INTRODUCTION

Many of the known visual systems in nature are characterized by a wide field of view allowing animals to keep the whole surrounding environment under control. In this sense, dragonflies are one of the best examples: their compound eyes are made up of thousands of separate light-sensing organs arranged to give nearly a 360° field of vision. However, animals with eyes on the sides of their head have high periscopy but low binocularity, that is their views overlap very little. Differently, raptors’ eyes have a central part that permits them to see far away details with an impressive resolution and their views overlap by about ninety degrees. Those characteristics allow for a globally wide field of view and for accurate stereoscopic vision at the same time, which in turn allows for determination of distance, leading to the ability to develop a sharp, three-dimensional image of a large portion of their view.

In mobile robotics applications, autonomous robots are required to react to visual stimuli that may come from any direction at any moment of their activity. In surveillance applications, the opportunity to obtain a field of view as wide as possible is also a critical requirement. For these reasons, a growing interest in omnidirectional vision systems (Benosman 2001), which is still a particularly intriguing research field, has emerged. On the other hand, requirements to be able to carry out object/pattern recognition and classification tasks are opposite, high resolution and accuracy and low distortion being possibly the most important ones. Finally, three-dimensional information extraction can be usually achieved by vision systems that combine the use of at least two sensors at the same time.

BACKGROUND

There are essentially two ways to observe a very wide area. It is possible to use many cameras pointed on non-overlapping areas or, conversely, a single camera with a wide field of view. In the former case, the amount of data to be analyzed is much bigger than in the latter one. In addition, calibration and synchronization problems for the camera network have to be faced. On the other hand, in the second approach the system is cheaper, easy to calibrate, while the analysis of a single image is straightforward. In this case, however, the disadvantage is a loss of resolution at which objects details are seen, since a wider field of view is projected onto the same area of the video sensor and thus described with the same amount of pixel as for a normal one. This was clear since the mid 1990s with the earlier experiments with omnidirectional vision systems. Consequently a number of studies on omnidirectional sensors “enriched” with at least one second source of environmental data arose to achieve wide fields of view without loss of resolution. For example some work, oriented to robotics applications, has dealt with a catadioptric camera working in conjunction with a laser scanner as, to cite
Hybrid Dual Camera Vision Systems

only few recent, in (Kobilarov 2006; Mei 2006). More surveillance application-oriented work has involved multi-camera systems, joining omnidirectional and traditional cameras, while other work dealt with geometric aspects of hybrid stereo/multi-view relations, as in (Sturm 2002; Chen 2003).

The natural choice to develop a cheap vision system with both omni-sight and high-detail resolution is to couple an omnidirectional camera with a moving traditional camera. In the sequel, we will focus on this kind of systems that are usually called “hybrid dual camera systems”.

Omnidirectional Vision

There are two ways to obtain omnidirectional images. With a special kind of lenses mounted on a standard camera, called “fisheye lenses”, it is possible to obtain a field of view up to about 180-degrees in both directions. The widest fisheye lens ever produced featured a 220-degrees field of view. Unfortunately, it is very difficult to design a fisheye lens that satisfies the single viewpoint constraint. Although images acquired by fisheye lenses may prove to be good enough for some visualization applications, the distortion compensation issue has not been solved yet, and the high unit-cost is a major drawback for its wide-spread applications.

Combining a rectilinear lens with a mirror is the other way to obtain omnidirectional views. In the so-called “catadioptric lenses” a convex mirror is placed in front of a rectilinear lens achieving a field of view possibly even larger than with a fisheye lens. Using particularly shaped mirrors precisely placed with respect to the camera is also possible to satisfy the single viewpoint constraint and thus to obtain an image which is perspectively correct. Moreover, catadioptric lenses are usually cheaper than fisheye ones. In Figure 1 a comparison between these two kinds of lenses can be seen.

OVERVIEW OF HYBRID DUAL CAMERA SYSTEMS

The first work concerning hybrid vision sensors is probably the one mentioned in (Nayar 1997) referred to as “Omnidirectional Pan/Tilt/Zoom System” where the PTZ unit was guided by inputs obtained from the omnidirectional view. The next year (Cui 1998) presented a distributed system for indoor monitoring: a peripheral camera was calibrated to estimate the distance between a target and the projection of the camera on the floor. In this way, they were able to precisely direct the foveal sensor, of known position, to the target and track it. A hybrid system for obstacle detection in robot navigation was described in (Adorni 2001) few years later. In this work, a catadioptric camera was calibrated along with a perspective one as a single sensor: its calibration procedure permitted to compute an Inverse Perspective Mapping (IPM) (Little 1991) based on a reference plane, the floor, for both images and hence, thanks to the cameras’ disparity, to detect obstacles by computing the difference between the two images. While this was possible only within the common field of view of the two cameras, awareness or even tasks such as ego-motion estimation were potentially pos-

Figure 1. Comparison between image formation in fisheye lenses (left) and catadioptric lenses (right)
sible thanks to the omni-view. This system was further improved and mainly tested in RoboCup\(^1\) applications, (Adorni 2002; Adorni 2003; Cagnoni 2007). In Figure 2 it is possible to see a pair of images acquired with such a system.

