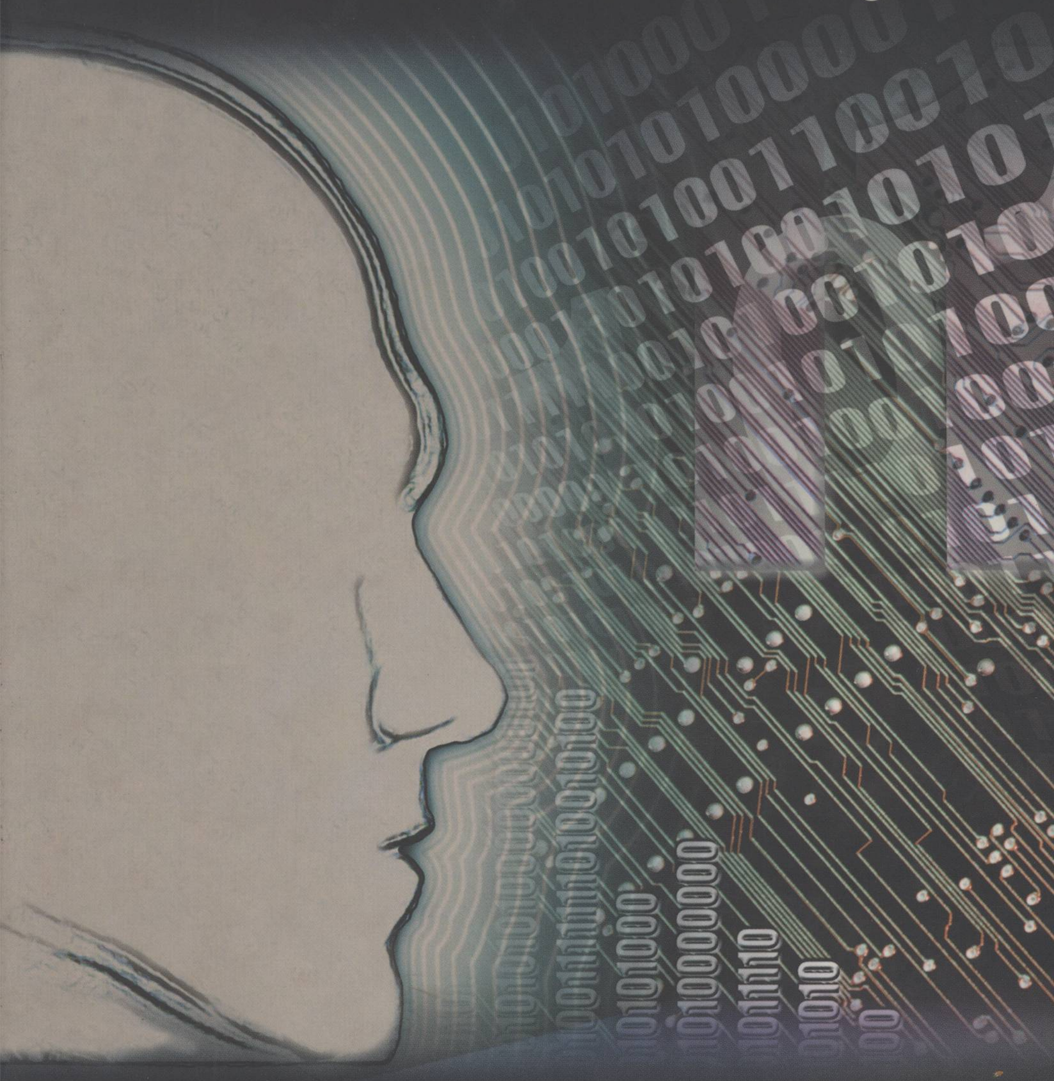
HUMAN & MACHINE PERCEPTION

Communication, Interaction, and Integration



EDITORS: SERGIO VITULANO

VITO DI GESÙ • VIRGINIO CANTONI

ROBERTO MARMO • ALESSANDRA SETTI

HUMAN & MACHINE PERCEPTION

Communication, Interaction, and Integration

Santa Caterina di Pittinuri, Oristano, Italy
September 6–9, 2004

EDITORS

SERGIO VITULANO

Dipartimento di Scienze Mediche e Internistiche, Università di Cagliari
Policlinico Universitario, Monserrato, 09042 Cagliari, Italy

