Chapter 4

Application Scenarios

The application program written for this thesis, called *OmniTracking*, uses video streams captured with an omnidirectional camera system. Its main aim is to detect moving objects and to track them as they move across the field-of-view of the camera. The omnidirectional cameras used by this program are catadioptric camera systems based on a paraboloidal mirror.

4.1 Indoor and Outdoor Environments

The program was designed and written with two applications in mind – to be used for an outdoor surveillance and object-tracking system, and for an indoor automatic meeting-based annotation application[1]. The reason for choosing these two scenarios is that they provide quite diverse conditions under which the program must work. Usually an outdoor environment:

- covers large areas and targets are relatively small in size and far away from the camera.
- lighting variations in outdoor scenes tend to be much more extreme than for indoor situations.
- outdoor environments tend to contain 'unimportant' motion that normal surveillance systems are not interested in (e.g. moving trees due to wind).
- and in general, outdoor environments are less structured.

On the other hand, in an indoor environment:

[1] For "round-table" type meetings, to be more precise.

- activity tends to take place in relatively small and enclosed areas, so targets appear much larger, closer to the camera and with a higher probability of interacting with each other (for example, there is a higher chance for the targets to occlude each other).
- lighting conditions do not vary that much.
- there is usually more control over the structure of an indoor environment and what happens inside it.

4.2 Datasets

Datasets from the PETS[2] workshops are used to test the application ([FERR03] and [FERR01]). One of the PETS2001 datasets was taken with an omnidirectional camera (dataset 4 camera 2) and shows a typical outdoor scene with different moving targets (people, cars, and even a bicycle).

frame #2450 from PETS2001 dataset 4 camera 2

frame #17105 from PETS-ICVS scenario C camera 3

Figure 4.1: Two different scenarios

The PETS-ICVS 2003 datasets are of an indoor environment with several people interacting in a meeting. Of the PETS-ICVS datasets, the one labelled 'Scenario C – Going to the white board' is of particular interest as it contains considerable subject motion. The others (example, 'Scenario D') are also used for testing the OmniTracking application. (If not specified, when a reference is made to the PETS-

[2] IEEE Workshops on *Performance Evaluation of Tracking and Surveillance (PETS).*

ICVS dataset in the rest of this document, it is to the 'Scenario C' dataset). Figure 4.1 shows a representative frame from these two datasets. The omnidirectional camera used for these datasets is a catadioptric camera with a paraboloidal mirror.

4.3 Omnidirectional Cameras – Advantages for the Program

The use of an omnidirectional camera for these two environments, offers several advantages as well as some disadvantages. These influence what the program is able to do and how it operates in general.

An omnidirectional-based surveillance system, unlike conventional cameras, does not suffer from limited fields-of-view and has no blind areas. This makes it the 'perfect' sensor for surveillance because of its omnidirectional awareness at all times and is able to track multiple objects in different parts of the scene simultaneously. Something similar can be attempted with conventional active cameras, for example by a PTZ (pan-tilt-zoom) camera[3]. But a PTZ still doesn't see the whole surroundings at the same time, is not able to track multiple objects in different parts of the scene and usually requires complex control strategies. A network of conventional cameras suffers from the *camera handoff* problem [TRIV02], when a target leaves the field-of-view of one camera and has to be re-identified by (handed over to) the other cameras.

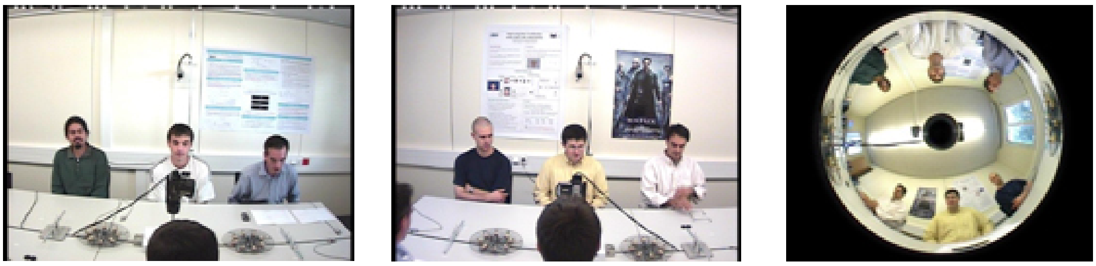
But on the other hand, an omnidirectional camera suffers from low resolution and is not suitable to do high-resolution behaviour analysis or object recognition. For most surveillance-based applications, this is not such a restriction after all. [ISHI98] says that omnidirectional cameras are best suited to act as "vision sensors" rather than as traditional "vision cameras".

