PCS 2007

26th PICTURE CODING SYMPOSIUM

7-9 NOVEMBER 2007
LISBOA, PORTUGAL

ABSTRACT BOOK

ISBN: 978-989-8109-04-0

www.pcs2007.org

TABLE OF CONTENTS

WELCOME FROM THE CHAIRMEN

Welcome to Lisboa for the 26th Picture Coding Symposium, PCS'2007 !

Picture Coding Symposium (PCS) is an international forum devoted specifically to advancements in visual data coding. Since 1969, PCS has provided the meeting place for the visual coding community: industry, research, academia and users.

After Cambridge - USA'69, Raleigh - USA'70, Lafayette - USA'71, Los Angeles - USA'72, Goslar - Germany'74, Asilomar - USA'76, Tokyo - Japan'77, Ipswich - England'79, Montreal - Canada'81, Davis - USA'83, Cesson Sévigné - France'84, Tokyo - Japan'86, Stockholm - Sweden'87, Torino - Italy'88, Cambridge - USA'90, Tokyo - Japan'91, Lausanne - Switzerland'93, Sacramento - USA'94, Melbourne - Australia'96, Berlin - Germany'97, Portland - USA'99, Seoul - Korea'01, Saint-Malo - France'03, San Francisco - USA'04 and Beijing - China'06, the Picture Coding Symposium takes place, in 2007, at Instituto Superior Técnico, in the city of Lisboa, Portugal.

And it was the time to pay us this visit ! Also, some five centuries ago, Portugal was one of the first European countries to sail the oceans in search of new horizons, new cultures, new trades and also new technologies. Today, image and video coding related technologies are also in the search of new horizons, new services and new applications to be provided to the customers, so that the technologies and capabilities being developed find themselves useful and successful. To reach this goal, the right blend between academia and industry is truly a need in these days, without which research and development will not open new frontiers.

More than everything, we expect you to enjoy your stay in Lisboa with us ! We hope you enjoy our city, our old monuments and also the new architecture, the great food, the friendly people, the flavors, tastes and smells of the most western part of Europe.

We also hope you enjoy PCS'2007 ! We hope you enjoy the sessions and the breaks, the talks by the 'big names' and the posters by the young and enthusiastic students, the discussions in the panels and also in the corridors, the days of work and the nights of entertainment; most of all, we expect you to enjoy the people and the friendly environment, and find it useful to learn, to exchange ideas and grow as experts in the various areas related to image and video coding.

To achieve this goal, we have put together a program that we believe very exciting and with high technical quality. The 140 papers accepted have been organized into 6 oral sessions and 12 poster sessions. The papers are coming from 34 countries which really shows the increasing success and global dimension of the Picture Coding Symposium. For curiosity, the top contributors were China with 46 papers, Japan with 31 papers, UK with 21 papers, South Korea with 18 papers, US with 17 papers, Germany with 14

papers, Italy with 12 papers and France with 10 papers.

To further increase the quality of the PCS papers, and for the first time in PCS, the accepted papers have gone through a second review to check that the reviewers' comments have been taken into account by the authors and have, in effect, contributed to improve the final quality of the PCS papers. Since reviewers are so important for the overall quality of any conference, for the first time, PCS will attribute a Best Reviewer Award with the precise intention to highlight the relevance of good reviewing.

Finally, and again with the intention to stimulate good research work, the Picture Coding Symposium will attribute a Best Paper Award, for the first time in the context of a Special Session, where four candidate papers previously selected by the PCS International Steering Committee will have the opportunity to show their excellence and contend to be recognized as the best.

In the PCS program, we will also have four outstanding invited talks in areas which are nowadays very hot, very relevant and thus moving very fast:

- *H.264/AVC and its Extensions: How Close is this Family?,* by Anthony Vetro, Mitsubishi Electric Research Labs, USA
- *From Picture Coding to Image Understanding: Finding the Object of Interest,* by Tsuhan Chen, Carnegie Mellon University, USA
- *DCT, Wavelets and X-lets: The Quest for Image Representation, Approximation and Compression,* by Martin Vetterli, EPFL, Switzerland and UC Berkeley, USA
- *Efficient Representation of Sound Images: Recent Developments in Parametric Coding of Spatial Audio,* by Jürgen Herre, Fraunhofer Institute for Integrated Circuits, Erlangen, Germany

Finally, because we also want time for lively discussions and controversial arguing, involving all the symposium participants, three panels on challenging topics have been organized involving people who, due to their experience, knowledge and vision can only teach us every time we listen to them:

- *Video Content Protection: Does it Really Matter?*, with Edward J. Delp (chairman), Purdue University, USA; Inald Lagendijk, Delft University of Technology, The Netherlands; Niels Rump, Rightscom, UK; Ian Burnett, University of Wollongong, Australia; and Touradj Ebrahimi, EPFL, Switzerland
- *Distributed Video Coding: Trends and Challenges*, with Touradj Ebrahimi (chairman), EPFL, Switzerland; Bernd Girod, Stanford University, USA; Martin Vetterli, EPFL, Switzerland and UC Berkeley, USA; Pier Luigi Dragotti, Imperial College London, UK; and Zixiang Xiong, Texas A&M University, USA
- *Multimedia: a World of Possibilities or just a Buzzword ?*, with Alan Hanjalic (chairman), Delft University of Technology, The Netherlands; Andrea Cavallaro, Queen Mary, University of London, UK; Anthony Vetro, Mitsubishi Electric

Research Labs, USA; Jürgen Herre, Fraunhofer Institute for Integrated Circuits, Erlangen, Germany; and Tsuhan Chen, Carnegie Mellon University, USA

But, life is more than technology, as interesting as it can be … So we prepared for you two nights of entertainment, culture, gastronomy and also a relaxing environment to chat with old and new friends and colleagues. In the first night of the Symposium, all participants are invited to a Welcome Reception. For the occasion, the best Portuguese Academic Musical group (Tuna), composed by IST students, will entertain us with lively and funny moments. The welcome reception will be held at the Belém Tower, very likely the most famous ex-libris of Lisboa. In the second night of the Symposium, PCS'2007 participants are invited to a lively Brazilian style dinner. The dinner will be preceded by a guided visit to the Oceanário de Lisboa. The Oceanário was build for the World Exhibition, held in Lisboa in 1998, and celebrates life on Earth with a stunning display of living creatures, evoking the complexity of diversity inhabiting the Global Ocean and the vital role which it plays on the health and evolution of our Planet.

We hope you will find this mix of activities and styles, exciting and stimulating, rewarding us with the conclusion: It was worthwhile to go to PCS'2007 in Lisboa !

As the Symposium Chairs, we would like to thank all of those who have contributed to the success of PCS'2007. First, the authors of the papers submitted to the symposium, who are the real motivation for organizing a Symposium. Second, the invited speakers, panel members, and session chairpersons, who accepted the invitation to be here, hence providing an important added value to the Symposium. Third, the members of the International Steering Committee and the Reviewing Committee, who provided the necessary expertise and performed an excellent work in shaping the symposium from the technical point of view. Also to our sponsors REN, TAP, CGD, Fundação Luso-Americana and Fundação Calouste Gulbenkian. Last, but definitely not the least, to all people involved in the organization of the Symposium, especially our students for their very hard and dedicated work.

On behalf of the PCS'2007 organizers, we warmly welcome you to join us and hope all of you will enjoy this Symposium in Lisboa.



Fernando Pereira, Paulo Lobato Correia and Luis Ducla Soares

WELCOME TO LISBOA

Centuries ago, an idyllic settlement took root on the hills, close to a river and not far from a vast ocean. The hills offered protection against the wrath of enemies and the river a mean of survival so that the settlement soon became a good place to live. Legend tells us Ulysses founded it during one of his famous journeys. Another chronicle reports that Elisha, grandson of Abraham, founded the city around 3259 BC. But the local conditions are sufficient proof that it was well populated as far back as Neolithic times. The tribes that lived there originally were Iberians, a stocky, stout, dark-haired race.

Then the Phoenicians arrived by sea (c. 1200 BC) and founded a port by the name of Alis Ubbo ("serene harbour"). After the Phoenicians, the Celts came (first half of sixth century BC) and mixed with the local tribes, resulting in a people known as Celt Iberians. Later, around 205 BC, the Romans invaded the city and its name changed to Olisipo.

Apparently it was also named Felicitas Iulia, after Julius Caesar, but this name was short lived. Time passed and once more new invaders reached that site on the hills with its calm river and mild climate. Among them were the Visigoths and the Moors.

It was only in 1147 that D. Afonso Henriques, who was to be the first king of the territory Portus Cale, later known as Portugal, conquered it with the help of thousands of Crusaders from northern Europe. King Afonso III was very fond of Lissabona, as it was called then, and a century later he made what is now Lisboa the capital of the country. It stands on the most westerly point of the European continent, where the river Tagus flows into the Atlantic Ocean. The climate is probably one of the mildest in Europe.

During the Portuguese Age of Discovery (15th and 16th centuries), the city was endowed with bigger and more luxurious buildings. In 1775, it was devastated by an earthquake, in which it is believed that about 30,000 lives were lost and more than 9000 buildings destroyed. The city was rebuilt rapidly for that age under the direction of the Marquis of Pombal. The downtown part of Lisboa has retained its eighteenth century character.

Over the centuries, Lisboa has witnessed a great deal of change and development although many traces of the city's rich and troubled past have been retained. In 1998, Lisboa hosted the last World Exhibition of the last century, Expo´98, which was subordinated to the theme of the oceans because of Portugal's importance as a nation involved in discovery and sea-voyages. After Expo'98, the Parque das Nações area emerged as the paradigm for a new Lisboa in the 21st century … These and a wealth of other attractions make Lisboa such a pleasant and hospitable city.

For more information, check:

- http://www.atl-turismolisboa.pt/
- http://www.cm-lisboa.pt/turismo/
- http://www.portugalvirtual.pt/_tourism/costadelisboa/index.html

ABOUT PORTUGAL

Portugal is one of the world's oldest countries and very likely the oldest in Europe. It became an independent nation in 1143, and has maintained its independence and national boundaries from the 13th century to this day, except for a brief period under Spanish rule between 1580 and 1640.

Portugal became a republic on the 5th October 1910 and is an established democracy today. Portugal's geographic location at the southwestern part of Continental Europe assures quick and easy access not only to the European market, but also to the eastern shore of the United States and the African Continent.

Population density is highest in the Lisboa area, in the northern city of Porto and in other coastal cities.

Area: 92.646-sq. km, including Madeira, Porto Santo and Azores.

Climate: Portugal has a mild weather, winters are never too cold and summers are always moderately hot. Lightweight clothing is suitable from June to September. A topcoat and umbrella are advisable for autumn and winter.

Language: Portuguese.

Location: Portugal is the westernmost country in Continental Europe. It is bounded by Spain on the north and east and by the Atlantic Ocean on the south and west.

Population: Approximately 10.300.000, of which 9.800.000 live in Continental Portugal, 240.000 in Azores, 250.000 in Madeira and Porto Santo.

Time Zone: GMT

For more information, check:

- http://www.visitportugal.com
- http://www.portugalvirtual.pt

HISTORY

Edition	Year	Location	Date
1.	1969	Cambridge / USA	April 2-4
2.	1970	Raleigh / USA	September 10-11
3.	1971	Lafayette / USA	October 28-29
4.	1973	Los Angeles / USA	January 22-24
5.	1974	Goslar / Germany	August 26-28
6.	1976	Asilomar / USA	August 28-30
7.	1977	Tokyo / Japan	August 29-31
8.	1979	Ipswich / England	July 9-11
9.	1981	Montreal / Canada	June 3-5
10.	1983	Davis / USA	March 28-30
11.	1984	Cesson Sévigné / France	July 3-5
12.	1986	Tokyo / Japan	April 2-4
13.	1987	Stockholm / Sweden	June 9-11
14.	1988	Torino / Italy	September 12-14
15.	1990	Cambridge / USA	March 26-28
16.	1991	Tokyo / Japan	September 2-4
17.	1993	Lausanne / Switzerland	March 17-19
18.	1994	Sacramento / USA	September 21-23
19.	1996	Melbourne / Australia	March 13-15
20.	1997	Berlin / Germany	September 10-12
21.	1999	Portland / USA	Apr 21-23
22.	2001	Seoul / Korea	April 25-27
23.	2003	Saint-Malo / France	April 23-25
24.	2004	San Francisco / USA	December 15-17
25.	2006	Beijing / China	April 24-26
26.	2007	Lisboa / Portugal	November 7-9

ORGANIZING COMMITTEE

GENERAL CHAIR

Fernando Pereira, IST-IT, Lisbon, Portugal

PROGRAM CHAIR

Paulo Lobato Correia, IST-IT, Lisbon, Portugal

PUBLICATIONS AND SPECIAL SESSIONS CHAIR

Luis Ducla Soares, ISCTE-IT, Lisbon, Portugal

MEMBERS

Paula Queluz, IST-IT, Lisbon, Portugal

Paulo Nunes, ISCTE-IT, Lisbon, Portugal

João Ascenso, ISEL-IT, Lisbon, Portugal

Catarina Brites, IST-IT, Lisbon, Portugal

Tomás Brandão, ISCTE-IT, Lisbon, Portugal

INTERNATIONAL STEERING COMMITTEE

Tsuhan Chen,
Carnegie Mellon University, USA

Chang Wen Chen,
Florida Institute of Technology, USA

Edward J. Delp,
Purdue University, USA

Touradj Ebrahimi,
EPFL, Switzerland

Pascal Frossard,
EPFL, Switzerland

Moncef Gabbouj,
Tampere University of Technology, Finland

Bernd Girod,
Stanford University, USA

Christine Guillemot,
INRIA, France

Yun He,
Tsinghua University, China

Yo-Sung Ho,
Gwangju Institute of Science and Technology, South Korea

Tom Huang,
Univ. of Illinois at Urbana-Champaign, USA

Aggelos K. Katsaggelos,
Northwestern University, USA

Shipeng Li,
Microsoft Research Asia, China

Joern Ostermann,
Technical Univ. of Hannover, Germany

Fernando Pereira,
IST-IT, Portugal

Takahiro Saito,
Kanagawa University, Japan

Mihaela van der Schaar,
UCLA, USA

Ralf Schaefer,
Fraunhofer - HHI, Germany

Thomas Sikora,
Technical University of Berlin, Germany

Masayuki Tanimoto,
Nagoya University, Japan

Murat Tekalp,
University of Rochester, USA

Lu Yu,
Zhejiang University, China

REVIEWING COMMITTEE

Tomasz Adamek, DCU, Ireland
Maria Teresa Andrade, INESC Porto-FEUP, Portugal
Yiannis Andreopoulos, QMUL, UK
Marc Antonini, CNRS, France
John Arnold, ADFA, Australia
Xavi Artigas, UPC, Spain
João Ascenso, ISEL-IT, Portugal
Pedro Assunção, IP Leiria-IT, Portugal
Luigi Atzori, University of Cagliari, Italy
Mauro Barni, UNISI, Italy
Jenny Benois-Pineau, LABRI, France
Jesus Bescos, UAM, Spain
Tomás Brandão, ISCTE-IT, Portugal
Catarina Brites, IST-IT, Portugal
Josep R. Casas, UPC, Spain
Andrea Cavallaro, QMUL, UK
Homer Chen, National Taiwan University, Taiwan
Miguel Coimbra, FCUP-IT, Portugal
Saman Cooray, DCU, Ireland
Paulo Lobato Correia, IST-IT, Portugal
Klaus Diepold, TUM, Germany
Safak Dogan, University of Surrey, UK
Pier Luigi Dragotti, Imperial College, UK
Frédéric Dufaux, EPFL, Switzerland
Adriana Dumitras, Apple, USA
Peter Eisert, Fraunhofer - HHI, Germany
Alex Eleftheriadis, Layered Media, USA
Sérgio Faria, IP Leiria-IT, Portugal
Anil Fernando, University of Surrey, UK
Mário Figueiredo, IST-IT, Portugal
Michael Frater, UNSW, Australia
Pascal Frossard, EPFL, Switzerland
Moncef Gabbouj, TUT, Finland
Daniele Giusto, University of Cagliari, Italy
Christine Guillemot, IRISA, France
Hermann Hellwagner, Klagenfurt University, Austria
Arianne Hinds, InfoPrint Solutions Company, USA
Michael Isnardi, Sarnoff Corporation, USA
Ebroul Izquierdo, QMUL, UK
André Kaup, Universität Erlangen-Nürnberg, Germany
Hemantha Kodikara Arachchi, University of Surrey, UK

Yiannis Kompatsiaris, ITI, Greece
R. (Inald) L. Lagendijk, TU Delft, The Netherlands
Haibo Li, Umea University, Sweden
João Magalhães, Imperial College, UK
Detlev Marpe, Fraunhofer - HHI, Germany
Ferran Marques, UPC, Spain
Jose Martinez, UAM, Spain
Pierangelo Migliorati, UNIBS, Italy
Adrian Munteanu, ETRO, Belgium
Noel Murphy, DCU, Ireland
António Navarro, UA-IT, Portugal
Paulo Nunes, ISCTE-IT, Portugal
Noel O'Connor, DCU, Ireland
Tobias Oelbaum, Technische Universität München, Germany
Fernando Pereira, IST-IT, Portugal
Fernando Perez-Gonzalez, Univ. Vigo, Spain
Béatrice Pesquet-Popescu, ENST, France
Armando Pinho, University of Aveiro, Portugal
Josep Prades, UPV, Spain
Ricardo Queiroz, Universidade de Brasilia, Brazil
Paula Queluz, IST-IT, Portugal
Iain Richardson, RGU, UK
Roberto Rinaldo, UNIUD, Italy
Nuno Roma, INESC-ID, Portugal
Paul Salama, Indiana University - Purdue University, USA
Philippe Salembier, UPC, Spain
Luis Salgado, UPM, Spain
Jorge Salvador Marques, IST-ISR, Portugal
Peter Schelkens, ETRO, Belgium
Heiko Schwarz, Fraunhofer - HHI, Germany
Eduardo Silva, Universidade Federal do Rio de Janeiro, Brazil
Aljoscha Smolic, Fraunhofer - HHI, Germany
Luis Ducla Soares, ISCTE-IT, Portugal
Thomas Stockhammer, Nomor, Germany
Karsten Suehring, Fraunhofer - HHI, Germany
Marco Tagliasacchi, Polimi, Italy
Francesc Tarrés, UPC, Spain
Christian Timmerer, ITEC - Klagenfurt University, Austria
Luis Torres, UPC, Spain
Alexis Tourapis, Dolby Laboratories, USA
Stefano Tubaro, Polimi, Italy
Deepak Turaga, IBM, USA

Rik Van de Walle, Univ. Gent, Belgium
David Varodayan, Stanford Univ., USA
Anthony Vetro, MERL, USA
Mathias Wien, RWTH Aachen, Germany
Stewart Worrall, University of Surrey, UK
Lu Yu, Zhejiang University, China

SPONSORS

TECHNICAL SPONSORS

Institute of Electrical and Electronics Engineers (IEEE)
The IEEE is a non-profit, technical professional association of more than 360,000 individual members in approximately 175 countries. The IEEE is a leading authority in technical areas ranging from computer engineering, biomedical technology and telecommunications, to electric power, aerospace and consumer electronics, among others.

European Association for Signal Processing (EURASIP)
EURASIP's goal is to improve communication between groups and individuals that work within the multidisciplinary, fast growing field of Signal Processing in Europe and elsewhere, and to exchange and disseminate information in the field all over the world.

Institution of Engineering and Technology (IET)
The IET is a new institution formed in spring 2006 by the coming together of the IEE and IIE. The IET aims to serve a global community engaged in engineering and technology, providing a knowledge network accessible whenever or wherever you choose.

SPONSORS

TAP Air Portugal (TAP)

TAP is a Portuguese airline market leader, which celebrated its 60th anniversary on March 14th of 2005; the date on which it also became part of the **Star Alliance**, the largest global alliance of airlines in the world.

With its Hub in Lisbon - a privileged access platform to Europe - at the crossroads with other Continents, the Company currently flies to 43 destinations, throughout 25 countries in Africa and North and South America; a region in which TAP stands out as being the leading operational European transportation company to Brazil.

TAP is the PCS'2007 official carrier.

Redes Energéticas Nacionais (REN)

REN is responsible for the electricity transmission in Portugal. REN was set up as a company in August 1994 in result of the split-up of EDP - Electricidade de Portugal, S.A. to which it already belonged as its Operational Department. Its history, however, goes back to 1947, when the pioneer of electricity transmission in Portugal, therefore its original predecessor, CNE - Companhia Nacional de Electricidade, S.A.R.L. was founded.

Nowadays, REN has diversified its activity areas, covering also the telecommunications and gas businesses.

Caixa Geral de Depósitos (CGD)

CGD has been at the forefront of Portuguese economic and social development over the last 127 years. It has been a benchmark operator in the Portuguese banking sector in terms of its support to households, companies and domestic institutions, since its formation on 10 April 1876. It is currently the parent company of a modern financial group, prepared to cater for the needs and expectations of millions of customers and the challenges of market globalisation.

Calouste Gulbenkian Foundation
The Calouste Gulbenkian Foundation is a Portuguese private institution of public utility whose statutory aims are in the fields of arts, charity, education and science. Created by a clause in Calouste Sarkis Gulbenkian's will, the Foundation's statutes were approved in 1956.

Luso-American Foundation
The Luso-American Development Foundation is a private, financially independent Portuguese institution.
Its main goal is to contribute towards Portugal's development by providing financial and strategic support for innovative projects by fostering cooperation between Portuguese and American civil society.

ORGANIZERS

Instituto Superior Técnico (IST)
IST is the engineering school of the Technical University of Lisbon. IST mission is to contribute to the development of society by promoting a higher education of outstanding quality in the areas of Engineering, Science & Technology, at the undergraduate and graduate levels, and by developing the Research & Development activities essential to ensure the highest standard of education.

Instituto de Telecomunicações (it)
it is a research institute whose mission is to create and disseminate scientific knowledge in the field of telecommunications.
it is actively involved in fundamental and applied research in telecommunications both at national and international level. Simultaneously **it** is committed to foster higher education and training, by hosting and tutoring graduate and postgraduate students.

REGISTRATION

REGISTRATION FEES

	Early Registration (before 21 Sept 2007)	Late registration
Student	200 €	300 €
Standard regular	450 €	550 €
Standard for IEEE, EURASIP, IET members	400 €	500 €

Registration fees include all conference sessions, conference CD Proceedings, refreshments, lunches, welcome reception and conference dinner.

Important: Each paper needs to be associated with a standard registration. Also, a standard registration can be associated to a maximum of 3 papers.

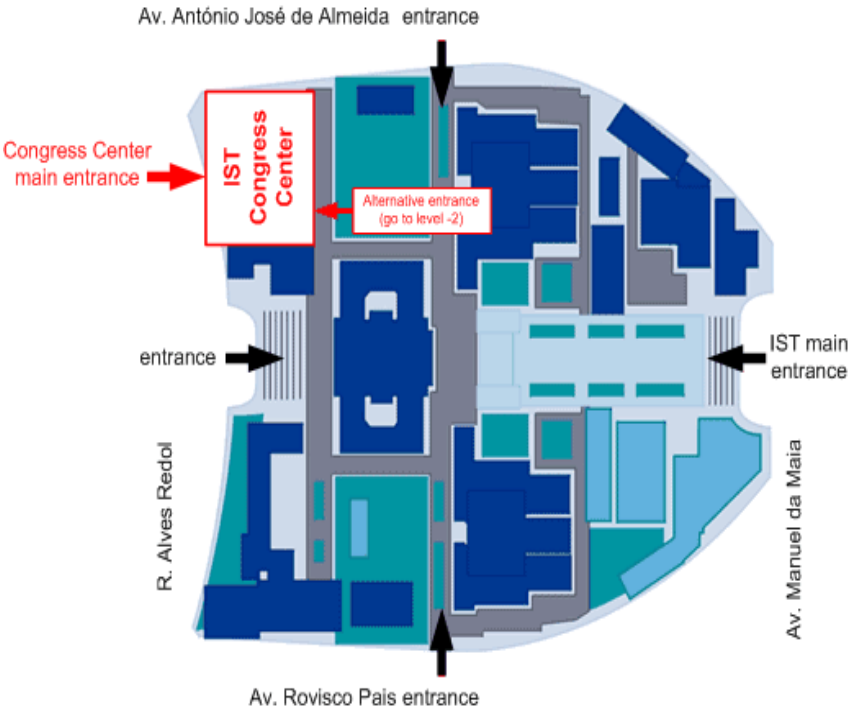
Refund policy: Processing fee of 50 EUR before October 5, 2007. No refunds after October 5, 2007.