VITO DI GESÙ

Dipartimento di Matematica ed Applicazioni, Università Di Palermo
Via Archirafi 34, 90123 Palermo, Italy

VIRGINIO CANTONI * ROBERTO MARMO * ALESSANDRA SETTI

Dipartimento Informatica, e Sistemistica, Università di Pavia, via Ferrata 1
27100 Pavia, Italy

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

HUMAN AND MACHINE PERCEPTION: COMMUNICATION, INTERACTION, AND INTEGRATION

ISBN 981-238-431-6

Printed in Singapore by World Scientific Printers (S) Pte Ltd

HUMAN & MACHINE PERCEPTION

Communication, Interaction, and Integration

PREFACE

The following are the proceedings of the Sixth International Workshop on Human and Machine Perception held in Santa Caterina di Pittinuri (Oristano), Italy, on September 06-09, 2004, under the auspices of the following Institutions: the Italian Ministry of the University and the Researches, the University of Cagliari, the Pavia University, and the Inter-Department Centers of Cognitive Sciences of Palermo University.

A broad spectrum of topics are covered in this series, ranging from computer perception to psychology and physiology of perception. The theme of this workshop on Human and Machine Perception was focused on *Communication, Interaction and Integration*. As in the past editions the final goal has been the analysis and the comparison of biological and artificial solutions.

The focus of the lectures has been on presenting the state-of-the-art and outlining open questions. In particular, they sought to stress links, suggesting possible synergies between the different cultural areas. The panel discussion has been conceived as a forum for an open debate, briefly introduced by each panelist, and mainly aimed at deeper investigation of the different approaches to perception and strictly related topics. The panelists were asked to prepare a few statements on hot-points as a guide for discussion. These statements were delivered to the participants together with the final program, for a more qualified discussion.

The number of participants to the workshop was limited to 70. Besides the 8 invited lecturers and panelists, more participants were admitted. Priority for these positions was given to young researchers who made significant contributions to the open discussions. Both the lectures and the contents of the panel discussions are included in these proceedings.

The workshop structure consisted of three modules each of one organized in four general talks. The first one, *Representing and coding for communication*, was grounded on general talks and a panel discussion to give the foundations. The second one, *Exploration, visualization and discovering in very large data set*, was focused on methodologies for data mining, classification and recognition. The third one, *Information exchange: Machine versus Machine,* was dedicated to the communication and interaction of hybrid complex systems.

In this edition we included, for the second time in the series, also a few solicited presentations from the audience.

The description of both natural and artificial structures and basic approaches represents a natural starting point for a workshop which intends to analyze the processes of systems communication, interaction and integration from different viewpoints. The lectures were organized by alternating the functional descriptions of natural and artificial information management and decision making related to perception, learning, knowledge acquisition and storage. Further inquiries concern the human ability to exploit spatial information for reasoning and communication purposes and the new technologies and facilities on this field.

Representing and coding for communication intends to addressed and to explore basic references for communication processes that allow natural or/and artificial systems to exchange knowledge and information to achieve a smart situated behaviour. Tools and modalities (languages, protocols, coding schemes, ect.) for handling perceptual information are discussed and compared. Advances in technology lead towards communications that are often mediated by computers and networks.

Exploration and discovering in very large dataset explores how multimedia technology supplies a new way of gaining understanding into large data structures, allowing the user to gain insight into the process under investigation. Data-mining provides a new tool consisting of automatic discovering of patterns, changes, correlations and anomalies through an interactive display and analysis of data up to the elicitation of information and knowledge. Progress in statistics and data analysis are also shown in relation with data retrieving and discovering.

