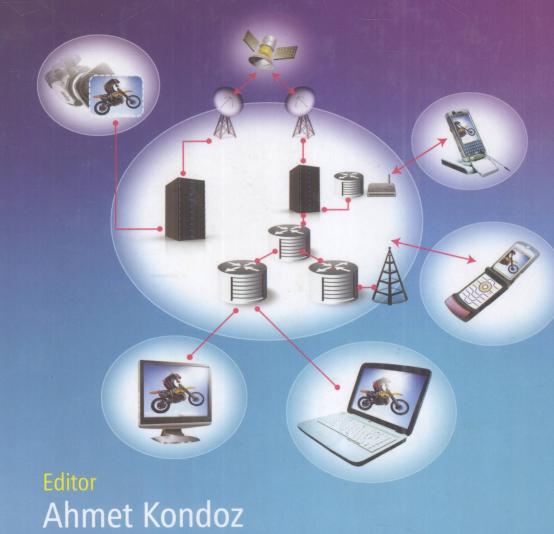
# Visual Media Coding and Transmission



WILEY

TN919.8 V834

## Visual Media Coding and Transmission

#### **Ahmet Kondoz**

Centre for Communication Systems Research, University of Surrey, UK







This edition first published 2009 © 2009 John Wiley & Sons Ltd.

Registered office

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

 $\bigcirc$ 1998,  $\bigcirc$ 2001,  $\bigcirc$ 2002,  $\bigcirc$ 2003,  $\bigcirc$ 2004. 3GPP<sup>TM</sup> TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them. They are subject to further modifications and are therefore provided to you 'as is' for information purposes only. Further use is strictly prohibited.

#### Library of Congress Cataloging-in-Publication Data

Kondoz, A. M. (Ahmet M.)

Visual media coding and transmission / Ahmet Kondoz.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-470-74057-6 (cloth)

 Multimedia communications.
 Video compression.
 Coding theory.
 Data transmission systems.
 Title. TK5105.15.K65 2009

621.382'1-dc22

2008047067

A catalogue record for this book is available from the British Library.

ISBN 9780470740576 (H/B)

Set in 10/12pt Times New Roman by Thomson Digital, Noida, India. Printed in Great Britain by CPI Antony Rowe, Chippenham, Wiltshire

# Visual Media Coding and Transmission

## VISNET II Researchers

#### UniS

Omar Abdul-Hameed Zaheer Ahmad Hemantha Kodikara Arachchi Murat Badem Janko Calic Safak Dogan Erhan Ekmekcioglu Anil Fernando Christine Glaser Banu Gunel Huseyin Hacihabiboglu Hezerul Abdul Karim Ahmet Kondoz Yingdong Ma Marta Mrak Sabih Nasir Gokce Nur Surachai Ongkittikul Kan Ren Daniel Rodriguez Amy Tan Eeriwarawe Thushara Halil Uzuner Stephane Villette Rajitha Weerakkody Stewart Worrall

#### HHI

Peter Eisert Jürgen Rurainsky Anna Hilsmann Benjamin Prestele David Schneider

Lasith Yasakethu

Philipp Fechteler Info Feldmann Jens Güther Karsten Grüneberg Oliver Schreer Ralf Tanger

#### **EPFL**

Touradj Ebrahimi Frederic Dufaux Thien Ha-Minh Michael Ansorge Shuiming Ye Yannick Maret David Marimon Ulrich Hoffmann Mourad Quaret Francesca De Simone Carlos Bandeirinha Peter Vaida Ashkan Yazdani Gelareh Mohammadi Alessandro Tortelli Luca Bonardi Davide Forzati

#### IST

Fernando Pereira João Ascenso Catarina Brites Luis Ducla Soares Paulo Nunes Paulo Correia José Diogo Areia José Quintas Pedro Ricardo Martins

#### **UPC-TSC**

Pere Joaquim Mindan José Luis Valenzuela Toni Rama Luis Torres Francesc Tarrés

#### **UPC-AC**

Jaime Delgado Eva Rodríguez Anna Carreras Rubén Tous

#### TRT-UK

Chris Firth
Tim Masterton
Adrian Waller
Darren Price
Rachel Craddock
Marcello Goccia
Ian Mockford
Hamid Asgari
Charlie Attwood
Peter de Waard
Jonathan Dennis
Doug Watson
Val Millington
Andy Vooght