VENUE

INSTITUTO SUPERIOR TÉCNICO (IST)

PCS 2007 will be held at the IST Congress Center.

The Congress Center is located in level *-2* of the Civil Engineering Building, inside the IST campus – see map below.



IST street address:

Instituto Superior Técnico
Av. Rovisco Pais
1049-001 Lisboa

ARRIVAL IN LISBON: FROM THE AIRPORT TO IST

IST is located in the center of Lisbon. For the precise location, see the map below:

Lisbon Airport is located about 5 km to the North of the IST area.
To reach IST, you may use either the Airport bus (AeroBus), the regular bus line 22 or a taxi (more info available in www.ana.pt).

AEROBUS

On a daily basis, buses on CARRIS line 91, called the **AeroBus**, runs between the Lisbon Airport and the city centre.

Service begins at 7:45 and ends at 20:15. AeroBuses depart every 20 minutes and are faster than regular buses, having fewer stops. The ticket may be purchased from the driver as you board the bus.

On-board ticket: 1.30€
Ticket for 1-day travel: 3.35€
Source: Carris (http://www.carris.pt)

Take the Bus at the sidewalk near the Arrivals, and exit the bus at the "Avenida da República" bus stop. Travel time to IST can be 20-40 minutes depending on the traffic. It takes about 10 minutes to walk from the bus stop to IST.

REGULAR BUSES

Alternatively, you may take the regular **bus line 22**, which has a bus stop located at one of the IST entrances. When compared with the AeroBus, it will probably take more time, but you'll be closer to IST when leaving the bus. The on-board ticket costs 1.30€.

TAXI

There are two taxi stands within the perimeter of the airport's main terminal, one at the arrivals floor and the other at the departures floor (above). **It is strongly suggested that you take a taxi at the departures floor.** Ask the airport's staff on how to go from the arrivals to the departures floor.

All taxis are metered.

- The fare on the taxi meter should read 2.00€ (daytime pick-up) or 2.50€ (night time pick-up).
- The transport of baggage or animals implies a surcharge of 1.60€.
- Taxis that are called by phone have an additional cost of 0.80€.
- Services on weekends and holidays, as well as night services from 21:00 to 6:00 cost an additional 20%.

You may get updated information and obtain a Taxi Voucher at the Tourism Information Counter in the atrium of the Airport Arrivals Hall.

Travel cost to IST should be around 10 € (depending on hour and if luggage boot is used) and it should take from 15 to 20 minutes to get there, depending on the traffic.

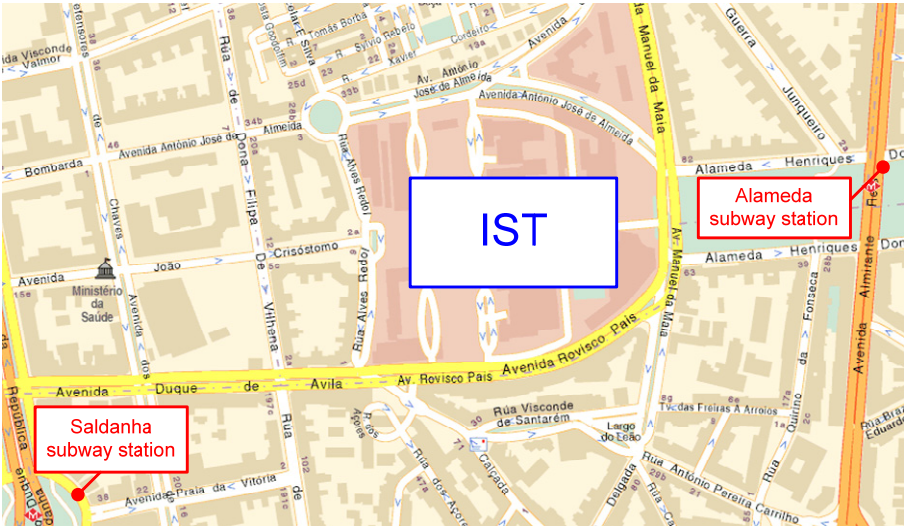
Most taxis are beige but some are black and green. If the customer is pleased with the service, a tip is usually given to the driver.

GETTING TO IST FROM OTHER LOCATIONS

Subway

Lisbon's subway system, called **Metro**, is probably the fastest way to travel between city locations. The network consists of 4 lines, identified by colors: Blue, Green, Red and Yellow (see map below).

From IST, the closest subway stations are:
ALAMEDA (5 minutes walk) – located at the junction between the GREEN and RED line;
SALDANHA (10 minutes walk) – located in the YELLOW line route.



Subway tickets can be purchase at ticket offices or at automatic vending machines (both are located at the subway station entrance halls). There are several types of tickets. The most popular are:

Single ticket (one zone): 0.75€
Return ticket (one zone): 1.35€
1-day network ticket Carris/Metro: 3.35€

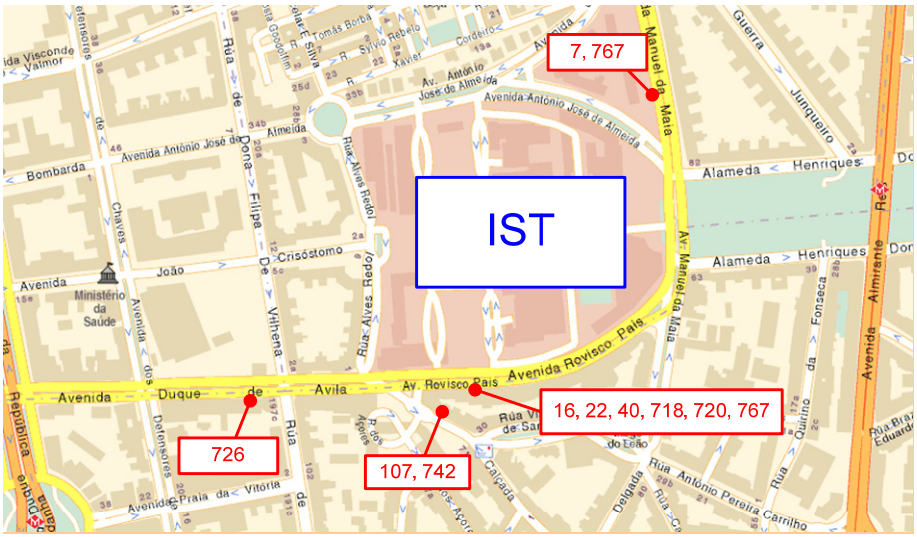
The latter is valid for buses and subway. This type of ticket requires the "7 Colinas" card, which costs 0.50€ and can be purchased at any Metro ticket office, or at the automatic vending machines.

You can get more information at Metro's website: http://www.metrolisboa.pt/

Bus
Several buses from Carris have stops located near IST:

Bus lines 16, 22, 40, 718, 720, 767 have stops at IST entrances.
Bus lines 7, 107, 726, 742 have stops located close to IST.

Check the map below to get an idea on the bus stops location.



Whatever means of transport you choose, IST is easily reachable. You can take any public transport that passes near “Alameda D. Afonso Henriques” or “Saldanha”.

HOTEL INFORMATION

RECOMMENDED HOTEL LIST

Holiday Inn ****
Av. António José de Almeida, 28-A
1000-044 Lisbon, Portugal
Phone: +351 21 0044000 | Fax: +351 21 7936672
Webpage: http://www.holiday-inn.com/lisbonprt

Getting to IST
Walking: This hotel is very close to IST. When leaving the hotel you'll see IST in the opposite corner of the square. Just cross the square and follow the street on the right (R. Alves Redol). The entrance for the IST congress center is located in the left side of that street (less than 100 meters from the hotel).

Hotel A. S. Lisboa ***
Av. Almirante Reis, 188
1000 - 055 Lisbon, Portugal
Phone: +351 21 8429360 | Fax: +351 21 8429374
Webpage: http://www.hotel-aslisboa.com/main_ing.htm

Getting to IST
Walking: When you exit the hotel, you should see IST at the top of the hill to the right. Take a 5 minute walk to IST along "Alameda D. Afonso Henriques". If you don't feel like going up the hill, you may also get on the bus lines 16 or 718, which have a nearby stop.

Comfort Inn Embaixador ***
Av. Duque de Loulé, 73
1050-088 Lisbon, Portugal
Phone: +351 21 3194000 | Fax: +351 21 3557596
Webpage: http://www.grupo-continental.com/hotels/ciembaixador.htm

Getting to IST
By bus: Get to a bus stop and get on buses 720 or 22. Both have a bus stop near an IST entrance. After getting in the bus, it should take about 10-20 minutes to reach IST, depending on traffic.
By subway: Go down the street (Av. Duque de Loulé) until you get to the large square Marquês de Pombal (around 200 meters from the hotel). In this square, find the nearest subway entrance and get the YELLOW line train to SALDANHA. The complete journey should take about 30 minutes but you may want to use this alternative if traffic is intense.

Hotel Príncipe Lisboa ***
Av. Duque d'Ávila, 201
1050-082 Lisbon, Portugal
Phone: +351 21 3592050 | Fax: +351 21 3592055
Webpage: http://www.hotelprincipelisboa.com

Getting to IST
By bus: Go to the nearest bus stop in the hotel street (Av. Duque d'Ávila) and take bus lines 718 or 742.
Walking: You may prefer a 15 to 20 minutes walk to IST. In that case, just follow the street (Av. Duque d'Ávila) to the east direction until you reach IST.

Altis Park Hotel ****
Av. Engenheiro Arantes de Oliveira, 9
1900-221 Lisbon, Portugal
Phone: +351 21 8434200 | Fax: +351 21 8460838
Webpage: http://www.altishotels.com/default_en.asp?pag=hotel_aph

Getting to IST
By bus: Get on bus line 40. Once in the bus, it should take about 10 to 20 minutes to reach IST, depending on traffic.
By subway: Find the nearby entrance to OLAIAS subway station and take a train to ALAMEDA. Then take a 5 minute walk to IST along Alameda D. Afonso Henriques (west direction).

Hotel Florida ****
R. Duque de Palmela, 34
1250-098 Lisbon, Portugal
Phone: +351 21 3576145 | Fax: +351 21 3141347
Webpage: http://www.hotel-florida.pt

Getting to IST
By bus: When leaving the hotel, go to your right until you reach R. Braancamp (it's quite close). Find the nearest bus stop and take bus line 22. Once you're in the bus, it should take less than 20 minutes to reach IST.
By subway: When leaving the hotel, go to your right until you reach R. Braancamp (it's quite close). Then turn right again into "Marquês de Pombal" square direction. Once you reach the square, you should easily find an entrance to the subway. Take a YELLOW LINE train to SALDANHA, and then walk to IST. Total traveling time should be about 20 minutes.

OTHER HOTELS
(Located at less than 30 minutes from IST)

Holiday Inn Continental ****
Rua Laura Alves, 9
1069-169 Lisbon, Portugal
Phone: +351-210 046 000 | Fax: 351-217 973 669
Webpage: http://www.holiday-inn.com/lisbon-contine

Getting to IST
By bus: When leaving the hotel, go to your left until you reach a large avenue (Av. de Berna). Find one of the bus stops located near the church in the opposite side of the street. Take bus lines 16 or 726 to IST. Once you're in the bus it should take about 10 minutes to reach IST (maybe more time if traffic is intense).

Hotel Roma ***
Av. de Roma, 33
1749-074 Lisbon, Portugal
Phone: +351 21 7932244 | Fax: +351 21 7932981
Webpage: http://www.hotelroma.pt/

Getting to IST
By bus: Take bus lines 7 or 767 to IST. Once you're in the bus it should take less than 10 minutes to reach IST.
By subway: Enter ROMA subway station and take the subway train to ALAMEDA station. Then walk to IST along Alameda D. Afonso Henriques (west direction).

Ibis Lisboa Saldanha **
Avenida Casal Ribeiro, 23
1000-090 Lisbon, Portugal
Phone: +351 21 3191690 | Fax: +351 21 3191699
Webpage: http://www.ibishotel.com/ibis/fichehotel/gb/ibi/2117/fiche_hotel.shtml

Getting to IST
By bus: When leaving the hotel, walk down the street until you find a bus stop. Take bus lines 22 or 720. Once you're in the bus, it should take less than 10 minutes to reach IST.
Walking: Cross the street (Av. Casal Ribeiro) and find the junction with Av. Defensores de Chaves (a V-shaped junction). Follow Av. Defensores de Chaves straight ahead until you reach Av. Duque de Ávila (it's the second crossroad that you find and you should see a construction site due to subway network expansion). Turn to

your right and go straight ahead through Av. Duque de Ávila until you reach IST. The whole walk should take less than 20 minutes.

The map below shows the hotels location:

Hotel location map

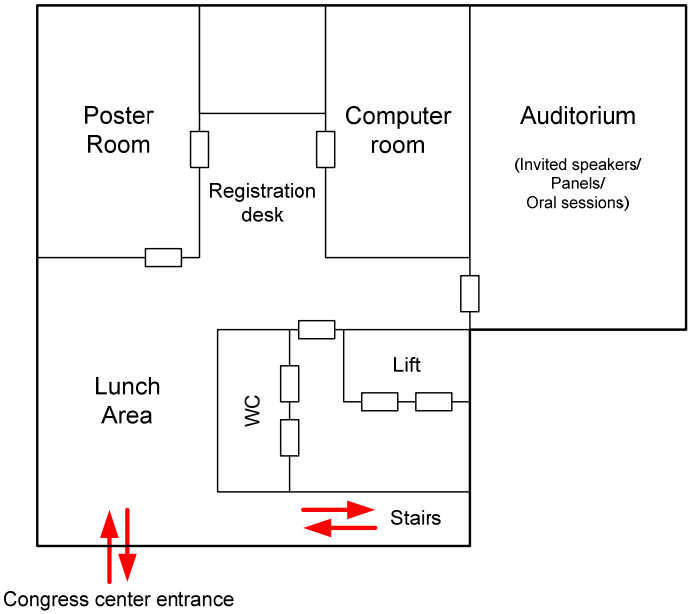
ON-SITE ACTIVITIES

REGISTRATION DESK

The registration desk opening times are:

- Tuesday, November 6, from 14:00 to 19:00.
- Wednesday and Thursday, November 7-8, from 8:00 to 18:00.
- Friday, November 9, from 9:00 to 16:00.

The registration desk is located at the IST Congress Center, see map below.



COFFEE BREAKS

There will be two coffee breaks per day, one in the morning, and one in the afternoon; see the program for the precise times.

The coffee breaks will be served in the lunch area of the IST Congress Center, see map above.

LUNCHES

Lunches will be offered during the three Symposium days. They are included in the registration fee and will be served at the IST Congress Center hall.

INTERNET ACCESS

There is a computer room equipped to allow Internet access. It will be opened during the Symposium days, from 9:00 to 18:00. The computer room is located at the IST Congress Center (room 02.3), see map in previous page.

A wireless network connection will also be available.

Login credentials:

Username: pcs2007
Password: mRyxYS

Configuration instructions:

1. Browse available wireless networks and select as SSID: "eduroam-guest";
2. Set IP to automatic (DHCP). This is usually the default setting, so you may probably skip this step;
3. Open your browser and try to access any external website. You will be automatically redirected to the page: https://wifi.ist.utl.pt/index_auth.php. Follow the link "Web based login" at the top of the page concerning short-time, conference and meetings accounts. Enter the above username/password when requested;
4. After step 3 you may freely browse and access the Internet. You may need to repeat the above steps if you close your browser or if the connection times out.

SOCIAL ACTIVITIES

WELCOME RECEPTION

Wednesday, November, 7, 18:30, Belém Tower
The first night of the Symposium, all participants are invited to a Welcome Reception. This will be a night of history, and entertainment, in a beautiful and relaxing environment to chat with old and new friends and colleagues. For the occasion, the best Portuguese Academic Musical group (Tuna), composed by IST students, will entertain us with lively and funny moments.

The welcome reception will be held at the Belém Tower, very likely the most famous ex-libris of Lisboa. The Belém Tower was declared a World Heritage Site by the UNESCO and is one of the world's most beautiful military constructions. Built as a stronghold to assure the defense of the entrance to the river Tagus, an identical fortress should have been constructed on the other side of the river, but this never happened. Recently restored, it is austere inside (but from the windows you can admire dazzling views of the city and river), in contrast with its outside architecture, profusely decorated with symbols of the Manueline style and others inspired by the Portuguese travels to discover the world.

Transportation will be arranged to go to the Welcome Reception. The buses will leave from the Symposium site at 18:30 and should return to the same place at about 22:00.

DINNER

Thursday, November, 8, 18:30, Parque das Nações

In the second night of the Symposium, PCS'2007 participants are invited to a lively Brazilian style dinner at the restaurant "Buffalo Grill" at Parque das Nações, Lisboa, Telephone: 218922740.

The banquet will be preceded by a guided visit to the Oceanário de Lisboa. The Oceanário was built for the World Exhibition, held in Lisboa in 1998, and celebrates life on Earth with a stunning display of living creatures, evoking the complexity of diversity inhabiting the Global Ocean and the vital role which it plays on the health and evolution of our Planet.

Transportation will be arranged to go to the Oceanário. The buses will leave from the workshop site at 18:30 and should return to the same place at about 23:00. The restaurant is within walking distance from the Oceanário (more precisely, on the other side of the small water front where the Oceanário is located).

GENERAL INFORMATION

ENTRY REQUIREMENTS IN PORTUGAL

For entering the Portuguese territory, the following requirements must be met:

1. Citizens of the European Union, Andorra, Iceland, Liechtenstein, Norway and Switzerland need an identity card (or passport) to enter Portugal.
2. Citizens from countries that signed an agreement with Portugal (e.g. USA, Canada) do not require a visa to come into Portugal, as long as they come as tourists. They need a valid passport and their stay must not be longer than 90 days.
3. Citizens from other countries are required a visa to come into Portugal. Visas can be obtained at the nearest Portuguese Embassy or Consulate.

Under the terms of the Convention Implementing the Schengen Agreement, flights between Schengen states are considered to be internal flights and passengers do not need to obtain another visa.

CURRENCY

Since 1 January 1999, the official currency in Portugal is the Euro (€). The Euro is also the official currency of the 13 European Union member states that are currently on the Eurozone: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Slovenia and Spain.

One Euro is divided up into 100 Cents. There are eight coins denominated in 1, 2, 5, 10, 20 and 50 Cents and 1 and 2 Euro values. The design on one side of the coins is common to all Eurozone countries (the European side) while the reverse side features a national symbol. All Euro coins can be used across the Eurozone, regardless of their national origin. The notes are distinguishable by both their color and size and are available in the following denominations: 5, 10, 20, 50, 100, 200 and 500 Euro.

BANKS

Banks are typically open to public from 8:30 to 15:00, from Monday to Friday.
Inside the IST campus, there are 2 banks: Caixa Geral de Depósitos (CGD) and Banco Português de Investimento (BPI), where you can easily deal with currency exchange, deposits and all other bank operations.

The use of ATM machines is widely spread in Portugal (often several machines can be found in the same street), and credit card use is also very common.
Most ATM machines accept the following cards: VISA/Plus, Mastercard, American Express, Eurocard, Cirrus, 4B, Bancontact/Mistercash, Bancomat3/Eufiserv, Clau and Eurocheque.

ATM machines work 24 hours a day and are commonly known as **Multibanco**; you can use your Multibanco card to withdraw money or pay for services (at shops, main train stations, etc). Check with your bank for details on debit and credit cards compatible with those used in Portugal.

POST OFFICES

During the week, Post Offices are open from 9:00 to 18:00. Most are closed on weekends (for more information, please visit http://www.ctt.pt).

SHOPPING

Shops are usually open from Monday to Friday, from 9:00 or 10:00 to 18:00 or 19:00. Some close for lunch, from 13:00 to 15:00. Several shops are also open on the weekends, especially on Saturday mornings.

Shopping malls are usually not much more expensive, and are open during the weekends. The biggest shopping malls in Lisboa are Colombo and Vasco da Gama (both are easily accessible by subway). Near IST, in the Saldanha square, there are smaller malls such as Atrium and Residence.

Big shopping areas can be found in the following areas of Lisboa: Chiado, Rossio, Rua Augusta (Baixa de Lisboa – downtown) and Avenida da Liberdade.

RESTAURANTS
Restaurants are usually open for lunch from 12:00 to 15:00 and for dinner from 19:00 to 22:00. A large number of restaurants have longer opening hours.
You can also have cheaper meals (6€ to 10€) on most "snack-bars" during the lunch time.

PHARMACIES
Working days: Open from 9:00 to 13:00 and 15:00 to 19:00.
Saturdays: Open from 9:00 to 13:00.
Pharmacies are identified by a green cross at the door, which is illuminated at night. When the pharmacy is closed, there will be a notice on the window showing the nearest pharmacy on duty, which is open 24 hours a day.

EMERGENCIES
In case of emergency, dial 112 at any time of the day or night and wherever you are in Portugal.

EU citizens have the same benefits as Portuguese citizens in accordance with Community Law. For access to health services, EU citizens who are not resident in Portugal should bring their European Health Insurance Card (more information is available at http://ec.europa.eu/employment_social/healthcard/index_en.htm).

ELECTRICAL POWER
Voltage: 230 Volt at a frequency of 50 Hertz.
All sockets follow European standards. To use American type plugs, a 230 volt transformer should be applied together with an adapter plug.

TELECOMMUNICATIONS
Public telephones can be found at airports, seaports, subway, bus and railway stations, in the streets, bars and restaurants. There are two types of telephones that you are likely to come across: coin and card operated.

International Calls: say, for example, that you want to call London: you dial directly 00, followed by the country code (44), then area code (206), and finally the telephone number.

SECURITY
Criminal rates in Lisbon are low. However, just like everywhere else, empty streets should be avoided (especially at night). You should also be aware of pickpockets in crowded areas. If you run into trouble, call the police using the 112.

WATER
It is perfectly safe to drink tap water in private homes and in public places.

PORTUGUESE FOR DUMMIES ;-)

The Portuguese alphabet is similar to the English one, but without the letters ‘k’, ‘w’, and ‘y’. Additionally, there is the symbol ‘ç’ and the orthographic accents ^, ~, ` and ´, which are used on some vowels.

The vowels can be pronounced differently depending on stress and on the use of markers. The following table depicts some examples:

Vowel	English approximation
a … stressed	**A**fter, **a**rmy
a … unstressed	**A**bout, **a**gain
e … stressed	S**e**t, g**e**t
e … unstressed	Mak**e**, tak**e**
i	S**ee**k, s**i**ck
o … stressed	C**au**ght, m**o**re
o … unstressed	T**o**,
u	M**oo**d, w**oo**d

To make different sounds, vowels can be combined, and the orthographic accents can be used. The next table shows the most frequent combinations, with English pronunciation examples – note that some of them are not exact, because equivalent sounds are not found in the English language.

