen through the cylinder is visible as well as its reflection on the mirror.

A sample image acquired by the OMNIVIEWS set-up is shown in Figure 15. As it is clearly seen the spacing of the horizontal lines of the grid is constant achieving the objective.

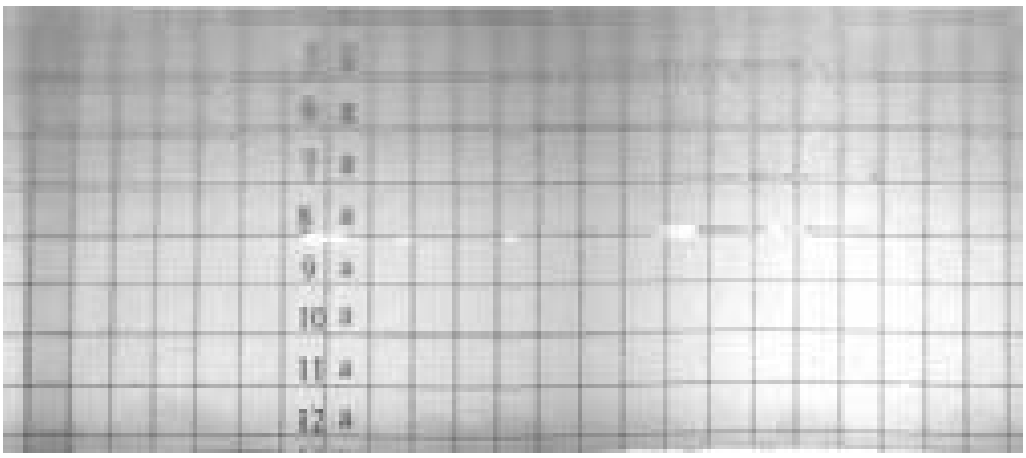


Figure 15: Image of the regular grid acquired by direct read-out from the OMNIVIEWS set-up.

5.1.4 Mirror Profile and Size

> ***The target here is to have a mirror of no more than 60 mm in diameter, which is the diameter of the best mirror that currently have been implemented by CTU.***

The mirror realized in the first phase is 60 mm in diameter as planned. Moreover we want to stress that this is by no means a minimum value. On the contrary we want to stress that, during the execution of the project, we investigated different technologies for the realization of the mirror and discovered a more promising alternative, which will allow, not only to produce mirror of smaller size, but also to considerably reduce its cost. A comparison between the two technologies is shown Figure 16. The "standard" technology is based on chrome coating of a brass core. With this solution the core of the mirror made from brass is manufactured on a numerically controlled lathe. Then, it is polished to optical quality and coated with a Chrome alloy. The price of one mirror is about 350 €. The advantage of the technology is that it is very flexible. Almost any shape can be manufactured and thus it is suited for the sensor development and testing. In the alternative "new" solution the core of the mirror is molded from glass. First, a mould is manufactured and then a number of cores can be molded. The cores are coated with Aluminium and covered by a protective anti oxidation layer. The cost of the mould is approximately 1000 €. The cost of one mirror is under 10 € for large series of mirrors. This technology is ideal for mass production of mirrors.

Figure 16: right: Hyperbolic mirror realized with chrome coating of a brass core; left: alternative technology based on aluminum coating of a glass core (shown here for an hyperbolic mirror).

The prototype realized successfully meet this criterion.

5.1.5 Lens Characteristics

The characteristics of the lenses used in the different demonstrations are included in the following table. For what concerns the effects of the "log-polar mirror" realized on the lens characteristics no difference has been found with respect to "standard" solutions. The prototype realized successfully meets this criterion.

5.1.6 Camera Costs

> ***… the cost of the camera, compared with traditional solutions, will be defined by the cost of the sensor. This cost is, independently of the sensor's layout, defined by the chip size. Our target here is to remain in the range of 1 cm^2, which is approximately the size of our current retina-like sensor as well as that of standard CMOS sensors.***

Having demonstrated that the design of the mirror can be done without increasing the size of the sensor the cost of the camera will be at least the same (if not less) of that based on a standard sensor. Moreover the cost of the hardware/software required to remap the image of a standard camera will be zero in our case. Consequently the overall cost of the log-polar mirror camera will be lower.
As to the mirror it will be possible to reduce its cost for mass production by using the technology described in section 5.1.4. The extra cost required for the glass mold (about 1,000 €) will be amortized already with a rather small number of pieces while the cost of each piece will go from about 350 € to about 10 €.