#### **TUB**

Thomas Sikora Zouhair Belkoura Juan Jose Burred Michael Droese Ronald Glasberg Lutz Goldmann Shan Jin Mustafa Karaman Andreas Krutz Amjad Samour

#### **TiLab**

Giovanni Cordara Gianluca Francini Skjalg Lepsoy Diego Gibellino

#### **UPF**

Enric Peig Víctor Torres Xavier Perramon

#### **PoliMi**

Fabio Antonacci Calatroni Alberto Marco Marcon Matteo Naccari Davide Onofrio Giorgio Prandi Riva Davide Francesco Santagata Marco Tagliasacchi Stefano Tubaro Giuseppe Valenzise

#### **IPW**

Stanisław Badura Lilla Bagińska Jarosław Baszun Filip Borowski Andrzej Buchowicz Emil Dmoch Edyta Dabrowska Grzegorz Galiński Piotr Garbat Krystian Ignasiak Mariusz Jakubowski Mariusz Leszczyński Marcin Morgoś Jacek Naruniec Artur Nowakowski Adam Ołdak Grzegorz Pastuszak

Andrzej Pietrasiewicz Adam Pietrowcew Sławomir Rymaszewski Radosław Sikora Władysław Skarbek Marek Sutkowski Michał Tomaszewski Karol Wnukowicz

#### **INECS Porto**

Giorgiana Ciobanu
Filipe Sousa
Jaime Cardoso
Jaime Dias
Jorge Mamede
José Ruela
Luís Corte-Real
Luís Gustavo Martins
Luís Filipe Teixeira
Maria Teresa Andrade
Pedro Carvalho
Ricardo Duarte
Vítor Barbosa

### **Preface**

VISNET II is a European Union Network of Excellence (NoE) in the 6th Framework Programme, which brings together 12 leading European organizations in the field of Networked Audiovisual Media Technologies. The consortium consists of organizations with a proven track record and strong national and international reputations in audiovisual information technologies. VISNET II integrates over 100 researchers who have made significant contributions to this field of technology, through standardization activities, international publications, conferences workshop activities, patents, and many other prestigious achievements. The 12 integrated organizations represent 7 European states spanning across a major part of Europe, thereby promising efficient dissemination and exploitation of the resulting technological development to larger communities.

This book contains some of the research output of VISNET II in the area of Advanced Video Coding and Networking. The book contains details of video coding principles, which lead to advanced video coding developments in the form of scalable coding, distributed video coding, non-normative video coding tools, and transform-based multi-view coding. Having detailed the latest work in visual media coding, the networking aspects of video communication are presented in the second part of the book. Various wireless channel models are presented, to form the basis for following chapters. Both link-level quality of service (QoS) and cross-network transmission of compressed visual data are considered. Finally, context-based visual media content adaptation is discussed with some examples.

It is hoped that this book will be used as a reference not only for some of the advanced video coding techniques, but also for the transmission of video across various wireless systems with well-defined channel models.

Ahmet Kondoz University of Surrey VISNET II Coordinator

## Glossary of Abbreviations

3GPP

3rd Generation Partnership Project

AA

Adaptation Authorizer

**ADE** 

Adaptation Decision Engine

**ADMITS** 

Adaptation in Distributed Multimedia IT Systems

**ADTE** 

Adaptation Decision Taking Engine

AE

Adaptation Engine

**AES** AIR

Adaptation Engine Stack Adaptive Intra Refresh

API

Application Programming Interface

**AOoS** 

Adaptation Quality of Service Aspect-Scale-Context

ASC AV

Audiovisual

AVC

Advanced Video Coding

**BLER** Block Error Rate

**BSD** 

Bitstream Syntax Description

**BSDL** 

Bitstream Syntax Description Language

CC

Convolutional Coding

CC

Creative Commons

CC/PP

Composite Capabilities/Preferences Profile

CD

Coefficient Dropping

CDN CIF

Content Distribution Networks Common Intermediate Format Context Broker Architecture

