MULTIMEDIA VIDEO-BASED SURVEILLANCE SYSTEMS
Requirements, Issues and Solutions

edited by
Gian Luca Foresti, Petri Mähönen and Carlo S. Regazzoni

Multimedia surveillance systems is an emerging field that includes signal and image processing, communications, and computer vision. *MULTIMEDIA VIDEO-BASED SURVEILLANCE SYSTEMS, Requirements, Issues and Solutions*, combines the most recent research results from these areas for use by engineers and end-users involved in the design of surveillance systems in the fields of transportation and services. The book covers emerging surveillance requirements, including new digital sensors for real-time acquisition of surveillance data, low-level image processing algorithms, and event detection methods. It also discusses problems related to knowledge representation in surveillance systems, wireless and wired multimedia networks, and a new generation of surveillance communication tools. Timely information is presented on digital watermarking, broadband multimedia transmission, legal use of surveillance systems, performance evaluation criteria, and other new and emerging topics, along with applications for transports and pedestrian monitoring. The information contained in *MULTIMEDIA VIDEO-BASED SURVEILLANCE SYSTEMS, Requirements, Issues and Solutions*, bridges the distance between present practice and research findings, and the book is an indispensable reference tool for professional engineers.

ISBN 0-7923-7927-6

9 780792 379270

SECS573
0-7923-7927-6

MULTIMEDIA VIDEO-BASED SURVEILLANCE SYSTEMS
Requirements, Issues and Solutions

Foresti, Mähönen, Regazzoni

KAP

MULTIMEDIA VIDEO-BASED SURVEILLANCE SYSTEMS

Requirements, Issues and Solutions

edited by
Gian Luca Foresti
Petri Mähönen
Carlo S. Regazzoni

Kluwer Academic Publishers

THE KLUWER INTERNATIONAL SERIES
IN ENGINEERING AND COMPUTER SCIENCE

MULTIMEDIA VIDEO-BASED SURVEILLANCE SYSTEMS

Requirements, Issues and Solutions

edited by

Gian Luca Foresti
University of Udine

Petri Mähönen
VTT, Technical Research Center of Finland and the University of Oulu

Carlo S. Regazzoni
University of Genova

Kluwer Academic Publishers
Boston/Dordrecht/London

Distributors for North, Central and South America:
Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, Massachusetts 02061 USA
Telephone (781) 871-6600
Fax (781) 871-6528
E-Mail <kluwer@wkap.com>

Distributors for all other countries:
Kluwer Academic Publishers Group
Distribution Centre
Post Office Box 322
3300 AH Dordrecht, THE NETHERLANDS
Telephone 31 78 6392 392
Fax 31 78 6546 474
E-Mail services@wkap.nl>

Electronic Services <http://www.wkap.nl>

Library of Congress Cataloging-in-Publication

Multimedia video-based surveillance systems : requirements, issues and solutions / edited by Gian Luca Foresti, Petri Mähönen, Carlo S. Regazzoni.
p. cm. -- (Kluwer international series in engineering and computer science ; SECS 573)
Includes bibliographical references and index.
ISBN 0-7923-7927-6
1. Multimedia systems. 2. Electronic security systems. 3. Closed-circuit television. I. Foresti, Gian Luca, 1965- II. Mähönen, Petri, 1963- III. Regazzoni, Carlo S. IV. Series.

QA76.575 .M85245 2000
621.389'28--dc21

00-058397

Printed on acid-free paper.

Printed in the United States of America

CONTENTS

Chapter 1.3

Requirements for Visual Perception of Automotive Environments

MASSIMO BERTOZZI
Dipartimento di Ingegneria dell'Informazione, Università di Parma, Italy - bertozzi@ce.unipr.it

ALBERTO BROGGI
Dipartimento di Informatica e Sistemistica, Università di Pavia, Italy - alberto.broggi@unipv.it

ALESSANDRA FASCIOLI
Dipartimento di Ingegneria dell'Informazione, Università di Parma, Italy – fascal@ce.unipr.it

1. INTRODUCTION

This contribution deals with vision-based surveillance systems installed on moving vehicles. Some important issues must be carefully considered in the design of a visual perception system for mobile automotive applications.