Information exchange: machine versus machine is addressed to information exchange among the living and how human being and machines may interact for improving performances and/or behavior. In particular, the usability of a given multi-media interface, in terms of performance interface evaluation, is investigated. Moreover, how machines exchange information, learn by experimenting, and improve their performances is discussed. Modalities of interaction among distributed and heterogeneous systems, coordination in multi-agent systems, control strategies and suitable hierarchies are compared.

The workshop dealt with most of these problems and the results of the presentations and discussions are herewith included even if the chapters of the book vary somewhat from the scheduled program, to take care of the positions that emerged from the debates and to the audience reactions.

Acknowledgements

The workshop, and thus indirectly this book, was made possible through the generous financial support of the universities involved, and the research organizations that are listed separately. Their support is gratefully acknowledged.

In particular we would like to thank Roberto Marmo and Alessandra Setti for the completion of this volume. We are also grateful for the assistance offered by World Scientific for the preparation of this volume.

Virginio Cantoni, Vito Di Gesù, Sergio Vitulano

SPONSORING INSTITUTIONS

The following institutions are gratefully acknowledged for their contributions and support to the Workshop:

- MURST, Ministero dell'Università e della Ricerca Scientifica e Tecnologica, Italy
- Dottorato di Ricerca in Tecnologie Biomediche Applicate alle Scienze odontostomatologiche, Seconda Università di Napoli, Italy
- Dottorato di ricerca in Alimenti e Salute: Biotecnologie e Metodologie Applicate alla Fisiopatologia Digestiva, Seconda Università di Napoli, Italy
- Università di Cagliari, Italy
- Università di Pavia, Italy
- Università di Palermo, Italy
- Provincia di Cagliari, Italy

CONTENTS

INFORMATION EXCHANGE: MACHINE VERSUS MACHINE

ONE EYE, TWO EYES... HOW MANY EYES? NOTES AND REFLECTIONS ON VISION SYSTEMS FOR MOBILE ROBOTICS*

GIOVANNI ADORNI

Dipartimento di Informatica, Sistemistica e Telematica,
Università di Genova,
Via all'Opera Pia 13,
16145 Genoa, Italy
E-mail: adorni@unige.it

STEFANO CAGNONI, MONICA MORDONINI

Dipartimento di Ingegneria dell'Informazione,
Università di Parma,
Parco Area delle Scienze 181a,
43100 Parma, Italy
E-mail: {cagnoni,monica}@ce.unipr.it

In this chapter we discuss some issues on mobile robot perception from the point of view of the structure of vision systems. Through some experiences and results from research projects we will approach the problem from the point of view of purposive vision, with particular regard to less conventional sensors and systems, such as omnidirectional and hybrid systems.

1. Introduction

Differentiation of living beings is the most outstanding result of million of years of natural evolution. The main characterizing feature of Darwinian evolutionary processes is having the fitness of an individual, with respect to the surrounding environment and to performing the actions required to survive in it, as both their goal and driving force. As a result, each species has developed its own specialized *apparatus*, which maximize performance in its specific environment. Sensory and locomotory systems are probably the ones for which species differentiation is most immediately evident.

* This work was partially supported by ASI (Italian Space Agency) within the grants "Coordination of the cluster Robotic Vision" and "Hybrid Vision System for Long Range Rovering", and by ENEA (National Institute for Alternative Energy) within the grant "Sensori Intelligenti".

For many living beings, among sensory systems, the vision system is the richest source of information which is used to realize where they are and what is around them, in order to plan what their immediate goal should be and what sequence of actions should be performed to reach it. A basic taxonomy of vision systems in animal species can be based, firstly, on the number of sensitive elements by which an eye is composed and on the number of eyes; secondly, on features such as their field of view and sensitivity to light or colour.

Virtually all vertebrates, along with some invertebrates such as molluscs, medusas and worms, have simple eyes, similar to humans'. Obviously, despite their similar physical structure, their performance generally differs, according to the environmental conditions in which they are expected to operate.