CoBrA

Context-Driven Adaptation of Mobile Services

CoDAMoS CoOL

Context Ontology Language

**CoGITO** 

Context Gatherer, Interpreter and Transformer using Ontologies

**CPU** 

Central Processing Unit

**CROSLOCIS** 

Creation of Smart Local City Services

CS/H.264/AVC Cropping and Scaling of H.264/AVC Encoded Video

CxP

Context Provider

DAML Directory Access Markup Language

DANAE Dynamic and distributed Adaptation of scalable multimedia content

in a context-Aware Environment

dB Decibel DB Database

DCT Discrete Cosine Transform

DI Digital Item

DIA Digital Item Adaptation
DID Digital Item Declaration

DIDL Digital Item Declaration Language

DIP Digital Item Processing

Distributed Systems and Computer Networks

DPRL Digital Property Rights Language
DRM Digital Rights Management

DS Description Schemes

EC European Community

EIMS ENTHRONE Integrated Management Supervisor

FA Frame Adaptor FD Frame Dropping

FMO Flexible Macroblock Ordering

FP Framework Program

gBS Generic Bitstream Syntax

HCI Human-Computer Interface HDTV High-Definition Television

HP Hewlett Packard

HTML HyperText Markup Language

IEC International Electrotechnical Commission

IETF Internet Engineering Task Force

IBM International Business Machines Corporation

iCAP Internet Content Adaptation Protocol

IPR Intellectual Property Rights
IROI Interactive Region of Interest

ISO International Organization for Standardization

IST Information Society Technologies

ITEC Department of Information Technology, Klagenfurt University

JPEG Joint Photographic Experts Group JSVM Joint Scalable Video Model

MDS Multimedia Description Schemes

MB Macroblock

MDS Multimedia Description Schemes
MIT Massachusetts Institute of Technology

MOS Mean Opinion Score

MP3 Moving Picture Experts Group Layer-3 Audio (audio file format/extension)

MPEG Motion Picture Experts Group

MVP Motion Vector Predictor

NAL Network Abstract Layer
NALU Network Abstract Layer Unit
NoE Network of Excellence

ODRL Open Digital Rights Language
OIL Ontology Interchange Language

OMA Open Mobile Alliance

OSCRA Optimized Source and Channel Rate Allocation

OWL Web Ontology Language

P2P Peer-to-Peer

PDA Personal Digital Assistance PSNR Peak Signal-to-Noise Ratio

QCIF Quarter Common Intermediate Format

QoS Quality of Service QP Quantization Parameter

RD Rate Distortion

RDF Resource Description Framework

RDB Reference Data Base
RDD Rights Data Dictionary
RDOPT Rate Distortion Optimization
REL Rights Expression Language

ROI Region of Interest

SECAS Simple Environment for Context-Aware Systems

SNR Signal-to-Noise Ratio

SOAP Simple Object Access Protocol

SOCAM Service-Oriented Context-Aware Middleware

SVC Scalable Video Coding

TM5 Test Model 5

UaProf User Agent Profile

UCD Universal Constraints Descriptor
UED Usage Environment Descriptions

UEP Unequal Error Protection

UF Utility Function

UI User Item

UMA Universal Multimedia Access

UMTS Universal Mobile Telecommunications System

URI Uniform Resource Identifiers

UTRAN UMTS Terrestrial Radio Access Network

VCS Virtual Collaboration System

VoD Video on Demand VOP Video Object Plane VQM Video Quality Metric

W3C World Wide Web Consortium WAP Wireless Access Protocol

WCDMA Wideband Code Division Multiple Access

WDP Wireless Datagram Protocol
WLAN Wireless Local Area Network
WML Website Meta Language

WiFi Wireless Fidelity (IEEE 802.11b Wireless Networking)