5.1.7 Image quality

> ***The project will be successful if we demonstrate that it is possible to create virtual images by simple reading out the pixels from the proposed sensor and to use such images in the aimed applications in an equivalent way as images from a moving***

standard camera can be used. In this project three application areas will be investigated: 1) Robot navigation; 2) Surveillance; 3) Telepresence/virtual-navigation

As far as robot navigation is concerned the image should have a FOV wide enough, in the vertical dimension, to see the ground plane around the vehicle (for obstacle avoidance) as well as the environment (for robot orientation). As a reference the vertical FOV should be of the order of 45 deg … and achieve an acceptable control frequency of the order of 10-20 Hz.

In the surveillance field success will depend almost exclusively on the possibility of visualizing the environment with a FOV and resolution sufficient to recognize the face of the persons as well as illegal actions (such as people stealing goods in a shop) without the need of steering the camera.

In the case of telepresence we intend to use the prototype to generate, not only full panoramic images, but also virtual perspective images of part of the environment. This can be used to allow virtual navigation in the environment without moving the camera. If the project will be successful this feature will be obtained by simply reading out parts of the panoramic images *without warping and mapping algorithms*. ***The result will be evaluated in terms of frame rate obtained (we expect at least*** *25 frames/s****) as well as computation required (we expect just partial read-out to simulate panning and tilting of the camera and, possibly, magnification to simulate forward motion).***

The quality of the images has to be considered in relation to the main advantage of our approach namely the possibility of obtaining panoramic images with a direct read-out. More specifically what we have demonstrated with the sample images and demonstration is that **the topology** of the images resulting from the coupling of a log-polar sensor and a matching mirror provides perfectly usable panoramic images without further warping of the image themselves and/or extra processing to be performed to compute the relevant features. This is a net increase in performance/cost ratio as opposed to a conventional approach. Of course there is still space for improvements particularly in relation to the resolution of the images, which is currently constrained by the available sensor. In a later section 7 we will provide examples of the kind of image resolution improvements we expect in a possible follow-up project.

Another factor affecting image quality derives from the quality of the mirror itself and particularly from the influence of two parameters:

1) Numerical Errors -As we do not have an analytic description of the mirror shape and the actual profile is obtained through numerical integration it is important to verify the influence of numerical integration errors in the overall process.
2) Sensitivity -As the Omniviews camera does not have a single center of projection, the linear mappings obtained between pixel distances and world distances is only valid for specific world surfaces (e.g. specific vertical cylinders or horizontal planes in our case. What happens if the radius of the cylinder is changed?

The analysis of the mirror profiles (as reported in Annex 4) is done by calculating a quality index, q(?). This quality index is defined as the ratio between the rate of variation of distances measured in the world (e.g. along vertical distances) and the corresponding rate in the image. For the perfect design process we should have q(?)=1. Computing q(?) involves numerically differentiating the data points of the mirror profile F(t). Here we use different discrete approximations to derivatives as well as the analytical expression for the mirror local slope. In the following Figure 17 we present some results obtained with different methods of derivations.

These results show two main aspects. Firstly, the influence of varying distance with respect to the desired mapping properties does not seem to be too important, which suggests that we are close to the situation of a single projection center. Secondly, the way how derivatives are computed is very important in terms of quality analysis. This is particularly relevant for the log-polar sensor where by increasing the sensor resolution towards the fovea imposes strong constraints on the way to compute derivatives of the mirror profile.

