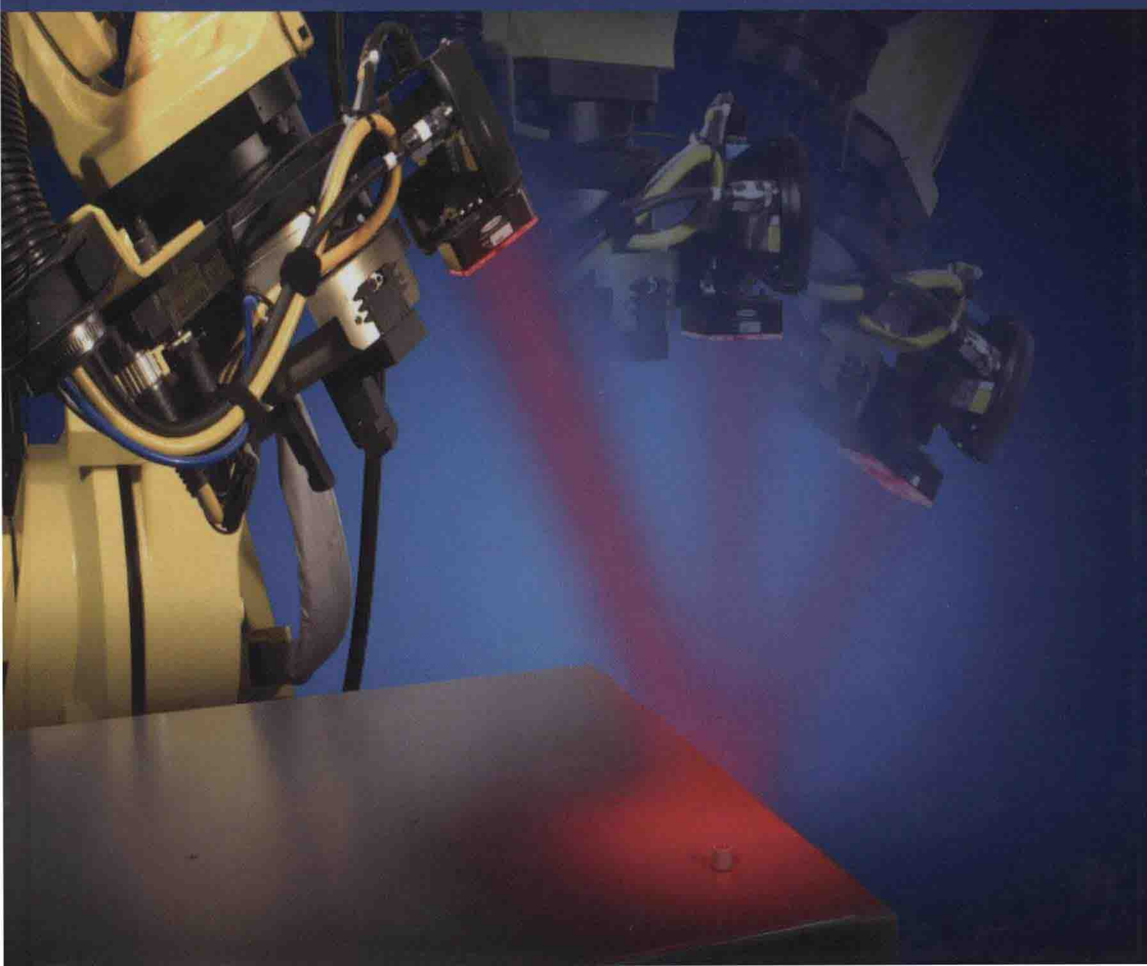
Fundamentals of Machine Vision

Contributors

Alejandro Nieto, David López Vilarino et al.

Edited and Compiled by Koros Press Editorial Board



Fundamentals of Machine Vision

Contributors

Alejandro Nieto, David López Vilarino, et al.

KOROS PRESS LIMITED

Fundamentals of Machine Vision

Contributors: Alejandro Nieto, David López Vilarino, et al.