For instance, cats, bats, owls, which usually act in dark environments, if not exclusively at night-time, have only rods, more sensitive and numerous than in human eyes, to obtain high sensitivity to light. In dolphins' eyes, as well, there are 7000 times as many rods as in humans, to cope with limited underwater lighting.

Fish eyes have usually a flat cornea and a spherical, non-deformable crystalline lens, to improve sight in the near field of view. On the other hand, birds' eyes are elongated in antero-posterior direction to produce larger images of distant objects.

Among invertebrates, one of arthropods' peculiar features is having composite eyes, made up of many (up to thousands) simple eyes. The final image which they perceive is formed as apposition or superposition of the single images which are perceived by the simple eyes.

Similarly to what has been happening in natural evolution through million of years, the choices made by designers of vision systems for mobile robots are mainly influenced by the tasks which the robots, on which such systems are going to be mounted, are expected to perform. In computer vision, the purposive vision approach aims at solving problems encountered in dealing with complex tasks or environments by identifying the goal of the task; this simplifies design by making explicit just that piece of information that is needed to achieve such a goal. As a matter of fact, mobile robotics is probably the field in which the purposive vision paradigm has been applied most frequently. Mobile robotics applications actually require that task-specific information be selected from the usually huge amount available, to satisfy the constraints of real-time or just-in-time processing imposed by the dynamic environments in which mobile robotics tasks are usually performed.

In designing robot vision systems the main degrees of freedom available to the designer are the same peculiarities that distinguish biological vision systems, namely the number of sensors and some of their features, of which the width of the field of view and the related parameters of distortion and resolution are perhaps the most relevant. Therefore, also in the case of robotic vision systems, a similar taxonomy to the one reported above for natural vision systems can be applied.

In this chapter, we aim at discussing the task of designing robot vision systems according to such a taxonomy and from the point of view of purposive vision applications. The chapter reviews examples of applications or preliminary experiences we developed in previous research projects, highlighting the task-oriented motivation of design, and discussing the results that such systems can achieve.

We will consider the case of single-sensor systems, as opposed to multi-sensor ones. As regards sensor types, we will consider traditional pin-hole and omnidirectional sensors in single or multi-sensor configurations. Particular attention will be given to the design of omnidirectional and hybrid omnidirectional/pin-hole systems.

2. Single-sensor Systems

Up to a few years ago, most computer vision applications relied on single-sensor systems. However, the choice very often did not depend on an objective evaluation of the specifications on the design of the vision system, but mostly on constraints imposed by the computation power that was available. As processor power increases, along with sensor technology, it is possible to design computer vision applications which are more and more complex, both intrinsically and from the point of view of the sensors by which information is acquired. Mobile robotics applications, in particular, have become more and more reliant on heterogeneous sensory systems (cameras, ultrasonic sensors, laser scanners, etc.), in which even the vision system is often made up of multiple, homogeneous or heterogeneous, components.

In the far-from-complete review of vision sensors made by this chapter, attention is principally focused onto multi-sensor vision systems. Therefore, discussion of single-sensor system will be limited to describing their essential features, considering them mainly as basic components of more complex systems. Also, since the main focus of this chapter is on less conventional sensor design, in describing conventional cameras we will limit our discussion to a very concise review of the main features which affect performance of mobile robot vision systems. In particular we will focus on the features which distinguish conventional from omnidirectional sensors, discussing the latter more amply, even if, also in their case, with the aim of introducing them as components of multi-sensor systems.

2.1. *Conventional Cameras*

Conventional cameras use the classic pin-hole principle, by which objects in a scene are projected onto a planar sensor after being focused on it by a lens, or simply after the light rays which they reflect pass through a tiny hole.

If the size and resolution of the sensor are set, the field of view of such a system is inversely dependent on the focal distance, i.e. the distance between the lens (hole) and the sensor onto which the image is projected. Conversely, the resolution of the image which can be obtained increases with the focal distance, since a smaller scene is projected onto the same sensitive area.

