

# INFORMATION SOCIETY TECHNOLOGIES (IST) PROGRAMME



**Project Number: IST-2001-32795**

**Project Title: Network of Excellence in Content-Based Semantic Scene Analysis and Information Retrieval**

**Deliverable Type: PU**

Deliverable Number: D4.2

Contractual Date of Delivery: 30.04.2005 (month 35 of the project)

Actual Date of Delivery: 30.04.2005

Title of Deliverable: **Contributions to standards final report**

Work-Package contributing to the Deliverable: WP 4

Nature of the Deliverable: RE

Leading Author: Raul Medina Beltran de Otalora (TUM)

Contributors: Raúl Medina Beltrán de Otálora (TUM), Stephan Herrmann (TUM), Josep R. Casas (UPC), Vassileios Mezaris (ITI), E. Cooke (DCU), P. Migliorati (UNIBS)

**Abstract:**

The MPEG and JPEG standards are strongly related with the research objectives of the SCHEMA Network of Excellence. The present document reports on the exchange of information between standards development and SCHEMA partners. Representatives of the SCHEMA NoE attend MPEG and JPEG2000 meetings and closely follow current standardization activities. Newest developments are reported back to the SCHEMA group at the periodic Technical Committee meetings. Contributions from SCHEMA partners to the MPEG and JPEG2000 standards are also reported in the document.

**Keyword List:** Standards, MPEG, JPEG2000, JPSearch

**\*Type: PU-public**

**Table of Contents**

1. Introduction and objectives .....	3
Structure of this report .....	3
Current developments in MPEG .....	3
Main MPEG activities in meetings #60 and #61 organized according to the different standards: .....	4
Main MPEG activities in meetings #62, #63 & #64 according to the different standards:.....	5
Main MPEG activities in meetings #65, #66 & #67 according to the different standards:.....	6
Main MPEG activities in meetings #68- #72 according to the different standards: .....	9
72 <sup>st</sup> MPEG meeting: Busan (Corea, 17-22 April 2005).....	9
Summary of MPEG activity and discussion of potential joint contributions from SCHEMA to MPEG.....	9
2. Contributions to MPEG standards .....	11
MPEG Input Documents .....	11
Activities related to MPEG-7 XM .....	17
3. Contributions to JPEG 2000.....	21
OpenJPEG .....	21
4. Conclusion .....	25
5. References, MPEG standards & parts and table of acronyms .....	26
References.....	26
MPEG standards and parts .....	26
MPEG acronyms .....	27

## 1. Introduction and objectives

The aim of WP4 is the constant exchange of information between the standards and all partners, by establishing and maintaining a communication flow between the Network of Excellence and the relevant standardization bodies (MPEG, JPEG2000). This document reports on the exchange of information between standards development and SCHEMA partners up to month 35 of the project duration (i.e. the completion of the project).

### **Structure of this report**

The document is organized according to both directions of the information flow (in-flow and out-flow) between the SCHEMA Network of Excellence and the relevant standardization bodies.

For the in-flow (towards SCHEMA), representatives of SCHEMA are closely following current standardization activities and attending MPEG and JPEG2000 meetings. Newest developments are reported back to the SCHEMA group at the Technical Committee meetings. The first section of the present document describes the newest developments reported in the relevant standardization bodies and comments on the possibility of potential joint contributions.

For the out flow (from SCHEMA to standardization bodies), SCHEMA partners are contributing to the standards development by means of specific contributions to MPEG and JPEG2000. These contributions are described in the remainder of the current report. Furthermore, the SCHEMA consortium is preparing a joint contribution to the standards community in terms of feedback regarding the usability of MPEG-7 part 6 (reference software), which is being implemented as part of the SCHEMA reference system in WP3.