Symbol	Approximation	Notes
à, á	**A**fter, **a**rmy	always spelled as stressed ‘a’
ã	**An**swer, **un**der	
ão	T**ow**n, r**ou**nd	
ãe	M**ai**n	
é	S**e**t	always spelled as stressed ‘e’
ê	S**e**nd	
ei	K**a**te, s**ay**	
í	S**i**ck	
ô	C**o**ld	
õe	B**oi**ng	
ou	B**oa**t	
ú	M**oo**d	

A very general stressing rule (keep in mind that there are many exceptions…):
If there's a vowel marked with ^, ~, ´ or `, the stress always falls into the corresponding syllable
else if the word ends in 'i' or 'u' the stress falls on the last syllable
else if the word ends in 'a', 'e' or 'o' the stress falls on the second-to-last syllable
else if the word ends in 'm', 'r', 'l', 'z' the stress falls on the last syllable

Most consonants sound like as if they were spelled in English. However, a few consonants (e.g. C, G, R, S) have additional rules that must also be considered.

Consonant	Rules	Example	Notes
B		**B**ase	
C	before a, o, u	**C**ase	
	before e, i	**C**ycle	
Ç		**C**ycle	only used before 'a', 'o' and 'u'
Ch		**Sh**ame	
D		**D**ay	
F		**F**og	
G	before a, o, u	**G**ame	
	before e i	**J**ane	
H		-	the H followed by vowel is always silent
J		**J**ournal	
L	General case	**L**and	
	At the end of a word	Tunne**l**	
LH		Mi**ll**ion	similar to how "torti**ll**a" is pronounced
M		**M**ore	
N		**N**ever	
NH		Ca**n**yon	similar to how "El ni**ñ**o" is pronounced
P		**P**ace	
Q		-	the Q is always followed by an 'u'
Qu	Before a, o	**Ko**ala	the 'u' is spelled
	Before e, i	**C**ase	the 'u' is mute
R	At the begin of the word	**R**ime	
	Other cases	Ca**r**, Pa**r**k	
RR		**R**ime	used between vowels
S	at the begin of the word	**S**un	
	between vowels	Ca**s**e	
	before a consonant	Co**s**mic	
SS		**S**un	used between vowels

T		**T**oy	
V		**V**alue	
X	At the begin of the word	Fi**sh**,	there are lots of exceptions…
	Between vowels	Ma**z**e	
	Before a consonant	E**x**pect	
Z	General case	**Z**oo,	
	At the end of the word	re**s**ort	
		Mi**s**ter	

The next table depicts a few Portuguese words for starters. To pronounce them correctly, you should consider the previous tables and check the spelling with local people.

English	**Português**
Greetings	**Cumprimentos**
Hello!	Olá!
Good bye	Adeus
See you later	Até logo
Good morning	Bom dia
Good afternoon	Boa tarde
Good evening Good night	Boa noite
How are you?	Como está?
Thank you	Obrigado (male speaker) Obrigada (female speaker)
Please	Por favor
Excuse me	Com licença
Food and drinks	**Comida e bebida**
Cow	Vaca
Pork	Porco
Lamb	Borrego
Chicken	Frango, galinha
Turkey	Peru
Water	Água
Juice	Sumo
Tea	Chá
Coffee	Café
Beer	Cerveja
Wine	Vinho

Other useful words	**Outras palavras úteis**
Where is …?	Onde é … ? (for locations) Onde está … ? (for objects or people)
Car	Carro
Bus	Autocarro
Train	Comboio
Entrance	Entrada
Exit	Saída
Open	Aberto
Closed	Fechado

PICTURE CODING SYMPOSIUM 2007 - SCHEDULE OVERVIEW

Wednesday, 7 November 2007

8:45-9:00	Opening Session		Auditorium
9:00-10:00	Keynote Session	*H.264/AVC and its Extensions: How Close is this Family?* by Anthony Vetro, Mitsubishi Electric Research Labs, USA	Auditorium
10:00-10:20	Coffee Break		Atrium
10:00-11:25	Poster Session	Coding Standards 1	Poster Room
10:00-11:25	Poster Session	Coding Standards 2	Poster Room
11:25-12:45	Special Oral Session	Best Paper Award	Auditorium
12:45-14:00	Lunch		Atrium
14:00-15:20	Oral Session	Scalable Video Coding	Auditorium
15:20-16:35	Panel Session	*Video Content Protection: Does it Really Matter?* chaired by Edward J. Delp, Purdue University, USA	Auditorium
16:35-16:55	Coffee Break		Atrium
16:35-18:00	Poster Session	Processing for Applications	Poster Room
16:35-18:00	Poster Session	Scalable Coding	Poster Room
18:30-22:00	Welcome Reception		Belém Tower

Thursday, 8 November 2007

8:30-9:30	Keynote Session	*From Picture Coding to Image Understanding: Finding the Object of Interest* by Tsuhan Chen, Carnegie Mellon University, USA	Auditorium
9:30-10:50	Oral Session	3D and Multiview Video Coding	Auditorium
10:50-11:10	Coffee Break		Atrium
10:50-12:00	Poster Session	Multiview Video Coding	Poster Room
10:50-12:00	Poster Session	Image Coding	Poster Room
12:00-13:00	Keynote Session	*DCT, Wavelets and X-lets: The Quest for Image Representation, Approximation and Compression* by Martin Vetterli, EPFL, Switzerland and UC Berkeley, USA	Auditorium
13:00-14:00	Lunch		Atrium
14:00-15:20	Oral Session	Distributed Source Coding	Auditorium
15:20-16:35	Panel Session	*Distributed Video Coding: Trends and Challenges* chaired by Touradj Ebrahimi, EPFL, Switzerland	Auditorium
16:35-16:55	Coffee Break		Atrium
16:35-18:00	Poster Session	Implementation	Poster Room
16:35-18:00	Poster Session	Distributed Video Coding	Poster Room
18:30-23:00	Visit to Oceanário de Lisboa and PCS'2007 Dinner		Parque das Nações

Friday, 9 November 2007

9:00-10:00	Keynote Session	*Efficient Representation of Sound Images: Recent Developments in Parametric Coding of Spatial Audio* by Jürgen Herre, Fraunhofer Institute for Integrated Circuits, Erlangen, Germany	Auditorium
10:00-10:20	Coffee Break		Atrium
10:00-11:25	Poster Session	Video Coding	Poster Room
10:00-11:25	Poster Session	Error Resilience	Poster Room
11:25-12:45	Oral Session	Quality Assessment	Auditorium
12:45-14:00	Lunch		Atrium
14:00-15:20	Oral Session	Network Aware Coding	Auditorium
15:20-16:35	Panel Session	*Multimedia: a World of Possibilities or just a Buzzword?* chaired by Alan Hanjalic, Delft University of Technology, The Netherlands	Auditorium
16:35-16:55	Coffee Break		Atrium
16:35-18:00	Poster Session	Analysis for Coding 1	Poster Room
16:35-18:00	Poster Session	Analysis for Coding 2	Poster Room

SESSIONS AND PAPER ABSTRACTS

Opening Session

Location: Auditorium
Time: 08:45 - 09:00
Chair: Fernando Pereira, *Instituto Superior Técnico - Instituto de Telecomunicações*

Keynote Session WedAM1: H.264/AVC and its Extensions: How Close is this Family?

Location: Auditorium
Time: 09:00 - 10:00
Chair: Fernando Pereira, *Instituto Superior Técnico - Instituto de Telecomunicações, Portugal*

H.264/AVC AND ITS EXTENSIONS: HOW CLOSE IS THIS FAMILY?

Anthony Vetro, *Mitsubishi Electric Research Labs, United States*

Abstract: H.264/AVC is a state-of-the-art video coding standard that has ushered in a new benchmark for video coding efficiency. The design offers a powerful set of coding tools and provisions for network-friendly representation of the video. Building on the success of the base specification, a number of extensions have been recently developed to meet the demands of various application needs. For instance, profession applications require coding of higher bit depths and color sampling formats. A scalable video representation is useful to serve a diverse set of display and networking environments or to satisfy dynamic delivery constraints imposed during transmission. Then there is multiview video coding, which aims to enable 3D video and free-view video applications. In this talk, I will provide a brief overview of the new coding tools that have been introduced in the various extensions and summarize their performance. To understand the intimate associations among this family of tools, I will then analyze the conceptual, architectural and performance relationships among them. I will also speculate on the potential business impact of these extensions and highlight the market relationships that exist. We will find that although the application space is quite broad, the current family of coding tools is rather tight-nit. I will close this talk by identifying emerging opportunities and some possibilities for new extensions, some of which might bring this family even closer and others that appear to be more divergent.

Poster Session WedAM2: Coding Standards 1

Location: Poster Room
Time: 10:00 - 11:25

A NOVEL TWO-PASS VBR CODING ALGORITHM FOR THE H.264/AVC VIDEO CODER BASED ON A NEW ANALYTICAL R-D MODEL

Jianfei Huang, *Graduate University of Chinese Academy of Science, China*; Jun Sun, Wen Gao, *Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, China*

Abstract: In this paper, the generalized Gaussian distribution is employed to model the 16(4*4) integer transform coefficients of the residue image in H.264 videos. Then the distortion-rate function of generalized Gaussian model is analyzed as well as a flexible rate-distortion (R-D) model is developed to approximate the actual R-D function. Finally, an effective variable bit rate (VBR) algorithm for H.264/AVC is proposed, which adopts two-pass encoding to achieve a constant video quality. Experimental results show that under the same average bit rate, the proposed algorithm achieves 94% to 99% reduction in PSNR variation compared to the model of JM9.8.

ONE DIMENSIONAL TRANSFORM FOR H.264 BASED INTRA CODING

Yumi Sohn, Woo-Jin Han, *Samsung Electronics, Co., South Korea*

Abstract: Two dimensional (2D) discrete cosine transform (DCT) are widely used as the transform method in most of video and image coding area. In general, 2D separable DCT is effective for smoothed regions of which components are highly correlated. When the block to be transformed has directional edges or a few values, however, 2D transform can yield worse performance than 1D transform. This drawback is taken care of by 1D transform with line prediction this paper proposed. Line prediction between each coefficient ahead of 1D transform can reduce the repetitive residue. If it is applied with an appropriate scan order, therefore, it can bring additional bit saving of transform coefficients in entropy coding. By experiments, it is shown that the proposed framework including 1D transform and line prediction improves the overall rate-distortion (R-D) coding performance up to nearly 8% in comparison with H.264.

DYNAMIC BASIC UNIT SIZE IN RATE CONTROL FOR REAL-TIME H.264 VIDEO CODING

Sergio Sanz-Rodríguez, Darío García-García, Manuel De Frutos-López, Jesús Cid-Sueiro, *Universidad Carlos III de Madrid, Spain*

Abstract: Several rate control (RC) schemes include the basic unit (BU) layer, where the quantization parameter (QP) value can be modified within a picture to get a fine adjustment to the target bits. The BU is a group of macroblocks (MBs) which share the same QP value, and its size is set previously to the encoding process. This paper describes a RC algorithm capable of detecting the instants in the sequence encoding process where a small BU size works efficiently and, for the rest of cases, use a large one to enhance quality. Our experimental results show a great robustness in both quality and buffer control for different kind of sequences and target bit rates.

TWO-PASS RATE CONTROL FOR CONSTANT QUALITY H.264/AVC HIGH DEFINITION VIDEO CODING

Dongdong Zhang, Zhenzhong Chen, King Ngi Ngan, *CUHK, China*

Abstract: In this paper, we propose a novel rate control algorithm that adopts two-pass encoding to provide constant quality for H.264/AVC high definition videos. In the first-pass encoding, the rate and distortion information of each frame is collected to model the frame scene complexity. In the second-pass encoding, a GOP-level bit allocation scheme is designed to determine the bit rate and the expected constant distortion of each frame based on the first-pass statistics. The quantization step (Q-step) of each frame can be resolved according to the relationship between distortion and Q-step. The experimental results show that our proposed algorithm can achieve more constant quality through the whole sequence, compared to traditional Joint Model (JM) rate control algorithm.

MODE DECISION FOR H.264/AVC BASED ON SPATIO-TEMPORAL SENSITIVITY

Yukihiro Bandoh, Kazuya Hayase, Seishi Takamura, Kazuto Kamikura, Yoshiyuki Yashima, *NTT, Japan*

Abstract: It is really important to use the proper prediction mode in the H.264 encoder, since there are many more modes than the conventional methods such as MPEG-2. The representative H.264 encoders like JM and JSVM use squared error as the criterion of distortion for mode decision. However, squared error does not always correspond to the distortion from the viewpoint of subjective quality. In this paper, we investigate a mode decision method based on the human visual system in order to improve H.264 encoder. Our method was implemented on reference software JSVM and compared with the original JSVM. Experiments show that our method can achieve average bit-rate savings of the order of 4.7 to 5.7% compared to the original JSVM. We confirm that both methods yield the reconstructed images with basically the same subjective image quality.

ON OVERRIDING H.264/AVC B-SLICE PREDICTED RESIDUE CODING

Nuno Rodrigues, *Inst. Telecomunicações/ESTG IP Leiria, Portugal*; Eduardo Silva, *PEE/COPPE/DEL/Poli, Univ. Fed. Rio de Janeiro, Brazil*; Murilo Carvalho, *TET/CTC, Univ. Fed. Fluminense, Brazil*; Sérgio Faria, *Inst. Telecomunicações/ESTG IP Leiria, Portugal*; Vitor Silva, *Inst. Telecomunicações/DEEC Univ. Coimbra, Portugal*

Abstract: The use of motion compensation (both forward and bidirectional) allied with transform-quantisation based coding of the residual predicted error has been ubiquitous in video coding standards. The most recent standard, H.264/AVC, maintains this paradigm, introducing even more efficient tools to determine and compress predicted residue information. In this paper we investigate the effects of eliminating the encoding of bidirectional prediction residue in H.264/AVC. Extensive experimental tests show gains of up to 50% of the bit rate for the same PSNR value, for B slices. An analysis of this procedure also reveals an interesting relation between H.264/AVC motion estimation and dictionary-based methods, namely vector quantisation and Lempel-Ziv encoders.

ENHANCED INTRA PREDICTION USING CONTEXT-ADAPTIVE LINEAR PREDICTION

Limin Liu, *Purdue University, United States*; Yuxin (Zoe) Liu, *Hewlett-Packard Laboratories, United States*; Edward Delp, *Purdue University, United States*

Abstract: Intra prediction is a fundamental prediction type in block-based video coding, which has been adopted by the most recent finalized standard H.264/AVC. H.264/AVC has specified several intra prediction modes that predict the current block through surrounding spatial neighboring pixels in a causal window along various directions. In this paper, we investigate the use of linear prediction for intra block coding and propose a context-adaptive intra prediction approach. Specifically, we use the least square prediction and derive the linear prediction coefficients using reconstructed data. The linear prediction coefficients implicitly embed the local texture characteristics and thus the intra prediction mode is adaptively adjusted according to the local context. No extra overhead is needed for signaling the coefficients since the decoder simply repeats the same deriving process. We take this context-adaptive linear prediction as an additional intra prediction mode along with existing H.264/AVC intra prediction modes, and choose the best mode through rate-distortion optimization. By turning on the new intra prediction mode, we will demonstrate an enhanced coding efficiency performance. The downside of the proposed approach is the increased computational complexity at both the encoder and the decoder.

IMPROVED ARITHMETIC CODING IN H.264/AVC USING CONTEXT-TREE WEIGHTING METHOD

Damian Karwowski, *Poznan University of Technology, Poland*

Abstract: In this paper, an improvement of coding efficiency of the Context-based Adaptive Binary Arithmetic Coding (CABAC) is proposed for applications in Advanced Video Codecs (H.264/AVC). In the paper, more accurate estimation technique for conditional probabilities of symbols is proposed. To achieve that, the well-known data modeling technique of the Context-Tree Weighting (CTW) is used. A novel method of incorporating CTW into the H.264/AVC codec is proposed. The compression efficiency of the modified H.264/AVC codec (with CTW) has been thoroughly tested and compared against the coding efficiency of the original H.264/AVC (with CABAC). Experimental results show that the use of CTW with CABAC allows for up to 2.5% compression efficiency increase with respect to the original CABAC.

DERIVATION FOR ADAPTIVE SCAN OF INTRA PREDICTION IN VIDEO CODING

Xiaoyang Wu, Sijia Chen, Qichao Sun, Lu Yu, *Zhejiang University, China*

Abstract: In this paper, a mathematical derivation for adaptive scan of intra prediction in video coding is investigated, which is based on the elliptical symmetric model (ESM) of natural image in spatial domain. The proposed scan scheme is applied into H.264 and AVS platforms, and experimental results show that the coding efficiency improvement is about 0.11 dB in H.264 and 0.14 dB in AVS.

JOINT BIT-ALLOCATION FOR MULTI-SEQUENCE H.264/AVC VIDEO CODING RATE CONTROL

Paulo Nunes, *ISCTE-IT, Portugal*; Grzegorz Pastuszak, Andrzej Pietrasiewicz, *Warsaw University of Technology, Poland*; Fernando Pereira, *IST-IT, Portugal*

Abstract: Jointly controlling the encoding of multiple video sequences (VSs) allows the dynamical allocation of the overall available resources (e.g., bit rate and buffer space) along time and between the various VSs to encode; this brings a better rate-distortion (RD) trade-off, notably, when compared with single VS encoding with a rigid allocation of resources. This paper proposes a new method for joint bit allocating taking into the account the frame coding complexity and the coding type weight of each frame to be encoded for each joint encoding time instant. In order to map bit allocations into quantization parameters, rho-domain RD models are kept independently for each VS. Since the ideal and actual coding results do not always match, feedback compensation mechanisms are also proposed for adjusting the frame bit allocations along time and between the various VSs to encode for each encoding time instant. The proposed method is compared with independent sequence encoding derived from JM reference model rate control solution for H.264/AVC video coding and multi-sequence encoding using a joint RD model.

Poster Session WedAM3: Coding Standards 2

Location: Poster Room
Time: 10:00 - 11:25

ADAPTIVE PICTURE FLIPPING CODING FOR ENHANCING H.264/AVC

Tomokazu Murakami, Isao Karube, Hiroaki Ito, Masashi Takahashi, *Hitachi, Ltd., Japan*

Abstract: We developed a new video coding method called adaptive picture flipping coding (APFC) that improves coding efficiency by utilizing the dependency of the spatial direction of images in intra/inter coding. This method involves flipping input pictures horizontally and/or vertically, and then selecting the best case by using a rate-distortion optimization method. Decoders can identify the selected flipping pattern by using a ”flip flag” in the bitstream so that they can decode the sequence in the right position. We implemented APFC onto a H.264/AVC based video encoder. The experimental results show that the bitrate obtained using this method is reduced up to 4.3% for natural image sequences and 7.6% for most typical artificial images.

FAST MULTIPLE REFERENCE FRAME SELECTION METHOD IN H.264 VIDEO ENCODING

Spyridon Kapotas, Athanassios Skodras, *Hellenic Open University, Greece*

Abstract: The new JVT H.264 video coding standard is becoming the dominant standard, mainly due to its high coding efficiency. Among all modules in the encoder, motion estimation plays a key role since it can significantly affect the output quality of an encoded sequence. However, motion estimation requires a significant part of the encoding time. One reason is because the H.264 allows motion estimation performing on multiple reference frames. In this paper, a new method for fast reference frame selection is proposed. The method performs a simple and quick test prior to motion estimation in order to choose the best reference frame

A CONTEXT-BASED ADAPTIVE FAST INTRA_4X4 PREDICTION MODE DECISION ALGORITHM FOR H.264/AVC VIDEO CODING

Kai Zhang, Qiang Wang, Qian Huang, Debin Zhao, *Institute of Computing Technology, Chinese Academy of Sciences, China*; Wen Gao, *Graduate University of the Chinese Academy of Sciences, China*

Abstract: In this paper, we propose a context-based adaptive fast Intra_4x4 prediction mode decision algorithm for H.264/AVC. Firstly, a 3-order Markov random field model is introduced to describe the spatial distribution of RD optimal Intra_4x4 prediction modes at picture level. Secondly, in a neighboring context specified by this model, only a small set of candidate modes are chosen for RDO calculation according to a constrained error probability criterion. Thirdly, to reduce candidate modes further, thresholds which can be estimated by linear regression are adopted to perform early terminations. Finally, the proposed algorithm can adapt itself well to diverse video sequences due to an ability of automatic adjustment, while the model and thresholds are initialized off-line by statistics. No apparent computational overheads are involved throughout the algorithm. Experiments show that the novel fast algorithm searches only 2 - 6 Intra_4x4 modes per block and reduces up to 55% encoding time with a penalty on coding efficiency less than 0.1 dB.

VIDEO QUALITY MODEL BASED ON A SPATIO-TEMPORAL FEATURES EXTRACTION FOR H.264-CODED HDTV SEQUENCES

Stéphane Péchard, Dominique Barba, Patrick Le Callet, *IRCCyn/IVC, France*

Abstract: As a contribution to the design of an objective quality metric in the specific context of High Definition Television (HDTV), this paper proposes a video quality evaluation model. A spatio-temporal segmentation of sequences provide features used together with the bitrate to predict the subjective evaluation of the H.264-distorted sequences. In addition, a subjective tests campaign have been conducted to provide the mean observer's quality appreciation and assess the model against reality. Existing video quality algorithms have been compared to our model. They are outperformed on every performance criterion.

FAST MODE DECISION FOR H.264/AVC BASED ON CLUSTERING OF MPEG-7 TEXTURE DESCRIPTOR VALUES

Nawat Kamnoonwatana, Dimitris Agrafiotis, Nishan Canagarajah, *University of Bristol, United Kingdom*

Abstract: As multimedia content is rapidly increasing everyday, the availability of indexing metadata such as MPEG-7 descriptors becomes increasingly important for accessibility purposes. However, the use of this indexing metadata in improving coding efficiency has rarely been explored. A novel fast mode decision algorithm for H.264/AVC encoders based on the use of an MPEG-7 descriptor is proposed in this paper. The descriptor is used to form homogenous clusters within a picture frame and a range of coding modes is decided for each macroblock based on the mode of an already coded macroblock that belongs to the same cluster. The experimental results show that the proposed algorithm is able to achieve an average of 35% time-saving when compared to the full search method and 15% time-saving when compared to the fast mode decision algorithm employed in the recent JM12.2 reference H.264 software encoder. In both cases, results yield only a small degradation in rate-distortion performance and a negligible lost in subjective quality.

MODE CONVERSION FOR H.264 TO MPEG-2 VIDEO TRANSCODER

Sandro Moiron, *I.T. - Instituto de Telecomunicações, Portugal*; Sérgio Faria, Pedro Assunção, *Instituto Politécnico de Leiria, Portugal*; Vitor Silva, *Universidade de Coimbra, Portugal*; António Navarro, *Universidade de Aveiro, Portugal*

Abstract: This paper proposes an efficient method for transcoding compressed video from H.264/AVC to MPEG-2. This transcoder aims at maintaining backward compatibility between these two standards which is an important feature in modern communication systems such as WiMAX or xDSL, where customer demand for high-quality of diversified rich multimedia contents continues to impose great challenges to the video delivery systems. Since both standards use a block based approach for intra and interframe coding, the proposed transcoder explores their similarities in order to reduce the computational complexity measured as processing time. The experimental results show a computational complexity reduction up to 60% without quality loss when compared with a reference transcoder comprised of a full decoder-encoder cascade.

MODELING OF H.264 VIDEO SOURCES PERFORMING BITSTREAM SWITCHING

Stefania Colonnese, Gianpiero Panci, Stefano Rinauro, Gaetano Scarano, *Dip. Infocom. Università di Roma "La Sapienza", Italy*

Abstract: The recent H.264 video coding standard introduces an innovative compression tool allowing fast bit-rate switching. The tool, based on the novel syntactic element Switching Pictures (SP), allows flexible rate adaptation and is expecially suitable for streaming applications in wireless networks. This work addresses the modeling of a H.264 source performing bit-rate switching using SP frames. The coded bitstream structure is modelled as a Markov chain with frame size distribution depending on the frame coding mode. The performances of the model are assessed by comparison of the cell loss rate of a fixed size buffer filled with a synthetic source imple-menting the model and by a real H.264 codec.