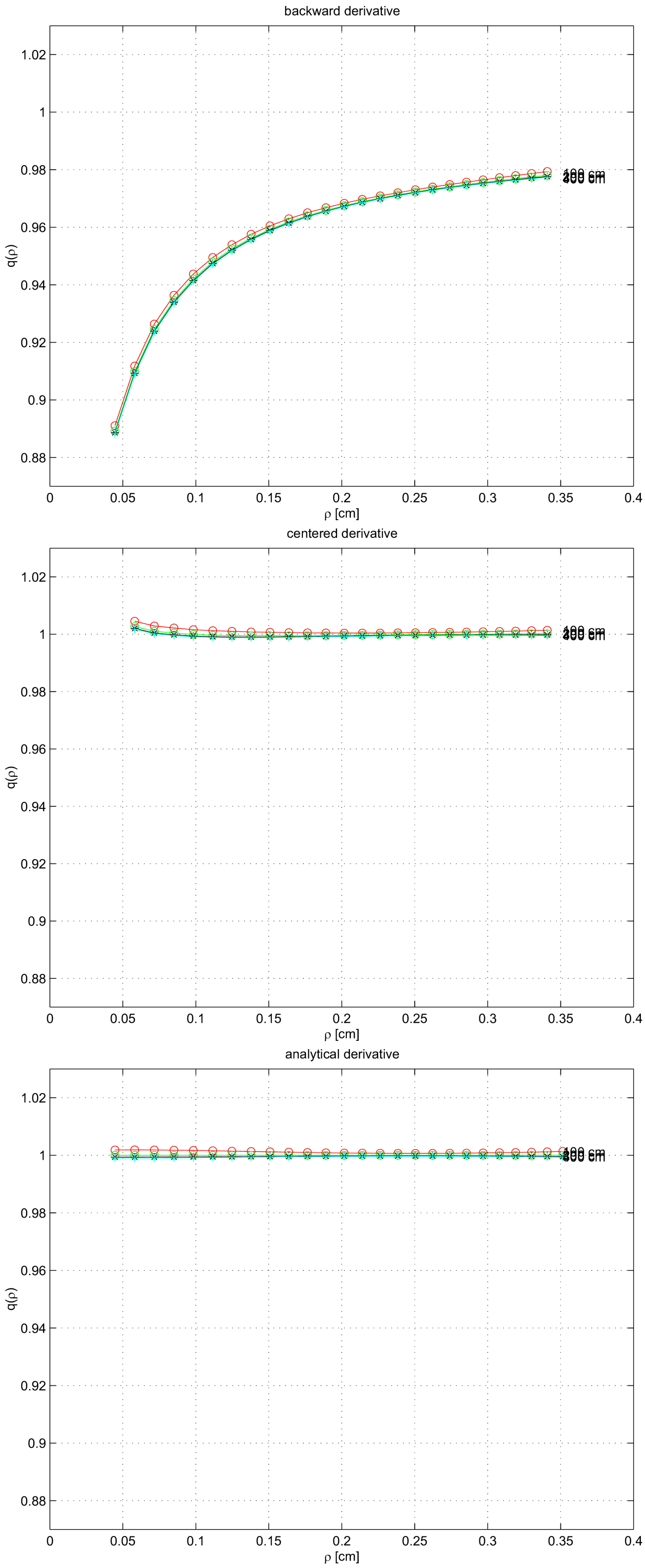


Figure 17: Analysis of the design criterion for different distances with a log-polar image sensor using different numeric approximations to the derivative: backward (left) and centered (middle) differences versus the usage of the analytic equation for F'(t) (rightmost).

6 Formal Deliverables

Reference to TA	Components	Type	Where
DI-1: Project Presentation	http://cmp.felk.cvut.cz/cmp/omniviews	S	Web
DI-2 Mirror Implementation	Mirror	H	Demo
	Panoramic Imaging with SVAVISCA camera: simulation and reality	R	Annex 1
	Mirror Design for an Omnidirectional Camera with a Uniform Cylindrical Projection when Using the SVAVISCA Sensor	R	Annex 2
	Real Uniform resolution of SVAVISCA sensor: experimental validation	R	Annex 3
	OMNIVIEWS Mirror Design Software Tools	R	Annex 4
DI-3 Software and Hardware Requirements Definition	Matlab code for mirror simulation and Design	S	CDROM
	Simulating Svavisca panoramic images of Agam Fiducials	R	Annex 5
	Localization Using SVAVISCA Panoramic Images of Agam Fiducials - Limits of Performance	R	Annex 6
	Motion Detection as a Application for the Omnidirectional camera	R	Annex 7
	Vision Algorithms for OMNIVIEWS camera	R	Annex 8
DI-4 Omnidirectional Camera	Hardware prototype	H	Demo
	Matlab code for image acquisition and preprocessing	S	CDROM
	C++ code for image acquisition and preprocessing (windows and Linux)	S	CDROM
DI-5 Final Report	This document plus Annexes	R	This document