Using omnidirectional cameras for meeting-based applications, offers similar advantages as for the outdoor surveillance applications. There is no need for multiple cameras or active cameras with their movements and distractions. In addition, the omnidirectional camera is normally placed at the centre of the meeting table and can acquire images showing all the faces of the participants – this mitigates a bit the low-resolution disadvantage of omnidirectional cameras as persons are nearer to the camera and it's possible to get close-ups of faces [ISHI98]. It has also been shown that

[3] These are known as "domes" in the security industry.

the added “context” provided by omnidirectional cameras is an important aspect of meeting-viewing application user-interfaces [RUI01].

For example, in the case of the PETS-ICVS datasets, in addition to the omnidirectional camera, there are two other high-resolution conventional cameras placed on opposite sides on the walls. These cameras have a limited field-of-view – while being able to show the participants when seated around the table, they cannot show other important events like when each participant goes to write on the white-board (in the ‘scenario C’ dataset). Figure 4.2 shows the output from all three cameras. Applications using the omnidirectional camera have the potential of detecting and tracking a participant from the moment he/she enters the room until they leave.



frame #11340 from PETS-ICVS dataset, scenario C cameras 1, 2 and 3 respectively.

Figure 4.2: Omnidirectional and conventional (linear) cameras

4.4 Omnidirectional Cameras – Restrictions for the Program

The use of catadioptric omnidirectional cameras places some restrictions and limitations on the algorithms implemented by the OmniTracking program. Catadioptric systems, because of their curved mirror surfaces, introduce *non-linear distortion* effects in the image they acquire. This means that special attention must be paid to what computer vision algorithms and techniques are used to process omnidirectional images. [DANI02] laments that some applications apply image operators on omnidirectional images “blindly” – that is, using the operators as if on a rectilinear (conventional) image. Many techniques in computer vision (like template-matching, optical flow) are based on assumptions of linear motion, affine deformations, rectangular windows, etc. – all these assumptions fail in the case of omnidirectional images.

One solution is to adapt or generalise the technique in question to remove or relax these assumptions. Such work has been done, for example, by [DANI02] (for optical flow) and [SVOB01] (for window-based correlation).

Another solution is to 'dewarp' the omnidirectional image to get a flat panoramic image (similar to Figure 3.14) with uniform resolution and then applying the operator on this panorama without violating its assumptions. This is usually fine for indoor environments, where one can assume that objects are constrained to moving horizontally and no two objects will be (appear) 'on top of each other' (and so occupying the same vertical columns). Examples of use of this method can be found in [STIE02; HUAN01; SOGO00], all for indoor environments. But working on 'dewarped' images is not suitable for outdoor environments where the movement of objects is less constrained. In addition, the dewarping process introduces noise (because of interpolation), which will affect later stages of processing.

For the OmniTracking application, most of the algorithms used by the program are low-level algorithms (example, motion detection) and therefore can be used on omnidirectional images without any problems. For some others, the algorithm is amended slightly to better approximate the underlying geometry of the omnidirectional image. And wherever there is no alternative but to use an existing technique directly on the omnidirectional image, even though it was designed for rectilinear images, attention was paid to try and minimise possibilities of error and inaccuracies in the result.

4.5 Objectives

To conclude, the aim of the OmniTracking program is to detect and track moving objects from both datasets (using the same set of algorithms for both the indoor and outdoor environments). The main aim is for objects to be tracked from the time when they first appear in the scene (or shortly thereafter), until they leave the scene, with object identity maintained throughout occlusion events. This in order for the omnidirectional-based system to truly offer an advantage over conventional cameras by being able to observe the objects for a longer time than possible with conventional systems.