COMBINED INTER-INTRA PREDICTION FOR HIGH DEFINITION VIDEO CODING

Xin Jin, King Ngi Ngan, *The Chinese University of Hong Kong, China*; Guangxi Zhu, *Huazhong University of Science and Technology, China*

Abstract: In rate-distortion optimized video coding, such as H.264, a macroblock uses either Inter or Intra prediction as its compensation method, which exploits temporal or spatial correlation alternatively for compression. In this paper, analysis and experimental results show that Inter and Intra prediction can be combined to generate a more accurate prediction especially for high definition video. A weighted combination method, denoted as CII_16x16, is proposed by combining Inter_16x16 with Intra_16x16_Horizontal. The weighted coefficient of the combined mode is further optimized. The implementation method is also discussed to provide a H.264 compatible scheme with minimal complexity increase both in the encoder and decoder. Experimental results show that by applying the combined mode to the H.264 rate-distortion optimized encoding flow, an average rate reduction of 1.8% can be achieved for high definition video coding.

AN ADAPTIVE COEFFICIENT SCANNING SCHEME FOR INTER-PREDICTION CODING

Qichao Sun, Xiaoyang Wu, Lu Yu, *Zhejiang University, China*

Abstract: Transform coefficient scan is an important procedure of video coding. In video coding standards such as MPEG2, H.264 and AVS, there are zig-zag scan for frame coding and field scan for field coding. In this paper, a novel adaptive coefficient scanning scheme for inter-prediction coding is proposed. This scheme includes scanning pattern generation based on DCT energy statistical calculation and macroblock level Rate-Distortion Optimized scanning pattern selection. Simulation results show that the proposed adaptive coefficient scanning scheme can achieve significant compression improvement.

POST-PROCESSING FOR FLICKER REDUCTION IN H.264/AVC

Yair Kuszpet, Dimitry Kletsel, *Technion, IIT, Israel*; Yair Moshe, *Signal and Image Processing Lab, Dept. of Electrical Engineering, Technion - IIT, Israel*; Avi Levy, *Intel Corporation, Israel*

Abstract: The H.264/AVC standard mitigates some of the most noticeable artifacts of former video coding standards, such as blocking and ringing. However, it exhibits a new temporal artifact, not prevalent in these standards: a noticeable discontinuity between an intra frame and its preceding inter frame, known as flicker. This paper proposes a post-processing scheme for dealing with the flicker artifact. A motion compensated version of the intra frame is generated based on its preceding inter frame, and is used to filter the intra frame in order to reduce the discontinuity. In addition, a non-reference flicker measure is proposed, and is used as a basis for an adaptive flicker-reduction technique. Subjective, as well as objective, results indicate that the proposed method does indeed significantly reduce the discontinuity, with an almost negligible drop in PSNR.

Oral Session WedAM4: Best Paper Award

Location: Auditorium
Time: 11:25 - 12:45
Chair: Fernando Pereira, *Instituto Superior Técnico - Instituto de Telecomunicações, Portugal*

11:25 **GENERALIZED LIFTING FOR SPARSE IMAGE REPRESENTATION AND CODING**
Julio Rolon, Philippe Salembier, *Technical University of Catalonia (UPC), Spain*

Abstract: This paper investigates the use of generalized lifting to increase the sparseness of wavelet decompositions with application to image representation and coding. As in the bandelet approach, the strategy consists in applying first a separable wavelet decomposition and then in processing the details subbands to further decorrelate the signal representation. For this second step, we use a generalized lifting [13] which allows nonlinear processing of the details subbands. In this paper, the generalized lifting design is based on the pdf of the details coefficients after the separable wavelet decomposition and its goal is to minimize the coefficients energy. Both separable and non separable approaches are investigated. The generalized lifting is shown to reduce significantly the energy and the entropy of the representation. Furthermore, a simple quantification and entropy coding strategy is used to compare the rate-distortion characteristics of wavelet, bandelet and the proposed approach based on generalized lifting. Promising results are demonstrated.

11:45 **WYNER-ZIV CODING OF STEREO IMAGES WITH UNSUPERVISED LEARNING OF DISPARITY**
David Varodayan, Yao-Chung Lin, Aditya Mavlankar, Markus Flierl, Bernd Girod, *Stanford University, United States*

Abstract: Wyner-Ziv coding can exploit the similarity of stereo images without communication among the cameras. For good compression performance, the disparity among the images should be known at the decoder. Since the Wyner-Ziv encoder has access only to one image, the disparity must be inferred from the compressed bitstream. We develop an Expectation Maximization algorithm to perform unsupervised learning of disparity at the decoder. Our experiments with natural stereo images show that the unsupervised disparity learning algorithm outperforms a system which does no disparity compensation. It is also superior to conventional compression using JPEG.

12:05 **DISTRIBUTED SOURCE CODING APPLICATION TO LOW-DELAY FREE VIEWPOINT SWITCHING IN MULTIVIEW VIDEO COMPRESSION**
Ngai-Man Cheung, Antonio Ortega, *Univ. of Southern California, Los Angeles, CA, United States*

Abstract: Multiview video coding (MVC) exploits the temporal and spatial redundancy between neighboring frames of the same view or that of adjacent views to achieve compression. Free viewpoint switching, however, poses challenges to MVC, as when a user is able to choose different playback paths it would become unclear to encoder which previously reconstructed frame would be available for decoding the current frame. Therefore,

to support free viewpoint switching in MVC, encoder would need to operate under uncertainty on the decoder predictor status. Extending our previous work on video compression with decoder predictor uncertainty, this paper proposes a MVC algorithm based on distributed source coding (DSC) to tackle the free viewpoint switching problem, where the encoder has access to several predictor candidates but there is uncertainty as to which one will be available at decoder to serve as predictor for the current frame. Since cross-view correlation could be much less significant than temporal correlation, a main challenge of the present DSC application is to achieve competitive compression efficiency. Some of the novelties of the proposed MVC algorithm are that it incorporates different macroblock modes and significance coding within the DSC framework, so that competitive coding performance can be achieved. Experimental results demonstrate the proposed DSC-based MVC algorithm can outperform solutions based on intra or closed-loop predictive (CLP) coding in terms of compression efficiency. In addition, the proposed method incurs a negligible amount of drifting, making it an attractive solution to facilitate low-delay, free viewpoint switching.

12:25 **JOINT DEPTH/TEXTURE BIT-ALLOCATION FOR MULTI-VIEW VIDEO COMPRESSION**

Yannick Morvan, Dirk Farin, *Eindhoven University of Technology, Netherlands*; Peter H.N. De With, *Eindhoven University of Technology / LogicaCMG, Netherlands*

Abstract: Multi-View display technology allows the presentation of a 3D video by showing simultaneously several views of the same scene. One approach to render these multiple views is to synthesize novel views using a Depth Image Based Rendering (DIBR) algorithm. Consequently, for the efficient transmission of 3D video signals, the compression of texture and also the depth images is required. Since the ratio between the depth and texture bit-rate is still an open question, we propose in this paper a novel joint depth/texture bit-allocation algorithm for the compression of multi-view video. The described algorithm combines the depth and texture rate-distortion (R-D) curves to obtain a single R-D surface that allows the optimization of the joint bit-allocation problem in relation to the obtained rendering quality. We subsequently discuss a fast hierarchical optimization algorithm that exploits the smooth monotonic properties of the R-D surface. The hierarchical optimization algorithm employs an orthogonal search pattern so that the number of image-compression iterations for measuring quality is minimized. Experimental results show an estimated gain of 1 dB compared to an ad-hoc selection of bit-rates. Besides this, our joint model can be readily integrated into an MVC H.264 coder because it yields the optimal compression setting with a limited computation effort.

Oral Session WedPM1: Scalable Video Coding

Location: Auditorium
Time: 14:00 - 15:20
Chair: Christine Guillemot, *IRISA, France*

14:00 **SELECTIVE INTER-LAYER PREDICTION IN SCALABLE VIDEO CODING**
Kai Zhang, *Institute of Computing Technology, Chinese Academy of Sciences, China*; Jizheng Xu, Feng Wu, *Microsoft Research Asia, China*; Xiangyang Ji, *Institute of Computing Technology, Chinese Academy of Sciences, China*; Wen Gao, *Graduate University of the Chinese Academy of Sciences, China*

Abstract: In the scalable video coding (SVC) standard, spatial scalable coding outperforms simulcast coding when programs with several display resolutions are needed. Nevertheless, it is not suitable for end devices which only need the high resolution, due to a serious performance loss on the high spatial layer compared with single layer coding. To tackle this dilemma, a selective inter-layer prediction (SIP) method is presented in this paper. SIP attains an optimal trade-off by disabling inter-layer prediction on a set of selected frames. Theoretically, this selection can be modeled as a 0-1 knapsack problem which can be solved by dynamic programming. Experimental results show that the proposed method can achieve significant gains up to 1 dB on the high spatial layer when the content of the low spatial layer is not needed, and can keep the loss unapparent even when it is. The SIP method has been adopted into the SVC reference software JSVM on the JVT 19th Meeting, held in Geneva.

14:20 **BUFFER-BASED CONSTANT BIT-RATE CONTROL FOR SCALABLE VIDEO CODING**
Tea Anselmo, Daniele Alfonso, *STMicroelectronics, Italy*

Abstract: The emerging Scalable Video Coding (SVC) is an extension of H.264/AVC standard, which shows improved coding efficiency and scalability functionalities with respect to previous scalable video coding approaches. These new functionalities make SVC suitable for meeting the increasing demand for multimedia data and the diversification of network requirements and terminal devices capabilities. To this purpose, Rate Control plays a key role in video coding process to fulfill all possible rate constraints. This paper presents a Constant Bit-Rate (CBR) control algorithm suitable for multiple layers coding. The proposed single-pass CBR scheme has been implemented into JSVM-8.0, the SVC reference software model, and experimental results show the effectiveness of the proposed algorithm for single and multiple layers coding with different configurations.

14:40 **A REDUNDANT FRAME STRATEGY TOWARDS ERROR ROBUSTNESS FOR THE H.264-BASED SCALABLE VIDEO CODING SYSTEMS**
A. C. Yu, *Imperial College London, United Kingdom*; Chaminda T.E.R. Hewage, Stewart T. Worrall, Hezerul A. Karim, S. Dogan, Ahmet M. Kondoz, *University of Surrey, United Kingdom*

Abstract: In this paper, a redundant frame strategy is introduced for the Scalable Video Coding (SVC)

systems. The proposed algorithm aims to provide error robustness for key frames transmitted in unreliable channels and error concealment at the decoder side. The technique employs correlated frames, the resultant pictures constructed from successive key frames. By transmitting additional bit overhead, the error-protection for the channel transmission is achieved. The performance of the algorithm is examined extensively with various objective and subjective evaluations in both error-free and error-prone environments. The simulation results show that the correlated frame can be used to recover the lost key frame, which prevents error propagation to surrounding non-key pictures.

15:00 **A LOW-COMPLEXITY APPROACH FOR INCREASING THE GRANULARITY OF PACKED-BASED FIDELITY SCALABILITY IN SCALABLE VIDEO CODING**

Heiner Kirchhoffer, Detlev Marpe, Heiko Schwarz, Thomas Wiegand, *Fraunhofer Institute for Telecommunications - Heinrich Hertz Institute, Germany*

Abstract: Packet-based fidelity scalability (PFS) is a desirable feature in many video coding or transmission applications. Any realization of PFS in a hybrid video coding approach, however, requires suitable concepts for controlling drift and for generating sufficiently small increments in bit rate in order to allow progressive refinements of perceptual quality relative to a given base layer quality. This paper addresses those problems in the context of the scalable video coding (SVC) extension of H.264/AVC. We present an algorithmically simple but yet remarkably well-performing method for packet-based fidelity scalability that is maximally consistent with the existing entropy coding design of H.264/AVC, allows sufficiently small increments in bit rate, and has been adopted as a normative element of SVC. We also discuss the benefits of the key picture concept of SVC in view of our proposed PFS approach. Experimental results are presented that demonstrate the effectiveness of our method for a few selected SVC conforming encoder configurations.

Panel Session WedPM2: Video Content Protection: Does it Really Matter?

Location: Auditorium
Time: 15:20 - 16:35
Chair: Ed Delp, *Purdue University, USA*

Inald Lagendijk, *Delft University of Technology, Netherlands*

Niels Rump, *Rightscom, UK*

Ian S. Burnett, *University of Wollongong, Australia*

Touradj Ebrahimi, *EPFL, Switzerland*

Poster Session WedPM3: Processing for Applications

Location: Poster Room
Time: 16:35 - 18:00

AWARENESS COMMUNICATION BASED ON FUNCTIONALLY LAYERED CODING

Masahiro Iwahashi, Takayuki Suzuki, Suriyon Tansuriyavong, *Nagaoka University of Technology, Japan*; Somchart Chokchaitam, *Thammasat University, Thailand*

Abstract: This paper proposes a new video coding system which transmits functionally layered data for awareness communication. The system can display awareness, which is defined as a blurred and semi-transparent person region, by transmitting minimum data necessary for displaying the awareness. The minimum data contains (1) person region (foreground) and (2) lower frequency band signals of the region for displaying blurred person and (3) higher (MSB side) bit-planes of the band signals for displaying blurred and semi-transparent person. Scenery without person (background) is previously transmitted. It becomes possible to reduce redundancy in data transmission since (2) and (3) above are implemented by the band decomposition and the bit-plane decomposition of the JPEG2000 (JP2K) international standard respectively.

MULTISTANDARD VIDEO DECOMPRESSION BASED ON A UNIFORM META FORMAT STREAM

Henryk Richter, Erika Müller, *University of Rostock, Germany*

Abstract: In this paper, we propose an alternative design approach for multistandard video decompression systems. The key component of our proposal is a self-contained and extensible meta data format for unified description of image processing and reconstruction operations. Dividing between the syntactic layer of the video coding schemes and the signal processing part, the meta data format offers the option of designing straightforward hardware processing units without the burden of multiple higher layer bitstream management considerations. The meta data concept provides additional benefits regarding the current trend in processor development towards multi-core chips. The specified syntax is tailored for low computational overhead. Since there is no backward channel required, the inter-process communication on software-driven multi-core and multiprocessor systems can be reduced to a minimum.

MRI IMAGE RECONSTRUCTION USING MULTIFRAME INTEGRATION

Krzysztof Malczewski, Ryszard Stasinski, *Poznan University of Technology, Poland*

Abstract: The MRI reconstruction based on super-resolution back-projection algorithm is presented in the paper. It is shown that the approach improves MRI spatial resolution in cases when PROPELLER sequences are used. The PROPELLER MRI method collects data in rectangular strips rotated around the origin of the k-space. Inter-strip patient motion is the premise for the use of super-resolution technique. Images obtained from sets of irregularly located frequency domain samples are combined into the high resolution MRI image. The super-resolution reconstruction replaces usually applied direct averaging of low-resolution images.

FULLY COMPRESSED-DOMAIN TRANSCODER FOR PIP/PAP VIDEO COMPOSITION

Nuno Roma, Leonel Sousa, *INESC-ID / IST, Portugal*

Abstract: An efficient architecture to perform Picture-In/And-Picture (PIP/PAP) composition in the compressed DCT-domain is proposed in this paper. One of the main innovative features of this architecture is a quite efficient least-squares motion re-estimation algorithm that is applied to improve the temporal prediction mechanism. Experimental results have shown that the proposed architecture may provide up to 1.3dB PSNR gain over a traditional DCT-domain approach, without any re-estimation of the composited motion vectors, with a reduction of about 33% on the output bit-rate. Furthermore, the presented DCT-domain approach does not impose any limitation on the composition setup, allowing each foreground video sequence to be placed over any arbitrary location of the background sequence.

DETECTION OF FADING-IN/OUT TEXT REGIONS FROM MPEG VIDEOS

Takaaki Kashio, *University of Fukui, Japan*

Abstract: This paper proposes a new detection technique for fading-in/out text regions from MPEG bitstreams based on a Markov model. We focus on the property that the luminance of a text region varies linearly in fading. This feature together with the other one are extracted directly from a target MPEG bitstream and they are fed into the Markov model to detect the status of text regions. The advantage of our approach lies in the fact that erroneous feature values do not directly affect the resulting text region detection due to a good behavior of the Markov model. A recall of 80% and a precision of 90% have been obtained for several sample video sequences.

NON-UBIQUITOUS WATERMARKING FOR IMAGE AUTHENTICATION BY REGION OF INTEREST MASKING

Huajian Liu, Martin Steinebach, *Fraunhofer SIT, Germany*

Abstract: In this paper we propose a novel watermarking scheme for image authentication with region of interest (ROI). The watermark is embedded in a non-ubiquitous way avoiding degrading the fidelity of selected image regions, while still providing an integrity protection for the whole image including these regions. The localization resolution of tampered areas remains equal inside and outside of the watermarked regions. Experimental results demonstrate the proposed scheme can precisely localize the manipulations both inside and outside the ROI(s).

NOISE PROCESSING FOR SIMPLE LAPLACIAN PYRAMID SYNTHESIS BASED ON DUAL FRAME RECONSTRUCTION

Aditya Mavlankar, David Chen, Sameh Zakhary, Markus Flierl, Bernd Girod, *Stanford University, United States*

Abstract: The Laplacian pyramid (LP) provides a frame expansion. Thus, there exist infinitely many synthesis operators which achieve perfect reconstruction in the absence of quantization. However, if the subbands are quantized in the open-loop mode then the dual frame synthesis operator, which is the pseudo-inverse of the analysis operator, minimizes the mean squared error (MSE) in the reconstruction. Note that this requires mod-

ification of the conventional simple synthesis scheme. For the open-loop mode, we propose novel quantization noise processing at the encoder that allows us to achieve the same performance as dual frame reconstruction and yet retain the simple synthesis scheme at the decoder. This has the advantage that the decoder can be simple in structure as well as be agnostic of whether the encoder was open-loop or closed-loop and achieves minimum MSE reconstruction for both cases. Experimental results show a gain of around 1 dB with the dual frame reconstruction compared to the simple synthesis operator. Furthermore, experiments confirm that this gain can also be obtained by retaining the simple synthesis operator and performing the proposed quantization noise processing at the encoder.

MODEL-BASED TRANSRATING OF H.264 INTRA-CODED FRAMES

Naama Hait, David Malah, *Technion - IIT, Israel*

Abstract: This paper presents a transrating (bit-rate reduction) algorithm for H.264 intra-coded frames via requantization. Previous works focused on adapting the input prediction modes to the lower bit rate and hence performed requantization using a one-pass algorithm. We propose a model-based algorithm for uniform requantization of the transform coefficients in intra-coded frames. The spatial prediction in such frames introduces block dependencies. We suggest a novel statistical-based closed-loop model for estimating the relation between the rate and the requantization step that overcomes the dependency problem. The performance of an overall transrating system for H.264 coded video, incorporating this work for intra-coded frames and our previous work for inter-coded frames is also examined.

Poster Session WedPM4: Scalable Coding

Location: Poster Room
Time: 16:35 - 18:00

INTER-SCALE PREDICTION OF MOTION INFORMATION FOR A WAVELET-BASED SCALABLE VIDEO CODER

Jens-Uwe Garbas, André Kaup, *University of Erlangen-Nuremberg, Germany*

Abstract: In this paper we present an efficient inter-scale motion vector prediction scheme for a wavelet based scalable video codec which uses inter-scale wavelet prediction. The scalable video codec considered in our work belongs to the class of 2D+t+2D schemes which applies 2-D spatial wavelet transform on the input data first, motion compensated temporal transform second and further spatial decomposition at the last stage. Although inter-scale wavelet coefficient prediction is proven to be very efficient in this coding scheme, inter-scale motion prediction has not yet been analyzed in this context. We propose a prediction scheme that enables exploitation of inter-scale motion redundancy and significantly enhances coding efficiency. PSNR increase of up to 1.45 dB and side information rate savings of up to 13 % can be denoted, particularly for fast moving sequences.

FROM AVC DECODER TO SVC: MINOR IMPACT ON A DATAFLOW GRAPH DESCRIPTION

Maxime Pelcat, Médéric Blestel, Mickaël Raulet, *IETR Image Group, France*

Abstract: The Joint Video Team (JVT) of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) has standardized a scalable video coding (SVC) extension of the H.264/AVC standard. Thanks to SVC, it is possible to partially transmit or decode a video bitstream resulting in various video qualities (spatial, temporal or reduced fidelity). Within our lab, the aim is to develop an SVC decoder that is easily reconfigurable for different computing architectures such as mono or multiprocessors. Our dataflow methodology enables such rapid prototyping and facilitates the extension from AVC to SVC. The MPEG Reconfigurable Video Coding (RVC) project uses the same kind of dataflow description to unify MPEG codecs in a single reconfigurable decoder. In this paper, ways to reuse AVC dataflow blocks are investigated while developing an SVC decoder such as in RVC. Necessary memory extensions and enhancement layer decoding optimisations are described as well as experimental results. Rapid prototyping on several architectures shows that the impacts of SVC extension are quite limited. Furthermore, both processing time and memory consumption remain reasonable.

SCALABLE MULTIPLE DESCRIPTION 3D VIDEO CODING BASED ON EVEN AND ODD FRAME

Hezerul Abdul Karim, Chaminda T.E.R. Hewage, A. C. Yu, Stewart T. Worrall, S. Dogan, Ahmet M. Kondoz, *University of Surrey, United Kingdom*

Abstract: Scalable multiple description video coding provides adaptability to bandwidth variations and receiving device characteristics and at the same time improves error robustness in multimedia networks. One promising application includes error resilient scalable video conferencing in a virtual collaboration system. In

this paper, a scalable multiple description video coding (MDC)is proposed for stereoscopic 3D video. Scalable MDC has been applied to 2D video for error resilience but not much on 3D video. The proposed algorithm enhances the error resilience of the base layer of H.264/SVC using even and odd frame based MDC. The performance of the algorithm is examined in error free environment and in mobile WiMax (IEEE802.16e) error prone environments. Simulation results show improved objective and 2D/3D subjective performance using the proposed scalable MDC in an IEEE802.16e network at high error rates compared to single description coding (SDC).

A WEIGHTED PREDICTION OF SPATIAL SCALABLE VIDEO CODING WITH INTER-LAYER INFORMATION

Kazuya Hayase, Yukihiro Bandoh, Seishi Takamura, Kazuto Kamikura, Yoshiyuki Yashima, *NTT, Japan*

Abstract: Weighted prediction (WP) is a very efficient tool to encode video scenes that contain brightness variation caused by fade. Scalable Video Coding (SVC) extension of H.264/AVC can apply the WP tool of H.264/AVC to each spatial layer. However, because the weighted parameter sets in the WP of SVC are assigned to every slice, coding efficiency is degraded if the brightness variation is non-uniform in the slice. We propose a new implicit mode WP for enhancement layers that can assign weighted parameter sets to every macroblock or macroblock partition without bit addition; the sets are derived by referring to reconstructed signals of the subordinate layer. Experiments show that the proposed implicit mode WP can achieve a significant coding gain (up to 8.22% with average of 2.23%) over white/black fade-in/out scenes versus the conventional WP of the SVC reference encoder.

LOW-COMPLEXITY POWER-SCALABLE MULTI-VIEW DISTRIBUTED VIDEO ENCODER

Li-Wei Kang, Chun-Shien Lu, *Academia Sinica, Taiwan*

Abstract: To meet the requirements of resource-limited video sensors, low-complexity video encoding technique is highly desired. In this paper, we propose a low-complexity power-scalable multi-view distributed video encoding scheme by using the correlations among video frames from adjacent video sensor nodes via robust media hashing extracted at encoder and using the global motion parameters estimated and fed back from the decoder. In addition, the proposed method is power-scalable, which is adaptive based on the available power supply of the video sensor. The power-rate-distortion behavior of the proposed scheme is also analyzed in order to maximize the video quality under limited sensor resource allocation.