Some recent work has concentrated on using dual camera systems for surveillance applications. In (Scotti 2005), when some alarm is detected on the omnidirectional sensor, the PTZ camera is triggered and the two views start to track the target autonomously. Acquired video sequences and other metadata, like object classification information, are then used to update a distributed database to be queried later by users. Similarly in (Yao 2006), after the PTZ camera is triggered by the omnidirectional one, the target is tracked independently on the two views, but then a modified Kalman filter is used to perform data fusion: this approach achieves an improved tracking accuracy and permits to resolve occasional occlusions leading to a robust surveillance system.

**FUTURE TRENDS**

Nowadays public order keeping, private property access control and security video surveillance are reasons for which we need to surveil wide areas of our environment. Surveillance is an ever growing market and automatic surveillance is an interesting challenge: many projects are oriented in this direction and in some of them an important role is already played by hybrid dual camera systems. The monitoring system installed between Eagle Pass, Texas, and Piedras Negras, Mexico, by engineers of the Computer Vision and Robotics Laboratory at the University of California, San Diego, affiliated with the California Institute for Telecommunications and Information Technology, is an example of a very complex surveillance system in which hybrid dual camera systems are involved (Hagen 2006). Because of the competitive cost, the compactness and the opportunities offered by these systems, they are likely to be used more and more in the future in intelligent surveillance systems.

Another field subjected to great interest is autonomous vehicle navigation. Even if at the moment there are still many problems to be solved before seeing autonomous public vehicles, industrial applications are already possible. Since omnidirectional visual servoing and ego-motion estimation can actually be implemented also using hybrid dual camera systems, and many more opportunities are offered by the presence of a second high-resolution view, their future involvement in this field is desirable.

**CONCLUSIONS**

The class of hybrid dual camera systems has been described and briefly overviewed. The joint use of a standard camera and of a catadioptric sensor provides
this kind of sensors with their different and complementary features: while the traditional camera can be used to acquire detailed information about a limited region of interest (“foveal vision”), the omnidirectional sensor provides wide-range, but less detailed information about the surroundings (“peripheral vision”).

Tracking of multiple objects/people relying on high-resolution images for recognition and access control or estimating object/people velocity, dimensions and trajectory are some examples of possible automatic surveillance tasks for which hybrid dual camera systems are suitable. Furthermore, their use in (autonomous) robot navigation, allows for accurate obstacle detection, egomotion estimation and three-dimensional environment reconstruction. With one of these sensors on board, a mobile robot can be provided with all the necessary information needed to navigate safely in a dynamic environment.

REFERENCES


**KEY TERMS**

**Camera Calibration:** A procedure used to obtain geometrical information about image formation in a specific camera. After calibration, it is possible to relate metric distances on the image to distances in the real world. In any case only one image is not enough to reconstruct the third dimension and some a priori information is needed to accomplish this capability.

**Catadioptric Camera:** A camera that uses in conjunction catoptric, reflective, lenses (mirrors) and dioptric, refractive, lenses. Usually the purpose of these cameras is to achieve a wider field of view than the one obtained by classical lenses. Even if the field of view of a lens could be improved with any convex surface mirror, those of greater interest are conic, spherical, parabolic and hyperbolic-shaped ones.

**Central Catadioptric Camera:** A camera that combines lenses and mirrors to capture a wide field of view through a central projection (i.e. a single viewpoint). Most common examples use paraboloidal or hyperboloidal mirrors. In the former case a telecentric lens is needed to focalize parallel rays reflected by the mirror and there are no constraints for mirror to camera relative positioning: the internal focus of the parabola acts as the unique viewpoint; in the latter case it is possible to use a normal lens, but mirror to camera positioning is critical for achieving a single viewpoint: it is essential that the principal point of the lens coincides with the external focus of the hyperboloid to let the internal one be the unique viewpoint for the observed scene.

**Omnidirectional Camera:** A camera able to see in all directions. There are essentially two different methods to obtain a very wide field of view: the older one involves the use of a special type of lens, usually referred to as fisheye lens, while the other one uses in conjunction rectilinear lenses and mirrors. Lenses obtained in the latter case are usually called catadioptric lenses and the camera-lens ensemble is referred to as catadioptric camera.

**PTZ Camera:** A camera able to pan left and right, tilt up and down, and zoom. It is usually possible to freely control its orientation and zooming status at a distance through a computer or a dedicated control system.

**Stereo Vision:** A visual perception process that exploits two different views to achieve depth perception. The difference between the two images, usually referred to as binocular disparity, is interpreted by the brain (or by an artificial intelligent system) as depth.

**Single Viewpoint Constraint:** To obtain an image with a non-distorted metric content, it is essential that all incoming principal light rays of a lens intersect at a single point. In this case a fixed viewpoint is obtained and all the information contained in an image is seen from this point.

**ENDNOTE**