Vis Comput
DOI 10.1007/s00371-008-0278-0

ORIGINAL ARTICLE

A dynamically reconfigurable stereoscopic/panoramic vision mobile robot head controlled from a virtual environment

Walter A. Aprile · Emanuele Ruffaldi · Edoardo Sotgiu · Antonio Frisoli · Massimo Bergamasco

Abstract We have built a mobile robotic platform that features an Active Robotic Head (ARH) with two high-resolution cameras that can be switched during robot operation between two configurations that produce respectively panoramic and stereoscopic images. Image disparity is used for improving the quality of the texture. The robot head switches dynamically, based on robot operation between the stereoscopic configuration and the panoramic configuration.

Keywords Panoramic computer vision · Stereoscopic computer vision · Robot head · Mobile robot · Telerobotics · Spherical mirror

1 Introduction

Outside of the field of autonomous robotics, omnidirectional imaging systems [1] are employed as the input for surveillance systems [2]. They also see employment, mostly after rectilinear reprojection, as visual input for the operator of teleoperated robots [4]. At the same time, an omnidirectional imaging system, be it catadioptric or multicamera, usually has limitations in resolution and the quality of the image.

We achieve hardware economy and increased simplicity of design by using our Active Robot Head in two different configurations: stereoscopic and omnidirectional. We experiment with automated switching from panoramic to stereoscopic mode to adapt to the different robot operator tasks, such as navigation, orientation, and inspection.

The ARH is mounted on a commercial differentially driven robot platform, and it has been designed to simulate human head and eye movements. For testing purposes, the Robot is tracked by a commercial Vicon MX optical motion capture system.

The system is being currently tested as a teleoperated system, but in the future we will be testing algorithms for autonomous navigation that make use of the availability of binocular images.

2 System architecture

In our current system, whose high level architecture is depicted in Fig. 1, the Active Robotic Head is mounted on a commercial robotic platform.

The platform hosts the drivers of the head motors and a small size factor Control PC that manages the image acquisition and network connectivity.

The platform is connected through an umbilical cord to the Master Control Unit.

2.1 The master control unit

The Master Control Unit handles user presentation and interface. Robot position and orientation data is currently ac-

W.A. Aprile (✉) · E. Ruffaldi · E. Sotgiu · A. Frisoli · M. Bergamasco
Scuola Superiore Sant'Anna, Pisa, Italy
e-mail: walter.aprile@sssup.it
url: http://www.percro.org/

E. Ruffaldi
e-mail: emanuele.ruffaldi@sssup.it

E. Sotgiu
e-mail: edoardo.sotgiu@sssup.it

A. Frisoli
e-mail: antony@sssup.it

M. Bergamasco
e-mail: m.bergamasco@sssup.it

W.A. Aprile
Delft University of Technology, Delft, The Netherlands

quired through a VICON MX System. The unit acquires image data from the Control PC and presents to the user a control environment, generated with the XVR ([3] and [7]) virtual reality toolkit.

Informal testing has shown that the platform is currently usable for simple navigation and inspection tasks.

2.2 The Vicon MX system

The Vicon MX is a commercial marker-based digital motion capture system: it is composed of a set of custom infrared CMOS camera with annular illuminators, capture hardware and driver software. The system is calibrated with a rigid frame. Tracking improves if the tracking software is provided with a rigid body model of the tracked object that, in our case, is the whole mobile robot top surface.