Firstly, these systems, when used for road and/or driver monitoring, require faster processing than other applications, since the vehicle speed is bounded by the processing rate, and, therefore, real-time performance constraints are to be strictly met.

Secondly, in general in the automotive field no assumptions can be made on key parameters, such as scene illumination or contrast, which are directly measured by the vision sensor. Hence, the subsequent processing must be robust enough to adapt to different environmental conditions (for example, sun, rain, fog). In addition to this, a moving vision system has to take into account dynamic and sometimes abrupt changes in the illumination or contrast (such as transitions between sun and shadow, or the entrance or exit from a tunnel). On the other hand, a static surveillance system may take advantage of a partial knowledge about the environment, a fixed background, and slowly varying conditions.

Furthermore, when the acquisition system is installed on a moving vehicle other key issues, such as the robustness to vehicle's movements and drifts in the camera's calibration, must be handled. Image stabilization is a critical problem as well.

Eventually, on-board perception systems are obviously safety critical devices, requiring a strong degree of reliability and tolerance both to hardware and software failures.

However, recent advances in both computer and sensor technologies promote the use of machine vision on-board of intelligent vehicles. The developments in computational hardware, such as a higher degree of integration and a reduction of the power supply voltage, permit to produce machines that can deliver a high computing power, with fast networking facilities at an affordable price. In addition to this, current cameras include new important features that permit the handling and solving of some basic problems directly at sensor level. The resolution of the sensors has been drastically enhanced and the camera dynamics extended, and in order to decrease the image acquisition and transfer time, new technological solutions can be found in CMOS sensors, such as the possibility of dealing with pixels independently.

2. REQUIREMENTS IN AUTOMOTIVE ENVIRONMENT

In order to be sold on the market, a perception system for the automotive environment must fit several hard requirements related both to technical and non-technical aspects. The following paragraphs sketch the most important ones.

2.1 Robustness

A commercial perception system to be installed on a standard vehicle must be robust in relation to different facets. First of all, any kind of road

perception system ranging from the simplest to the most complex have to face the difficult task of handling different environmental conditions and their evolving. For example, it must be smart enough to adapt to different scenarios (flat or hilly paths), road (highways, extra-urban, or urban ones), traffic (different volumes of vehicles and obstacles), illumination (day, night, sunset, sunrise), and weather (sun, fog, rain, or even snow). Also, unexpected or sudden changes of these conditions must be handled.

The number of faults must be reduced to zero, since even rare faults lead users to mistrust the system, i.e. to ignore warnings or take over the commands in the case of an automatic system.

It should be clear that not only the algorithms have to be robust, but also the hardware system (sensors, actuators, computing engines,...) has to withstand mechanical and thermic stresses, such as vibrations or high temperatures; namely, in other words, the whole system must be *automotive-engineered.*

2.2 Reliability

Since on-board perception systems are safety critical devices, a strong degree of reliability is mandatory. This entails that the project has to be thorough and rigorous during all its phases, from the requirements specification to the design and implementation. An extensive phase of testing and validation is of paramount importance.

2.3 Costs

The cost issue presents no problems for high-cost vehicles like trucks or busses for which an expensive safety system can be regarded as an investment for the vehicle's security and safety. Contrariwise for most of the car market the cost issues is a key parameter and requires a specific engineering phase aimed at the reduction of the market price through a careful redesign of the complete platform. It has been estimated that for marketing reasons such a system should cost no more than standard optionals.

The development costs due to the system re-engineering will be compensated by the large car market.

From another point of view, not only do production costs of perception systems have to be carefully considered and reduced as much as possible, but the operative costs need to be kept low as well. In fact another cardinal parameter for the car market is power consumption that should not affect vehicle's performance.

