Li-minn Ang · Kah Phooi Seng
Li Wern Chew · Lee Seng Yeong
Wai Chong Chia

Wireless Multimedia Sensor Networks on Reconfigurable Hardware

Information Reduction Techniques

Springer

Wireless Multimedia Sensor Networks on Reconfigurable Hardware

Li-minn Ang • Kah Phooi Seng • Li Wern Chew
Lee Seng Yeong • Wai Chong Chia

Wireless Multimedia Sensor Networks on Reconfigurable Hardware

Information Reduction Techniques

Springer

Li-minn Ang
Kah Phooi Seng
School of Engineering
Edith Cowan University
Joondalup, Western Australia
Australia

Lee Seng Yeong
Wai Chong Chia
Dept. of Computer Science and Networked Systems
Sunway University
Selangor, Malaysia

Li Wern Chew
Intel Architecture Group
Intel Corporation
Penang, Malaysia

ISBN 978-3-642-38202-4 ISBN 978-3-642-38203-1 (eBook)
DOI 10.1007/978-3-642-38203-1
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013950611

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Grace Ang Yi-en, our blessed daughter for all the joy you bring.
–Li-minn Ang and Kah Phooi Seng

To my parents for your unconditional love and never ending support.
–Li Wern Chew

To my parents and loved ones for your unceasing support and love.
–Lee Seng Yeong

To my parents, thank you for all your unconditional support.
–Wai Chong Chia

Acknowledgements

We would like to express our deepest appreciation to all the people who contributed to the work presented in this book. We would like to thank our former Ph.D. student Christopher Ngau Wing Hong, in particular for his help on some of the work described in Chaps. 4 and 5.

Li-minn Ang and Kah Phooi Seng gratefully acknowledge support from the Centre for Communications Engineering Research at Edith Cowan University, Australia. Lee Seng Yeong and Wai Chong Chia gratefully acknowledge support from the School of Computer and Networked Systems at Sunway University, Malaysia. Li Wern Chew gratefully acknowledges support from the Intel Architecture Group at Intel, Malaysia. We gratefully acknowledge our Springer editor Ronan Nugent and the publication team for their diligent efforts and support towards the publication of this book. The support of the Malaysian eScience Fund research grants (01-02-12-SF0054, 01-02-16-SF0026) is gratefully acknowledged.

Contents

List of Figures

List of Tables

Acronyms

AIM	Attention based on information maximization
CS	Center surround
DWT	Discrete wavelet transform
EX	Instruction execution
EXMEM	Instruction execution and memory access
FOV	Field of view
FPGA	Field-programmable gate array
ID	Instruction decode
IF	Instruction fetch
MCH	Multimedia cluster head
MEM	Instruction memory access
MIPS	Microprocessor without interlocked pipeline stages
MSB	Most-significant bit
MSF	Medium spatial frequencies
NCC	Normalized cross-correlation
RAM	Random-access memory
RANSAC	Random sample consensus
SAD	Sum of absolute difference
SIFT	Shift invariant feature transform
SIG	Significant
SOT	Spatial orientation tree
SPIHT	Set partitioning in hierarchical trees
SSD	Sum of square difference
SW	Slepian–Wolf
VA	Visual attention
WB	Instruction write back
WBSME	Wavelet-based saliency map estimator
WMN	Wireless multimedia node
WMSN	Wireless multimedia sensor network
WSN	Wireless sensor network
WZ	Wyner–Ziv

Chapter 1
Introduction

Abstract This book describes how reconfigurable hardware technology like field programmable gate arrays can be used to offer a cost-efficient and flexible platform for implementation in wireless multimedia sensor networks (WMSNs). A main focus of the book is towards the development of efficient algorithms and architectures for information reduction techniques such as event detection, event compression and multi-camera processing for hardware implementation in WMSNs.

1.1 Overview

The field of wireless sensor networks (WSNs) has seen much success in diverse applications ranging from environmental monitoring, smart homes, healthcare to industrial and defence applications. Traditional WSNs capture scalar data such as temperature, vibration, pressure or humidity. Motivated by the success of WSNs and also with the emergence of new technology in the form of low-cost image sensors, researchers have proposed combining image and audio sensors with WSNs to form wireless multimedia sensor networks (WMSNs). The use of image sensors in WMSNs increases the range of potential applications because compared to scalar sensors, image sensors are able to provide more information which can be used for visual processing tasks such as detection, identification and tracking. On the one hand, WMSNs can be seen as an extension of traditional WSNs where multimedia sensors have replaced scalar sensors. On the other hand, the use of multimedia sensors in WMSNs brings with it a different set of practical and research challenges. This is because multimedia sensors, particularly image sensors, generate a high amount of data that would have to be processed and transmitted within the network. The main issue is that sensor nodes have limited battery power and hardware resources.