6.1 Software

Software tools for MATLAB have been realized to simulate omnidirectional images and to design mirrors for both standard and log-polar cameras:

1) *svsim.m*: to simulate a SVAVISCA sensor used with "standard" mirrors to produce SVAVISCA-like images (available through the web)
2) *mirror uniform pix density.m:* To design mirrors for uniform-resolution cameras – Listing included in DI-2
3) *mirror non uniform pix density.m*: To design mirrors for SVAVISCA camera – Listing included in DI-2
4) Matlab routine to acquire images in real-time from the OMNIVIEWS set-up

The more general software tool developed in MATLAB and briefly described in section 4.2 will be included in a CD-ROM.

6.2 Publications[3]

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*

[3] References marked with (*) can be accessed through the WEB from the OMNIVIEWS web page: http://cmp.felk.cvut.cz/cmp/omniviews/

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.(poster)*

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, B.Micusik. Mirror Design for an Omnidirectional Camera with a Space Variant Imager. *Omnidirectional Vision Applied to Robotic Orientation and Nondestructive Testing Workshop. Budapest, 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*. To appear.

10) (*) H. Bakstein and T. Pajdla. 3D Reconstruction from 360 x 360 mosaics. *Proceedings of the CVPR'01 conference, Hawaii, USA. December 2001.* To appear.

7 Future Outlook

The current OMNIVIEWS consortium is convinced that the assessment phase has been successfully completed and that the results obtained justify the submission of a follow-up project. At this stage the detailed project structure is not ready but a general overview of the intended goals is given below.

7.1 Objectives

The main objective of a follow-up project will be to design, realize and demonstrate a miniature omnidirectional camera based on a log-polar sensor and a matching optical components (mirror plus lens). The design of such components will be possible on the basis of the experience acquired so far and also by using the software tools specifically developed during the assessment phase.

The size reduction will be accompanied by an increase in performance obtained, mainly but not only, through an increase in the overall resolution of the omnidirectional camera. Work in this aspect, along with the related optical issues, will be carried out by using forefront microelectronics and micro optical technology. We do not intend to use “established” solutions (such as, for example, the ones used for the current sensor and mirror) but we are interested in testing solutions that, even if feasible nowadays, are still considered risky.

Another objective of the second phase is to realize a more industrial prototype considering not only the technological aspects but also the assembly and manufacturing issues related to the production of a cost-effective device. We intend to do that in all three technological areas involved, namely: electronics, optics and software. The consortium will be extended to include both industrial and academic partners providing specific expertise.

Furthermore we intend to focus on a very limited number of application areas on the basis of technological criteria as well as potential market size and benefit to society.

