

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

© Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

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

1	Introduction	1
1.1	Overview	1
1.2	Book Organisation	3
2	Wireless Multimedia Sensor Network Technology	5
2.1	Introduction	5
2.2	WMSN Network Technology	6
2.2.1	Structure of a WMSN Network	7
2.2.2	WMSN Network Architecture	9
2.3	WMSN Node Technology	13
2.3.1	Structure of a WMSN Node	13
2.3.2	Sensing Unit	14
2.3.3	Processor Unit	19
2.3.4	Communication Unit	23
2.3.5	Power Unit	27
2.3.6	Frame Memory	31
2.3.7	Localisation Unit	31
2.4	Information Processing in WMSN	32
2.4.1	Multi-Camera Network Models	32
2.4.2	Collaborative Object Detection, Tracking and Recognition	33
2.4.3	Information Reduction in WMSN	35
2.5	Applications of WMSNs	37
2.6	Chapter Summary	38
3	Hardware Technology and Programming Languages for Reconfigurable Devices	39
3.1	Introduction	39
3.2	Hardware Technology for Reconfigurable Hardware	40
3.2.1	System-Level Architectures	41
3.2.2	Reconfigurable-Level Models	43

3.2.3	Functional Block Architectures	46
3.2.4	Soft-Core Microprocessor Architectures	52
3.3	Programming Languages for Reconfigurable Devices	53
3.3.1	Hardware Description Languages (HDLs)	54
3.3.2	HDL Design Flow	54
3.3.3	HDL Abstraction Models	56
3.3.4	Advantages of HDLs	57
3.3.5	Examples of HDLs	58
3.4	Chapter Summary	68
4	FPGA Wireless Multimedia Sensor Node Hardware Platforms	69
4.1	Introduction	69
4.2	WMSN Hardware Platforms	71
4.2.1	Low-Cost WMSN Platform Using Celoxica RC10	72
4.2.2	Medium-Cost WMSN Platform Using Celoxica RC203E	73
4.2.3	Digi XBee ZB Zigbee Wireless Module	74
4.2.4	Platform Architecture	75
4.2.5	Memory Architecture	75
4.3	WMSN Processor Architecture	77
4.3.1	Structure of a MIPS Microprocessor and Customisations ...	77
4.3.2	Strip-Based MIPS Processor Architecture	80
4.4	Discrete Wavelet Transform and Handel-C Implementation	90
4.4.1	Lifting-Based DWT	91
4.4.2	Hardware Architecture for DWT	94
4.4.3	Stage 0: Data Copy	94
4.4.4	Stage 1: Discrete Wavelet Transform	94
4.5	Hardware Platform Implementations	102
4.6	Chapter Summary	102
5	Single-View Information Reduction Techniques for WMSN Using Event Detection	105
5.1	Introduction	105
5.2	Computational Visual Attention and Models	108
5.2.1	High-Level and Low-Level Features	108
5.2.2	Bottom-Up and Top-Down Approaches	108
5.2.3	Centre-Surround Process	109
5.2.4	Saliency Map	110
5.2.5	Visual Attention Models	111
5.3	Low-Complexity and Low-Memory VA Model	122
5.3.1	Integer-Based MSF Model for Hardware Implementation	123
5.3.2	Lowering Memory Requirements Through Strip-Based Processing	125
5.4	Hardware Implementation on the WMSN Platform	129
5.4.1	Stages 0 and 1: Data Copy and DWT	129
5.4.2	Stage 2: Feature Emphasis	129