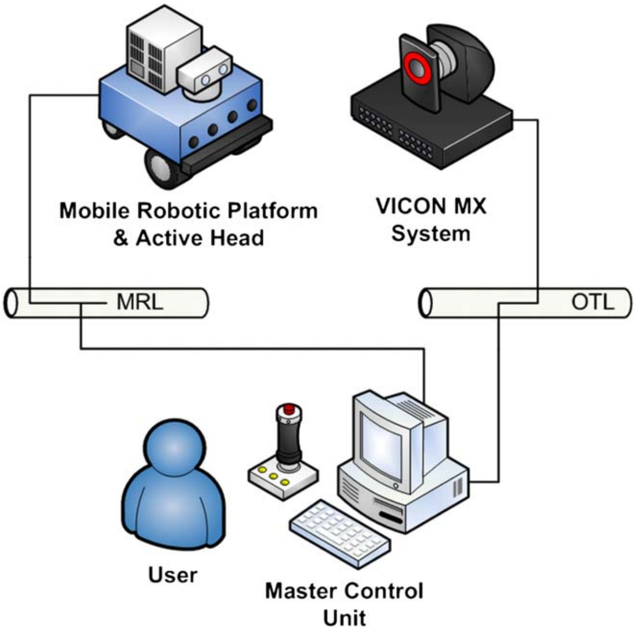


Fig. 1 General system architecture

2.3 The active robotic head

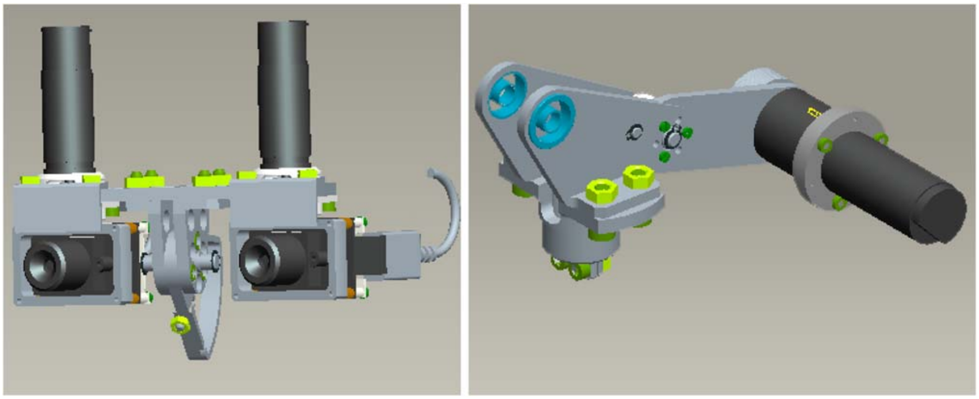
The Active Robotic Head (ARH, Fig. 2) has been designed in order to simulate the pan and tilt movements of the human head and eyes (PointGrey Firefly cameras).

It is composed of two subsystems:

1. *Vision module*: a common-elevation (Helmholtz) configuration (Fig. 3) in which both cameras tilt up and down about a common elevation axis and verge independently about axes perpendicular to the elevation plane. The first degree of freedom (DoF) is driven by a tendon transmis-

Fig. 2 The active robotic head

Fig. 3 Vision module and pan DOF



sion, while the second one is driven directly by a DC motor.

2. *Base module*: a 2 DoF pan/tilt serial mechanism, that extends the Vision Module field of view in both directions. The rotations of the Base Module are directly driven for the pan DoF (Fig. 3) and tendon driven for the tilt DoF.

The motors (model 1727-024 by Faulhaber) of both tilt movements are coupled with Harmonic Drive gear characterized by a maximum continuous torque of 300 N mm. The rest of joints are actuated by DC motors (model 1524-024 by Faulhaber) coupled with Faulhaber gears, model 15/1, that guarantee maximum continuous torques of 100 N mm. Obtained speeds and angular ranges of motion are summarized in Table 1.

Table 1 Axis performance of the ARH

	DoF	Angular range (°)	Max. speed (°/s)
Head	Tilt	±30°	80
	Pan	±100°	400
Eyes	Tilt	±30°	200
	Verge	±45°	400

The implemented law of control is characterized by a decentralized structure of 5 independent SISO subsystems. Each of them is based on the error between the desired and the actual angular position of each joint. The nonlinear effects (inertia and friction) due to the coupling of the joints for each configuration of the robot are treated as external noise that is rejected by the high gain of the control.

3 The panoramic vision system

The head is augmented by a spherical mirror, rigidly affixed to the head frame at the edge of the eyes field of view, in order to intrude as little as possible on stereoscopic operation. The head can rapidly switch between two configurations.

Fig. 4 Stereoscopic and panoramic configurations of the head with spherical mirror