7.2 Sensor's technology and resolution

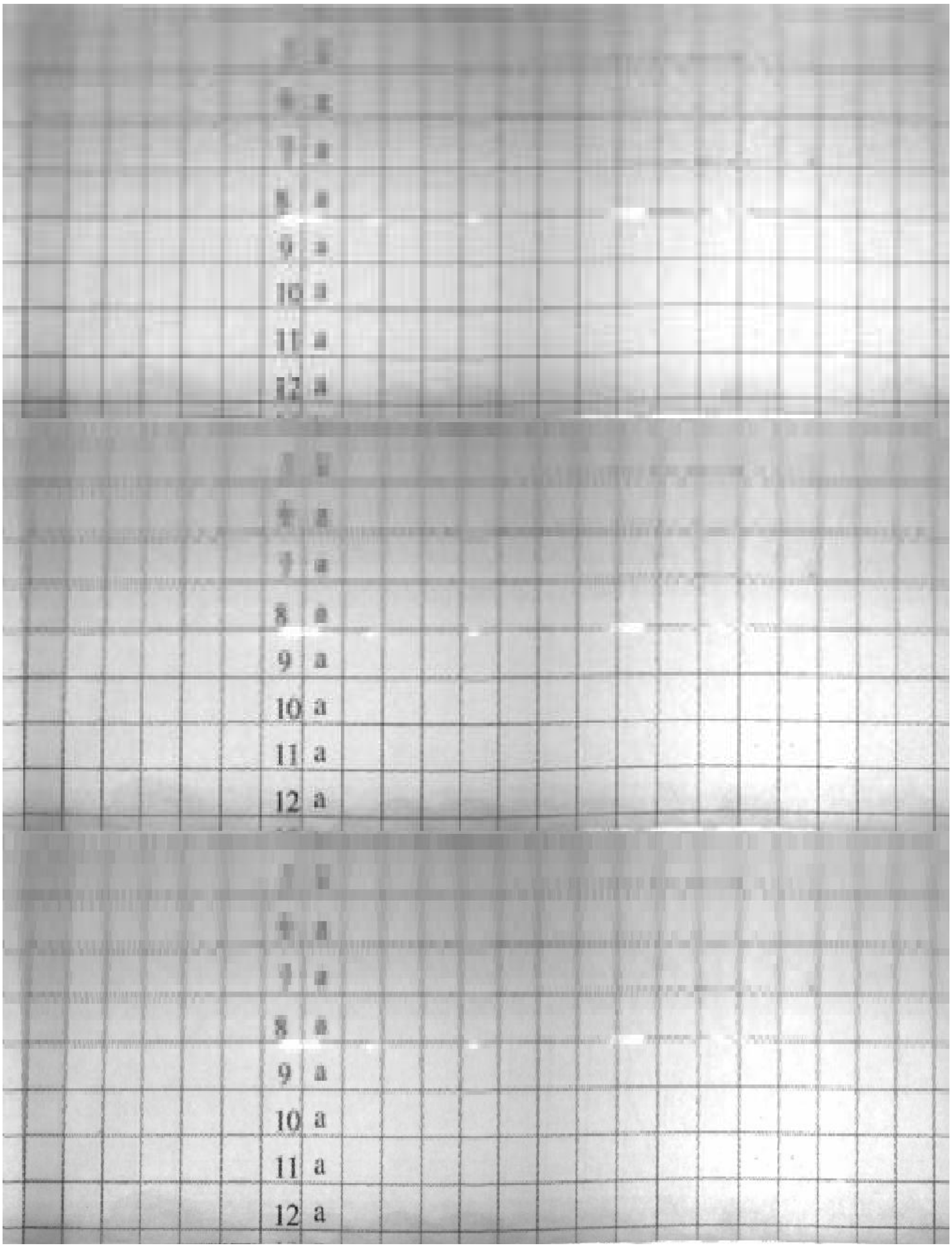
Being the key component of the project we will need to investigate if alternative solutions to the CMOS technology used for the SVAVISCA sensor is available. On the other hand, also considering the current trend we expect to continue using CMOS. The main reason being the fact that, because of the rapid increase in market penetration of CMOS-based visual sensor, we can take advantage of the corresponding improvement in the technology resulting, among other things, in smaller pixels and more efficient use of the silicon surface. The main parameter to be considered here is the pitch of the CMOS fabrication. Let's consider, as a reference, our past experience with CMOS log-polar sensor. The first ever realized CMOS log-polar sensor was based on a technology with a pitch of 0.7 µm. This resulted in a sensor of about 8 mm in diameter with approximately 8,000 pixels. The SVAVISCA sensor used in the assessment phase was realized with a CMOS 0.35 µm technology allowing, with a slightly smaller sensor 7.2 mm, a fourfold increase in number of pixels (about 33,000). This improvement was obtained by "simply" exploiting the forefront technology available at that time which allowed, moreover, an improvement in the overall geometry and the addition of "on-chip" processing. Since then the technology has constantly improved and there are foundries in Europe offering through Multi project Wafer services technologies based on 0.18 µm and 0.13 µm. In order to test, qualitatively, how the adoption of this technology will affect image quality, sensors with an increased number of pixels were simulated. In particular we made the hypothesis of adopting either a 0.25 µm or a 0.18 µm technology and maintaining the overall size of the sensor.. The expected minimum pixel's size is show in Table 2 along with the data of the current SVAVISCA array. Also included, for reasons that will be explained below, is the "equivalent constant resolution array" parameter. This is the size of the constant resolution array that would be required to obtain log-polar images identical to those obtained directly from a log-polar sensor. This parameter (as explained in section 5.1.1) is obtained by dividing the sensor's size (constant in our case and about 7.2 mm) by the size of the smallest pixels in the log-polar array.