XML eXtensible Markup Language XrML eXtensible rights Markup Language

XSLT eXtensible Stylesheet Language Transformations

## Contents

VISNET II Researchers				xii	
Preface Glossary of Abbreviations				X	
				xvi	
1	Intr	oductio	duction		
2	Vid	eo Cod	ing Principles	-	
	2.1		luction	-	
	2.2	Redu	ndancy in Video Signals	5	
	2.3		amentals of Video Compression	8	
		2.3.1	Video Signal Representation and Picture Structure	8	
		2.3.2	Removing Spatial Redundancy	ç	
		2.3.3	Removing Temporal Redundancy	14	
		2.3.4	Basic Video Codec Structure	16	
	2.4	Advar	nced Video Compression Techniques	17	
		2.4.1	Frame Types	17	
			MC Accuracy	19	
			MB Mode Selection	20	
			Integer Transform	21	
			Intra Prediction	22	
			Deblocking Filters	22	
		2.4.7	Multiple Reference Frames and Hierarchical Coding	24	
			Error-Robust Video Coding	24	
	2.5		Codec Standards	28	
			Standardization Bodies	28	
			ITU Standards	29	
			MPEG Standards	29	
			H.264/MPEG-4 AVC	31	
	2.6		sment of Video Quality	31	
		2.6.1	Jest Jest Land Land Brantation	31	
	2 7	2.6.2	Objective Performance Evaluation	32	
	2.7 Conclusions			35	
	References			36	

vi Contents

3	Scal	able Video Coding	39
	3.1	Introduction	39
		3.1.1 Applications and Scenarios	40
	3.2	Overview of the State of the Art	41
		3.2.1 Scalable Coding Techniques	42
		3.2.2 Multiple Description Coding	45
		3.2.3 Stereoscopic 3D Video Coding	47
	3.3	Scalable Video Coding Techniques	48
		3.3.1 Scalable Coding for Shape, Texture, and Depth for 3D Video	48
		3.3.2 3D Wavelet Coding	68
	3.4	Error Robustness for Scalable Video and Image Coding	74
		3.4.1 Correlated Frames for Error Robustness	74
		3.4.2 Odd–Even Frame Multiple Description Coding	
		for Scalable H.264/AVC	82
		3.4.3 Wireless JPEG 2000: JPWL	91
		3.4.4 JPWL Simulation Results	94
		3.4.5 Towards a Theoretical Approach for Optimal Unequal	
		Error Protection	96
	3.5	Conclusions	98
	Refe	rences	99
4		ributed Video Coding	105
	4.1	Introduction	105
		4.1.1 The Video Codec Complexity Balance	106
	4.2	Distributed Source Coding	109
		4.2.1 The Slepian–Wolf Theorem	109
		4.2.2 The Wyner–Ziv Theorem	110
		4.2.3 DVC Codec Architecture	111
		4.2.4 Input Bitstream Preparation – Quantization and Bit Plane Extraction	112
		4.2.5 Turbo Encoder	112
		4.2.6 Parity Bit Puncturer	114
		4.2.7 Side Information	114
		4.2.8 Turbo Decoder	115
		4.2.9 Reconstruction: Inverse Quantization	116
		4.2.10 Key Frame Coding	117
	4.3	Stopping Criteria for a Feedback Channel-based Transform	
		Domain Wyner–Ziv Video Codec	118
		4.3.1 Proposed Technical Solution	118
		4.3.2 Performance Evaluation	120
	4.4	Rate-distortion Analysis of Motion-compensated Interpolation	
		at the Decoder in Distributed Video Coding	122
		4.4.1 Proposed Technical Solution	122
		4.4.2 Performance Evaluation	126
	4.5	Nonlinear Quantization Technique for Distributed Video Coding	129
		4.5.1 Proposed Technical Solution	129
		4.5.2 Performance Evaluation	132