### **Current developments in MPEG**

Until the completion of this final report, the Moving Picture Experts Group (MPEG – ISO/IEC JTC1/SC29/WG11) held 13 meetings. The main results presented at these MPEG meetings are listed below (please check the MPEG acronyms table at the end of this document):

#### **60<sup>th</sup> MPEG meeting: Fairfax VA (US, 6-10 May 2002)**

Main documents released:

- CD on Advanced Video Coding, JVT: H-264/MPEG-4 part 10
- FDIS MPEG-21 part 2: Digital Item Declaration (DID)
- WD MPEG-21 part 7: Digital Item Adaptation (DIA)

Open calls potentially related to SCHEMA NoE objectives:

- Requirements for Persistent Association (content—ID/D)
- MPEG-7 systems (coding efficiency)

#### **61<sup>st</sup> MPEG meeting: Klagenfurt (Austria, 22-26 July 2002)**

Documents released:

- FCD Advanced Video Coding (JVT)
- FDIS MPEG-21 part 3: Digital Item Identification & Description (DII)
- CD MPEG-21 part 5: Rights Expression Language (REL)
- CD MPEG-21 part 6: Rights Data Dictionary (RDD)

Open calls:

- Lossless Audio Coding (44-192kHz)
- MDS request for Registration Authority

**Main MPEG activities in meetings #60 and #61 organized according to the different standards:**

- **MPEG-2 and -4 (audio-visual coding of media objects)**  
The MPEG group has carried out work in Systems extensions in order to adopt the Advanced Video Coding (JVT) functionalities in the multiplex of data streams for advanced text and 2D graphics
- **MPEG-7 (multimedia content description)**  
In the Systems part, a number of efficient and robust coding techniques have been developed for descriptors, whereas in Multimedia Description Framework, linguistic aspects were the main focus. There is also an on-going work on conformance testing for MPEG-7.
- **MPEG-21 (framework for multimedia delivery and consumption)**  
This was the main new standard being finalized at the time the SCHEMA Network started. Three parts of the standard were in the process of completion. In particular: Part-5 "Rights Expression Language" (REL), Part-6 "Rights Data Dictionary" (RDD) and Part-7 "Digital Item Identification & Description" (DIA).

**62<sup>nd</sup> MPEG meeting: Shanghai (China, 21-25 October 2002)**

Documents released:

- Extensions to MPEG-4 IPMP
- MPEG ad hoc on 3DAV, MP4 File Format, MPEG-4 Audio extensions

Open calls:

- Font format representation and font compression technology
- Digital Item Processing:
- Base Operations (DIBO), Method Language (DIL)

**63<sup>rd</sup> MPEG meeting: Awaji Island (Japan 9-13 December 2002)**

Documents released:

- FCD ISO 14496-10 Advanced Video Coding (AVC) / ITU-T Recommendation H.264
- FCD REL/RDD
- MPEG-21 Testbed: MPEG-4 over IP (8), IPMP ext, DIA/streaming

Main activities:

- MPEG-2 and -4: work in Systems extensions finished and Advanced Video Coding (JVT) specification already frozen.
- MPEG-21: Rights Data Dictionary (RDD) and Rights Expression Language (finished). On-going work on Digital Item Identification & Description (DIA)

**64<sup>th</sup> MPEG meeting: Pattaya (Thailand, 10-14 March 2003)**

Documents released:

- FDIS ITU-T Rec. H.264 /ISO 14496-10 Advanced Video Coding (AVC)
  - Improvement in compression performance for general coding
  - Applications: Video Communication, DigitalTV, HDTV, video streaming, DVD...
- MPEG-4 High Efficiency AAC (HE-AAC)
- MP4 file format
- MPEG-7 Reference Software extensions (FPDAM)

Open calls:

- Evidence on Scalable Video Coding Advances (MPEG-4)
- Extended Sample Bit Depth and Chroma Format in AVC (JVT)