CMOS Technology	Minimum pixel size (µm)	Equivalent constant resolution array
0.35 µm (current)	6.8 µm	1060x1060
0.25 µm	5 µm	1440x1440
0.18 µm	3.6 µm	2000x2000

Table 2: Minimum pixel's size and total number of pixels for the CMOS technology used for the SVAVISCA sensor (0.35 µm) and for CMOS technology now feasible.

It worth point out that the pixels arrangement values are still to be fully verified (this will be done in the final version of the proposal) However, considering that with 0.18 µm technology the pixel size becomes about half the pixel's size of the current technology we expect to be able to design a sensor with approximately 220x500 pixels (about 110,000 pixels). This sensor, compared with an equivalent array of 2000x2000, will provide an even bigger advantage in terms of computational savings (see section 5.1.1). In particular from about 110,000 readout operations required by OMNIVIEWS technology to 3 Million readouts plus 3 Million additions and 110,000 divisions required for a "conventional" camera with 2000x2000 pixels.

Images showing a comparison between the current resolution and the one obtainable with updated CMOS technology are shown in Figure 18 and Figure 19. Images shown in Figure 18b and Figure 19b, simulating log-polar sensors of size 220x504, were obtained by acquiring raw images with constant resolution camera (OSCAR) of resolution 768x576. Images shown in Figure 18c and Figure 19c, simulating log-polar sensors of size 440x1008, were obtained by acquiring raw images with constant resolution camera (PULNIX) of resolution 1000x1000. It is worth noting that these simulated images do not show, in the uppermost part, the actual resolution of a "real" log-polar sensor (as it will be explained below) but are shown here to demonstrate the improved quality obtained in the lower part corresponding to the peripheral part of the log-polar sensor.



A

B

C

Figure 18: Real and simulated OMNIVIEWS images from the set-up shown in Figure 14. A) Image acquired with the current prototype sensor composed of 252x110 pixels; B) Image obtained with the current OMNIVIEWS mirror but with a simulated log-polar sensor composed of 504x220 pixels; C) same ad B) with a simulated sensor composed of 1008x440 pixels.

A

B

C

Figure 19: Real and simulated OMNIVIEWS images from an office scene. A) Image acquired with the current prototype sensor composed of 252x110 pixels; B) Image obtained with the current OMNIVIEWS mirror but with a simulated log-polar sensor composed of 504x220 pixels; C) same ad B) with a simulated sensor composed of 1008x440 pixels.

In order to understand why the simulations shown do not fully present the quality of the images it is necessary to consider how the simulated images are obtained as explained in section 5.1.1 and the "equivalent constant resolution array" values reported in Table 2. In practice, as we are bounded by the array of the constant resolution camera used in the simulation (maximum is 1000x1000 for the PULNIX camera), not even the SVAVISCA sensor can be truthfully simulated. The practical result of this is that resolution of in the upper part of the images shown in Figure 18b-c and Figure 19b-c is lower than it would be for a real sensor (in practice the same pixel in the constant resolution array is mapped to many pixels in the log-polar one). This undersampling is plotted in Figure 20. The main point of the graph is to show that with the constant resolution array of the PULNIX camera part of the simulated images are undersampled not only in the fovea (corresponding to the flat part of the three curves, but also in the peri-foveal region (which corresponds to the uppermost part of the simulated panoramic images). In the case of the upper curve (1008x508) the effective size of the simulated pixel is 5 times bigger than the size of a "real" pixel. In other words, the resolution in the upper part of a real image would be 5 times better than what is shown by the simulations of Figure 18c and Figure 19c.