Published by Koros Press Limited

www.korospress.com

United Kingdom

Fundamentals of Machine Vision

ISBN: 978-1-78163-932-0

British Library Cataloguing in Publication Data
A CIP record for this book is available from the British Library

Printed in the United Kingdom

Fundamentals of Machine Vision

List of Contributors

Alejandro Nieto
Centro de Investigación en Tecnoloxías da Información (CITIUS) University of Santiago de Compostela Spain

David López Vilarino
Centro de Investigación en Tecnoloxías da Información (CITIUS) University of Santiago de Compostela Spain

Víctor Brea Sánchez
Centro de Investigación en Tecnoloxías da Información (CITIUS) University of Santiago de Compostela Spain

Yiping Tang
College of Information Engineering, Zhejiang University of Technology, Hangzhou, China

Caiguo Chen
College of Computer Science, Zhejiang University of Technology, Hangzhou, China

Robin Gruna
Fraunhofer Institut of Optronics, System Technologies and Image Exploitation IOSB, Karlsruhe

Stephan Irgenfried
Institute for Process Control and Robotics (IPR), Karlsruhe Institute of Technology KIT, Germany

M. Tornow
Otto-von-Guericke University of Magdeburg Germany

M. Grasshoff
Otto-von-Guericke University of Magdeburg Germany

N. Nguyen
Otto-von-Guericke University of Magdeburg Germany

A. Al-Hamadi
Otto-von-Guericke University of Magdeburg Germany

B. Michaelis
Otto-von-Guericke University of Magdeburg Germany

Juan Yan
College of Mechanical Engineering, Shanghai University of Engineering Science, Shanghai, China

Huibin Yang
College of Mechanical Engineering, Shanghai University of Engineering Science, Shanghai, China

Erik Hultman
Division for Electricity, Uppsala University, Box 534, 751 21 Uppsala, Sweden

Mats Leijon
Division for Electricity, Uppsala University, Box 534, 751 21 Uppsala, Sweden

Deokjin Joo
Department of Electrical Engineering and Computer Science, Seoul National University, Seoul, Korea

Ye-seul Kwan
Division of EcoScience, Ewha Womans University, Seoul, Korea

Jongwoo Song
Department of Statistics, Ewha Womans University, Seoul, Korea

Catarina Pinho
CIBIO, Centro de Investigac¸a˜o em Biodiversidade e Recursos Gene´ticos, Universidade do Porto, Vaira˜o, Portugal

Jody Hey
Department of Biology, Temple University, Philadelphia, Pennsylvania, United States of America

Yong-Jin Won
Division of EcoScience, Ewha Womans University, Seoul, Korea

Xiang Zhang
State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Hangzhou 310027, China
School of Computer, Hangzhou Dianzi University, Hangzhou 310018, China

Zhangwei Chen
State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Hangzhou 310027, China

Tomasz Kosicki
Norwegian University of Science and Technology (NTNU) in Trondheim, Norway
PPM AS in Trondheim, Norway

Trygve Thomessen
PPM AS in Trondheim, Norway

Preface

The text *Fundamentals of Machine Vision* provide imaging-based automatic inspection and analysis for such applications as automatic inspection, process control, and robot guidance in industry. Machine vision is the ability of a computer to "see." A machine-vision system employs one or more video cameras, analog-to-digital conversion (ADC), and digital signal processing (DSP). The resulting data goes to a computer or robot controller. Machine vision is similar in complexity to voice recognition. First chapter addresses an analysis of different computing paradigms and platforms oriented to image processing. Second chapter presents a new Omni-Directional Tilt Sensor (ODTS), which consists of the LED light, transparent cone-shaped closed container, mercury, camera, embedded systems. In third chapter, we show how machine vision for automated visual inspection can greatly benefit from reflectance measuring and modeling in the context of image analysis and synthesis. In fourth chapter, a high speed calculation of depth maps from stereo images based on FPGAs has been introduced. Fifth chapter describes industrial sorting system, which is based on robot vision technology, introduces main image processing methodology used during development, and simulates algorithm with Matlab. Sixth chapter presents a production method using industrial robots for automation of cable winding of electric machine stators. In seventh chapter, we discuss the usage of the current automated feature extractor and species identification for future study of evolution and speciation of cichlids, which will have implications for other animal studies. Eighth chapter proposes a novel solution for a stereo vision machine based on the System-on-Programmable-Chip (SoPC) architecture. The aim of last chapter is to qualitatively investigate whether audio-visual cognitive Info-communication.