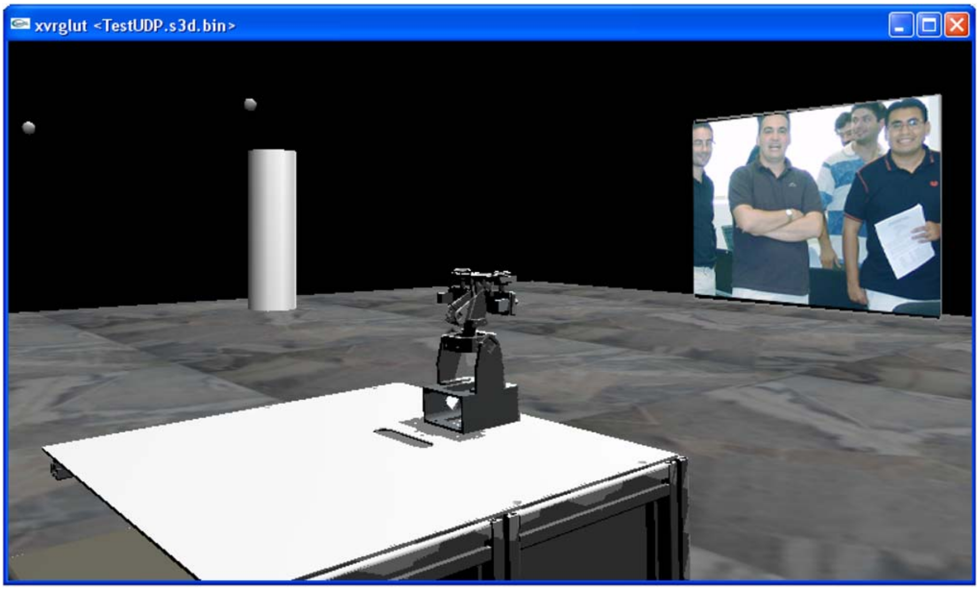


Fig. 5 The operator's virtual environment. The *gray spheres* represent the Vicon cameras

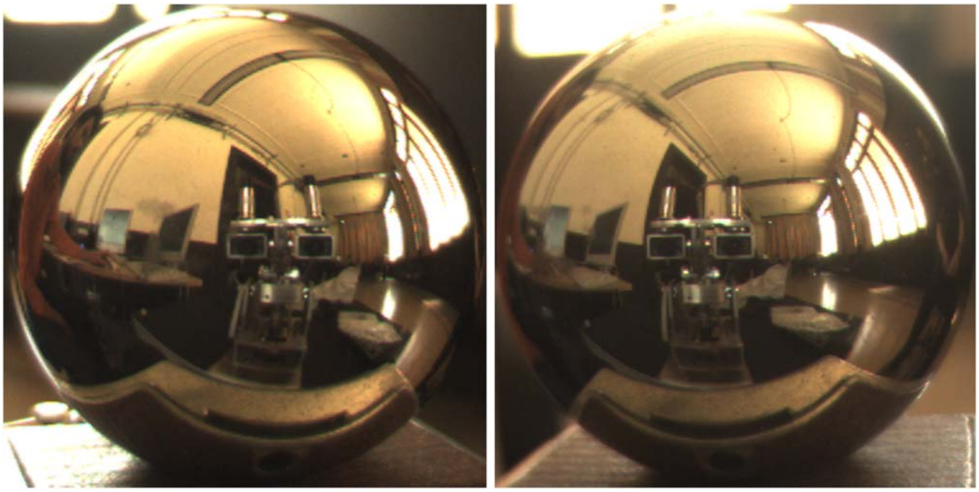


Fig. 6 The two mirrorball images from the left and right cameras in the panoramic configuration