There are currently three options a designer has to implement algorithms in hardware: hardwired logic or application-specific integrated circuits (ASICs), software programmed microprocessors and reconfigurable computing. The

L. Ang et al., *Wireless Multimedia Sensor Networks on Reconfigurable Hardware*,
DOI 10.1007/978-3-642-38203-1_1, © Springer-Verlag Berlin Heidelberg 2013

advantage of using ASICs is that they can be designed specifically to meet the requirements for a particular application to give fast and efficient computations. The disadvantage is that any modifications to the functionality would require the entire IC to be replaced which would be a costly process. Software programmed microprocessors provide flexibility by allowing changes to the software instructions to be executed. However, this comes at a cost of performance, and, typically microprocessor-based designs are slower than ASICs due to the high overheads required for processing each instruction. The third approach is to use reconfigurable hardware which provides higher performance than software programmed microprocessors while also maintaining a higher level of flexibility than ASICs. A reconfigurable computer has the facility to make significant changes to its datapath and control. The hardware can be specially configured to perform an application-specific task such as image processing or pattern matching, and after completion the hardware can be reconfigured to perform another specific task. Currently, a popular platform for reconfigurable hardware technology is the field programmable gate array (FPGA).

The energy-efficient processing and transmission of the data within the WMSN is of primary importance to maximise the lifetime of the overall network. Depending on the application, the network should be active for a duration of time ranging from weeks to years without the need for battery replacement. To save energy in transmission, information reduction algorithms can be applied to minimise the amount of data to be transmitted. Two different visual information processing approaches can be employed to reduce the image data. The approaches can be divided into single-view approaches and multi-view approaches. Single-view approaches attempt to reduce the image data from each individual sensor node. Two techniques can be used here: event detection and event compression. The first technique uses event detection to reduce data by only transmitting frames in the network when significant events are detected. Image frames which are not significant are discarded. For example, a surveillance application could use a face event detector to decide which image frames to send to the base station. Image frames which do not contain faces do not need to be transmitted. However, the face detector would need to have low computational complexity to meet the energy requirements in the sensor node. There is a trade-off between energy required for processing and energy required for transmission. On the one hand, using an event detector in the sensor node requires more computational power. On the other hand, this could result in a saving of transmission power when frames are discarded. The other advantage of an event detector is that it could also serve as an early stage for visual pre-processing.

To perform the facial recognition process, the central computer would need to perform at least two stages. The first stage is to locate the face location in the image and the second stage would then be to perform the recognition task by comparing the facial features with a stored database. To reduce the large amount of image data for processing by the central computer, the event detector performs the face detection task, and the location(s) of the face(s) is then communicated to the central computer to perform the facial recognition task. The second technique is to perform event compression on the image frames which have to be transmitted. Event compression approaches rely on image and video compression algorithms to remove redundancy from the data. These algorithms range from current standards like JPEG, MPEG-x and H.26x to newer techniques using distributed video coding and compressive sensing. While single-view approaches attempt to reduce the scene data from each individual sensor node, multi-view approaches perform the data reduction by aggregating the data from different sensor nodes. This is possible because of the overlapping field of views (FOVs) captured by different nodes. Regions of image frames that overlap are discarded prior to transmission. However, the implementation of these information reduction algorithms would require higher computational complexity, leading to higher requirements for processing. For cost-efficient implementation in WMSNs, the algorithms would need to have low computational and memory complexity.

1.2 Book Organisation

Multimedia sensors can comprise both audio and visual sensors. For simplicity, we will use the term multimedia sensor networks even though the focus of the book is on visual sensors. The book material will be of interest to university researchers, R&D engineers, electronics, communications, and computer engineers, and graduate students working in signal and video processing, computer vision, embedded systems and sensor networks. Although some of the material presented in the book has been previously published in conference proceedings and articles [27–30, 97], most of the material has been rewritten, and algorithm descriptions have been included for practical implementations.

The remainder of the book is organised into seven chapters. Chapters 2 and 3 present background material on WMSNs and current technology. Chapter 2 discusses the emergence and architectures of WMSNs and describes the components and technology for a wireless multimedia sensor node, information processing approaches for WMSNs and various applications for WMSNs. Chapter 3 discusses the technology available and the advantages of reconfigurable hardware for WMSNs and the range of FPGA technology, internal architectures and families available. The chapter also presents an overview of programming languages for reconfigurable devices and discusses the range of available languages from lower-level languages like VHDL and Verilog to higher-level languages based on C-based languages like SystemC, SpecC and Handel-C.