A NOVEL CLASSIFIED RESIDUAL DCT FOR HYPERSPECTRAL IMAGES SCALABLE COMPRESSION

Jing Zhang, Guizhong Liu, *Xi'an Jiaotong University, China*

Abstract: Transform-based lossy compression has a huge potential for image compression. In this paper, we propose a scalable lossy compression algorithm using transform technology for hyperspectral images. The novelty of this paper lies in the classified residual DCT (Discrete Cosine Transform) as a spectral decorrelator. The classified residual DCT is an improvement of the traditional DCT, which makes the performance of DCT more close to the performance of KLT that is considered as the optimal transform for data compression in a

statistical sense. After 2D wavelet transform in the spatial domain and classified residual DCT in the spectral domain, an appropriate 3D-SPIHT image coding scheme is applied to the transformed coefficients, which makes the bit stream have scalable property. Experiments show that our proposed algorithm is capable of providing a high compression performance.

RATE CONTROL FOR SPATIAL SCALABLE CODING IN SVC

Long Xu, Wen Gao, Xiangyang Ji, Debin Zhao, *ICT,CAS,China, China*; Siwei Ma, *Department of Electrical Engineering, University of Southern California, China*

Abstract: In this paper, we present a new rate control scheme for spatial and coarse-grain-SNR (CGS) scalable coding in SVC. Firstly, a rate-distortion (R-D) model is provided for I/P/B frames according to mode analysis. Secondly, an efficient hierarchical bit allocation and two-pass refinement of quantization parameter (QP) are proposed. Thirdly, a standard deviation prediction between inter layers is presented. The experiments show that the proposed rate control scheme can achieve an average PSNR improvement of 0.3-0.7dB on average in terms of PSNR against the anchor scheme using hierarchical B frame coding with fixed QP. Meanwhile, the mismatch between target bit rate and real coded bit rate does not exceed 2%.

THE INFLUENCE OF BITRATE ALLOCATION TO SCALABILITY LAYERS ON VIDEO QUALITY IN H.264 SVC

Andreas Unterweger, *Salzburg University of Applied Sciences, Austria*; Herbert Thoma, *Fraunhofer Institute for Integrated Circuits, Germany*

Abstract: After the finalization of the scalable video coding standard, H.264 SVC faces a wide range of usage areas. In an effort to save bandwidth (and therefore costs) while satisfying the customers' need for high perceptual quality, it is necessary to find an optimal tradeoff between base and enhancement layer bit rate. This paper will analyse the influence of the possibilities H.264 SVC offers to maximize the video quality for a given total bit rate. This will be achieved by varying the number of spatial and quality layers in the SVC stream and the bit rate spent for each layer. The bit rate distributions considered ideal in terms of overall quality will be outlined and compared with the distributions analysed.

JOINT CONTENT AND TEMPORAL CORRELATION ADAPTIVE MCTF

Jianru Xue, Shugang Guo, *Institute of Artifical Intelligence and Robotics, Xi'an Jiaotong University, China*; Xuguang Lan, Nanning Zheng, *Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, China*

Abstract: In this paper, we propose an adaptive lifting-based Motion- Compensated Temporal Filtering (MCTF) framework which can remove both long and short range temporal redundancies by choosing wavelet kernels with different support sets according to temporal correlations in the video sequence. The framework can also separate noise and sampling artifacts from the reconstructed frames by embedding a spatial predicted update step into each lifting step. To support the adaptive MCTF, we also use an improved hierarchical variable size block matching algorithm for motion trajectory estimation. Experimental results confirm that the proposed framework improves both the rate-distortion performance and the visual quality of the reconstructed video significantly.

ORTHOGONALITY OF DWT AND CORRELATION OF DISTORTION

Shogo Muramatsu, Atsuyuki Adachi, Minoru Hiki, Hisakazu Kikuchi, *Niigata University, Japan*

Abstract: This work shows the significance of orthogonality of discrete wavelet transforms (DWTs) in its application to video coding. As well, a novel design method of orthogonal symmetric non-separable DWTs with a lattice structure is proposed. It is known that DWTs employed in JPEG2000 cause flickering artifacts in its low bit-rate video coding applications. Through simulations that reconstruct a picture from random transform coefficients, it is pointed out that the orthogonality of DWTs influences to spatial correlation of distortion, which visually causes the annoying flickering artifacts. Unlike biorthogonal DWTs, orthogonal ones preserve the whiteness of subbands in the reconstructed picture and suppresses the flickering artifacts. The relation between the orthogonality and correlation of distortion is also numerically evaluated through experiments with real motion pictures.

LOW COMPLEXITY LOW DELAY FGS CODING WITH ADAPTIVE REFERENCE

Xianglin Wang, *Nokia Inc, United States*

Abstract: In this paper, a fine granularity SNR scalable video coding scheme is presented. Targeted for low delay applications requiring high coding efficiency, temporal prediction is incorporated into the FGS layer of closed-loop P frames with the prediction formed adaptively from enhancement layer and base layer references. This generally requires motion compensation at both enhancement layer and base layer for decoding a FGS frame, which results in high computation complexity when multiple FGS layers are involved. The coding scheme presented in this paper can efficiently reduce the complexity in such cases. With this method, two motion compensations are sufficient for decoding a FGS frame regardless of which FGS layer is to be decoded, while preserving good coding efficiency.

Keynote Session ThursAM1: From Picture Coding to Image Understanding: Finding the Object of Interest

Location: Auditorium
Time: 08:30 - 09:30
Chair: Paulo Correia, *Instituto Superior Técnico - Instituto de Telecomunicações, Portugal*

FROM PICTURE CODING TO IMAGE UNDERSTANDING: FINDING THE OBJECT OF INTEREST

Tsuhan Chen, *Carnegie Mellon University, United States*

Abstract: From Discrete Cosine Transform to 3D model-based coding, the progress of picture coding goes hand-in-hand with the progress of image understanding. Among recent image understanding techniques, topic models have become a popular approach to object discovery, i.e., extracting the ”object of interest” from a set of images, in a completely unsupervised manner. In this talk, we will outline this approach, and extend it from still images to motion videos, using a novel spatial-temporal framework that models both the appearance and the motion of the object of interest. The spatial and temporal models are tightly integrated so that motion ambiguities can be resolved by appearance, and appearance ambiguities can be resolved by motion. This framework finds application in video retrieval (e.g., Google Video or YouTube), video surveillance, and of course, picture coding.

Oral Session ThursAM2: 3D and Multiview Video Coding

Location: Auditorium
Time: 09:30 - 10:50
Chair: Bernd Girod, *Stanford University, USA*

09:30 **TIME-CONSTANT HISTOGRAM MATCHING FOR COLOUR COMPENSATION OF MULTI-VIEW VIDEO SEQUENCES**
Ulrich Fecker, Marcus Barkowsky, André Kaup, *University of Erlangen-Nuremberg, Germany*

Abstract: Significant advances have recently been made in the coding of video data recorded with multiple cameras. However, luminance and chrominance variations between the camera views may deteriorate the performance of multi-view video codecs and renderers. In this paper, the usage of time-constant histogram matching is proposed to compensate these differences in a pre-filtering step. It is shown that the usage of histogram matching prior to multi-view video coding leads to significant gains for the coding efficiency of both the luminance and the chrominance components. Histogram matching can also be useful for image-based rendering to avoid incorrect illumination and colour reproduction resulting from miscalibrations in the recording setup. It can be shown that the algorithm is further improved by additionally using RGB colour conversion.

09:50 **MULTI-VIEW VIDEO CODING VIA VIRTUAL VIEW GENERATION**
Erhan Ekmekcioglu, Stewart T. Worrall, Ahmet M. Kondoz, *University of Surrey, Centre for Communications Systems Research, United Kingdom*

Abstract: In this paper, a multi-view video coding method via generation of virtual picture sequences is proposed. Pictures are synthesized for the sake of better exploitation of the redundancies between neighbouring views in a multi-view sequence. Pictures are synthesized through a 3D warping method to estimate certain views in a multi-view set. Depth map and associated colour video sequences are used for view generation and tests. H.264/AVC coding standard based MVC draft software is used for coding colour videos and depth maps as well as certain views which are predicted from the virtually generated views. Results for coding these views with the proposed method are compared against the reference H.264/AVC simulcast method under some low delay coding scenarios. The rate-distortion performance of the proposed method outperforms that of the reference method at all bit-rates.

10:10 **3-D STRUCTURE ASSISTED REFERENCE VIEW GENERATION FOR H.264 BASED MULTI-VIEW VIDEO CODING**
Burak Ozkalayci, O. Serdar Gedik, A. Aydin Alatan, *METU, Turkey*

Abstract: A 3D geometry-based multi-view video coding (MVC) method is proposed. In order to utilize the spatial redundancies between multiple views, the scene geometry is estimated as dense depth maps. The dense depth estimation problem is modeled by using a Markov random field (MRF) and solved via the belief propagation algorithm. Relying on these depth maps of the scene, novel view estimates of the intermediate

views of the multi-view set is obtained with a 3D warping algorithm, which also performs hole-filling in the occlusion regions. The proposed MVC method, based on H.264 standard, encodes a number of reference views in a hierarchical manner and the generated novel predictions are employed in the hierarchical coding scheme. The proposed MVC method is tested against the well-known JMVM compression algorithm, yielding comparable performances, while additionally providing 3D structure information of the observed scene.

10:30 **LOSSY COMPRESSION OF VOLUMETRIC MEDICAL IMAGES WITH 3D DEAD ZONE LATTICE VECTOR QUANTIZATION**

Gaudeau Yann, *CRAN - Nancy-University, France*; Jean-Marie Moureaux, *CRAN- Nancy University, France*

Abstract: This paper presents a new lossy coding scheme based on 3D Wavelet Transform and Lattice Vector Quantization for volumetric medical images. The main contribution of this work is the design of a new codebook enclosing a multidimensional dead zone during the quantization step which enables to better account correlations between neighbour voxels. Furthermore, we present an efficient rate-distortion model to simplify the bit allocation procedure for our intra-band scheme. Our algorithm has been evaluated on several CT an MR image volumes. At high compression ratios, we show that it can outperform the best existing methods in terms of rate-distortion trade-off. In addition, our method better preserves details and produces thus reconstructed images less blurred than the well-known 3D SPIHT algorithm which stands for a reference.

Poster Session ThursAM3: Multiview Video Coding

Location: Poster Room
Time: 10:50 - 12:00

VIEW INTERPOLATION PREDICTION FOR MULTI-VIEW VIDEO CODING

Cheon Lee, Kwan-Jung Oh, Yo-Sung Ho, *Gwangju Institute of Science and Technology, South Korea*

Abstract: Since multi-view video is a collection of videos captured by a multiple camera array, the volume of data is huge. Various algorithms have been developed for multi-view video coding. In this paper, we propose two methods: efficient view interpolation and 'VIP P-picture' coding. The view interpolation method includes initial disparity estimation, variable block-based matching, and pixel-level disparity estimation operations. The 'VIP P-picture' coding method is an additional motion estimation process at the encoder. The proposed view interpolation method improves objective quality of generated images up to about 1 4 dB, and the 'VIP P-picture' coding method increases the average coding gain about 0.66 dB in well synthesized sequences.

OPTIMAL SUBBAND BIT ALLOCATION FOR MULTI-VIEW IMAGE CODING WITH DISPARITY COMPENSATED WAVELET LIFTING TECHNIQUE

Wuttipong Kumwilaisak, Pongsak Lasang, *KMUTT, Thailand*

Abstract: This paper presents the optimal subband bit allocation based on a new rate distortion (R-D) model for multi-view image coding with disparity-compensated wavelet lifting. First, the distortion prediction of the reconstructed multi-view image with lifting scheme is presented. A new rate distortion model combining the exponential and power model is developed. Then, the analyzed prediction error and rate distortion model are used in the optimal bit allocation framework. The bit allocation framework allocates bits to all subbands with the goal to minimize distortion of the reconstructed multi-view images. Low-pass and high-pass subbands are compressed by SPIHT [5] with the optimal bit solution. We verify the proposed method with several test multi-view images. Simulation results show that the bit allocation based on the proposed method provides close results to the exhaustive search method in both allocated bits and PSNR. It also outperforms the uniform bit allocation over a wide range of target bit rate.

DIAGONAL INTERVIEW PREDICTION FOR MULTIVIEW VIDEO CODING

Ping Yang, Yun He, *Tsinghua University, China*

Abstract: Multi-view video receives many attentions in these years, because it can support a wide range of applications, such as 3D video communication and free view point video. To improve the coding efficiency, multi-view video coding not only employs temporal predictions but also predictions between different views. This paper investigates the parallel computing, low delay and scalability problems induced from the temporal-spatial predictions, proposes a diagonal interview prediction (DIP) structure. It is shown in the test results that the DIP can exploit the correlations between different views largely, it can support parallel computing, low delay and scalable decoding features, and it obtains a high coding gain especially for the large motion cases.

VIEW GENERATION FOR FTV IN CIRCULAR CAMERA ARRANGEMENT

Takeshi Uemori, Tomohiro Yendo, Toshiaki Fujii, Masayuki Tanimoto, *Nagoya University, Japan*

Abstract: There are many researches about the methods to generate free viewpoint image in linear camera arrangement, but it is rare in circular camera arrangement. In this paper, we propose a novel method to generate free viewpoint image for FTV in circular camera arrangement. This method is based on the Ray-Space method, which is one of the Image Based Rendering (IBR) techniques. We can generate arbitrary views by interpolating ray data of the Ray-Space that is constructed by collecting all images of a sequence captured by cameras arranged circularly but not enough in number. Moreover, by analyzing the trajectory of a ray in the Ray-Space, our method can be applied to the case using perspective cameras or the case that the object is not at the center of the circle or it is not small enough compared to the circle.

MULIRESOLUTION, ADAPTIVE VECTOR QUANTIZATION AND PERCEPTUAL BASED MULTIVIEW IMAGE CODEC

Akbar Sheikh Akbari, Nishan Canagarajah, David Redmill, David Bull, *Bristol University, United Kingdom*

Abstract: This paper presents a multiresolution adaptive vector quantization and perceptual based multiview image coding scheme. It decorrelates the input views into a number of subbands using a lifting based wavelet transform. The coefficients in the same subbands of different views are divided into vectors and then joined together. The resulting vectors are then vector quantized using an adaptive vector quantization scheme. Perceptual weights are designed for different viewing distances and used in the vector selection and bit allocation stages of the adaptive vector quantization technique. In order to evaluate the performance of the proposed codec, two sets of multiview test images were coded using the proposed codec with and without employing perceptual weights and the monoview vector quantization coding algorithm. Results indicated that the proposed codec with and without using perceptual weights significantly outperform the basic vector quantization technique. Results also showed that the proposed technique with perceptual weights gave superior objective and subjective image quality compared to the algorithm without perceptual weights.

BUFFER REQUIREMENT ANALYSES FOR MULTIVIEW VIDEO CODING

Ying Chen, *Tampere University of Technology, Finland*; Ye-Kui Wang, *Nokia Research Center, Finland*; Moncef Gabbouj, *Tampere University of Technology, Finland*

Abstract: Multiview video coding (MVC), which is becoming an extension of H.264/AVC, is currently under development by the Joint Video Team (JVT). Compared to H.264/AVC, the main new compression tool in MVC is inter-view prediction, which, among others, causes a substantial increase of the decoded picture buffer (DPB) size. Therefore to have an efficient buffer management for MVC is highly desirable. In this paper, we provide analyses of minimum buffer requirements for typical MVC coding structure with two coding methods, view-first coding and time-first coding. The analysis results are helpful in designing reference picture management or reference picture marking methods.

TWO NOVEL METHODS FOR FULL FRAME LOSS CONCEALMENT IN STEREO VIDEO

Cagdas Bilen, Anil Aksay, Gozde Akar, *Middle East Technical University, Turkey*

Abstract: Streaming media applications often suffer from packet losses in wired or wireless IP links. In order to get reasonable degree of quality in case of packet losses, it is necessary to have error concealment tools at the decoder. Even though several research has been done on monoscopic video, very few studies are found in the literature for stereoscopic error concealment. In this paper we propose novel full frame loss concealment algorithms for stereoscopic sequences. The proposed methods use redundancy and disparity between the two views and motion information between the previously decoded frames to estimate the lost frame. The results show that, the proposed algorithms outperform the monoscopic methods when they are applied to the same view as they are simulcast coded

Poster Session ThursAM4: Image Coding

Location: Poster Room
Time: 10:50 - 12:00

A NOVEL LOSSLESS IMAGE COMPRESSION APPROACH "COOPERATIVE PREDICTION"

Cihan Topal, Ömer Nezih Gerek, *Anadolu University, Turkey*

Abstract: Conditional predictive coders (such as LOCO, CALIC, etc.) split the prediction rule into logical cases (channels) and produce prediction residuals for each case. It is a known fact that the distributions of these separate channels usually exhibit sharp, but mean-shifted shapes. If the mean-shift amount for each channel is determined and compensated for, the overall prediction error provides smaller entropy with a sharper distribution. In this work, several prediction rules are tested for obtaining sharp and possibly mean-shifted or skewed individual prediction channel outputs. The overall prediction output was not considered as the optimization criteria. By compensating for the shifts of each channel mean, very sharp and symmetric distributions are sought at each channel, so that the combination of these channels provides an overall sharp prediction error distribution. It is shown that the proposed method provides better compression results than the celebrated LOCO which is the well-known efficient lossless compression algorithm.

A GENERALIZATION OF ZEROTREE CODING ALGORITHMS

Luca Cicala, *CIRA (Italian Aerospace Research Center), Italy*; Giovanni Poggi, *Univ. "Federico II" - Naples, Italy*

Abstract: Despite the release of the JPEG-2000 standard, wavelet-based zerotree coders keep being object of intense research because of their conceptual simplicity and excellent performance. In this work we show that any zerotree coder can be described by specifying the involved data structures (typically, order-k zerotrees) and a very limited set of tree decomposition rules. This simple grammar allows one to easily design and implement new zerotree coders, with performance sometimes superior to existing algorithms.

IMPROVING CALIC PERFORMANCE ON BINARY IMAGES

Sherif Moursi, Mahmoud El-Sakka, *The University of Western Ontario, Canada*

Abstract: Context-based Adaptive Lossless Image Codec (CALIC) is one of the most efficient lossless encoding techniques for continuous-tone images. However, its performance is considerably downgraded on images with fewer and widely separated grey levels. As a result of this, CALIC may provide lower compression rates in binary images. In this paper we provide an improved version of CALIC that gives better compression performance in binary images, without negatively affecting the performance on other type of images.

L-INFINITY PROGRESSIVE IMAGE COMPRESSION

Armando Pinho, António Neves, *University of Aveiro, Portugal*

Abstract: This paper presents a lossless image coding approach that produces an embedded bit-stream optimized for L-infinity-constrained decoding. The decoder is implementable using only integer arithmetic and is able to deduce from the bit-stream the L-infinity error that affects the reconstructed image at an arbitrary point of decoding. The lossless coding performance is compared with JPEG-LS and JPEG2000. Operational rate-distortion curves, in the L-infinity sense, are presented and compared with JPEG2000.

IMPROVED REDUNDANCY REDUCTION FOR JPEG FILES

Matthias Stirner, Gerhard Seelmann, *HTW Aalen, Germany*

Abstract: In this paper several methods are presented that allow improved compression of JPEG image files. This is mainly achieved through segmented entropy coding in reference to the EOBs (End-of-Block) of the 8x8 DCT transformed macroblocks. All of the algorithms discussed here are implemented in the freely available JPEG compression software 'packJPG'. Our approach reduces baseline JPEG image file sizes by an average of 15%, based on tests with the Kodak image set and 800 randomly selected JPEG files. This is an improvement of about 5% over a recent proposal for the JPEG standard.

PIECEWISE SMOOTH IMAGE COMPRESSION USING NORMAL DISPLACEMENT MAPS

Ward Van Aerschot, Maarten Jansen, Adhemar Bultheel, *K.U.Leuven, Belgium*

Abstract: This paper applies the idea of normal mesh techniques, utilized in CG rendering applications of smooth manifolds in 3d space, to piecewise smooth functions defined on the plane. The nonsmoothness of these functions is located along a smoothly varying curve in the domain. The nonlinear nature of the proposed method allows to deal with the 'regularity' of the curve. The smooth curve is approximated by a polyline existing of triangle edges of the adaptively created nested triangulations. The proposed method can be used in transform coders generating sparse representations of piecewise smooth colored images. We state a full compression algorithm to encode images where all information is located in one smooth contour.

A MULTIRESOLUTION APPROACH FOR THE CODING OF EDGES OF STILL IMAGES USING ADAPTIVE ARITHMETIC CODING

Guillaume Jeannic, Vincent Ricordel, Dominique Barba, *IRCCyN / IVC / University of Nantes, France*

Abstract: An edge coding scheme based on chain code representation in an multiresolution image coding context is presented. Our method enhances the coding scheme that describes the source structure with Markov models by using also an a priori knowledge from the previous decoded resolution images. Experiments using adaptive arithmetic coding have shown up to a 5% improvement for the bitrate compared to a Markovian scheme.

MULTISCALE RECURRENT PATTERNS AND GENERALISED SIDE-MATCH APPLIED TO IMAGE COMPRESSION

Eddie Filho, *Centro de Ciência, Tecnologia e Inovação do Pólo Industrial de Manaus, Brazil*; Úrsula Abecassis, *Fundação Centro de Análise, Pesquisa e Inovação Tecnológica, Brazil*; Waldir Júnior, *Universidade Federal do Amazonas - CETELI, Brazil*; Eduardo Silva, *Universidade Federal do Rio de Janeiro, Brazil*; Murilo Carvalho, *Universidade Federal Fluminense, Brazil*

Abstract: The Side-Match Multidimensional Multiscale Parser (SM-MMP) is a coding method based on the approximate multiscale pattern matching concept, where the dictionary is built considering smoothness constraints around block boundaries. This assumption can favor the coding of smooth signals, resulting in superior-quality reconstructed representations. In this work, a generalized framework for side-match is presented, in which the match attempt with neighboring blocks is performed in a hierarchical way and with a greater degree of restriction. An improved dictionary usage strategy is also presented, which employs elements from the causal neighborhood of the input block. The simulations performed on gray-scale images show that the proposed method is effective, presenting superior performance when compared to its predecessor.

AN EFFICIENT COMPRESSION ALGORITHM FOR HYPERSPECTRAL IMAGES BASED ON A MODIFIED CODING FRAMEWORK OF H.264/AVC

Guizhong Liu, Fan Zhao, *Xi'an Jiaotong University, China*

Abstract: In this paper, an efficient compression algorithm for hyperspectral images is proposed, which is based on a modified coding framework of H.264/AVC. In virtue of the flexible and diverse prediction modes of H264/AVC, the most suitable ones are assigned for the macroblocks (16*16 pixel regions of a band) of the hyperspectral images other than for the whole band images. Only the 4*4 mode is employed for the intra-band prediction in view of the fact that correlation coefficients of pixels separated by not more than four pixels in the spatial domain are greater than 0.65 at most cases. After the optimal reference band is determined by the fast reference band selecting algorithm, the inter-band prediction mode is determined then. Thus, a modified coding scheme is proposed to speed up the implemental process with the fast reference band selection algorithm, the integer DCT and the quantification which just needs multiplication and bit-shifts operations. Several AVIRIS images are used to evaluate the proposed algorithm. Compared with the state of- the-art 3D-based compression algorithms, the proposed algorithm achieves the best compression performance at different rates.

A MULTI-LATTICE DIRECTION-ADAPTIVE DEARTIFACTING FILTER FOR IMAGE & VIDEO CODING

Oscar Divorra Escoda, Peng Yin, Cristina Gomila, *Thomson Corporate Research, United States*

Abstract: Recent video coding strategies, such as H.264/AVC, incorporate an in-loop deblocking filter in order to palliate the effects of quantization noise. This assumes a smooth model of the signal which makes it unable to deal with the quantization noise introduced in textured regions and/or near image edges. In order to overcome that, alternative approaches based on sparsity-based denoising using DCT filtering assume a piecewise-statitionary model of video pictures. However, non-i.i.d. characteristics of quantization noise and structural similarity between DCT artifacts and the DCT used for filtering may make some artifacts remain after filtering under strong compression conditions. This paper presents a study on a filter for quantization noise reduction using sparse decompositions on several lattice samplings of pictures (e.g. quincunx sampling).