Images acquired by conventional cameras are affected by two kinds of distortions: perspective effects and deformations that derive from the shape of the lens through which the scene is observed. Both kinds of distortion hamper direct measures, since distances in pixel are not proportional to actual distances in the real world.

In mobile robotics applications, being able to reconstruct the surrounding scene with high accuracy is a critical task in navigation, especially as obstacle detection and motion planning are concerned.

After introducing omnidirectional sensors in the next section, we introduce the Inverse Perspective Transform (IPT) as a general tool to, firstly, compensate for the classical perspective distortion in pin-hole images and, more generally, to compensate for any measurable distortion in images acquired by any sort of vision sensor. IPT can be effectively used to detect obstacles in mobile robotics application based on stereo vision systems, as shown in Section 3.2.

2.2. *Omnidirectional Cameras*

Being able to see behind one's shoulders has probably been one of the most desirable supernatural faculties for humans (possibly the most desirable after being able to fly), resulting in several mythological creatures, such as Janus, the divinity with two faces, and Argus Panoptes, the giant with a hundred eyes.

In applications involving dynamic environments such as mobile robotics, widening the field of view is of major importance, to be able to enhance awareness about what is happening in the surroundings and to obviate the need for active cameras, that require complex control strategies. In this section we will consider a particular class of omnidirectional sensors, catadioptric sensors, which consist of a conventional camera that acquires the image reflected on a convex mirror.

Catadioptric omnidirectional sensors suffer from two main limitations. The most relevant one is that the near field, which is the least distorted part of the image, is partially obstructed by the reflection of the camera on the mirror. A further limitation is that the accumulation of camera and system distortions makes it quite difficult either to find the resulting distortion law and to compensate for it, or to design a mirror profile that can achieve a good trade-off between width of the field of view, image resolution and distortion (see [5] for a discussion of the problem). To reduce this effect, mirrors that produce a non-distorted image of a reference plane have been recently described [13,16].

Omnidirectional systems are therefore very efficient as concerns detection of target position, but critical from the point of view of the accuracy with which the target is detected. For these reasons, either the omnidirectional vision sensor has been integrated with a different kind of sensor, to make object detection and robot self-localization more precise and robust (see, for example [9,12,15,19], or arrays of omnidirectional sensors have been used to implement triangulation algorithms [20]. In the following section we discuss the problem of designing omnidirectional systems from the point of view of mobile robotics applications.

2.2.1. *Design of catadioptric omnidirectional systems*

One of the most important issues in designing a catadioptric omnidirectional vision system is choosing the geometry of the reflecting surface of the mirror [14]. Different mirror profiles have been proposed in literature [21].

In particular, adopting the solution proposed by [13] one can design a mirror which preserves the geometry of a plane perpendicular to the axis of symmetry of the reflecting surface, i.e., an isometric mirror: the mirror acts as a computational sensor, capable of providing distortion-free images automatically, thus eliminating the need for further processing, if the distortion caused by the camera lens is neglected. Other researchers propose the use of multi-part mirrors [8], composed of a spherical mirror providing better resolution in the proximity of the robot to accurately localize significant features in the environment, merged with a conic mirror that extends the field of view of the system, to perceive objects at a greater distance at the price of a loss in resolution. This composite mirror shape permits to image objects that are closer to the sensor with respect to classical off-the-shelf conical mirrors that are often used to achieve omnidirectional vision. This design has been suggested by the requirements of RoboCup,[a] in which robots need to see the ball up to the point at which they touch it.

---

[a] Visit http://www.robocup.org for information on the RoboCup robot soccer competition.