Chapter 4 presents the specification for the FPGA WMSN hardware platforms. The chapter discusses the instruction set architecture and processor architecture for the datapath and control. A specification for the WMSN processor in Handel-C is included. Chapter 5 presents a case study for implementing visual event detection on the FPGA WMSN processor using visual attention. The chapter gives a background to event detection and discusses algorithms for visual saliency. The chapter also contains the description of the hardware modules and programs in assembly language. Chapter 6 presents a case study for implementing visual event compression on the FPGA WMSN processor using wavelet compression. The chapter gives a background to event compression and discusses algorithms for visual coding. The chapter also contains the description of the hardware modules and programs in assembly language. Chapter 7 presents a discussion for implementing multi-camera approaches such as visual event fusion in WMSNs. The chapter gives a background to event fusion, feature detection techniques and discusses algorithms and implementations for image stitching on the WMSN.

Chapter 2
Wireless Multimedia Sensor Network Technology

Abstract This chapter presents background material for wireless multimedia sensor network (WMSN) technology. The chapter will describe the general structure for a WMSN and various architectures and platform classifications for WMSNs. The chapter will also discuss the various components in a WMSN node such as the sensing, processing, communication, power and localisation units. The efficient processing of information in a WMSN is of primary importance, and the chapter will discuss various multi-camera network models and information reduction techniques such as event detection and event compression. The chapter concludes with a discussion of applications of WMSNs.

2.1 Introduction

The emergence of wireless multimedia sensor networks (WMSNs) is an evolutionary step for wireless sensor networks as audio and visual sensors are integrated into wireless sensor nodes. It has been a focus of research in a wide variety of areas including digital signal processing, communication, networking and control systems. WMSNs are able to store, correlate and fuse multimedia data originating from several camera input sources. The main applications for WMSNs are those that benefit from distributed and multi-camera vision systems. Deploying multiple, low-cost visual sensors both improves the coverage of the application and provides for a more robust operation. Also, multiple cameras provide redundancy to improve its reliability and usability. A single point failure will not cause a system failure, nor an obstruction or occlusion. Furthermore, multiple visual sources provide the flexibility to adaptively extract data depending on the requirements of the application. A multi-resolution description of the scene and multiple levels of abstraction can also be provided. A typical application for a WMSN would be as a surveillance and monitoring system. The WMSN provides several advantages over traditional monitoring and surveillance systems which include [36]:

L. Ang et al., *Wireless Multimedia Sensor Networks on Reconfigurable Hardware*,
DOI 10.1007/978-3-642-38203-1_2,

- *Enlarging the view.* Viewpoints from multiple cameras can provide a close-up view of an event either through the images captured by a camera nearer the scene or by engaging a node with a more advanced camera such as a pan-tilt-zoom (PTZ) camera. In such a system, an event detected by a node with a lower resolution camera can signal another node with a PTZ camera to detect and track the event.
- *Enhancing the view.* The use of multiple cameras can also enhance the view of an event by providing a larger field of view (FOV) or by using cameras with different capabilities such as mixing cameras for the visible and infrared spectrum in the network. Such systems are very useful when the view is obscured or when there is little or no illumination in the scene.
- *Providing multiple viewpoints for the same event.* When a single camera is considered for a surveillance application, the coverage of the application is only limited by the FOV of a fixed camera or the field of regard (FOR) of a PTZ camera. This is limiting as parts of a scene may often be obscured especially in monitoring areas of high object density such as in public transportation hubs.

These advantages come at the cost of an increase in the data generated in the network which in turn increases the energy consumption in the network. To ensure that the typical battery-powered WMSN lifespan is not significantly affected by this, the amount of data routed through the network can be reduced with the use of in-network processing techniques such as to remove redundancy from multi-camera systems, selective transmission of the data and compressing the data. These processing tasks can be performed at the node, cluster or distributed throughout the network. The use of error detection and correction can also help reduce the likelihood of a costly retransmission. In this chapter, a broad coverage on WMSN technology will be provided as the background information for understanding the WMSN design, its architectures, challenges and design considerations. These design considerations are strongly dependent on the application. Aspects such as deployment density, cost, size, geographic location and purpose determine the components for the implementation of a specific WMSN.

2.2 WMSN Network Technology

This section will describe the layout of a typical WMSN as shown in Fig. 2.1. The network typically consists of a large number of sensor nodes deployed in a region of interest and one or more base stations. The base station or sink acts as the main network controller or coordinator. In this role, its primary function is to coordinate the functions of the nodes. It also collects information gathered by the nodes to be stored or further processed. The sink also serves as a gateway to other networks. The sinks are normally located close to the nodes to avoid high energy-consuming long-range radio communications. The energy consumption in the WMSN will determine the network lifespan. The energy consumed for communication is much higher