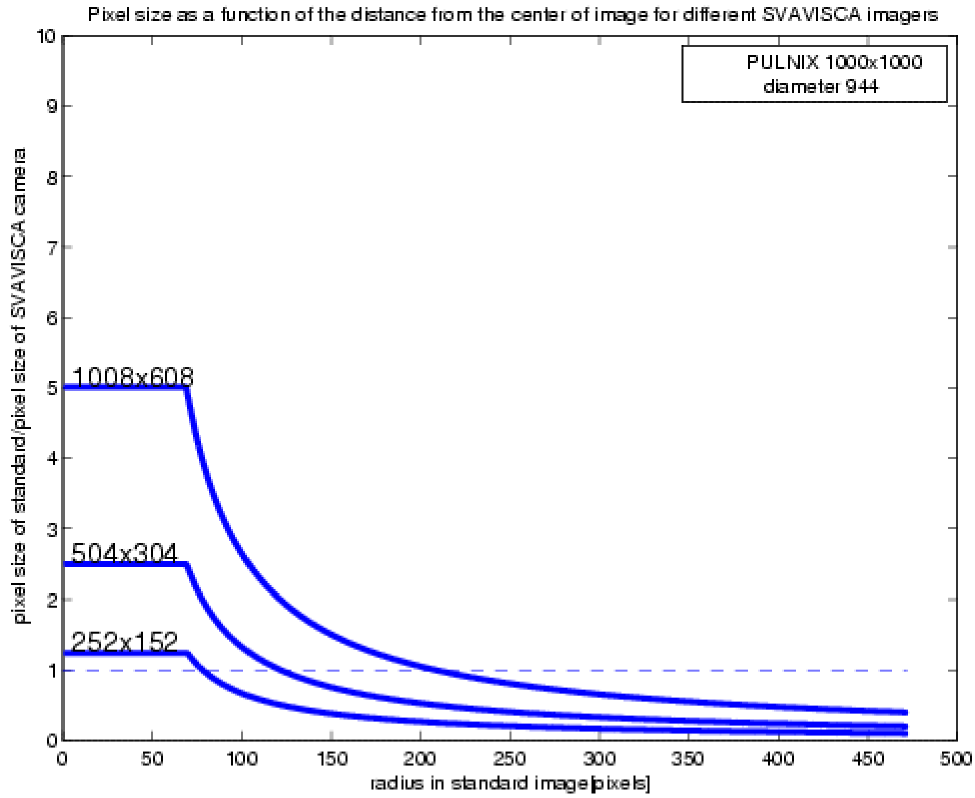


Figure 20: Plot of the ratio between the size of the pixel of a constant resolution camera and a log-polar camera. The curve "252x152" refers to the SVAVISCA sensor while the other two curves refer to the simulated sensors described in the text.

It is also worth noting that, even with the currently available sensor (252x152) the resolution, for the same sensor's size, is better in the fovea than what is possible with a 1000x1000 pixels constant resolution array.

7.3 Mirror and optical technology

As far as the mirror and optics is concerned, in the follow up project we intend to proceed along different lines.

From the assessment phase is has become evident that for some applications is would be advantageous to design "mixed" mirror. By this we mean mirror that reflects with different laws the foveal and the peripheral part of the log-polar image. For example, for navigation applications, the centermost part of the mirror could be designed to produce a "bird's eye" view of the scene (to provide the controller with "short-range" image of the ground plane as seen from above to identify obstacles in the vicinity of the robot), while the peripheral part could be designed in a way similar to the current prototype (to provide the controller with a "long-range" image of the environment surrounding the vehicle, such as the presence and direction of corridors, doors etc.). An example of the geometry of such mirror is shown in Figure 21. The field of view for each part of the sensor is determined by the corresponding parameters a and b which determine the vertical/horizontal segments that must be mapped onto the image.