Contents

List of Contributors .... *vii*

Preface .... *xi*

Chapter 1 **Towards the Optimal Hardware Architecture for Computer Vision ... 1**
Alejandro Nieto, David López Vilarino and Víctor Brea Sánchez

Chapter 2 **Design of Omni-Directional Tilt Sensor Based on Machine Vision... 35**
Yiping Tang, Caiguo Chen

Chapter 3 **Reflectance Modeling in Machine Vision: Applications in Image Analysis and Synthesis........ 55**
Robin Gruna and Stephan Irgenfried

Chapter 4 **Fast Computation of Dense and Reliable Depth Maps from Stereo Images........ 89**
M. Tornow, M. Grasshoff, N. Nguyen, A. Al-Hamadi and B. Michaelis

Chapter 5 **Research on Workpiece Sorting System Based on Machine Vision Mechanism........ 125**
Juan Yan, Huibin Yang

Chapter 6 **Utilizing Cable Winding and Industrial Robots to Facilitate the Manufacturing of Electric Machines........ 139**
Erik Hultman, Mats Leijon

Chapter 7 **Identification of Cichlid Fishes from Lake Malawi Using Computer Vision 175**
Deokjin Joo, Ye-seul Kwan, Jongwoo Song, Catarina Pinho, Jody Hey, and Yong-Jin Won

Chapter 8 **Sad-Based Stereo Vision Machine on a System-on-Programmable-Chip (Sopc) 219**
Xiang Zhang and Zhangwei Chen

Chapter 9 **Cognitive Human-Machine Interface Applied in Remote Support for Industrial Robot Systems** 237

Tomasz Kosicki, and Trygve Thomessen

Citations 263

Index 265

Chapter 1

TOWARDS THE OPTIMAL HARDWARE ARCHITECTURE FOR COMPUTER VISION

Alejandro Nieto, David López Vilarino
and Víctor Brea Sánchez

Centro de Investigación en Tecnoloxías da Información (CITIUS) University of Santiago de Compostela Spain

INTRODUCTION

Computer Vision systems are experiencing a large increase in both range of applications and market sales (BCC Research, 2010). From industry to entertainment, Computer Vision systems are becoming more and more relevant. The research community is making a big effort to develop systems able to handle complex scenes focusing on the accuracy and the robustness of the results. New algorithms provide more advanced and comprehensive analysis of the images, expanding the set of tools to implement applications (Szeliski, 2010). Although new algorithms allow to solve new problems and to approach complex situations with a high degree of accuracy, not all the algorithms are adequate to be deployed in industrial systems.

Parameters like power consumption, integration with other system modules, cost and performance limit the range of suitable platforms. In most cases, the algorithms must be adapted to achieve a trade-off solution and to take advantage of the target platform. Conventional PC-based systems are constantly improving performance but their use is still limited to areas where portability, power consumption and integration are not critical. In case of highly complex algorithms, with an irregular execution flow, complex data representation and elaborated patterns to access to data, a significant gain is not achieved when moving to an ad hoc hardware design. In this case, a high-end CPU and a GPU with general-purpose capabilities (GPGPU) is a flexible and very powerful combination that will outperform other options (Castano-Díez et al., 2008).

However, when a conventional system does not meet the requirements of the application, a more aggressive planning is needed. For instance, migrating the application to a dedicated device such as a DSP, an FPGA or a custom chip (Shirvaikar & Bushnaq, 2009). At this point, the designers have to consider alternatives as to reduce operating range, accuracy and robustness of the results, or to remove expensive operations in order to simplify the hardware that will be implemented (Kolsch & Butner, 2009). PC-based systems enable a great flexibility at cost of performance so pure software-based algorithms hardly match pure hardware implementations. This is a serious limitation because it can compromise the efficiency of the application. This is the reason why the industry is making great efforts to develop novel architectures that enable greater flexibility to adapt the algorithms without compromising the quality of the results.