2.4 Size and Design

The requirements of the car market are very specific: besides the cost and performance issues, also car styling should not suffer from the installation of new hardware and –in particular– new sensors for visual perception. Therefore these systems need to be compact in size and the sensors need to be installed in a position that does not cause disturbance. Although for many sensors their positioning does not present aesthetical problems (for example, radars are typically positioned on the vehicle's front bumpers), the installation of cameras may be considered as an additional problem: some approaches require the cameras to be installed inside the driving cabin behind the windscreen in a high position.

2.5 User Interface

A user-friendly interface is needed to control perception systems allowing the driver to switch on the system, resume control, and modify behavioral parameters (speed, route, smoothness, or alertness in driving) on the fly, with just a few operations. The new generation of steering wheels already incorporate handles for cellular phones, or radio; they are also expected to become the front-end for perception systems too.

Since results of perception system have to be fed to the driver, also the use of visual signals (leds, lights, control panels, or even on-board monitors), mechanical feedbacks (steering wheel, seat, or pedal vibrations), and vocal messages (warnings) must all be carefully evaluated. Anyway, the system –especially a driving assistance one– must not flood the driver with a large quantity of information in order not to reduce the driver's attention.

2.6 Integration With Off-Board Systems

The integration of data acquired and processed on-board with data coming from external sources can be exploited to get a more robust detection and to modify important driving parameters. For example, the knowledge of traffic conditions, such as average vehicles speed or information about traffic jams and alternative routes, can be of basic importance for route planning tasks and can be used to modify the driving behavior. The knowledge about the presence of a traffic jam on the way can in fact change the priority used to achieve the goal: instead of reaching the destination in the minimum time, it can be switched to a minimum fuel consumption strategy.

Some cities already have an experimental traffic control center which can deliver real-time information to drivers. Nowadays these information are shown on road signs, but an integration with on-board navigation systems is

already under study. This implies the use of communications between vehicles and earth stations. Obviously one-way communication can be beneficial to all vehicles, but a two ways information exchange can also be used to update the status of the traffic control center. Instead of relying only on fixed surveillance stations, information can be further gathered by a large amount of moving agents, installed on vehicles.

The prompt delivery of traffic information to the control center thanks to systems installed onto each vehicle can increase the quality and robustness of traffic data and therefore provide an improved service to drivers.

3. THE RATIONALE OF VISUAL PERCEPTION IN THE AUTOMOTIVE FIELD

For a perception system the choice of sensors is of paramount importance. Several kind of sensors are used in the robotic filed. The most common ones are tactile, acoustic, laser, radar, and vision sensors.

Tactile sensors

In automotive applications tactile sensors, such as bumpers, are barely used. In fact their use is reduced to situations where the speed is sufficiently low so as not to cause serious damages when hitting obstacles (i.e. indoor robotics or, at least, parking).

Acoustic sensors

They are *active sensors*, in fact, they emit a specific signal and measure its reflections and alterations. They feature a low cost, but, unfortunately, also a very reduced detection range, therefore limiting their use in the automotive field.

Laser-based radars

Also laser-based sensors are *active sensors*. Based on the Doppler effect, laser-based radars detect the distance of objects by measuring the travel time of a signal emitted by the sensors themselves and reflected by the object. A slow scanning speed and low spatial resolution are their primary drawbacks. In addition, security issues, especially potential retina damages, discourages a widespread adoption of such sensors.

Radar-based sensors

Another kind of *active sensors* is millimeter-wave radars. They share the same functioning principle with laser-based sensors, Despite a bigger cost, millimeter-based radars are more robust to rain and fog than laser-based sensors. Unfortunately, as in the case of laser-based sensors, they features a low spatial resolution and slow scanning speeds too.

Vision-based sensors
Vision-based sensors are *passive sensors*, namely they do not alter the environment when acquiring data, therefore featuring advantages with respect to *active sensors* like radar, laser, acoustic, and tactile devices. The primary advantage in the use of machine vision is that it permits to detect visual information therefore being of paramount importance in several automotive applications (i.e. traffic signs recognition, obstacle identification,...) without any modification to current transport infrastructures. Unfortunately, vision sensors do not measure quantities in a direct way, requiring complex processing. Moreover, they are less robust than millimeter-wave radars in foggy, night, or direct sun-shine conditions.