5.4.3	Stage 3: Feature Combination	135
5.4.4	Stage 4: Upsampling and Across Level Combination	136
5.4.5	Stage 5: Salient Pixel Computation	145
5.4.6	Stage 6: Information Reduction Using Salient Patches	146
5.5	Hardware Setup and Results	152
5.6	Chapter Summary	156
6	Single-View Information Reduction Techniques for WMSN Using Event Compression	159
6.1	Introduction	159
6.2	Visual Compression and Models	161
6.2.1	Spatial and Temporal Redundancy	161
6.2.2	Lossy and Lossless Compression	162
6.2.3	Evaluation of Image Quality	162
6.2.4	Image Compression Models	163
6.3	SPIHT Coding with Improvements	176
6.3.1	Low-Memory Implementation of SPIHT Coding	177
6.3.2	Modified SPIHT Algorithm	184
6.3.3	Listless SPIHT-ZTR Coding	190
6.4	Hardware Implementation on the WMSN Platform	193
6.4.1	Stages 0 and 1: Data Copy and DWT	195
6.4.2	Stage 2: Upward Scanning Significance Data Collection ...	196
6.4.3	Stage 3: Downward Scanning Listless SPIHT-ZTR Coding	198
6.5	Hardware Setup and Results	198
6.6	Chapter Summary	206
7	Multiple-View Information Reduction Techniques for WMSN Using Image Stitching	207
7.1	Introduction	207
7.2	Joint-Sources Compression	208
7.2.1	Slepian–Wolf and Wyner–Ziv Theorems	210
7.2.2	Collaborative Correspondence Analysis	213
7.2.3	Comparison of Joint-Sources Compression	214
7.3	Feature Extraction	215
7.3.1	Harris Corner Detector	215
7.3.2	Harris–Laplace Detector	216
7.3.3	Laplacian-of-Gaussian (LoG) Detector	216
7.3.4	Features from Accelerated Segment Test (FAST) Detector	217
7.3.5	Scale Invariant Feature Transform (SIFT)	217
7.3.6	Comparison of Feature Extraction Methods	218
7.4	Overview of Image Stitching	219
7.4.1	Feature Points Descriptor and Image Matching	220
7.4.2	Transformation Parameters and Image Stitching	221
7.4.3	Post-Processing	222

- 7.5 Analysis of Feature Extraction Methods on Subsampled Images ... 223
 - 7.5.1 Minimum Resolution of Subsampled Images 223
 - 7.5.2 Number of Feature Points 224
- 7.6 Proposed Framework 226
 - 7.6.1 Stage 1: Send Images to Aggregation Node 228
 - 7.6.2 Stage 2: Correspondence Analysis and Image Stitching 229
 - 7.6.3 Stage 3: Generate Data for Visual Nodes 229
 - 7.6.4 Stage 4: Discard the Overlapping Regions 233
 - 7.6.5 One-Pass Downward Scanning SPIHT Coding 233
- 7.7 Simulation Results 239
 - 7.7.1 Relationship Between Overlapping Regions
and Number of Output Bits 240
- 7.8 Hardware Implementation 241
- 7.9 Experimental Results 243
- 7.10 Chapter Summary 247

- A Appendix** 249

- References** 271

- Index** 279

List of Figures

Fig. 2.1	Typical WMSN layout	7
Fig. 2.2	General structure of a WMSN	8
Fig. 2.3	Multi-hop homogeneous flat network architecture	10
Fig. 2.4	Hierarchical single-tier single-hop network architecture	11
Fig. 2.5	Hierarchical single-tier multi-hop network architecture.....	11
Fig. 2.6	Hierarchical multi-tier multi-hop architecture	11
Fig. 2.7	State diagram for Wasp mote operational modes.....	13
Fig. 2.8	Components of a typical WSN node	14
Fig. 2.9	A typical CMOS sensor arrangement (Bayer pattern filter).....	16
Fig. 2.10	Front-illuminated vs. back-illuminated CMOS sensor	17
Fig. 2.11	The ECM microphone (foil type or diaphragm type) (a) Photo of ECM. (b) Cross-section of ECM	18
Fig. 2.12	Processor unit	19
Fig. 2.13	Different memory types.....	21
Fig. 2.14	A six-transistor CMOS SRAM	22
Fig. 2.15	Communication types	24
Fig. 2.16	ZigBee layers and specifications	26
Fig. 2.17	ZigBee network topologies	30
Fig. 2.18	Position calculation using angle-of-arrival.....	32
Fig. 2.19	Multi-camera network models	33
Fig. 2.20	General model for background subtraction	34
Fig. 3.1	WSN reconfigurable configurations	41
Fig. 3.2	Reconfigurable configuration for embedded processor.....	42
Fig. 3.3	Single-context reconfigurable model.....	44
Fig. 3.4	Multi-context reconfigurable model.....	44
Fig. 3.5	Partial reconfigurable model.....	45
Fig. 3.6	Pipeline reconfiguration model.....	45
Fig. 3.7	Overview of a FPGA architecture.....	47
Fig. 3.8	Xilinx Spartan-3 architecture.....	47
Fig. 3.9	Arrangement of Slices with the CLB.....	48