Obviously, to design the profile of a mirror it is always necessary to first define the requirements for the particular application for which the catadioptric vision system is meant; thus a general rule for generating the optimum mirror profile does not exist. For example, a vision system which is meant only for obstacle avoidance probably does not require that the robot recognize very distant objects; the opposite is true for a localization system which relies on the recognition of significant features distributed on the walls or extracted from the surrounding environment [9]. Once the requirements for a particular application have been specified, to design a mirror profile with the desired characteristics, a differential equation needs to be solved, which can be inferred by applying the laws of linear optics. The following specifications, suggested by considerations related with both the optics and the mechanics of the system, are to be met in mobile robot vision systems:

1. minimizing encumbrance and weight of the mirror by minimizing its radius. If one considers that the whole vision system is typically made up of a camera and a transparent cylinder holding the mirror, by reducing the mirror weight (and, if possible, the cylinder height) the mechanical stability of the whole system is increased. Notice also that, for a given focal length of the camera, decreasing the radius of the mirror also requires that the distance between the camera and the reflecting surface be decreased as well as, consequently, the height of the support;
2. designing profiles which reduce the area of the image containing useless information, e.g., the reflection of the robot on the mirror;
3. radially extending the field of view up to where relevant information can be still extracted;
4. keeping the resolution of the most relevant part of the image as high as possible, to provide the clearest possible view of the objects that lie close to it and require that the robot take strictly real-time decisions in dealing with them.

The rationale behind these basic requirements is clear if one observes Figure 1, in which a cheap but clearly non-optimal catadioptric surface (the convex face of a ladle!) has been used to acquire an omnidirectional view of a RoboCup field, in which a ball is placed at a distance of 1 *m* from the sensor.

In such an image, the most evident problems are:

1. poor resolution already at a relatively short distance: the ball is hardly visible, despite being at a distance of about 1 meter;
2. extension of the field of view beyond the area of interest: useless low-resolution information (very distant objects) is visible;
3. the image region occupied by the reflection of the robot body on the mirror is very large.

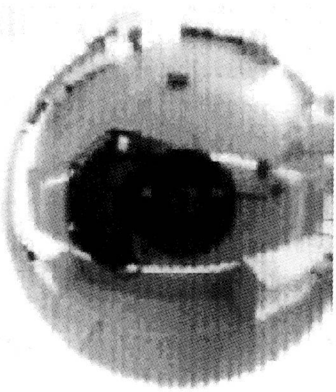


Figure 1. A 'home-made' omnidirectional sensor obtained using a kitchen ladle.

According to the above-mentioned guidelines, we have designed a mirror which is very similar to the one in Figure 2. The mirror is obtained as a revolution surface by revolving the profile around the z-axis, it has a radius of 5 cm and it is about 2.69 cm high.

The profile can be decomposed in two sections: the first section (from point A to point B) can be modeled as an arc of circumference, while the second section (from point B to point C) is a straight line. The tangent to the circumference in point B has the same slope as the straight line, preserving the continuity in the first order derivative. Observing the resulting surface (see Figure 3), it is straightforward to see that it is composed by a conic mirror jointed to an "almost spherical" mirror. On one hand, the property of the conic part to reflect objects which are far from the robot is very useful to detect objects of interest (other robots, openings, signs on walls, etc.). Obviously, resolution decreases with object distance from the robot, allowing only for an approximate perception of the distance of interesting features. However, if focusing on a particular feature

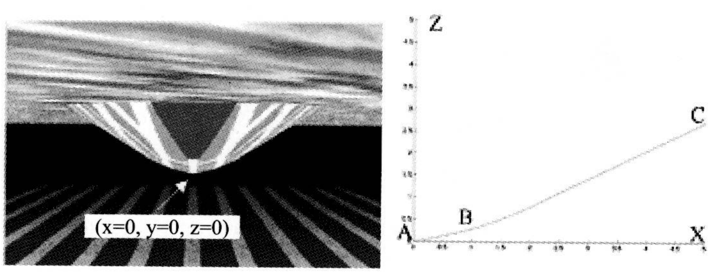


Figure 2. Catadioptric omnidirectional sensor for mobile robotics applications and its generating profile.