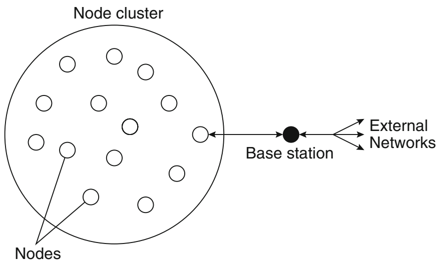

Fig. 2.1 Typical WMSN layout

than that for sensing and computation and grows exponentially with the increase of transmission distance. Therefore it is important that the amount of transmissions and transmission distance be kept to a minimum to prolong the network lifespan. This is one of the main motivations for in-network processing—the reduction of the information required for transmission for the efficient use of energy. In [60, 91], it was reported that the transmission of data can take 1,000–10,000 times more energy than processing, and this difference will increase as processor technology improves.

To reduce the transmission distance, a multi-hop short-distance communication scheme is preferred, and in most sensor networks, this is how it is implemented. In a multi-hop communication network, a sensor node transmits data towards the sink via one or more intermediate nodes. The architecture of a multi-hop network can be organised into two types: flat and hierarchical. The next sections will describe the structure of a WMSN and its various architectures and classifications.

2.2.1 Structure of a WMSN Network

Figure 2.2 shows a general structure of a WMSN consisting of four main components: wireless multimedia node (WMN), wireless cluster head (WCH), wireless network node (WNN) and base station. There is a decreasing information flow from the WMNs to the base station. The captured scene data are processed and transformed to useful event data. The focus for the upper levels of the network (WMN and WCH) is on information processing, and the focus for the lower levels of the network (WNN) is on wireless network communications. The primary theme for both upper and lower network levels is to achieve energy efficiency within the constraints of the battery-powered nodes.

2.2.1.1 Wireless Multimedia Node

The WMNs form the end points of the network. Each WMN consists of a camera or audio sensor, processing unit, communication unit and power unit. The camera

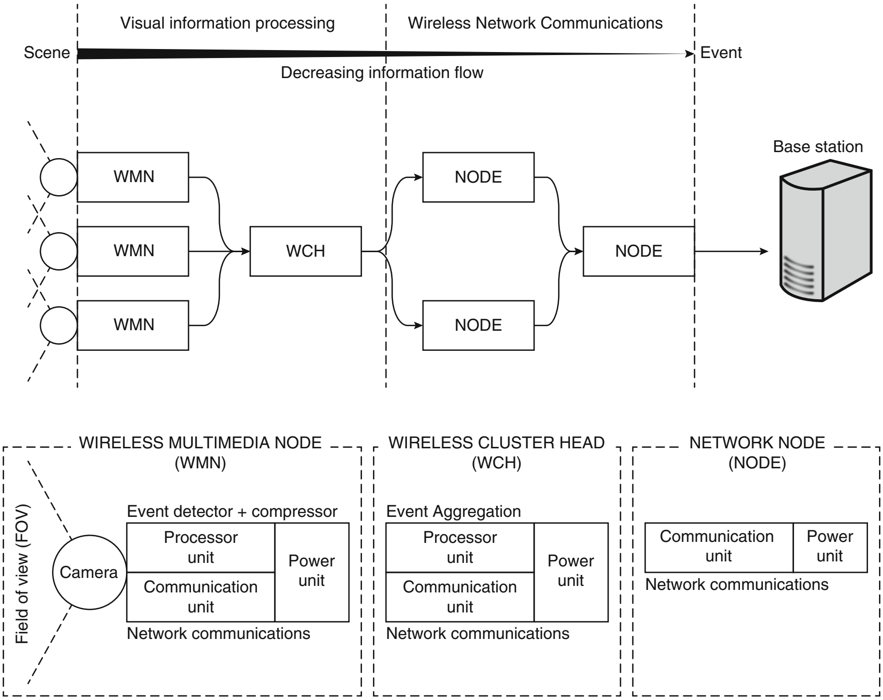

Fig. 2.2 General structure of a WMSN

sensor has a FOV of the scene. A captured scene is called an image frame. The processing unit performs the visual processing to reduce the high amount of scene data. Two approaches can be used. The first approach uses event detectors to identify useful events in the scene data. If an event is not detected, then the image frame is discarded and there is no need to transmit the frame through the network. The second approach uses event compressors to reduce the data for image frames that have to be transmitted through the network. Various image and video information processing techniques can be used, and this will be briefly described in Sect. 2.4 and discussed in detail in Chaps. 5 and 6. The communication unit transmits the compressed data to other nodes. Each WMN receives its power supply from a power unit which is mostly battery-powered.

2.2.1.2 Wireless Cluster Head

The WCHs receive data from several WMNs. Each WCH consist of a processing unit, communication unit and power unit. Each WCH receives data from several WMNs. The FOVs of WMNs may overlap, and the WCH can further reduce the