This allows for direction adaptive filtering in order to exploit signal components that are neither vertically nor horizontally oriented. As shown in the results, this helps improving PSNR as well as to be more robust to the visual impact of non-i.i.d. noise on filtered pictures.

A NEW CHAIN CODING BASED METHOD FOR BINARY IMAGE COMPRESSION AND RECONSTRUCTION

Saif Zahir, *UNBC, Canada*

Abstract: In this paper, we present a hybrid chain coding based scheme for contours and binary image compression and reconstruction. The proposed scheme comprises of a lossless and a lossy parts. This scheme is designed in such a way to generate extensive number of replicate links in the contours that can be assembled according to our (n10, 5) rule that can be highly compressed. Furthermore, for the lossy part, we introduce a new line processing technique to smooth the contours while maintaining high image quality. The experimental results show that the proposed method surpasses all published chain coding methods including FCC, DCC, DCC-8, VCC, CRCC, and L&Z. In addition, this scheme produced significantly higher compression ratio than WinZip, G3, G4, JBIG1, and JBIG2 standards.

TOWARDS LOSSLESS COMPRESSION USING IMAGE HIERARCHIES

Jorge Alejandro Butron-Guillen, Richard Harvey, *University of East Anglia, United Kingdom*

Abstract: This paper considers lossless image compression for grey-scale connected-set regions. We show that images can be analysed as a hierarchy or tree of connected-set regions. This tree represents a scale-space decomposition of the image and, for complicated images, could contain a great number of nodes. We examine different morphological implementations of skeleton algorithms which can be used for coding regions. We also discuss the possibilities for compression by removing redundancies found between parent-child relations in the nodes of the sieve tree. Although skeletons tend not to be used as an efficient region coding method in general compressors, here we show that they could represent an advantage for certain applications where the images contain large flat-zones or large untextured regions.

LOW-COMPLEXITY LOSSLESS COMPRESSION OF HYPER- AND ULTRA-SPECTRAL IMAGES BASED ON SLEPIAN-WOLF CODING

Enrico Baccaglini, *Politecnico di Torino, Italy*; Mauro Barni, Luca Capobianco, Andrea Garzelli, *University of Siena, Italy*; Enrico Magli, *Politecnico di Torino, Italy*; Filippo Nencini, *University of Siena, Italy*; Raffaele Vitulli, *ESA-ESTEC, Netherlands*

Abstract: Distributed source coding makes it possible to develop compression algorithms with a low-complexity encoder, while most of the signal modeling is moved to the decoder. This structure is an excellent match to the remote sensing scenario, in which the on-board processing units have limited computational capabilities. Moreover, remote sensing images do not exhibit motion, which greatly simplifies the design of a distributed coder. In this paper we propose a new algorithm for lossless compression of remote sensing images, based on distributed source coding. The objective of this algorithm is to achieve very low-complexity encoding, with performance as close as possible to a full-complexity coder. The complexity reduction is obtained by coding each spectral channel separately, whereas high coding efficiency is achieved through joint decoding. Experimental results on hyperspectral and ultraspectral images show that the proposed algorithm has significantly better performance than JPEG-LS, with similar complexity. We also provide profiling results on a Leon-2 architecture.

Keynote Session ThursPM1: DCT, Wavelets and X-lets: The Quest for Image Representation, Approximation and Compression

Location: Auditorium
Time: 12:00 - 13:00
Chair: Luis Ducla Soares, *Instituto Superior de Ciências do Trabalho e da Empresa - Instituto de Telecomunicações*

DCT, WAVELETS AND X-LETS: THE QUEST FOR IMAGE REPRESENTATION, APPROXIMATION AND COMPRESSION

Martin Vetterli, *EPFL, Switzerland and UC Berkeley, United States*

Abstract: Expansion of signals in orthonormal bases is central to signal and image processing. From the KLT and its approximation, the DCT, basic transform coding has been very successful. Over the last 15 years or so, wavelets have appeared as a powerful alternative to the more traditional Fourier like representations, having impact for example on image coding standards, like JPEG2000. We first briefly review Fourier and wavelet bases, and address approximation theoretic properties, in particular the interesting behavior of certain simple non-linear approximation schemes for piecewise smooth signals. We extend this to compression schemes, indicating the basic difference between approximation and compression. We then move to the ”real” problem, namely schemes suited for true two-dimensional signals, with objects having smooth 1-dimensional singularities, or contours. We review recent constructions in this area, including curvelets, contourlets, directionlets as well as signal adaptive schemes. The challenge of constructing generic two-dimensional bases that have optimal approximation behavior is described, and the various proposals contrasted. In particular, the proof that contourlets can achieve the optimal $1/M^2$ NLA rate will be briefly outlined. We end by pointing out areas of current research. First, a challenge is certainly finding practical schemes. This entails dealing with finite size data, as well as sampled and possibly noisy data. Only this will allow using new bases for ”real” compression tasks. In addition, new types of imagery start appearing, where true multidimensional processing will be required, like for example plenoptic images. The applicability of directional analysis in such cases will be discussed. This talk is based on work done with a number of collaborators, in particular B.Beferull-Lozano (UValencia), M.Do (UIUC), P.L.Dragotti (Imperial),L.Sbaiz (EPFL), P.Vandewalle (EPFL) and V.Velisavljevic (DTelekom).

Oral Session ThursPM2: Distributed Source Coding

Location: Auditorium
Time: 14:00 - 15:20
Chair: Luis Torres, *Technical University of Catalonia, Spain*

14:00 **ON EXTRAPOLATING SIDE INFORMATION IN DISTRIBUTED VIDEO CODING**
Stefan Borchert, Ronald Westerlaken, *TU Delft, Netherlands*; Rene Klein Gunnewiek, *Philips, Netherlands*; Inald Lagendijk, *TU Delft, Netherlands*

Abstract: The ongoing research in Distributed Video Coding (DVC) for low complexity encoding is trying to bridge the substantial performance gap to well known state-of-the-art coders. We introduce our true motion based extrapolation scheme and compare its performance to other state of the art systems in the field of DVC. The results of the extrapolation based approach display nearly the same performance as interpolation based ones. These results are also significantly better than the results from other state of the art extrapolation approaches. Furthermore we study the inluence of the motion estimation part by investigating an interpolation approach, based on the same scheme. Finally, we study the losses incurred by using only past frames to represent the motion in the current one.

14:20 **DISTRIBUTED VIDEO STREAMING USING COMPLETE COMPLEMENTARY SEQUENCES**
Chadi Khirallah, Vladimir Stankovic, Lina Stankovic (Fagoonee), *Lancaster University, United Kingdom*; Yang Yang, Zixiang Xiong, *Texas A&M University, United States*

Abstract: In many distributed video streaming applications multiple terminals stream correlated video data to a central station to be processed. The fact that those terminals may be placed within a short range of each other in a time-varying environment, results in a high level of interference, multipath fading and noise effects. One classical solution to reduce those effects is to employ the well-known spread spectrum technique; however, this leads to a substantial increase in the required bandwidth and usually makes the system not acceptable for real-time wireless video communications. In this paper we provide a novel spreading scheme that reduces the required bandwidth by exploiting correlation among different terminal observations of a video source without performance penalty. Results obtained show reduction in a terminal transmission rate of approximately 1 Mbit/sec per terminal for the same reconstructed video quality.

14:40 **DYNAMIC COMPLEXITY CODING: COMBINING PREDICTIVE AND DISTRIBUTED VIDEO CODING**
Stefaan Mys, Jürgen Slowack, Jozef Skorupa, Peter Lambert, Rik Van de Walle, *Multimedia Lab - UGent, Belgium*

Abstract: In this paper a new video coding paradigm called dynamic complexity coding is presented. It combines conventional predictive coding techniques and distributed video coding concepts to create a codec that is able to adapt to the possibly varying complexity constraints imposed on both encoder and decoder. The

main idea is to share the complex motion estimation computations in a dynamic way: if the encoder has more resources available than the decoder, it will do most of the motion search, and vice versa. We present and discuss a dynamic complexity codec that uses a spatial approach, having three modes of operation: predictive video coding mode, hybrid video coding mode using a checkerboard pattern and distributed video coding mode.

15:00 **SPATIAL MODELS FOR LOCALIZATION OF IMAGE TAMPERING USING DISTRIBUTED SOURCE CODES**

Yao-Chung Lin, David Varodayan, Bernd Girod, *Stanford University, United States*

Abstract: Media authentication is important in content delivery via untrusted intermediaries, such as peer-to-peer (P2P) file sharing. Many differently encoded versions of a media file might exist. Our previous work applied distributed source coding not only to distinguish the legitimate diversity of encoded images from tampering but also localize the tampered regions in an image already deemed to be inauthentic. An authentication decoder was supplied with a Slepian-Wolf encoded image projection as authentication data. A localization decoder required only incremental localization data beyond the authentication data since we use rate-adaptive distributed source codes. We extend the localization decoder with 1D and 2D spatial models to exploit the contiguity of the tampered regions. Our results show that the spatial decoders save 10% to 17% of authentication plus localization data size and offer greater confidence in tampering localization.

Panel Session ThursPM3: Distributed Video Coding: Trends and Challenges

Location: Auditorium
Time: 15:20 - 16:35
Chair: Touradj Ebrahimi, *EPFL - Swiss Federal Institute of Technology, Switzerland*

Bernd Girod, *Stanford University, USA*

Martin Vetterli, *EPFL, Switzerland and UC Berkeley, USA*

Pier Luigi Dragotti, *Imperial College London, UK*

Zixiang Xiong, *Texas A&M University, USA*

Poster Session ThursPM4: Implementation

Location: Poster Room
Time: 16:35 - 18:00

EFFICIENT AND ACCURATE IMPLEMENTATION OF IMAGE SCALING IN THE FREQUENCY DOMAIN

Arianne T. Hinds, Nenad Rijavec, Joan L. Mitchell, *InfoPrint Solutions Company, United States*

Abstract: Image scaling, particularly by powers of two, has previously been shown as implemented efficiently by direct manipulation of coefficients in the frequency domain. Typically, these coefficients are computed via the 8x8 DCT, or other block transforms, for their compressed storage in popular standard formats such as JPEG or MPEG. Many existing implementations of these image scaling algorithms use coarse fixed-point approximations for the floating point constants required by their definitions. This paper demonstrates application of a new methodology to approximate the floating point constants needed to perform scaling down of an image in the frequency domain. Our fixed-point implementation is a multiplierless and relatively precise (as shown by PSNR) implementation of the algorithm and is suitable for applications, such as high speed and high resolution color printing, where the visual quality of the resulting scaled image cannot be compromised.

PIPELINING ARCHITECTURE DESIGN OF THE H.264/AVC HP@L4.2 CODEC FOR HD APPLICATIONS

Kiwon Yoo, *Samsung Electronics, South Korea*; Kwanghoon Sohn, *Yonsei University, South Korea*

Abstract: This paper presents the macroblock/slice level pipeline structure for an H.264/AVC HP@L4.2 codec. In H.264/AVC, level 4.2 (L4.2) in high profile (HP) describes the encoding/decoding capability of 1920x1088@64p sequence/ bitstream of up to 62.5 Mbps. To meet this tremendous specification, the novel hardwired architecture of the H.264/AVC codec is also presented. It supports both encoding and decoding and shares commonly used hardware modules. In our system, the video subsystem including the H.264/AVC codec is classified into four principle functions: video coding, memory management, reference cache-buffer control, and top control. With regard to H.264/AVC processing, the video coding function comprises eight modules. These modules are arranged as a six-stage macroblock pipeline for the encoder and a four-stage macroblock pipeline for the decoder. With the proposed schemes adopted, a software C model and an FPGA platform were developed for verification. The simulation results indicate that our design approach successfully performs the real-time encoding/decoding of the H.264/AVC HP@L4.2 sequence/bitstream at an operating frequency of 266MHz.

NEW LZW DATA COMPRESSION ALGORITHM AND ITS FPGA IMPLEMENTATION

Wei Cui, *Beijing Institute of Technology, China*

Abstract: This paper presents a new LZW data compression algorithm that partitions conventional single large dictionary into a dictionary set that consists of several small address space dictionaries. As doing so the dictionary set not only has small lookup time but also can operate in parallel. Simulation results show that the proposed algorithm has better compression ratio for image data than conventional LZW algorithm and DLZW (dynamic LZW) algorithm, has competitive performance for text data with DLZW algorithm. In addition, a

parallel VLSI architecture for implementing the new algorithm is proposed, and it is realized using FPGA XC4VLX15-10. The experiment results show that the chip can yield a compression rate of 198.4 Mbytes/s, it is about 6.9 times the compression rate of implementing conventional LZW, and 3.2 times the compression rate of implementing DLZW.

A COST-EFFICIENT RESIDUAL PREDICTION VLSI ARCHITECTURE FOR H.264/AVC SCALABLE EXTENSION

Yi-Hau Chen, Tzu-Der Chuang, Chuan-Yung Tsai, Yu-Jen Chen, Liang-Gee Chen, *National Taiwan University, Taiwan*

Abstract: In this paper, we propose a cost-efficient residual prediction hardware architecture to support inter-layer prediction in the state-of-art H.264/AVC scalable extension. Several residual prediction schemes are analyzed in hardware architecture and coding performances, and an integer motion estimation (IME)-simplified scheme is adopted. Then, the linearity of Hadamard transform is introduced to achieve data sharing for residual prediction in fractional motion estimation (FME) architecture. An Hadamard-free residual prediction FME architecture is proposed with 40% cost saving compared to direct implementation of duplicating FME module. The proposed architecture is implemented with 86K gates at 220 MHz by UMC 90nm technology for encoding HDTV720p 30fps. The proposed design concept can be also applied in other FME VLSI designs and software acceleration for supporting residual prediction.

ALGORITHM AND ARCHITECTURE DESIGN FOR INTRA PREDICTION IN H.264/AVC HIGH PROFILE

Tzu-Der Chuang, Yi-Hau Chen, Chen-Han Tsai, Yu-Jen Chen, Liang-Gee Chen, *National Taiwan University, Taiwan*

Abstract: In this paper, we propose a novel two-stage intra prediction algorithm and hardware architecture that can support H.264/AVC high profile for 1080p HD size. The proposed DCT-based open-loop intra prediction algorithm can parallel predict each sub block with quality loss. With reconfigurable 8-pixel parallelism processing elements, the proposed architecture can process intra prediction and reconstruction with almost 100% hardware utilization. The proposed architecture was implemented by UMC 90 nm technology with 100k gate counts at 223MHz. It is the first hardware architecture that can real-time encode 1080p HD sequence with H.264/AVC high profile.

FLEXBLE ARCHITECTURE OF PROCESSOR OPTIMIZED FOR MULTIMEDIA APPLICATIONS

Adam Luczak, Olgierd Stankiewicz, *Poznan University of Technology, Poland*

Abstract: Many modern multimedia applications exploiting complex coding standards like AVC/H.264 or AAC-HE, require processing that implies synergy of both software and hardware solutions for real-time operation. This paper presents a new architecture of processor designed to achieve efficiency and convenience in such cases. The paper proposes an efficient processing architecture based on the use of fast and lightweight ”hardware function calls”. This approach reduces hardware-to-software bottlenecks, allowing a smooth transition from software to hardware solution. The presented concepts were implemented on Xilinx Virtex4-SX35 FPGA. It is worth noticing that the core is small enough to be treated as a robust soft-core. In this paper,

specificities on the developed instruction set are provided as well as on the module inter-connection. Results of tests conducted using standard implementations of H.264/AVC and AAC algorithms are also presented.

AREA-EFFICIENT QUANTIZATION ARCHITECTURE WITH ZERO-PREDICTION METHOD FOR AVS ENCODERS

Ke Zhang, Yunpeng Zhu, Lu Yu, *Institute of Information and Communication Engineering, Zhejiang University, China*

Abstract: This paper presents a VLSI architecture of forward and inverse quantization for AVS encoders based on zeroprediction method. By analyzing the quantization process, we explore a simple formulation to predict whether the transformed coefficients will be quantized to zero. The coefficient predicted to be zero is skipped without quantization. With the zero-prediction method, the proposed architecture achieves more than 250% speedup in median-quality video with 4.2% (320gates) hardware cost overhead for zero-prediction. Overall hardware cost of quantization is 7.5k logic gates at the clock frequency constraint of 200MHz. The proposed architecture is suitable to cost and power aware video applications.

PARALLEL VARIABLE LENGTH DECODING WITH INVERSE DISCRETE COSINE TRANSFORM ON VLIW DSP

Shau-Yin Tseng, *ITRI/STC, Taiwan*

Abstract: This paper proposes an algorithm that uses the parallel computational ability to process Variable Length Decoding, Inverse Zigzag, Inverse Quantization and Inverse Discrete Cosine Transform simultaneously. When this method is used, combined clock cycles for VLD, IZ, IQ and IDCT are roughly 35% shorter than those resulting from conventional methods. Moreover, this effect is especially pronounced for that this assay is easily applied to other compression standards such as MPEG-1, MPEG-2, MPEG-4, H.261, H.262, H.263 and H.264.

VLSI ARCHITECTURE OF H.264 RDO-BASED BLOCK SIZE DECISION FOR 1080 HD

Ryoji Hashimoto, Kimiya Kato, *Osaka University, Japan*; Gen Fujita, *Osaka Electro-Communication University, Japan*; Takao Onoye, *Osaka University, Japan*

Abstract: Hardware architecture of Rate-Distortion Optimization (RDO) is proposed, which is dedicated to H.264 block size decision of 1080 HD. To achieve high encoding efficiency of H.264, RDO for block size decision is indispensable but suffers from enormous computational costs since distortion and the number of coded bits can be determined only after completing the whole encoding processes of the block. The proposed approach reduces the computational costs by the approximation of bit amount in entropy coding. In addition, four parallel seven stage codec pipeline enables high speed calculation of distortion originated from residual. As a result, the proposed architecture, which can be implemented by 14K gates, achieves real-time processing of HDTV(1920x1080) frames at a rate of 30 fps in 120MHz operation, where 0.5 dB of PSNR is gained in comparison to conventional approaches.

A HARDWARE-ORIENTED INTRA PREDICTION SCHEME FOR HIGH DEFINITION AVS ENCODER

Man-Lan Wong, Yi-Lun Lin, Homer H. Chen, *National Taiwan University, Taiwan*

Abstract: Audio and Video Coding Standard - Part 2 (AVS-P2) is a video coding standard developed by the AVS Workgroup of China. In this paper, an intra prediction scheme for high definition (HD) AVS encoder is proposed to reduce the local buffer storage and computational complexity in hardware. Simulation results show that our scheme induces very little performance degradation in general frame structure. It can be easily and efficiently implemented for practical applications.

A 158 MS/S JPEG 2000 CODEC WITH A BIT-PLANE AND PASS PARALLEL EMBEDDED BLOCK CODER

Masayuki Miyama, Yuusuke Inoie, Takafumi Kasuga, Ryouichi Inada, *Kanazawa University, Japan*; Masashi Nakao, *EIZO Nanao, Japan*; Yoshio Matsuda, *Kanazawa University, Japan*

Abstract: This paper describes a 158 MS/s JPEG 2000 codec with an embedded block coder (EBC) based on a bit-plane and pass-parallel architecture. The EBC contains bit-plane coders (BPCs) corresponding to each bit-plane in a code-block. An upper BPC transfers significance states and sign bits to a lower BPC via a first-in-first-out buffer (FIFO) to synchronize BPCs themselves. The upper and the lower bit-plane coding overlap in time with a 1-stripe and 1- column gap. The bit-modeling passes in the bit-plane also overlap in time with the same gap. These methods support not only vertically causal mode, but also regular mode, which enhances the image quality. Furthermore, speculative decoding is adopted to increase throughput. This codec LSI was designed using 0.18 um process. The core area is 4.7 x 4.7 mm2 and the frequency is 160 MHz. It is applicable to a wireless PC display.

LOW-POWER HIGH-THROUGHPUT MQ-CODER ARCHITECTURE WITH AN IMPROVED CODING ALGORITHM

Alireza Aminlou, Maryam Homayouni, Mahmoud Reza Hashemi, Omid Fatemi, *University of Tehran, Iran*

Abstract: In this paper, a high speed architecture with a well designed pipeline is presented for the arithmetic encoder in JPEG2000 algorithm. The coding algorithm has also been improved by changing the renormalization and byteout extraction steps. The proposed algorithm has been implemented with a four stage pipelined architecture in VHDL. The new design has reduced, and in some cases removed data dependencies. The proposed architecture has been synthesized on a Virtex2 FPGA and its power consumption, working frequency and encoding rate were analyzed. Simulation results indicate that the proposed design is able to encode a CIF video sequence at 47 frames/s.

Poster Session ThursPM5: Distributed Video Coding

Location: Poster Room
Time: 16:35 - 18:00

RCPT BASED DISTRIBUTED VIDEO CODING OVER IEEE 802.11G WLANS: THE ROLE OF THE PUNCTURING PERIOD ON THE SYSTEM'S OPERATION

Pierre Ferre, Dimitris Agrafiotis, David Bull, *University of Bristol, United Kingdom*

Abstract: In this paper, the role of the puncturing period on the operation of rate compatible punctured turbo (RCPT) code based distributed video coding systems is studied in the context of transmission over IEEE 802.11g Wireless Local Area Networks. Reasonably realistic network conditions are considered, with the impact of the Medium Access Control layer retransmissions taken into account. The effect of the puncturing period under such conditions is evaluated by means of rate/distortion results, incurred overhead and cumulative delay. The maximum number of requests permitted is derived for various packet error rates and is shown to be independent of the content. Increasing the puncturing period facilitates a refinement of the codec's rate/distortion performance at the cost of increased parity bit requests, which in turn lead to an increase in delays. A packet error rate of 10% with a puncturing period of 64 would only support 61 requests before delays build up.

TRANSFORM DOMAIN RESIDUAL CODING TECHNIQUE FOR DISTRIBUTED VIDEO CODING

Murat Badem, Hemantha Kodikara Arachchi, *University of Surrey, United Kingdom*

Abstract: Due to its lightweight encoder architecture, the Distributed Video Coding (DVC) concept has been seen as an attractive alternative to its conventional counterparts for a number of applications. Exceptionally low computational complexity has been achieved by moving redundancy exploitation to the decoder. However, going against this norm, redundancy exploitation techniques such as DCT transform and frame difference have been utilized at the encoder end at the expense of a slight increment in computational cost at the encoder. This paper proposes a novel residual quantization technique for DCT transform based inter-frame error coding for DVC. The proposed technique minimizes the entropy of a given video frame by taking the pixel-wise difference between the current frame and a reference frame before DCT transformation. Subsequently an improved quantization technique is proposed to take advantage of small transform coefficients. Experimental results show that the proposed technique significantly improves the objective quality.

A NEW APPROACH TO DISTRIBUTED CODING USING SAMPLING OF SIGNALS WITH FINITE RATE OF INNOVATION

Varit Chaisinthop, Pier Luigi Dragotti, *Imperial College London, United Kingdom*

Abstract: This paper proposes a new approach to distributed video coding. Distributed video coding is based on the concept of decoding with side information at the decoder. Such a coding scheme employs a low-complexity encoder and the load of computational complexity is shifted to the decoder side. This property makes it well suited for low-power devices such as mobile video cameras. The uniqueness of our approach lies in the combined use of discrete wavelet transform (DWT) and the concept of sampling of signals with finite

rate of innovation (FRI), which allow us to shift the task of motion estimation to the decoder side. Unlike the currently existing practical coders, we do not employ any traditional channel coding technique. Our preliminary results show that, for a simple video sequence with a uniform background, the proposed coding scheme can achieve a better PSNR than JPEG2000-intraframe coding at low bit rates.