Contents

	4.6	Symm	netric Distributed Coding of Stereo Video Sequences	134			
		4.6.1	Proposed Technical Solution	134			
		4.6.2	Performance Evaluation	137			
	4.7	Studying Error-resilience Performance for a Feedback Channel-based					
		-	form Domain Wyner–Ziv Video Codec	139			
			Proposed Technical Solution	139			
		4.7.2	Performance Evaluation	140			
	4.8		ling the DVC Decoder for Error-prone Wireless Channels	144			
			Proposed Technical Solution	145			
		4.8.2	Performance Evaluation	149			
	4.9	Error Concealment Using a DVC Approach for Video					
			ning Applications	151			
			Proposed Technical Solution	152			
		4.9.2		155			
	4.10	Concl	·	158			
				159			
	11010	References					
5	Non-	norma	tive Video Coding Tools	161			
	5.1	Introd		161			
	5.2	Overv	iew of the State of the Art	162			
		5.2.1	Rate Control	162			
		5.2.2	Error Resilience	164			
	5.3	Rate Control Architecture for Joint MVS Encoding and Transcoding					
		5.3.1		165 165			
		5.3.2		166			
		5.3.3		169			
		5.3.4	Conclusions	171			
	5.4		llocation and Buffer Control for MVS Encoding Rate Control	171			
		5.4.1		171			
		5.4.2		172			
		5.4.3	1	177			
		5.4.4	Conclusions	179			
	5.5		al Rate Allocation for H.264/AVC Joint MVS Transcoding	179			
	5.5	5.5.1		179			
		5.5.2	ů ů	180			
		5.5.3		181			
		5.5.4	Conclusions	182			
	5.6		o-temporal Scene-level Error Concealment for Segmented Video	182			
	5.0	5.6.1	Problem Definition and Objectives	182			
		5.6.2	Proposed Technical Solution	183			
		5.6.3	Performance Evaluation	187			
		5.6.4	Conclusions				
	5.7		tegrated Error-resilient Object-based Video	188			
	5.1		g Architecture	100			
		5.7.1	Problem Definition and Objectives	189			
		5.7.1	Proposed Technical Solution	189 189			
		2.1.4	i roposca iechnicai solullon	109			

		5.7.3	Performance Evaluation	195
			Conclusions	195
	5.8		bust FMO Scheme for H.264/AVC Video Transcoding	195
		5.8.1	J	195
			Proposed Technical Solution	195
		5.8.3	J	197
		5.8.4		198
	5.9		lusions	199
	Refe	erences		199
6	Tra	nsform	-based Multi-view Video Coding	203
	6.1		luction	203
	6.2	MVC	Encoder Complexity Reduction using a Multi-grid	
		Pyran	nidal Approach	205
		6.2.1	Problem Definition and Objectives	205
			Proposed Technical Solution	205
		6.2.3	Conclusions and Further Work	208
6.3 Inter-view Prediction using Reconstructed Disparity				
		Inform	mation	208
			Problem Definition and Objectives	208
			Proposed Technical Solution	208
			Performance Evaluation	210
			Conclusions and Further Work	211
	6.4		-view Coding via Virtual View Generation	212
			Problem Definition and Objectives	212
			Proposed Technical Solution	212
		6.4.3	Performance Evaluation	215
		6.4.4	The state of the s	216
	6.5		delay Random View Access in Multi-view Coding Using	
		a Bit	Rate-adaptive Downsampling Approach	216
		6.5.1	Problem Definition and Objectives	216
		6.5.2	Proposed Technical Solution	216
			Performance Evaluation	219
		6.5.4	Conclusions and Further Work	222
	Refe	erences		222
7	Intr	oductio	on to Multimedia Communications	225
	7.1		uction	225
	7.2	State	of the Art: Wireless Multimedia Communications	228
		7.2.1	QoS in Wireless Networks	228
		7.2.2	Constraints on Wireless Multimedia Communications	231
		7.2.3	Multimedia Compression Technologies	234
		7.2.4	Multimedia Transmission Issues in Wireless Networks	235
		7.2.5	Resource Management Strategy in Wireless Multimedia	
			Communications	239