Fig. 3.10	Xilinx Spartan-3 slice structure.	49
Fig. 3.11	Altera Stratix II FPGA architecture.	50
Fig. 3.12	Altera Stratix II ALM structure.	51
Fig. 3.13	Functional block architecture for ALU array.	51
Fig. 3.14	Functional block architecture for processor array.	52
Fig. 3.15	Xilinx MicroBlaze soft-core processor architecture.	53
Fig. 3.16	General design flow using hardware description language.	55
Fig. 3.17	Abstraction levels of a hardware description language.	56
Fig. 3.18	Sequential and parallelism in Handel-C.	66
Fig. 3.19	C-to-FPGA design flow.	67
Fig. 4.1	Overview of the Celoxica RC10 FPGA board.	72
Fig. 4.2	Overview of the Celoxica RC203E FPGA board.	73
Fig. 4.3	Overview of the Digi Xbee ZigBee mesh development kit.	74
Fig. 4.4	Hardware platform architecture for the RC10 WMSN.	75
Fig. 4.5	Illustration of the STRIP RAM with its partitions.	76
Fig. 4.6	Simplified architecture for a standard five-stage pipelined MIPS processor.	77
Fig. 4.7	Datapath design for strip-based MIPS processor.	83
Fig. 4.8	IF datapath for strip-based MIPS processor.	84
Fig. 4.9	ID datapath for strip-based MIPS processor.	85
Fig. 4.10	Address branching unit in the ID stage.	86
Fig. 4.11	EXMEM datapath for strip-based MIPS processor.	88
Fig. 4.12	WB datapath for strip-based MIPS processor.	89
Fig. 4.13	Hardware architecture for the lifting-based 5/3 DWT.	92
Fig. 4.14	Symmetric extension through reflection for row filtering and column filtering.	97
Fig. 4.15	Arrangements of DWT coefficients in the STRIP_RAM—1D DWT. (a) 2-D DWT arrangement (level 3). (b) Desired 1-D DWT arrangement in STRIP_RAM.	100
Fig. 4.16	Arrangements of DWT coefficients in the STRIP_RAM—lifting DWT. (c1) Lifting-based DWT filter output (level 1). (c2) 1-D DWT arrangement (level 1). (d1) Lifting-based DWT filter output (level 2). (d2) 1-D DWT arrangement (level 2). (e1) Lifting-based DWT filter output (level 3). (e2) 1-D DWT arrangement (level 3).	101
Fig. 5.1	Examples of visual salience caused by low-level features.	109
Fig. 5.2	Centre-surround process using the sliding window method.	110
Fig. 5.3	Input image for VA processing and its saliency map.	111

Fig. 5.4	Visual comparison of saliency maps for VA models using simple scene images	119
Fig. 5.5	Visual comparison of saliency maps for VA models using complex scene images	121
Fig. 5.6	An overview of the integer-based MSF	123
Fig. 5.7	Diamond-structured filter used in the CS process of the MSF 5/3 VA model	125
Fig. 5.8	Strip-based processing approach applied to VA model.....	126
Fig. 5.9	Visual comparison of saliency maps for various strip sizes	128
Fig. 5.10	Computing the CS filter start address	135
Fig. 5.11	Feature combination of orientation sub-band maps	138
Fig. 5.12	Four neighbourhood averaging in the upsampling method.....	139
Fig. 5.13	Obtaining image patches from the captured image	146
Fig. 5.14	Overview of information reduction using salient patches	150
Fig. 5.15	Hardware setup for VA implementation.....	154
Fig. 5.16	Data reduced images that are captured, processed and transmitted by the WMSN	155
Fig. 6.1	Overview of single-view event compression process	161
Fig. 6.2	Types of image compression algorithms	163
Fig. 6.3	Baseline JPEG—Compression and Decompression.....	164
Fig. 6.4	Tree structure and parent–children dependencies of sub-bands in EZW coding	166
Fig. 6.5	Flow chart for EZW coding	167
Fig. 6.6	Parent–offspring dependencies of SPIHT SOT.....	168
Fig. 6.7	JPEG 2000 compression	171
Fig. 6.8	EBCOT coding	172
Fig. 6.9	SPECK coding methodology	173
Fig. 6.10	Process of pyramidal coding. (a) Original image Lenna. (b) Gaussian pyramid image. (c) Gaussian interpolation image. (d) Laplacian pyramid image	174
Fig. 6.11	Process for segmentation-based coding	175
Fig. 6.12	SPIHT coding in STRIP_BUFFER	178
Fig. 6.13	New spatial orientation tree structures	180
Fig. 6.14	SOT Degree-1, 2 and 3	184
Fig. 6.15	Two combinations in modified SPIHT algorithm.....	185
Fig. 6.16	Listless modified SPIHT for strip-based implementation—sorting and refinement	191
Fig. 6.17	Listless modified SPIHT for strip-based implementation—Combi 1	192
Fig. 6.18	Listless modified SPIHT for strip-based implementation—Combi 2	193
Fig. 6.19	Hardware architecture of the proposed strip-based image coder	194
Fig. 6.20	Architecture for SPIHT-ZTR encoder module	195