DISTRIBUTED VIDEO CODING USING BLOCK BASED CHECKBOARD PATTERN SPLITTING ALGORITHM

Hongbin Liu, *Harbin Institute of Technology, China*; Xiangyang Ji, *Institute of Computing Technology, Chinese Academy of Sciences, China*

Abstract: This paper presents a novel Distributed Video Coding (DVC) scheme in which the decoder is able to generate side information (SI) with high prediction quality for Wyner-Ziv (WZ) frame decoding. In the proposed method, at encoder, each WZ frame is split into two parts based on checkerboard pattern and then separately coded. At decoder, one part of the WZ frame will be firstly reconstructed, and then Boundary Matching Algorithm (BMA) [12] is used to improve motion compensation of the other part, which can obtain an improved SI of it. As a result, different from the conventional DVC scheme which usually only exploits temporal correlations among the neighboring frames to generate SI, the proposed method is also able to furthermore improve SI quality by the aid of the partially decoded WZ frame. Our experiments are performed based on Transform Domain WZ coding architecture (TDWZA), and the results demonstrate that the proposed method is able to effectively improve the DVC performance.

A DISTORTION CONTROL ALGORITHM FOR PIXEL-DOMAIN WYNER-ZIV VIDEO CODING

Toni Roca, *Universidad Politécnica de Valencia, Spain*; Marleen Morbée, *Ghent University, Belgium*; Josep Prades, *Universidad Politécnica de Valencia, Spain*; Edward Delp, *Purdue University, United States*

Abstract: In contrast to conventional video coding, Wyner-Ziv video coders perform simple intra-frame encoding and complex inter-frame decoding. This feature makes this type of coding suitable for applications that require low-complexity encoders. In this paper, we present a model of the coding distortion introduced by pixel-domain Wyner-Ziv video coders. Our distortion model can be used to determine the value of coding parameters under certain coding constraints. Specifically, we show how our model can be used to select the quantization step size of each video frame so that a target distortion can approximately be met. Experimental results show that, even though the accuracy of the distortion predictions is limited by the restricted computational capacity of Wyner-Ziv encoders, the described distortion constraints can be approximately fulfilled by using our model.

A STATISTICAL MODEL FOR A MIXED RESOLUTION WYNER-ZIV FRAMEWORK

Bruno Macchiavello, *Universidade de Brasilia, Brazil*; Debargha Mukherjee, *HP Labs, United States*; Ricardo L. De Queroz, *Universidade de Brasilia, Brazil*

Abstract: In this paper we present a rate distortion analysis and a statistical model in order to select coding parameters for memoryless coset codes, for a spatial scalability based mixed resolution Wyner-Ziv framework. The mixed resolution framework, used in this work, is based on full resolution coding of the key frames and

spatial 2-layer coding of the intermediate non-reference frames where the spatial enhancement layer is Wyner-Ziv coded. The framework enables reduced encoding complexity through reduced spatial-resolution encoding of the non-reference frames. The quantized transform coefficients of the Laplacian residual frame are mapped to cosets and sent to the decoder. A correlation estimation mechanism that guides the parameter choice process is proposed based on extracting edge information and residual error rate in co-located blocks from the low resolution base layer.

ADAPTIVE KEY FRAME RATE ALLOCATION FOR DISTRIBUTED VIDEO CODING

Ugo Ciracì, Marco Dalai, Riccardo Leonardi, *University of Brescia, Italy*

Abstract: In the context of Distributed Video Coding (DVC), rate allocation among sources represents an important problem to solve. While in the information theoretical setting of Distributed Source Coding (DSC) the statistical correlation between sources is usually assumed to be known, in practical DVC systems there is no a priori knowledge of the underlying statistics of visual data. This lack of information makes it difficult to deal with the problem of rate allocation in practical DVC codecs. In this paper we focus on the problem of how to distribute the rate between differently encoded parts of the video sequence in a DVC system. Namely, we propose an adaptive rate allocation scheme for the encoding of the key frames depending on an estimation of the local motion activity of the sequence.

THE DISCOVER CODEC: ARCHITECTURE, TECHNIQUES AND EVALUATION

Xavi Artigas, *Technical University of Catalonia, Spain*; João Ascenso, *ISEL-IST, Portugal*; Marco Dalai, *University of Brescia, Italy*; Sven Klomp, *Leibniz Universität Hannover, Germany*; Denis Kubasov, *INRIA Rennes, France*; Mourad Ouaret, *Ecole Polytechnique Fédérale de Lausanne, Switzerland*

Abstract: Distributed Video Coding is becoming more and more popular among the research community, because of its interesting theoretical contributions and because there are still many open problems waiting to be solved. This paper introduces the codec architecture and the associated tools adopted by DISCOVER (DIStributed COding for Video sERvices), a European project which has been devoted to the advancement of Distributed Video Coding for two years. Along with the general description and pointers to references with more detailed information, this paper also presents some of the results obtained with the DISCOVER codec. An extended performance analysis and the codec's executable file are both publicly available on the project's web site www.discoverdvc.org.

Keynote Session FriAM1: Efficient Representation of Sound Images: Recent Developments in Parametric Coding of Spatial Audio

Location: Auditorium
Time: 09:00 - 10:00
Chair: Ian Burnett, *University of Wollongong , Australia*

EFFICIENT REPRESENTATION OF SOUND IMAGES: RECENT DEVELOPMENTS IN PARAMETRIC CODING OF SPATIAL AUDIO

Jürgen Herre, *Fraunhofer Institute for Integrated Circuits (IIS), Erlangen, Germany*

Abstract: Like with pictures, humans talk about a ”sound image” when they try to characterize an acoustic scene containing salient spatial aspects. This talk will review the basic aspects of stereophonic / multi-channel audio that determine the perceived sound image and will outline how these aspects can be represented efficiently. One of the most remarkable innovations in this context was the recent development of the ”Spatial Audio Coding” (SAC) approach. Exploiting the human perception of spatial sound, such coding schemes are capable of transmitting high quality surround sound using bitrates that have been used for carrying traditional two-channel stereo audio so far. The talk will outline the underlying ideas and describe the architecture of the recently finalized ”MPEG Surround” specification. Equipped with a set of attractive capabilities, the technology enables introduction of surround sound into existing distribution infrastructures while retaining full compatibility to mono or stereo receivers. Finally, an outlook is provided of a next technology generation envisaged for standardization within ISO/MPEG allowing for bit-efficient and backward compatible coding of several sound objects.

Poster Session FriAM2: Video Coding

Location: Poster Room
Time: 10:00 - 11:25

OPTIMAL MULTIPLE DESCRIPTION WAVELET VIDEO CODING

Moyuresh Biswas, Michael Frater, John Arnold, Mark Pickering, *University of New South Wales, Australia*

Abstract: We propose an optimal Multiple Description (MD) video coding technique with 3D-Set Partitioning in Hierarchical Tree (3D-SPIHT) algorithm. Multiple Description Coding (MDC) technique generates multiple bitstreams of the video content and has useful applications in packet video networks and error-prone networks. We describe a tw-description MDC system that embraces rate-distortion optimization to generate two balanced, efficient descriptions. Experimental results on a variety of test sequences show that the proposed method achieves a significant improvement in performance compared to existing methods.

USE OF ADAPTIVE RESIZING IN 3-D DCT DOMAIN FOR VIDEO CODING

Jin Li, Jarmo Takala, Moncef Gabbouj, *Tampere University of Technology, Finland*; Hexin Chen, *Jilin University, China*

Abstract: This paper proposes an adaptive resizing algorithm in DCT domain for 3-D DCT based video codec. An 8x8x8 cube is resized to three modes along temporal dimension: a single 8x8 block, a downsized 8x8x4 cube and two 8x8x4 cubes. The mode selection is based on the local motion activity and determined after 2-D DCT on each block. In addition, the proposed algorithm even simplifies the computational complexity for sequences with low motion activity. Experimental results show that the proposed algorithm can improve the coding efficiency for different types of video sequences. Best performance can be expected for those with low motion activity. Moreover, it outperforms other variable size of 3-D DCT schemes. Potential applications could be for portable digital devices with restrict battery lifetime and other areas with restrict real-time requirement.

EXTENDED LAGRANGE MULTIPLIER SELECTION FOR HYBRID VIDEO CODING USING INTERFRAME CORRELATION

Xiang Li, *University of Erlangen-Nuremburg, Germany*; Norbert Oertel, Andreas Hutter, *Siemens Corporate Technology, Germany*; André Kaup, *University of Erlangen-Nuremburg, Germany*

Abstract: The Lagrange multiplier based rate-distortion optimization has been proved an effective technique in hybrid video codec design. To improve the coding efficiency, an efficient Lagrange multiplier selection method is presented in this paper. As an extension to our previous work, the Laplace distribution based rate model is further refined by taking correlations among successive frames into account, so that the proposed method shows a superior performance over previous algorithms. Compared with the most recent reference software of H.264/AVC, a gain of 0.26dB on average, 1.3dB at maximum, was achieved.

COMPLEXITY CONTROL OF H.264 BASED ON A BAYESIAN FRAMEWORK

Chaminda Kannangara, Yafan Zhao, Iain Richardson, *The Robert Gordon University, United Kingdom*; Maja Bystrom, *Boston University, United States*

Abstract: A computational complexity control algorithm is proposed for an H.264 encoder running on a processor/power constrained platform. Computational savings are achieved by early prediction of skipped macroblocks prior to motion estimation. The early prediction is based on a Bayesian framework which accurately predicts skipped macroblocks using probability models. Complexity control is achieved by calculating a pre-skip decision threshold that corresponds to a given target complexity. Results show that the algorithm can effectively control the encoding complexity whilst maintaining good rate distortion performance.

MOTION-ADAPTED WEIGHTED LIFTING SCHEME FOR MCWT VIDEO CODERS

Marc Antonini, Marie Andree Agostini, *Laboratory I3S - UNSA - CNRS, France*

Abstract: Motion compensated wavelet transform (MCWT) coding is emerging as a promising framework for efficient and scalable compression of video. Moreover, the motion-compensated lifting scheme is known to be an efficient way to implement wavelet transforms for applications dedicated to video coding. In order to increase the performances of a wavelet-based video coder, we minimize the influence of some badly estimated motion vectors on the motion-compensated wavelet transform. We propose in this paper a novel and adaptive method for the implementation of the lifting scheme. More precisely, we closely adapt the lifting steps to the motion by sampling the mother scaling function of a multiresolution basis according to the energy of the motion. Then, we compute the new lifting operators. First experiments of this approach give promising results.

EFFICIENT BIT-PLANES BASED METHOD FOR COMPRESSION OF 3D-DCT COEFFICIENTS

Andriy Bazhyna, Karen Egiazarian, *Institute of Signal Processing, Tampere University of Technology, Finland*; Nikolay Ponomarenko, *Department 504, National Aerospace University, Ukraine*

Abstract: In this paper, a new efficient method for compression of 3D-DCT coefficients is presented. The method views at values of coefficients as a set of bit-planes. The bits of each bit-plane are classified into a number of probability models depending on the context of the bit. The use of contexts allows us to efficiently encode cubes with different localization of significant coefficients. This adaptation is performed automatically during the coding process and does not require any additional computations as well as preliminary assumptions on a nature of input data. The experimental results have shown superiority of the proposed method for a 3D-DCT based video compression in comparison to the some previously proposed approaches.

LAR VIDEO: HIERARCHICAL REPRESENTATION FOR LOW BIT-RATE COLOR IMAGE SEQUENCE CODING

Erwan Flécher, Marie Babel, Olivier Déforges, Véronique Coat, *IETR - Image and Remote Sensing Group, France*

Abstract: LAR video is a low complexity system for low bit-rate color image sequence encoding. It aims to propose a joint solution for coding and representation of the frame content. In particular, it allows to provide

a compressed description of both chromatic components and motion information at a region level without region partition encoding. Initialy proposed in the LAR coder, used principle has proved to be efficient for still color image encoding. Resulting from a hierarchical spatio-temporal segmentation, a Partition Tree (PT) is transmitted to the decoder with a controlled coding cost. Presented results show interesting performances considering both content representation and compression ratios.

CONTENT-ADAPTIVE VIDEO CODING COMBINING OBJECT-BASED CODING AND H.264/AVC

Andreas Krutz, Matthias Kunter, Michael Droese, *TU Berlin, Germany*; Michael Frater, *University of New South Wales, Germany*; Thomas Sikora, *TU Berlin, Germany*

Abstract: In recent years advanced video codecs have been developed, such as standardized in MPEG-4. The latest video codec standardized, the H.264/AVC, provides compression performance superior to previous standards, but is based on the same basic motion-compensated-DCT architecture. However, for certain kinds of videos, it has also been shown that it is possible to outperform the H.264/AVC using an object-based video codec. The challange now is to develop a general-purpose object-based video coding system. In this paper, we present an automated approach to separate a video scene into shots that are coded either with an object-based codec or the common H.264/AVC. Using this idea of applying different video codecs for different kinds of content, we achieve a higher coding gain for the whole video scene considered. For the first experimental evaluation, we consider a football sequence.

Poster Session FriAM3: Error Resilience

Location: Poster Room
Time: 10:00 - 11:25

ENHANCING ERROR RESILIENCE IN WIRELESS TRANSMITTED COMPRESSED VIDEO SEQUENCES THROUGH A PROBABILISTIC NEURAL NETWORK CORE

Reuben A. Farrugia, Carl James Debono, *University of Malta, Malta*

Abstract: Video compression standards commonly employed in the delivery of real-time wireless multimedia services regularly adopt variable length codes (VLCs) for efficient transmission. This coding technique achieves the necessary high compression ratios at the expense of an increased system's vulnerability to transmission errors. The more frequent presence of transmission errors in wireless channels requires video compression standards to accurately detect, localize and conceal any corrupted macroblocks (MBs) present in the video sequence. Unfortunately, standard decoders offer limited error detection and localization capabilities posing a bound on the perceived video quality of the reconstructed video sequence. This paper presents a novel solution which enhances the error detection and localization capabilities of standard decoders through the application of a Probabilistic Neural Network (PNN). The proposed solution generally outperforms other error detection mechanisms present in literature, as it manages to improve the standard decoder's error detection rate by up to 95.74%.

A NEW ERROR CONCEALMENT SCHEME FOR WHOLE FRAME LOSS IN VIDEO TRANSMISSION

Jialue Fan, Xudong Zhang, Yu Chen, *Tsinghua University, China*

Abstract: In video streaming applications, packet loss usually leads to the loss of an entire video frame. Based on the correlations of motion vectors and mode information in consecutive frames, this paper proposes a novel bi-directional error concealment scheme. The algorithm can achieve finer result especially in high motion scenes by pixel level processing. Simulation results show that the proposed method outperforms the existing methods both on the PSNR and visual quality. Also, it is very efficient in stopping temporal error propagation. Moreover, the SKIP mode judgment in the algorithm can reduce the computational complexity effectively without significant loss of video quality.

ORTHOGONALITY DEFICIENCY COMPENSATION FOR IMPROVED FREQUENCY SELECTIVE IMAGE EXTRAPOLATION

Jürgen Seiler, Katrin Meisinger, André Kaup, *Chair of Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Germany*

Abstract: This paper describes a very efficient algorithm for image signal extrapolation. It can be used for various applications in image and video communication, e. g. the concealment of data corrupted by transmission errors or prediction in video coding. The extrapolation is performed on a limited number of known samples and extends the signal beyond these samples. Therefore the signal from the known samples is iteratively projected onto different basis functions in order to generate a model of the signal. As the basis functions are not orthogonal with respect to the area of the known samples we propose a new extension, the orthogonal-

ity deficiency compensation, to cope with the non-orthogonality. Using this extension, very good extrapolation results for structured as well as for smooth areas are achievable. This algorithm improves PSNR up to 2dB and gives a better visual quality for concealment of block losses compared to extrapolation algorithms existent so far.

SUBJECTIVE EVALUATION OF ERROR CONTROL STRATEGIES FOR MOBILE VIDEO COMMUNICATION

Waqar Zia, Tobias Oelbaum, Klaus Diepold, *Institute for Data Processing, Technische Universität München, Germany*

Abstract: Robust video communication for hand-held devices is a demanding problem because of a complexity constrained environment. Effective error control techniques suitable for H.264/AVC in such a complexity constrained conversational environment are proposed and evaluated by extensive subjective testing. Based on the results of the tests, suitable error control techniques are identified for the given application scenario that ensure an enhanced quality of service (QoS).

ERROR CONCEALMENT FOR THE SCALABLE EXTENSION OF H.264/MPEG-4 AVC

Maryse Stoufs, Adrian Munteanu, Jan Cornelis, Peter Schelkens, *Vrije Universiteit Brussel, Belgium*

Abstract: Error concealment is of crucial importance when transmitting digital video material over error-prone channels. Indeed, loss of information can result in severe degradation of the video quality if precautions are not taken to contain the effect and propagation of errors. In this paper, we present an error concealment mechanism that supports quality scalability in packet-based transmission of video encoded with the recently developed scalable extension of H.264/MPEG-4 AVC, also known as Scalable Video Coding (SVC). Experimental results show that our error concealment yields significantly better average video quality results than the non-normative error-concealment tool of the reference SVC-codec. Also, the visual results demonstrate the importance of error concealment with support for quality scalability.

DETECTION OF VISUAL IMPAIRMENTS IN THE PIXEL DOMAIN

Luca Superiori, Olivia Nemethova, Markus Rupp, *Vienna University of Technology, Austria*

Abstract: In this article we investigate the utilization of the corrupted IP packets (indicated by checksum) at H.264/AVC decoders. The position of the error within a packet is pre-localized by a syntax analysis. The impairments remaining after syntax analysis are further detected in the pixel domain by means of a voting system, using difference frames and knowledge about the artifact appearance. The blocks considered as erroneous are concealed. The proposed detection of artifacts in the pixel domain considerably improves the error detection performance achieved by the syntax analysis. The combined system clearly outperforms the typically employed slice discarding in terms of quality measured by mean squared error.

ERROR RESILIENCE PERFORMANCE EVALUATION OF A DISTRIBUTED VIDEO CODEC

Claudia Tonoli, Marco Dalai, Pierangelo Migliorati, Riccardo Leonardi, *University of Brescia, Italy*

Abstract: Distributed Video Coding (DVC), one of the most active research field in the video coding community, is based on the combination of Slepian-Wolf coding techniques with the idea of performing the prediction at the decoder side rather than at the encoder side. Besides its main property, which is flexible allocation of computational complexity between encoder and decoder, the distributed approach has other interesting properties. One of the most promising DVC characteristics is its intrinsic robustness to transmission errors. In this work we have evaluated the error resilience performance of a video codec based on the DVC scheme proposed by Stanford, and we have carried out a preliminary comparison with traditional H.264 encoding, showing that at high error probabilities and high bitrates the distributed approach can also outperform the traditional one.

UNEQUAL ERROR PROTECTION BASED ON FLEXIBLE MACROBLOCK ORDERING FOR ROBUST H.264/AVC VIDEO TRANSCODING

Matteo Naccari, Giovanni Bressan, Marco Tagliasacchi, *Politecnico di Milano, Italy*; Fernando Pereira, *IST-IT, Portugal*; Stefano Tubaro, *Politecnico di Milano, Italy*

Abstract: This paper proposes an error resilient transcoding scheme to perform unequal error protection for H.264/AVC coded video sequences over error prone channels. In order to protect those macroblocks which impact mostly on the distortion introduced at the decoder, a slice partitioning algorithm is designed on the basis of information that is available as the output of the entropy decoder. Hence, the proposed error resilient transcoding algorithm does not require full decoding, which includes inverse transform and motion compensation, and it is indicated to work in those devices where the computational resources are scarce (routers and switching transmitter stations). The proposed method has been evaluated against a classic forward error correction scheme with equal error protection of the transmitted content. Experimental results on real video test sequences show gains of up to 3 dB with respect to equal error protection.

ERROR CONCEALMENT BY REGION-FILLING FOR INTRA-FRAME LOSSES

Mengyao Ma, *Hong Kong University of Science and Technology, China*; Oscar C. Au, *Hong Kong University of Science and Technology, Hong Kong*; Liwei Guo, Yan Chen, *Hong Kong University of Science and Technology, China*; S.-H. Gary Chan, *Hong Kong University of Science and Technology, Hong Kong*

Abstract: In this paper, we propose an Error Concealment algorithm for INTRA-frame losses over packet loss channels. The novelty is that not only the INTRA-frame but also the subsequent INTER-frames are refined using the received INTRA-MBs. Simulation results are given to demonstrate the performance of the proposed algorithm.

SELECTIVE MULTIPLE DESCRIPTION CODING OF MOTION VECTOR WITH H.264/AVC DATA PARTITIONING

Jungyoup Yang, Byeungwoo Jeon, *Sungkyunkwan University, South Korea*

Abstract: In this paper, we propose a selective MV-MDC (Multiple Description Coding of Motion Vector) scheme with the data partitioning in the H.264/AVC standard. The proposed MD (Multiple Description) encoder separates the MV (Motion Vector) into two parts of equal priority each of which is transmitted through an independent packet. Additionally, to reduce the size of bitstream, the proposed MD encoder decides either MD or normal mode by using a loss-aware MD mode selection process. In case of the MD mode, the proposed MD decoding scheme utilizes two matching criteria to find an accurate MV when one of the MV descriptions is lost. Simulation results show that compared to simply duplicated bitstream transmission, the proposed MV-MDC scheme reduces a large amount of data without serious visual quality loss of reconstructed picture.

ADAPTIVE UNEQUAL ERROR PROTECTION BASED ON WYNER-ZIV CODING

Liang Liang, *Purdue University, United States*; Paul Salama, *Indiana Univ.-Purdue Univ., United States*; Edward Delp, *Purdue University, United States*

Abstract: Compressed video is very sensitive to channel errors. A few bit losses can derail the entire decoding process. Therefore, protecting compressed video is always necessary for reliable visual communications. In recent years, the Wyner–Ziv lossy coding theorem has been applied for error resilience and achieved superior improvement over conventional techniques. In our previous work, we proposed an unequal error protection method for protecting data elements in a video stream, via a Wyner-Ziv codec. This paper describes an improved technique by adapting the parity bit rates of the protected video information to the video content. The parity bit rates are assigned to the motion information and the transform coefficients according to their impact on the visual quality of each frame. This results in an efficient way of improving the quality of the decoded video when it has been corrupted by transmission errors.

STUDYING ERROR RESILIENCE PERFORMANCE FOR A FEEDBACK CHANNEL BASED TRANSFORM DOMAIN WYNER-ZIV VIDEO CODEC

José Pedro, *IST-IT, Portugal*; Luís Soares, *ISCTE - IT, Portugal*; Catarina Brites, *IST-IT, Portugal*; João Ascenso, *ISEL-IT, Portugal*; Fernando Pereira, *IST-IT, Portugal*; Carlos Bandeirinha, Frédéric Dufaux, Touradj Ebrahimi, *EPFL, Switzerland*

Abstract: Wyner-Ziv (WZ) video coding is an emerging video coding paradigm based on two major Information Theory results: the Slepian-Wolf and Wyner-Ziv theorems. One of the most interesting and used WZ video coding architectures makes use of a feedback channel (FC) to perform rate control at the decoder; in this context, the Slepian-Wolf coding module is typically based on turbo coding with puncturing. Because WZ coding is not based on the prediction loop used in conventional video coding but rather on a statistical approach where a decoder estimation of the frame to be coded is 'corrected' by the encoder, it provides intrinsic error resilience capabilities. This paper intends to study the error resilience performance of a feedback channel based transform domain WZ codec using appropriate scenarios and conditions, notably in comparison with the best performing H.264/AVC standard.