Computer Vision applications are often divided into three stages: low, medium and high level processing. Low level operations are quite simple and repetitive but applied over a large set of data with a much reduced data dependent program flow, so massive parallelism is essential for performance. In the medium level stage, temporal and task parallelism is key as the data set is smaller and the program flow is quite data-dependent. High level operations are performed over complex data representations but reduced data sets and with high precision as requirement. An efficient Computer Vision system must deal with all these stages. The selection of the computation model will determine the performance of the device in each one of the stages. So, which is the optimal paradigm for such applications? Increase the number of concurrent instructions? Increase the number of data elements processed simultaneously? A combination of both?

Taking into account the application requirements, a suitable platform to build the system must be selected. Besides general parameters (performance, cost and integration), there is a set of factors that restricts the computing paradigms and the devices that can be selected. For instance, a critical parameter is the way in which the data are transferred to the device because I/O operations are one of the bottlenecks in high-performance systems. Data type and representation will affect the computational units. The program flow will constrain the inner connections between these units and the storage elements.

This chapter addresses an analysis of different computing paradigms and platforms oriented to image processing. Previously, a representative set of Computer Vision algorithms covering the three levels of processing is reviewed. This study will lead us to observe the algorithms in terms of a set of common characteristics: operations, data type, program flow, etc. This is critical to design new hardware architectures in order to maximize performance.

The analysis from the hardware point of view will highlight the best features of the most used computing paradigms in order to establish a relationship between the type of operation, data, programming model and hardware architecture. An efficient architecture for Computer Vision must combine all the selected features. The analysis of the characteristics of the different algorithms will lead us naturally to an optimized general-purpose hardware architecture for Computer Vision.

THE PROBLEM OF COMPUTER VISION

Traditionally, Computer Vision (CV) applications include building blocks from three computation levels: low-, mid- and high-level vision computing. The type of operations, the data representation and the flow execution of programs depend deeply on the considered level of this hierarchy. Nevertheless, current CV-algorithms are composed of many different processing steps regarding the type of data and the way these are computed which makes difficult to classify them only in one subgroup. Following, a rather rough classification of widely used CV-algorithms is made, keeping in mind the data domain and the complexity of the involved operations.

Low-Level Vision

After image acquisition, some preprocessing steps are often required. These are intended to provide reliable input data for subsequent computing stages. Some typical operations are noise reduction, color balancing, geometrical transformation, etc. Most of these operations are based on point or near-neighborhood operations. Point operations are performed at pixel-level in such a way that the output only depends on the value of any individual pixels from one or several input images. With this type of operation it is possible to modify the pixel intensity to enhance parts of the image, by increasing contrast or brightness. Equally, simple pixel-to-pixel arithmetic and Boolean operations also enable the construction of operators as alpha blending, for image combination or color space conversion. Neighborhood operations take also into account the value of adjacent pixels. This operation type is the basis of filtering, binary morphology or geometric transformation. They are characterized by simple operations, typically combining weighted sums, Boolean and thresholding processing steps.

After preprocessing stages, useful information has to be extracted from the resulting images. Common operations are edge detection, feature extraction or image segmentation. Edges are usually defined as step discontinuities in the image signal so finding local maxima in the derivative of the image or zero-crossings in the second derivative are suitable to detect boundaries. Both tasks are usually performed by the convolution of the input image with spatial filtering masks that approximate a first or second derivative operator.

Feature points are widely used for subsequent computing steps in multiple CV-applications.

Basically, a feature represents a point in the image which differs from its neighborhood. One of the benefits of local features is the robustness against occlusion and the ability to manage geometric deformations between images when dealing with viewpoint changes. In addition, they improve accuracy when, in the same scene, objects are at different planes, (i.e. at different scales). One of the most popular techniques is that proposed by Harris (Harris & Stephens, 1988) to detect corners. It is widely used due to its strong invariance to rotation, image noise and no large illumination changes. It uses the local auto-correlation function, which describes the gradient