Fig. 6.21	Wavelet coefficients arrangement in STRIP_BUFFER	195
Fig. 6.22	Significant coefficient at each bit-plane in the STRIP_BUFFER	196
Fig. 6.23	Visual compression experimental setup	202
Fig. 6.24	Visual compression back-end processing	203
Fig. 6.25	Results of visual compression (outdoor)	204
Fig. 6.26	Results of visual compression (indoor)	205
Fig. 7.1	The effect of applying image stitching to images captured by two image sensors	209
Fig. 7.2	(a) The encoding of X with Y present at both encoder and decoder. (b) The encoding of X with Y only present at the decoder	211
Fig. 7.3	Three-bit input can form eight codewords, and these eight codewords can be separated into four different cosets	211
Fig. 7.4	DCT step sizes	213
Fig. 7.5	Multiview visual node setup	214
Fig. 7.6	The Gaussian scale-space representation	216
Fig. 7.7	The Difference-of-Gaussian (DoG)	217
Fig. 7.8	The computation process of the DoG adopted in SIFT	218
Fig. 7.9	SIFT candidate point and its neighbours	218
Fig. 7.10	Image stitching steps	220
Fig. 7.11	RANSAC line estimation	222
Fig. 7.12	Number of feature points extracted from Image A and B of Set 1–4	227
Fig. 7.13	Four stage WMSN information reduction	228
Fig. 7.14	Distribution of wavelet coefficients at different decomposition levels	229
Fig. 7.15	Identification of overlapping regions	230
Fig. 7.16	Efficiency of SPHIT coding—effects of high-frequency components	232
Fig. 7.17	Binary map and tree pruning data	233
Fig. 7.18	Single-pass downward scanning SPHIT coding flowchart	236
Fig. 7.19	Single-pass SPHIT subroutines	237
Fig. 7.20	Single-pass SPHIT subroutines— <i>DESC</i> significance determination	238
Fig. 7.21	Single-pass SPHIT subroutines— <i>GDESC</i> significance determination	239
Fig. 7.22	Modified MIPS with tree pruning	242
Fig. 7.23	Hardware experiment setup	244
Fig. 7.24	Sample hardware stitching result	245
Fig. 7.25	Sample binary maps for stitching	246
Fig. 7.26	Result of number of mapping and refinement bits received from two visual nodes	247

List of Tables

Table 2.1	Comparison of mote platforms	12
Table 2.2	Operational modes for Waspote	13
Table 2.3	SHT1x, 2x and 7x series of humidity sensors from Sensirion ...	15
Table 2.4	Comparison of NAND and NOR flash memory operating specifications	23
Table 2.5	Comparison of RF communication standards	24
Table 2.6	ZigBee PHY layer IEEE 802.15.4 specifications, 2003–2012	28
Table 2.7	ZigBee and 802.15.4 network node types	30
Table 3.1	Resource differences between Spartan-3 FPGA and Spartan-6 FPGA	49
Table 3.2	Comparison of Xilinx MicroBlaze and Altera Nios II soft-core processors	54
Table 3.3	VHDL built-in data operators	59
Table 3.4	Verilog built-in data operators	60
Table 3.5	Comparison of SystemC and SpecC in terms of design modelling	63
Table 3.6	Handel-C data types	64
Table 3.7	Handel-C operators	65
Table 4.1	The structure of the MIPS instruction code in R-format, I-format and J-format	78
Table 4.2	Extended instruction set for strip-based MIPS processor	81
Table 4.3	Extended instruction set for strip-based MIPS processor	81
Table 4.4	Machine code for R-format instructions	82
Table 4.5	Machine code for I-format instructions	82
Table 4.6	List of instruction opcodes and their control signals	90
Table 4.7	Description and function of major control signals	90
Table 4.8	New address calculations of DWT coefficients to be rearranged in STRIP_RAM	102