Oral Session FriAM4: Quality Assessment

Location: Auditorium
Time: 11:25 - 12:45
Chair: Inald Lagendijk, *Delft University of Technology, Netherlands*

11:25 **BLIND PSNR ESTIMATION OF VIDEO SEQUENCES USING QUANTIZED DCT COEFFICIENT DATA**
Tomás Brandão, *ISCTE-IT, Portugal*; Maria Paula Queluz, *IST-IT, Portugal*

Abstract: This paper proposes a no-reference PSNR estimation method for video sequences subject to lossy DCT-based encoding, such as MPEG-2 encoding. The proposed method is based on DCT coefficient statistics, which are modeled by Laplace probability density functions, with parameter λ. The distribution's parameter is computed from the received quantized data, by combining maximum-likelihood with linear prediction estimates. The resulting coefficient distributions are then used for estimating the local error due to lossy encoding. Since no knowledge about the original (reference) sequences is required, the proposed method can be used as a no-reference metric for evaluating the quality of the encoded video sequences.

11:45 **AN EFFICIENT REDUCED-REFERENCE VIDEO QUALITY METRIC**
Irwan Prasetya Gunawan, *University of Glasgow, United Kingdom*; Mohammed Ghanbari, *University of Essex, United Kingdom*

Abstract: We present an efficient reduced-reference video quality assessment method based on local harmonic strength (LHS) feature. LHS is based on harmonics gain and loss information extracted from picture that correspond with the two most prominent compression artifacts, blockiness and blurriness. Efficiency is achieved through further reduction of the reduced-reference overhead. For the LHS based method, we used quantization of the LHS feature, temporal sampling, and block selection with segmentation-based block classification method. A typical reduced-reference low overhead data rate (around 0.3-1 kbps) with good prediction performance (0.82- 0.87 Pearson correlations) for broadcast type video sequences in VQEG Test Phase-I datasets has been achieved by the proposed method. With such low bit rate overhead data, the method proposed in this paper can be useful for continuous monitoring of video quality applications.

12:05 **A GENERIC METHOD TO INCREASE THE PREDICTION ACCURACY OF VISUAL QUALITY METRICS**
Tobias Oelbaum, Klaus Diepold, Waqar Zia, *Technische Universität München, Germany*

Abstract: A new simple, effective and generic method is presented that allows to increase the prediction accuracy of visual quality metrics. This method is based on the observation, that the correlation between subjective results and objective metrics is very high for one special sequence or image coded at several bit rates with the same codec. This can be used to decrease the absolute error between the predicted visual quality and the actual visual quality and therefore increase the overall correlation between prediction results and results from subjective tests. Using the described method PSNR is extended to PSNR+. Comparing this new metric PSNR+ to two popular image quality metrics shows that the prediction accuracy can be increased significantly.

12:25 **MULTI-RESOLUTION STRUCTURAL DEGRADATION METRICS FOR PERCEPTUAL IMAGE QUALITY ASSESSMENT**

Ulrich Engelke, Hans-Jürgen Zepernick, *Blekinge Institute of Technology, Sweden*

Abstract: In this paper, a multi-resolution analysis is proposed for image quality assessment. Structural features are extracted from each level of a pyramid decomposition that accurately represents the multiple scales of processing in the human visual system. To obtain an overall quality measure the individual level metrics are accumulated over the considered pyramid levels. Two different metric design approaches are introduced and evaluated. It turns out that one of them outperforms our previous work on single-resolution image quality assessment.

Oral Session FriPM1: Network Aware Coding

Location: Auditorium
Time: 14:00 - 15:20
Chair: Rik van de Walle, *Gent University, Belgium*

14:00 **DESIGN OF MULTIPLE DESCRIPTIONS WITH SYMBOL-BASED TURBO CODES OVER NOISY CHANNELS WITH PACKET LOSS**
Mehrdad Valipour, Farshad Lahouti, *University of Tehran, Iran*

Abstract: Multiple descriptions (MD) with symbol-based turbo (SBT) codes are proposed, where the decoder exploits both non-uniformity of descriptions and their dependencies. A distortion-power adaptive system is obtained by setting an entropy constraint for quantizer design, which together with the MD index assignment (IA), control the level of redundancy at the MD source coder output. This is in turn exploited by the source and channel decoders for robust transmission in presence of noise and packet loss (PL). At the source coder, the IA is designed for an M-description vector quantizer using an efficient simulated annealing algorithm. Through sensitivity analysis, it is shown that in contrast to noisy channels without PL, it is better that MD is designed for the operating channel SNR in the presence of PL. Numerical results indicate improved performance in comparison to the prior art.

14:20 **ASYMMETRIC MULTIPLE DESCRIPTION CODING USING LAYERED CODING AND ERASURE CODES**
Jacco Taal, Mark Barzilay, Inald Lagendijk, *Delft University of Technology, Netherlands*

Abstract: With Asymmetric Multiple Description Coding we address the transmission case where there is a plethora of communication paths between sender and receiver, all with a different bandwidth and different levels of reliability (in terms of packet loss rate). For this scenario we have several possibilities for coding and transmission at our disposal, namely Round-Robin transmission, Layered Coding, Traditional (Symmetric) Multiple Description Coding, and Asymmetric MDC. Asymmetric MDC is a form of MDC, where each description is matched to a channels bandwidth and packet-loss rate. Redundancy is introduced and divided over the descriptions such that the average distortion under these lossy conditions is minimized. This paper first discusses the concept of AMDC as a generic case of MDC and introduces an implementation of AMDC based on erasure codes and layered coding. The optimal allocation of erasure codes and redundancy to each description is then found by optimization algorithms. We compare different algorithms and evaluate their complexity and performance.

14:40 **JOINT SOURCE AND CHANNEL CODING OF MESHGRID-REPRESENTED OBJECTS**
Dan Cernea, Adrian Munteanu, Alin Alecu, Jan Cornelis, Peter Schelkens, *VUB, Belgium*

Abstract: MeshGrid is a scalable 3D object representation method which is part of MPEG-4 AFX. This paper proposes a new approach for optimized protection of MeshGrid-represented objects against transmission errors occurring over error-prone channels. An unequal error protection approach is followed, to cope with the

different error-sensitivity levels characterizing the various resolution and quality layers produced by MeshGrid. The protection level to be employed for each layer is determined by solving a joint source and channel coding problem. The L-infinite distortion metric is used, and its superiority over the classical L-2 norm is proven. In this context, a novel fast algorithm for solving the optimization problem is conceived. The proposed joint source and channel coding approach offers resilience against transmission errors, enables a real-time implementation, and preserves all the scalability features and animation capabilities that characterize MeshGrid.

15:00 **A RHO-DOMAIN RATE CONTROLLER FOR MULTIPLEXED VIDEO SEQUENCES**

Giuseppe Valenzise, Marco Tagliasacchi, Stefano Tubaro, Luca Piccarreta, *Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy*

Abstract: This paper addresses the problem of multiplexing pre-encoded video sequences to be transmitted across a bandwidth constrained channel. At each time instant, the available bit budget needs to be optimally allocated to the sequences. We seek for a solution that minimizes the output distortion variance, in order to keep the quality of the reconstructed sequences as similar as possible, by formulating the problem in the rho-domain. In addition, by enabling a shared video buffer, we show that it is possible to smooth the overall video quality along time. Experimental results on H.264/AVC compressed video data validate the performance of the proposed algorithm.

Panel Session FriPM2: Multimedia: a World of Possibilities or just a Buzzword?

Location: Auditorium
Time: 15:20 - 16:35
Chair: Alan Hanjalic, *Delft University of Technology, Netherlands*

Andrea Cavallaro, *Queen Mary, University of London, UK*

Anthony Vetro, *Mitsubishi Electric Research Labs, USA*

Jürgen Herre, *Fraunhofer IIS, Erlangen, Germany*

Tsuhan Chen, *Carnegie Mellon University, USA*

Poster Session FriPM3: Analysis for Coding 1

Location: Poster Room
Time: 16:35 - 18:00

SCALABLE FOVEATION IMAGE CODING WITH FUZZY SYSTEM

Wuttipong Kumwilaisak, Phakorn Thamcheewan, *KMUTT, Thailand*

Abstract: This paper presents a method of the scalable foveation image coding with a fuzzy system. Our objective is to maximize foveated wavelet image quality index (FWQI). With the foveated visual sensitivity model, image regions are first prioritized. The image regions closed to the foveation points have higher priorities than those, which are far away from the foveation points. The discrete wavelet transform (DWT) is utilized to transform the considering image to the wavelet domain. The wavelet coefficients corresponding to different image regions are weighted differently based on their priorities. To achieve the objective, we use a fuzzy logic system and an iterative method to select the image coding parameters, which are a number of weighted wavelet coefficients (NWC) and a quantization parameter (QP). There are two inputs and one output in the proposed scheme. The inputs are a set of NWC and targeted bit per pixel used to encode an image. The output is an interval of the potential QP depended on a specific value of NWC. The QP providing the maximum FWQI given a value of NWC is selected. We iteratively search for the value of NWC and QP providing the highest FWQI. SPIHT codec is used to generate the scalable bitstream of the discrete wavelet coefficients. Our simulation results show that the proposed scheme provides better image quality compared to previous work in both objective and subjective quality.

RATE-DISTORTION MODELING AND ITS APPLICATION TO QUALITY LAYER ASSIGNMENT IN SVC/H.264 FINE-GRANULAR SNR SCALABLE VIDEOS

Jun Sun, Wen Gao, *Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, China*; Debin Zhao, *Department of Computer Science, Harbin Institute of Technology, China*

Abstract: Fine-Granular SNR scalable (FGS) technologies of H.264/AVC-based scalable video coding (SVC) provide a flexible and effective foundation to accommodate different and variable network capacities. To support efficient quality extraction, it is important to obtain the rate-distortion (R-D) or distortion-rate (D-R) function of each FGS packet. In this paper, firstly, we analyze the sub-bitplane technology of SVC FGS coding, and conclude the MSE-based D-R function should be linear within a FGS layer, which also explains the slow-start phenomena of PSNR-based D-R function. Consequently, a piece-wise linear model is proposed to describe the no-drift R-D function of FGS EL. Secondly, considering both the picture prediction structure and the reference picture correlation, we investigate the SVC drift of hierarchical B pictures and propose a simple and effective distortion model to estimate the reconstructed frame distortion with drift. At last, based on above analysis and models, a new model-based quality layer assignment algorithm is proposed to achieve equivalent coding efficiency as the SVC test model with significantly reduced complexity.

FAST SHOT DETECTION FOR HIGH QUALITY LOW DELAY H.264 VIDEO CODING

Pau Usach, Jorge Sastre, Valery Naranjo, *Universidad Politécnica de Valencia, Spain*; Joaquín M. López, *Telefónica I+D, Spain*

Abstract: This paper deals with the detection of shot changes in order to improve H.264 compression efficiency. This improvement is achieved by inserting intra frames when cuts occur and coding the rest of the frames using inter-frame prediction. In previous works, the proposed algorithm has demonstrated to be a fast and robust method for low delay and very low bitrate video coding, based on the comparison of the number of intra-coded macroblocks with two thresholds, one fixed and the other adaptive. In this paper, the optimization of the algorithm to be applied to high quality and low delay video coding is discussed and the results with Enhanced Definition Television format sequences (EDTV) are presented. This algorithm is also compared with another recent method based on the same measure, with results favouring our approach.

WINDOWED REFERENCE PICTURE SELECTION FOR H.264 TRANSMISSION ERROR RECOVERY

Pat Mulroy, Mike Nilsson, *BT, United Kingdom*

Abstract: For conversational real time applications we are concerned with low latency, high efficiency techniques for rapid recovery from transmission errors and in this paper we propose the use of 'windowed' references and standard compliant reference remapping instructions within H.264 to aid in recovery of unreliably transmitted multi-reference inter predicted video. Key benefit of our approach is that the amount of reference picture memory required for our windowed reference picture selection technique is reduced for feedback channels suffering from high round trip time.

ON-LINE ADAPTIVE VIDEO SEQUENCE TRANSMISSION BASED ON GENERATION AND TRANSMISSION OF DESCRIPTIONS

Juan Carlos San Miguel, Jose María Martínez-Sánchez, *Universidad Autonoma de Madrid, Spain*

Abstract: This paper presents a system to transmit the information from a static surveillance camera in an adaptive way, from low to higher bit-rate, based on the on-line generation of descriptions. The proposed system is based on a server/client model: the server is placed in the surveillance area and the client is placed in a user side. The server analyzes the video sequence to detect the regions of activity (motion analysis) and the corresponding descriptions (mainly MPEG-7 moving regions) are generated together with the textures of moving regions and the associated background image. Depending on the available bandwidth, different levels of transmission are specified, ranging from just sending the descriptions generated to a transmission with all the associated images corresponding to the moving objects and background.

VIDEO OBJECT SEGMENTATION BASED ON ACCUMULATIVE FRAME DIFFERENCE

Bing Leng, Qionghai Dai, *Tsinghua University, China*

Abstract: This paper addresses the problem of extracting video objects from head-shoulder video sequences. A method based on accumulative frame difference is proposed. First, a preliminary motion analysis is performed to each block of the frame and the blocks with fast moving edges are detected. Then, for each block, we accumulate frame difference with a different amount of frames, based on its motion attributes. After thresholding and post processing, the objects are obtained. Experimental results demonstrate that the proposed method can eliminate the expanded changed region and thus achieves a significantly improved segmentation result.

A MODEL OF R-D PERFORMANCE EVALUATION FOR RATE-DISTORTION-COMPLEXITY EVALUATION OF H.264 VIDEO CODING

Mo Wu, Søren Forchhammer, *DTU, Denmark*

Abstract: This paper considers a method for evaluation of Rate-Distortion-Complexity (R-D-C) performance of video coding. A statistical model of the transformed coefficients is used to estimated the Rate-Distortion (R-D) performance. A model frame work for rate, distortion and slope of the R-D curve for inter and intra frame is presented. Assumptions are given for analyzing an R-D model for fast R-D-C evaluation. The theoretical expressions are combined with H.264 video coding, and confirmed by experimental results. The complexity frame work is applied to the integer motion estimation.

GENERATING ANTIPERSISTENT VBR VIDEO TRAFFIC

Mehdi Rezaei, *Tampere University of Technology, Finland*; Imed Bouazizi, *Nokia Research Center, Tampere, Finland*; Moncef Gabbouj, *Tampere University of Technology, Finland*

Abstract: A novel model for antipersistent variable bit rate (VBR) video traffic is proposed. Antipersistent VBR video bit streams are a subset of VBR video bit streams in which the bit rate is controlled by a rate controller when the video is encoded. Statistical properties of antipersistent VBR video bit streams are very different from those of uncontrolled VBR video bit streams that are encoded with a constant quantization parameter. Discriminating controlled and uncontrolled VBR, an accurate model is proposed for controlled VBR video traffic. The proposed model is built based on the interaction of video encoder and video bit stream that is controlled by the rate controller. The model parameters depend on the encoding and rate control parameters and also depend on the properties of video content. The proposed model not only captures the long-range dependent (LRD) and short-range dependent (SRD) properties of video traffic, but also it can include some properties related to the content and the encoding parameters into the synthetic generated video traffic. This is valuable when generated traffics are used for simulations in which the network behavior is studied based on the content properties and encoding parameters. The proposed model has been validated successfully by several measures.

OPTIMIZATION OF LOW-DELAY WAVELET VIDEO CODECS

Marek Domanski, *Poznan University of Technology, Poland*; Andrzej Poplawski, *University of Zielona Gora, Poland*

Abstract: The paper deals with lifting implementations of the motion-compensated temporal filter banks in the (t+2D) wavelet codecs. The paper is focused on the problem of optimum setup of the analysis and synthesis filters with the maximum coding efficiency possible for a given delay constrain. Systematic exhaustive experiments have been performed in order to find optimum setup as well as coding efficiency for a given delay constrain.

Poster Session FriPM4: Analysis for Coding 2

Location: Poster Room
Time: 16:35 - 18:00

PERCEIVED QUALITY OF THE VARIATION OF THE VIDEO TEMPORAL RESOLUTION FOR LOW BIT RATE ENCODING

Quan Huynh-Thu, *Psytechnics Ltd, United Kingdom*; Mohammed Ghanbari, *Essex University, United Kingdom*

Abstract: We conducted a subjective quality assessment experiment to measure the impact of video frame rate decimation and variation in relation with impairment duration but also with content motion and texture. We found that for intermediate and high frame rate values, quality was similar independently from the duration of the frame rate decimation. On the other hand, for very low frame rates, quality decreased as the duration of the frame rate decimation increased. Our results also do not confirm the traditional thinking of higher motion content requiring a higher frame rate to produce a given level of quality. Our observations indicate that for a given frame rate, perceived quality does not necessarily increase with decreasing motion speed and that a reduction of the temporal resolution over the entire video does not lead necessarily to a significant loss of quality.

OBJECT BOUNDARY BASED MOTION PARTITION FOR VIDEO CODING

Jianle Chen, SangRae Lee, Kyo-Hyuk Lee, Woo-Jin Han, *Digital Media R&D Center, Samsung Electronics Co., Ltd., South Korea*

Abstract: In the H.264/MPEG-4 AVC video coding standard, motion compensation can be performed by partitioning macroblocks into square or rectangular regions to improve inter prediction efficiency. However, current H.264 MB partition set is not optimal because the rigid square or rectangular divisions cannot match the boundary shape of a moving object well. In this paper, we analyze this problem in detail and propose an object-boundary-based motion partition scheme to overcome the problem. The proposed scheme generates the motion partition map for current macroblock by segmenting corresponding area in the reference picture. Due to continuity of the object boundary shape in the video sequence, this technique allows the shapes of the partition map to match better the boundaries of moving objects. Experimental results show that the proposed motion partition method improves coding efficiency of inter P picture of H.264 coding system with average 6.79% bit saving. The coding gain of test sequence with distinct moving objects is up to 10.63% bit saving.

FAST LONG-TERM MOTION ESTIMATION FOR HIGH DEFINITION VIDEO SEQUENCES BASED ON SPATIO-TEMPORAL TUBES AND USING THE NELDER-MEAD SIMPLEX ALGORITHM

Olivier Brouard, Fabrice Delannay, Vincent Ricordel, Dominique Barba, *IRCCyN, France*

Abstract: Multi-frame motion estimation is a new specificity of the video coding standard H.264/MPEG-4 Advanced Video Coding (AVC) to improve compression performances. These reference frames can be short or long-term. However, a long-term motion estimation needs to have a good initial search point to be effective. In this paper, we propose a new method for a fast long-term motion estimation with high definition (HD)

sequences. We use an implicit motion model. First we describe the multi-resolution motion estimation based on spatio-temporal tubes. These tubes permits to obtain a good initial search point for the long-term motion estimation that follows. Then, the search is refined using the Nelder-Mead Simplex method. The global approach allows to reduce computational cost and improves the accuracy of motion estimation.

ROBUST MOTION SEGMENTATION FOR HIGH DEFINITION VIDEO SEQUENCES USING A FAST MULTI-RESOLUTION MOTION ESTIMATION BASED ON SPATIO-TEMPORAL TUBES

Olivier Brouard, *IRCCyN/IVC, France*; Fabrice Delannay, Vincent Ricordel, Dominique Barba, *IRCCyN, France*

Abstract: Abstract Motion segmentation methods are effective for tracking video objects. However, objects segmentation methods based on motion need to know the global motion of the video in order to back-compensate it before computing the segmentation. In this paper, we propose a method which estimates the global motion of a High Definition (HD) video shot and then segments it using the remaining motion information. First, we develop a fast method for multi-resolution motion estimation based on spatio-temporal tubes. So we get a homogeneous motion vectors field (one vector per tube). From this motion field, we use a robust approach to estimate the parameters of the affine model that characterizes the global motion of the shot. After back-compensation of the video shot global motion, the remaining motion vectors are used to achieve the motion segmentation and extract the video objects.

HIERARCHICAL MOTION ESTIMATION USING RECURSIVE LMS FILTERS

Mejdi Trimeche, Jani Lainema, *Nokia Research Center, Finland*; Moncef Gabbouj, *Tampere University of Technology, Finland*

Abstract: In this paper, we present a hierarchical motion estimation algorithm that is based on adaptive LMS filters. The algorithm is an extension of an earlier work [1], which uses an adaptive 2-D LMS filter to match the intensity values while passing through the image pixels according to a Hilbert scanning pattern, the algorithm adapts the corresponding set of FIR coefficients. The peak value in the resulted coefficient distribution points to the localized displacement that happens between two consecutive frames. We extend the algorithm to use mirrored scanning and we apply it hierarchically across diadic spatial resolutions. The obtained displacement at each resolution level is mapped to the next level, which reduces the search area and improves the precision of the matching process. The algorithm is particularly useful for tracking slowly varying motion, such as affine or rotational motion, and is also robust to noise. We also show an example application for motion compensated sharpening of video frames using the proposed LMS filtering without explicit computation of the displacement vectors.

ESTIMATION OF GLOBAL MOTION USING A MODIFIED LORENTZIAN WEIGHT FUNCTION

Nafisa Tarannum, Michael Frater, Mark Pickering, *UNSW, Australia*

Abstract: Global motion estimation (GME) is an extensively used tool for many important video processing applications including mosaicing, image registration, video compression and segmentation. The presence of large foreground objects or other distortions often hinders the process of GME. The main challenges in this area

are the combining of the motions of the background and foreground objects and the use of user-defined thresholds that are often sequence dependent. We propose an improved algorithm that solves the above-mentioned problems by an automatic block-based weight function based on the Lorentzian estimator. This algorithm automatically ignores the effect of large foreground objects without the requirement of any user-defined threshold. Experimental results on different test sequences show the superior performance of this technique over some recent methods.

MOTION ANALYSIS OF 3-D DUAL-TREE DISCRETE WAVELET TRANSFORM

Jingyu Yang, *Department of Automation, Tsinghua University, China*; Yao Wang, *Electrical and Computer Engineering Dept, Polytechnic University, Brooklyn, NY 11201, China*; Wenli Xu, Qionghai Dai, *Department of Automation, Tsinghua University, China*

Abstract: 3-D dual-tree discrete wavelet transform (DDWT) is attractive for video representation since it isolates motion along different directions in its directional subbands, and has been employed for video coding. In this paper, we analyze its efficiency in capturing motion activities. Similar analysis is then extended to 3-D anisotropic DDWT (ADDWT). Experiments verify the effectiveness of our analysis. Based on the motion analysis, we realize that 3-D DDWT/ADDWT can not adapt to intense motion activities, and propose a new transform structure for high motion sequences. It is shown that the new scheme brings about 1dB coding gain.

FAST AND ACCURATE GLOBAL MOTION ESTIMATION BY COMBINING PIXEL SUBSAMPLING AND OUTLIER REJECTION

Hussein Alzoubi, W. David Pan, *Dept. of ECE, University of Alabama in Huntsville, United States*

Abstract: Tools based on global motion estimation (GME) have been adopted by the MPEG-4 standard. Estimating global motion parameters tends to be computationally intensive, due to the involvement of all the pixels in the calculation. If real time video coding using GME is required, then fast and accurate GME techniques are critical. Our previous work has shown that fast GME could be achieved by choosing a small subset of pixels for GME, by following certain subsampling patterns. However, GME based on these subsampling methods often suffer from loss in motion estimation accuracy. On the other hand, the accuracy of GME depends heavily on outlier rejection, which seeks to detect and eliminate those pixels (outliers) that do not follow the general pattern of the global motion. In the MPEG-4 standard, about 10% of the outlier pixels are skipped in the GME. Therefore, interestingly, outlier rejection could also lead to modest reduction in the computational complexity of the GME. In this paper, we show how to achieve faster and more accurate GME, by integrating pixel subsampling into the outlier rejection process of a conventional GME method. Simulation results showed that, at a subsampling ratio of 1/4, the proposed method could achieve about three times of speedup over the conventional GME method that uses outlier rejection alone, while attaining an average estimation accuracy slightly higher than the conventional method.

AUTHOR INDEX