4. MACHINE VISION

Obviously, the use of active sensors, involving the measurement of the alteration of signals emitted by the sensors themselves, features many advantages over the use of machine vision:

- active sensors measure quantities of interest in direct way. As an example, a radar device directly return the distance or relative speed of an object. Conversely, vision can detect distance or relative speed only trough a complex processing of sequences, or –more generally– sets, of images;
- active sensors acquires a small amount of data, therefore requiring less performing computing resources.

The main drawback in the use of active sensors in the automotive arena – besides the pollution of the environment– is related to potential interference amongst the same type of sensors in a scenario where several vehicles will be equipped with such sensors. In fact, active sensors may interfere with each other, thus decreasing their reliability and usefulness. This problem becomes even greater in outdoor unstructured environments, in which a large number of vehicles could be moving simultaneously, as –for example– in the case of autonomous vehicles traveling on intelligent highways. In addition, the maximum signal level must comply with some safety rules and must be lower than a safety threshold.

Thus, with a widespread use of autonomous sensing agents, the use of passive sensors obtains important advantages over active ones. Since, amongst the passive sensors, only cameras are of use in the automotive field, this is a case in which vision becomes of paramount importance. Obviously, while the use of other sensing agents would have the advantage of extending sensing capabilities besides human possibilities, machine vision fails in the same situations in which humans cannot see (e.g. in foggy conditions or during the night with no specific illumination). Hence, the use of vision allows the building of a system able to act as the human driver: for example, an active safety system that helps the driver in case of failure. Since the early stages of vision (low-level image processing) are computationally demanding, the availability of a low-cost computing engine is mandatory. Anyway recent developments in computational hardware (i.e. higher degree of integration and reduction of the power consumption) allow the access to machines that can deliver a high computational power at an affordable price. In addition, most of currently available general-purpose processors features extensions expressly tailored to exploit the intrinsic parallelism of the processing of visual and audio (i.e. the Intel's MMX or AMD's 3DNow! extensions [9]).

In a similar way, new cameras generation permits preliminary basic image processing during the acquisition of images; moreover, the extension of camera dynamics allows the removal of the processing required to adapt the acquisition parameters to specific light conditions. The resolution of the sensors has been drastically enhanced, and in order to decrease the acquisition and transfer time, new technological solutions can be found in CMOS sensors, such as the possibility of dealing with pixels independently as in traditional memories. Another key advantage of CMOS-based sensors is that their integration on the processing chip seems to be straightforward.

As a result, not only did this advanced technology promote improved hardware devices, but also triggered off renewed interest in the techniques for the processing of iconic information, generally addressed by the field of Artificial Intelligence, who deals with image interpretation and –more generally– with perception, when the fusion of data coming from other sensors is also integrated.
Nonetheless, when designing a vision system for automotive applications, some important characteristics must be carefully considered.

- the processing rate of visual perception systems for the automotive field bounds the vehicle speed; therefore, these systems require a faster processing than other applications. Unfortunately, as far as real-time processing is concerned, images represent a large amount of data. As a result, specific computer architectures and processing techniques must be devised in order to achieve real-time performance. Nevertheless, since the success of such apparata is strictly related to their cost, the computing engines cannot be based on expensive processors. Therefore, either off-the-shelf components [3] or ad-hoc dedicated low-cost solutions must be considered [5].

- another key issue is that, in an outdoor environment like the automotive one, no assumptions can be made on parameters measured by the vision sensor, i.e. the illumination or contrast of the scene. Therefore, the vision algorithms must be robust enough to tolerate these changes and dynamically adapt to different environmental conditions, such as sun (high brightness and contrast due to shadows), rain (extremely high contrast due to reflections), fog (low contrast). Also vision system movements and drifts in its calibration must be handled as well.