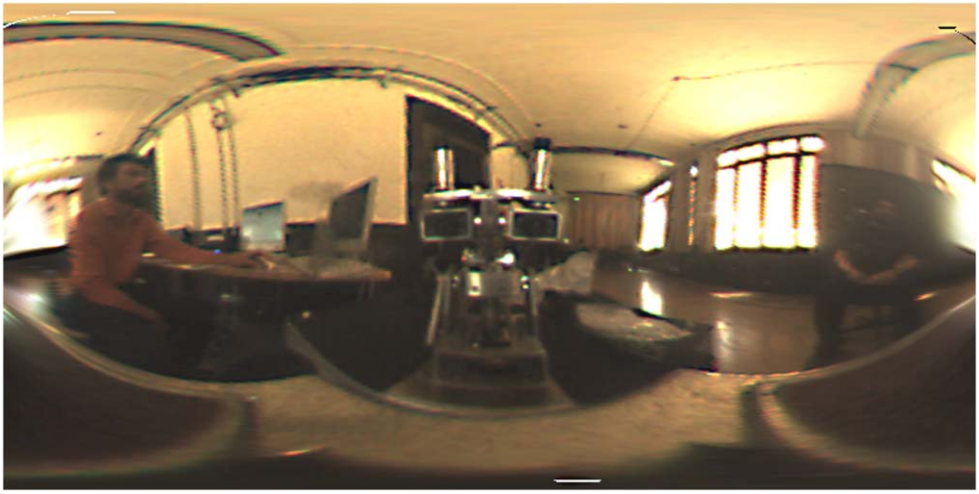


Fig. 7 The remapped and blended equirectangular image, ready for texture mapping

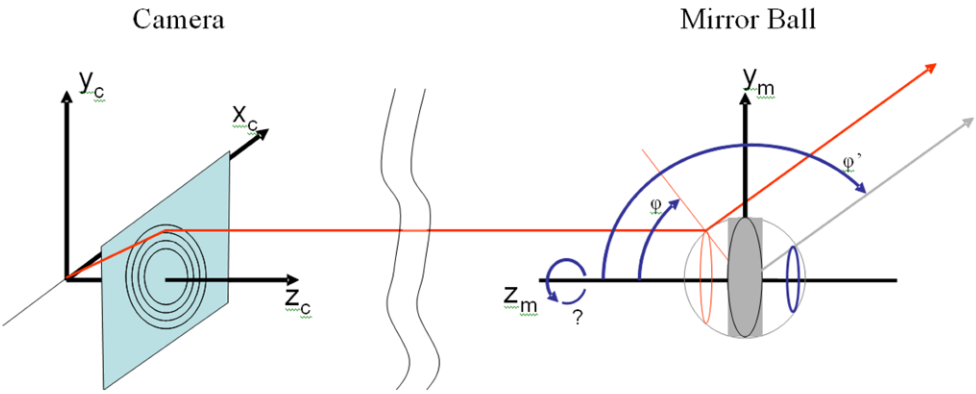


Fig. 8 Remapping mirrorball images to equirectangular images

In the stereoscopic configuration, the two cameras have their axes approximately parallel. In the panoramic configuration, the base module points front and downwards while the cameras rotate on their individual verge axes to converge on the spherical mirror and image a large part of the visible sphere around the robot (minus the robot body and ARH occlusion).

3.1 Head control

The stereo vision head and the mobile platform are controlled by the user from a virtual environment developed in the XVR platform that represents the real environment where the robot model is displayed. The two images coming from the robotic head are placed in the virtual environment relative to the mobile robot model. The robot is controlled via an analog 4-axis joystick. Joystick button and axis combinations allow control of the robot movement, of the various degrees of freedom of the head and of the stereoscopic/panoramic mode.

During operation a model of the robot (Fig. 5) is displayed in the operator's virtual environment.

The tracking cameras positions as supplied by the tracking system are displayed for orientation purposes. The views of the cameras appear in front of the robot.

During panoramic operation, the virtual environment point of view (POV) coincides with the POV of the robot head.

Since the two spherical mirror images (Fig. 6) have been taken from different points, the holes in the field of view (inherent to catadioptric systems) sited on the camera–mirror axes and opposite to the camera fall in two different points of the visible sphere (see [5] for a detailed discussion).

The difference in the location of the holes allows real-time patching with reasonable quality (Fig. 7), since all the parameters of the vision system (intrinsic and extrinsic) are statically determined at setup time. OpenGL texture warping operation is used to remap the spherical image to an equirectangular projection that is then texture-mapped on an environmental sphere that surrounds the virtual environment POV. This view is used for navigation, and it allows software simulation of a PTZ camera without cinematic constraints. Such an environment can be used (as in [6]) as the basis for an uncluttered and usable robot interface. Despite the fact that a spherical mirror violates the single point of view constraint (defined in [1]), in practice the image is not disrupted.

3.2 Shader-based image warping

Images from the cameras are in the so-called mirrorball projection, but texture mapping requires images in equirectangular projections. A custom OpenGL shader program remaps the spherical mirror view to an equirectangular projection.

We have employed a raytracing algorithm to derive an image map: we have sent approximately 10,000 rays from the unit ray circle centered on the origin of the camera image plane to the mirror ball, and computed the θ and φ polar coordinate angles (Fig. 8). Pixels not hit by the cast rays were interpolated from their neighbors. The mapping is based on the approximation that the camera is orthographic and that the sphere is of negligible size. The resolution and overall quality of the resulting images are limited by the surface irregularities of the sphere, a 50 mm stainless steel ball bearing, but it is sufficient for navigation.

Acknowledgements This work is being carried out in the context of EU Project "Decision In Motion" (FP6-IST), with the objective of building a testing platform for neuromorphic vision algorithms.