Table 4.9	Hardware platform summary for Celoxica RC10 WMSN hardware platform	103
Table 5.1	Memory requirements against the number of strips for the MSF 5/3 implementation	127
Table 5.2	Number of MIPS instructions and clock cycles for visual attention data reduction	156
Table 6.1	Prediction functions in lossless JPEG	164
Table 6.2	Tree structure	179
Table 6.3	Parent–children relationship of the original and the proposed new spatial orientation tree (SOT) structures for a three-scale SOT decomposition. Parent node is at coordinate (i, j) and its first descendant node is at coordinate (x, y)	181
Table 6.4	SPHIT memory requirements—strip vs traditional	183
Table 6.5	The percentage (%) of occurrence of possible outcomes of the $SIG(k, l)$ and $DESC(k, l)$ bits for various standard greyscale test images of size 512×512 pixels under Combination 1, $DESC(i, j) = 1$ and $GDESC(i, j) = 1$. Node (i, j) is the root node and (k, l) is the offspring of (i, j)	186
Table 6.6	The percentage (%) of occurrence of possible outcomes of the ABCD for various standard greyscale test images of size 512×512 pixels under Combination 2, $DESC(i, j) = 1$ and $GDESC(i, j) = 0$. ABCD refers to the significance of the four offspring of node (i, j)	187
Table 6.7	Performance of modified SPHIT (Deg-0)	190
Table 7.1	Comparison of different approaches used for feature extraction	219
Table 7.2	Object transformations and their parameters	221
Table 7.3	Stitching test sets	224
Table 7.4	Definition of stitching quality	225
Table 7.5	The evaluation results of using the subsampled images for stitching	226
Table 7.6	The simulation results obtained from applying the proposed framework to the first frame of video sequence “Ballroom” and “Vassar”	240
Table 7.7	The percentage of overlap (%) and the reduction in number of bits that are required to transmit the non-overlapping region	241
Table 7.8	New address calculation rules	242
Table 7.9	Number of MIPS instructions used to process the image data	243

Table 7.10	Total number of clock cycles required to encode all the 16 strips of an image frame.....	243
Table 7.11	Resources utilisation of the soft-core processor	244
Table 7.12	Battery lifespan	246

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

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]:

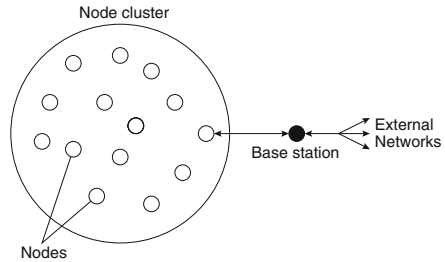
- *Enlarging the view.* Viewpoints from multiple cameras can provide a close-up view of an event either through the images captured by a camera nearer the scene or by engaging a node with a more advanced camera such as a pan-tilt-zoom (PTZ) camera. In such a system, an event detected by a node with a 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

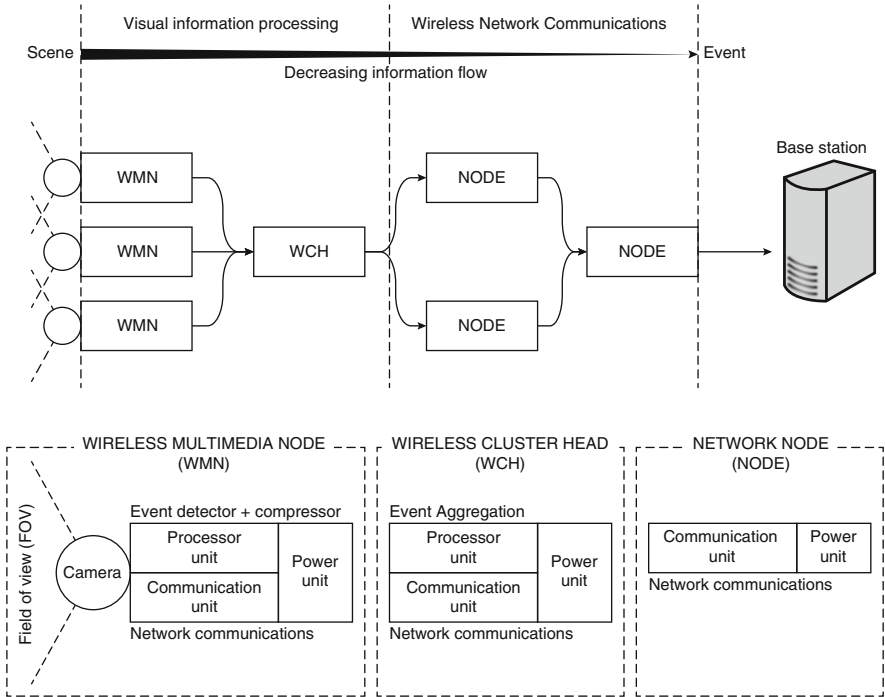


Fig. 2.2 General structure of a WMSN

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