Even if computer vision is extremely complex and highly demanding, thanks to the great deal of information that it can deliver, it has been widely employed to deal with a large number of tasks in the automotive field. These tasks include: Road Following (which involves *Lane Detection* [10] and *Obstacle Detection* [8]), *Platooning* (the automatic following of a preceding vehicle [23]), *Vehicle Overtaking* [19,2], *Automatic Parking* [15], *Driver Status Monitoring* [14], *Road Surveillance*, *License Plate Recognition*, *Queue Detection*,... To accomplish the tasks that allow a vehicle to autonomously drive different quantities must be measured and/or patterns recognized before the closing of the control loop, for example: the relative position of the vehicle with respect to the lane and the check for obstacles on the path or for known road signs [1] for *Road Following*; the recognition of specific vehicle's characteristics and the computation of the time-to-impact for *Platooning*; the sensing of multiple lanes [21] as well as obstacle detection for *Vehicle Overtaking* and *Collision Avoidance*; the distance amongst close vehicles for *Automatic Parking*; the status of the driver's eyes or head for *Driver Status Monitoring*.

A visual perception system can be divided in two distinct modules: image acquisition and processing system.

4.1 Image acquisition

Several parameters must be evaluated for the design and choice of an image acquisition device. Before anything else, the size and the number of bit/pixel of images must be accurately chosen; Also the use of color or grey-level images must be evaluated.

Other parameters are related to the choice of monocular vs stereo vision and the sensors' angle of view. Some systems adopt a multi-camera approach, by using more than one camera with different viewing angles (e.g. fish eye or zoom) or a custom camera with multi-focal capabilities [17].

Also parameters intrinsic to the sensor itself must be considered. Although the frame rate is generally fixed for low-cost CCD-based devices (25 or 30 Hz), the dynamic of the sensor is of paramount importance. Different approaches have been studied to achieve high dynamic ranges, ranging from the design of CMOS-based cameras with a logarithmically compressed dynamic [18], to the interpolation and superimposition regarding values of two subsequent images taken from the same camera [12].

A specific problem in the automotive field is related to perception systems installed on-board of moving vehicles. In this case image stabilization becomes a key problem. Some expensive devices allow to perform images stabilization directly at sensor level. Another solution is to perform it in a preliminary processing phase [7,13,22], trading the complexity and cost of the sensor for the complexity and cost of processing system. Also the calibration for system installed on-board of moving vehicles is a key parameter in order to create a mapping between image pixels and the 3D world.

Finally, the acquisition system must be robust enough not to interfere with the on-board electronics, especially with telecommunication devices.

4.2 Image processing systems

Hardware for visual perception in the automotive field need to fit two fundamental requirements: it has to be sufficiently small and cheap to allow its integration into commercial vehicles; at the same time it has to be powerful enough to support real-time processing.

In the early years of research on perception systems, several custom solutions were proposed, based on ad-hoc, *special-purpose* hardware. At that time, in fact, commercial processors were not able to deliver sufficient computational power. Generally, these solutions were based on the SIMD paradigm and composed of a number of simple processors working simultaneously on different pixels [6].

Anyway, as soon as commercial hardware began to fit the computational requirements, *general-purpose* solutions, based on off-the-shelf components, begun to be considered. In this phase the first MIMD systems [20], composed of a rather small number of powerful, independent processors, were built and installed on prototype vehicles.

Current trends, however, are moving towards a mixed architecture, in which a powerful processor is aided by specific hardware such as boards and chips implementing optical flow computation, pattern-matching, convolution, and morphological filters. In the same manner, also the SIMD capabilities of the last-generation CPUs are being widely exploited to boost up the performance.

5. DISCUSSION AND PERSPECTIVES

A number of groups researching on intelligent vehicle applications have integrated their solutions for visual perception into their prototypes. Many of these experimental results were presented and demonstrated during important international events [11,16,4].

Anyway, though for this kind of applications computing power do not seem to be a problem any more, still some problems remain regarding data acquisition. It has been shown that the main difficulties encountered during the demos were due to light reflections and non perfect conditions for image acquisition (wet road, direct sunshine on the cameras, tunnels and bridges' shadows). As a common framework for the next years of research a great deal of work will be addressed towards the enhancement of sensor's capabilities and performance, including the improvement of gain control and sensitivity in extreme illumination conditions.