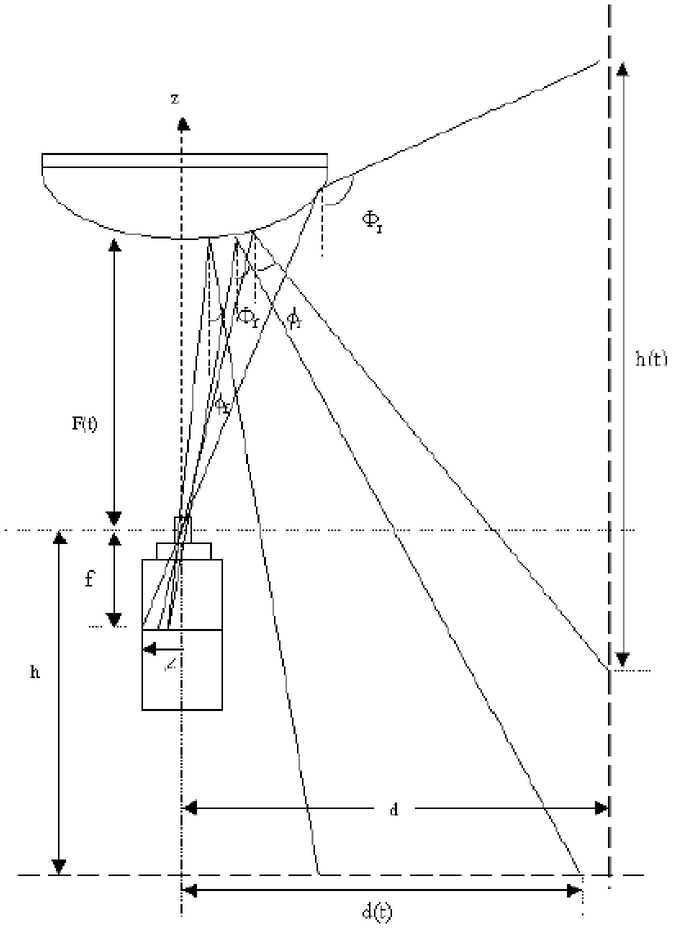


Figure 21: Geometry of a “mixed mirror” (see text).

As it is explained in Annex 4, using the geometry depicted in Figure 21, we can relate the desired fields of view specified by a user and the corresponding minimum and maximum distances observed on the pavement (to be mapped onto the fovea) and vertical directions (to be mapped by the retina).

The second direction of the project will be aimed at testing different technologies for mirror fabrication. For example the “aluminum on glass” technology described in section 5.1.4.

Finally we intend to re-analyze the overall optical arrangement (mirror plus lens) to see what is the best technological solution to reduce the overall size of the optical components. This will involve the participation of a partner specializing on optical components. For example we will investigate the possibility to integrate the mirror and lens into one catadioptric component.

7.4 Application areas

In general we believe the two application areas investigated during the first phase (surveillance and robot navigation) are all very promising including their extension to remote image transmission and teleoperation. The possibility of addressing, even in a possible follow-up project, a set of different application areas derives also from the generic mirror design tools developed up to the possibility of designing mixed-mirrors for specific applications. An application area which is slightly different from the one addressed so far but that does not require substantial modification of the design principle is related to the inspection of cylindrical pipes through the use of panoramic endoscopes. In these particular applications the aim is that of inspecting the internal surfaces of “pipes” by inserting an optical device in the chamber of the pipes. In this applications the interest for panoramic images of the kind analyzed in OMNIVIEWS is obvious in that the surface to inspect is cylindrical and totally surrounds the probe. The central part of the image is not very important in this case.
More specifically two possibilities are foreseen: 1) inspection of sewage systems; 2) inspection human’s body cavities. We address here the second case because, for obvious reasons, it is the most demanding in terms

of size, quality and social impact. Until very recently endoscopes for human use were based on flexible probes composed of optical fibers transferring the images to a TV camera at the end of the probe (as well as the illumination to the tip of the probe). The advancement of the probe was manually controlled with the help of the intrinsic compliance of the optical fiber pipe. More recently a different method has been proposed based on a miniaturized camera contained inside a container so small that can be actually swallowed by the patient. Images are transmitted through the body to an external recorder. Of course the major problem here is to reduce the size as much as possible but still obtaining the temporal and spatial resolution required to the physician. In these devices the image of the internal surface is obtained by using very short focal length lens that introduce appreciable distortions in the image. Moreover, particularly for trans-body transmission, the bandwidth of the channel has to be reduced as much as possible. The alternative solution based on an on-board recorder of images is even more demanding in terms of reducing the amount of information acquired. Finally there are obvious energy consumption problems due to the need of reducing battery size as much as possible. In conclusion then we see this specific application as very promising because the OMNIVIEWS approach not only is ideal for the kind of images required (omnidirectional images normalized for cylindrical surfaces) but also the coupling with a log-polar sensor helps in reducing enormously the amount of information acquired, processed and transmitted.