References

1. Baker, S., Nayar, S.K.: A theory of catadioptric image formation. In: Proceedings of the 6th International Conference on Computer Vision, Bombay, India. IEEE Computer Society, Washington (1998)
2. Boult, T.E.: Remote reality via omnidirectional imaging. In: DARPA Image Understanding Workshop (IUW), pp. 1049–1052 (1998)
3. Carrozzino, M., Tecchia, F., Bacinelli, S., Bergamasco, M.: Lowering the development time of multimodal interactive application: the real-life experience of the XVR project. In: Proceedings of ACE 2005, Valencia, Spain (2005)
4. Fiala, M.: Pano-presence for teleoperation. In: International Conference Intelligent Robots and Systems, IEEE/RSJ (2005)
5. Geyer, C., Daniilidis, K.: Catadioptric projective geometry. Int. J. Comput. Vis. **43**, 223–243 (2001)
6. Ricks, B.W., Nielsen, C.W., Goodrich, M.A.: Ecological displays for robot interaction: a new perspective. In: International Conference on Intelligent Robots and Systems IEEE/RSJ, Sendai, Japan (2004)
7. VRMedia: Extreme VR (2007). http://www.vrmedia.it

Walter A. Aprile holds a degree in Computer Science from the Università degli Studi di Milano and is currently completing a Ph.D. in the "Virtual Reality and Telerobotic Systems" (VRTS) research division of the PERCRO (PERCeptual Robotics) laboratory of Scuola Superiore Sant'Anna in Pisa. He has worked at Carnegie Mellon University, Instituto Latinoamericano de la Comunicaciòn Educativa, Interaction Design Institute Ivrea and Interaction Design Lab. He is interested in all aspects of panoramic images, with particular reference to their perceptual effects and he works as Assistant Professor at Delft University of Technology.

Emanuele Ruffaldi is Assistant Professor at Scuola Superiore Sant'Anna, Pisa in Applied Mechanics in which he obtained his Ph.D. in 2006 discussing a thesis on perceptually inspired Haptic Algorithms. He is member of the research division VRTS "Virtual Reality & Telerobotic Systems" at PERCRO (PERCeptual ROBotics) Laboratory of Scuola Superiore Sant'Anna, directed by prof. Massimo Bergamasco. The main project in which he is currently involved is the SKILLS IP, aimed to capture and transfer human skills. His research interests are in the field of haptics, human gesture analysis and system integration for virtual reality applications.

Edoardo Sotgiu graduated in Mechanical Engineering at the University of Pisa. He currently works as a collaborator in the Virtual Reality & Telerobotic Systems Division of Percro Laboratory at Scuola Superiore Sant'Anna. He is currently involved with the Body Extender powered exoskeleton project, the EchoDev ecography probe positioning 5DoF robotic system and the PEC-BPPE robotic cell for fatigue testing of a helicopter door. His main competencies concern the following fields: statics and dynamics parameters identification for robotic systems, synthesis of position and force control on haptic interfaces and robotic systems, development of bilateral teleoperator control architectures.

Antonio Frisoli is assistant professor at Scuola Superiore Sant'Anna, faculty of engineering, and head of the "Virtual Reality and Telerobotic Systems" (VRTS) research division of the PERCRO laboratory. In 2002 he received his Ph.D. in Robotics "summa cum laude" in the area of "Teleoperation in Virtual Environments" at Scuola Superiore Sant'Anna. He currently teaches "Mechanics of Robots" at the Faculty of Mechanical Engineering of the University of Pisa. He is involved in several European and national research projects, such as the Enactive NoE, DECISION IN MOTION and PRESENCCIA. His research interests range in the field of Virtual Reality, Haptic Interfaces and Haptic Rendering, Human Computer Interaction, Theoretical Kinematics, Robots Modeling and Control of Mechanical Systems. At PERCRO laboratory he is working on the control of a robotic full body extender for human performance augmentation.

Massimo Bergamasco is Full Professor of Mechanics of Robots and Virtual Environment at Scuola Superiore Sant'Anna. He is currently the Coordinator of the PERCRO laboratory. His research deals with the study and development of haptic interfaces for the control of the interaction between humans, Interaction Design and Virtual Environments. His research activities include both theoretical research as well as applied research into different fields. Among these: Medicine, Surgery and Rehabilitation; Simulators; Art and Cultural Heritage; Education; Engineering.