Therefore, a long period of exhaustive tests and refinement must precede the availability of these systems on the general market, and a fully automated highway system with intelligent vehicles driving and exchanging information is not expected for another two or more decades.

For the time being, complete automation will be restricted to special infrastructures such as industrial applications or public transportation. Then, automatic vehicular technology will be gradually extended to other key transportation areas such as the shipping of goods, for example on expensive trucks, where the cost of an autopilot is negligible with respect to the cost of the vehicle itself and the service it provides. Finally, once technology will have been stabilized and the most promising solution and best algorithms frozen, a massive integration and a widespread use of such systems will also take place with private vehicles.

Nevertheless, not all problems related to automatic vehicle driving that still remain to be solved are of technical nature: there are some aspects that must be considered and carefully evaluated in the design of such systems. First of all, prior to having an automatic driving system sold and installed on a commercial vehicle, all the legal aspects related to the responsibility in case of faults and incorrect behavior of the system must be solved. Secondly, in case no specific roads are built and dedicated to automatic vehicles only, the possibility of driving on a motorway along with automatic vehicles must

be considered and its impact on human drivers evaluated. Although technical aspects seem to have a higher importance, these problems must be dealt with and solved as well, since they represent the basics and prerequisites on which future automatic highways will rely.

ACKNOWLEDGEMENT

This work was partially supported by the Italian National Research Council (CNR) under the frame of the Progetto Finalizzato Trasporti 2 and MADESS Project.

REFERENCE

[1] G. Adorni, M. Gori, and M. Mordonini. Just-in-Time Sign Recognition in Image Sequences. *Real Time Imaging Journal*, 5(2):95–107, Apr. 1999.

[2] P. H. Batavia, D. A. Pomerleau, and C. E. Thorpe. Overtaking Vehicle Detection using Implicit Optical Flow. In *Procs. IEEE Intl. Conf. on Intelligent Transportation Systems'97*, page 329, Boston, USA, Nov. 1997.

[3] M. Bertozzi, A. Broggi, A. Fascioli, and S. Tommesani. Addressing Real-Time Requirements of Automatic Vehicle Guidance with MMX Technology. In J. R. et al., editor, *Procs. 4th Intl. Workshop on Embedded HPC Systems and Applications (EHPC'99) - Second Merged Symposium IPPS/SPDP 1999*, pages 1407–1417, Apr. 1999.

[4] Broggi, M. Bertozzi, and A. Fascioli. The 2000 km Test of the ARGO Vision-Based Autonomous Vehicle. *IEEE Intelligent Systems*, 14(1):55–64, Jan.–Feb. 1999.

[5] Broggi, M. Bertozzi, F. Gregoretti, R. Passerone, C. Sansoè, and L. Reyneri. A Dedicated Image Processor Exploiting Both Spatial and Instruction-Level Parallelism. In C. C. W. Jr., editor, *Procs. IEEE Computer Architectures for Machine Perception*, pages 106–115, Boston, Oct. 1997.

[6] Broggi, G. Conte, F. Gregoretti, C. Sansoè, and L. M. Reyneri. The Evolution of the PAPRICA System. *Integrated Computer-Aided Engineering Journal - Special Issue on Massively Parallel Computing*, 4(2):114–136, 1997.

[7] Z. Duric and A. Rosenfeld. Image Sequence Stabilization in Real-Time. *Real Time Imaging*, 2(5):271–284, Oct. 1996.

[8] W. Enkelmann. Obstacle detection by evaluation of optical flow field from image sequences. In *Procs. European Conf. on Computer Vision*, volume 427, pages 134–138, Antibes, France, Apr. 1990. LNCS, Springer Verlag.

[9] Intel Corporation. *Intel Architecture MMX Technology Developers' Manual*. Intel Corporation, 1997. Available at http://www.intel.com.

[10] K. Kluge. Extracting Road Curvature and Orientation from Image Edge Points without Perceptual Grouping into Features. In *Procs. IEEE Intelligent Vehicles Symposium'95*, pages 109–114, Detroit, Oct. 1995.