Contents

	<b></b>	~				
		Conc		244		
	Ref	erences		244		
8	Trouble of the state of the sta					
	8.1			247		
	8.2		S/EGPRS Channel Simulator	247		
		8.2.1		247		
		8.2.2	200 200 200 200 Perception	250		
		8.2.3	January English Medical Bescription	252		
		8.2.4	zini zini zini zini zini zini zini zini	256		
		8.2.5	====== Sinte Edyer Simulator	261		
		8.2.6	Interface Bata I tow Model	268		
		8.2.7	- San time Oblant Emiliary	270		
	0.2	8.2.8	Conclusion	271		
	8.3		S Channel Simulator	272		
		8.3.1	UMTS Terrestrial Radio Access Network (UTRAN)	272		
		8.3.2	UMTS Physical Link Layer Model Description	279		
		8.3.3	Model Verification for Forward Link	290		
		8.3.4	UMTS Physical Link Layer Simulator	298		
		8.3.5	Performance Enhancement Techniques	307		
		8.3.6	UMTS Radio Interface Data Flow Model	309		
		8.3.7	San	312		
	8.4		Conclusion	313		
	0.4	WINIA	AX IEEE 802.16e Modeling	316		
			Introduction	316		
		0.4.2	WIMAX System Description	317		
		8.4.4	Physical Layer Simulation Results and Analysis	323		
	0.5		Error Pattern Files Generation	324		
				328		
		rances	$\operatorname{Idix}$ . $\operatorname{E}_{b}/\operatorname{N}_{o}$ and $\operatorname{DPCH}_{-}\operatorname{E}_{c}/\operatorname{I}_{o}$ Calculation	329		
	Kele	rences		330		
9						
<ul> <li>8.5 Conclusions</li> <li>8.6 Appendix: E<sub>b</sub>/N<sub>o</sub> and DPCH_E<sub>c</sub>/I<sub>o</sub> Calculation</li> <li>References</li> <li>Enhancement Schemes for Multimedia Transmission over</li> <li>Wireless Networks</li> </ul>						
	9.1			333		
		9.1.1	The Indiovisial Requirements	333		
		9.1.2	Video Transmission over Mobile Communication Systems	335		
			Circuit-switched Bearers	339		
		9.1.4	Packet-switched Bearers	348		
		9.1.5	Video Communications over GPRS	350		
		9.1.6	GPRS Traffic Capacity	351		
		9.1.7	Error Performance	354		
		9.1.8	Video Communications over EGPRS	357		
		9.1.9	Traffic Characteristics	357		
			Error Performance	358		
		9.1.11	Voice Communication over Mobile Channels	359		

ix

		9.1.12	Support of Voice over UMTS Networks	360
		9.1.13	Error-free Performance	361
			Error-prone Performance	362
			Support of Voice over GPRS Networks	362
			Conclusion	363
	9.2	Link-le	evel Quality Adaptation Techniques	365
		9.2.1	Performance Modeling	365
		9.2.2		367
			Distortion Modeling	368
		9.2.4		368
		9.2.5	Energy-optimized UEP Scheme	369
		9.2.6	Simulation Setup	370
			Performance Analysis	372
			Conclusion	373
	9.3	Link A	Adaptation for Video Services	373
		9.3.1		374
		9.3.2	Link Adaptation for Real-time Video Communications	379
		9.3.3	Link Adaptation for Streaming Video Communications	389
		9.3.4	Link Adaptation for UMTS	396
		9.3.5	Conclusion	402
	9.4	User-c	entric Radio Resource Management in UTRAN	403
		9.4.1	Enhanced Call-admission Control Scheme	403
		9.4.2	Implementation of UTRAN System-level Simulator	403
		9.4.3	Performance Evaluation of Enhanced CAC Scheme	410
	9.5	Conclu	asions	411
	Refe	rences		413
10	Oual	lity Ont	timization for Cross-network Media Communications	44=
IV	10.1	Introdu	unitzation for Cross-network Media Communications	417
			c Inter-networked QoS-optimization Infrastructure	417
	10.2	10.2.1	State of the Art	418
			Generic of QoS for Heterogeneous Networks	418
	10.3	Implen	nentation of a QoS-optimized Inter-networked Emulator	420
	10.5	10 3 1	Emulation System Physical Link Layer Simulation	422
			Emulation System Frystcat Link Layer Simulation Emulation System Transmitter/Receiver Unit	426
			QoS Mapping Architecture	428
			General User Interface	428
	10.4		nances of Video Transmission in Inter-networked Systems	438
	10.4	10/1	Experimental Setup	442
			Test for the EDGE System	442
			Test for the UMTS System	443
				445
	10.5	Conclu	Tests for the EDGE-to-UMTS System	445
		ences	210112	452
	Kelel	CHCCS		453