**Main MPEG activities in meetings #62, #63 & #64 according to the different standards:**

- **MPEG-2 and –4 (coding of audio-visual objects)**  
The Final Draft International Standard for Advanced Video Coding (in the joint group with ITU-T named JVT) and the Third edition of MPEG-4 Visual have been released. For audio, work progressed for MPEG-2 AAC (bandwidth extension) and extensions to Ref. Software.
- **MPEG-7 (multimedia content description)**  
Amendments and extensions to the Reference Software (FPDAM) have been released
- **MPEG-21 (framework for multimedia delivery and consumption)**  
For Part–1: "Vision, Technology, Strategy", edition 2 of Technical Report was released. Part–3, CISAC was appointed as Registration Authority "Digital Item Identification & Description"

**65<sup>th</sup> MPEG meeting: Trondheim (Norway, 21-25 July 2003)**

Documents released:

- FDIS MPEG-4, Part 9: Reference Hardware Description
- CD MPEG-4, Part 17: Streaming text format

Open calls:

- Call for Proposals on Lossless Coding of 1-bit Audio Signals

Main activities:

- Completion of Parts 5 and 6 of MPEG-21, the Rights Expression Language (REL) and its associated Rights Data Dictionary (RDD)
- MPEG-2 and –4: MPEG Systems support for AVC, 'Error-resilient Simple Scalable Profile' in part 2. Also MPEG Reference Hardware.

**66<sup>th</sup> MPEG meeting: Brisbane (Australia, 20-24 October 2003)**

Documents released:

- MPEG-4, Systems: Synchronized Streams (audio, subtitling, 2D animated graphics) for AVC delivery to MPEG terminals
- MPEG-4, Audio: completed verification tests of bandwidth extension tool (Spectral Band Replication) for High Efficiency AAC with significant coding gain over MPEG-4 AAC
- MPEG-4, AFX: Animation Framework eXtension, rich set of 3D tools operating at the geometry, modeling and biomechanical level plus compression, streaming and seamless integration

Open Calls:

- Call for Proposals on Scalable Video Coding Technology (due Feb 1st, 2004)
  - algorithms and tools providing high flexibility in bit rate, frame size and frame rate adaptation at bitstream level, along with high compression efficiency (near state of the art single layer coding performance)
- Call for Requirements on MPEG-21 architecture and Intellectual Property Management and Protection: feedback on requirements for MPEG-21 specs
- Call for Comments on 3DAV: interest of industry in standardization of 3DAV technology

**67<sup>th</sup> MPEG meeting: Waikoloa HI (USA, 8-12 December 2003)**

Documents released:

- MPEG-4, AVC: completion of verification tests for AVC.
- MPEG-4, Audio: High Quality Parametric Audio Coding supporting pitch and tempo scaling

- FDIS MPEG-21, Part 7: Digital Item Adaptation, metadata information that facilitates adaptation of digital items and media resources for different MPEG-21 usage environments
- CD MPEG-21, Part 10: Digital Item Processing, specification of operators, methods and a language allowing functional operations to be carried out on digital items by MPEG-21 peers

Open Calls:

- Call for Proposals on MPEG Graphics API
- Call for Proposals on MPEG-21 Event Reporting
- Draft CFP on Lightweight Scene Representation

**Main MPEG activities in meetings #65, #66 & #67 according to the different standards:**

- **MPEG-2 and -4 (coding of audio-visual objects) Systems:** Synchronized Streams for AVC delivery. Audio: verification tests for HE-AAC bandwidth extension (Spectral Band Replication) and HQ parametric coding. AVC: Verification tests for AVC. And finally for AFX: 3D tools
- **MPEG-7 (multimedia content description)**  
Part-6: Reference Software Extensions (FDAM: 03/12, AMD:04/03)
- **MPEG-21 (framework for multimedia delivery and consumption)**  
Part-7: FDIS Digital Item Adaptation (DIA)  
Part-10: CD Digital Item Processing (DIP)

**68<sup>th</sup> MPEG meeting: Munich (Germany, 15-19 March 2004)**

Main achievements:

- MPEG-2 & -4: Extensions to Intellectual Property Management and Protection (IPMP) completed, along with corresponding reference software and conformance tests
- MPEG Video Coding: received 14 complete specifications and 12 discrete technology responses to the Dec 2003 CFP on Scalable Video Coding (SVC).
- MPEG-4, Advanced Text and Graphics: OpenType® font format representation for font data streaming. MicroType® font compression technology for OpenType TrueType fonts
- MPEG-4, X3D Interactive Profile complete from collaboration with Web3D consortium
- MPEG-21: Responses to the Dec 2003 CFP on MPEG-21 Intellectual Property Management and Protection (IPMP) have lead to a refined CFP. Added components for Digital Item Adaptation and Evaluation Tools for Persistent Association

Open Calls:

- Call for Proposals on MPEG-21 IPMP
- Call for Proposals on Lightweight Application Scene Representation LASR
- Call for Proposals on Spatial Audio Coding, coding of multichannel audio signals at rates comparable to those for coding 1 or 2 audio signals
- Draft CFP on Symbolic Music Representation, integrating Symbolic Music Representation into multimedia applications

**69<sup>th</sup> MPEG meeting: Redmont WA (USA, 19-23 July 2004)**

Main Achievements:

- MPEG-4: Video coding AVC/H.264 extensions
  - High Profile: Fidelity Range Extension (FRExt) including new profiles for high definition content
    - Hi10P: 10 bit coding suitable for film material

- Hi422: broadcasting and editing capabilities
- Hi444: full RGB color resolution & 12 bit precision for professional production and graphics environments
- Ability to carry alpha channel data and to analyze and synthesize the representation of film grain for high quality consumer apps.
- Simple Profile: Addition of two levels
  - Level 4a: documents VGA resolution
  - Level 5: extends simple profile to SD (standard definition) TV resolution
- MPEG21
  - LASeR: Lightweight Application Scene Representation (<30KB)
    - Leverages functionality of industry standards (MPEG-4 BIFS, W3C SVG, Macromedia Flash). E.g. W3C SVG T1.2 LASeR compressed SVG stream delivers rich graphic contents @10kbps

Open Calls:

- Call for Proposals on Symbolic Music Representation
  - Seeking solutions for the Symbolic Music Representation format
- MPEG-compatible technology to support symbolic representations of music (Symbolic Music Representation format) synchronized with other MPEG object types

Status of current standards activities and documents released:

- **MPEG-2 and -4 (coding of audio-visual objects)**
  - MPEG-4, Part 2 (Visual), Amd.2:
    - New levels in Simple Profile (FCD 04/10)
  - MPEG-4, Part 3 (Audio), Amds.4,5,6:
    - Lossless & scalable Audio coding (FCD 04/07 and 04/10)
  - MPEG-4, Part 5 (Software), Amd.6,7,8:
    - AVC, FRExt and AFX Reference Software (FCD 04/07)
  - MPEG-4, Part 10 (AVC), Amd.1:
    - FRExt (High Profile, FCD 04/07)
  - MPEG-4, Part 20:
    - Lightweight Application Scene Representation (FCD 04/07)
- **MPEG-7**
  - MPEG-7, Part 5 (MDS), Amd.2:
    - MDS user preference extension (FCD 04/07)
  - MPEG-7, Part 6 (Ref. Software), Amd.1:
    - Ref. Software extension (FCD 04/07)
  - MPEG-7, Part 7 (Conformance Testing), Amd. 1:
    - Conformance extensions (FCD 04/07)
  - MPEG-7, Part 9 (Profiles):
    - 1st Ed. FCD 04/10
  - MPEG-7, Part 10 (Schema definition):
    - 1st Ed. FCD 04/10
- **MPEG-21**
  - MPEG-21, Part 2 (DID): 2nd Ed. (FCD 04/07)
  - MPEG-21, Part 4 (IPMP Framework): 1st Ed (CD 04/10)

- MPEG-21, Part 8 (Reference Software): 1st Ed (FCD 04/07)
- Digital Item Processing, Scalable Video Coding, Event Reporting, Binary Formats...

### **70th MPEG meeting Palma de Mallorca (Spain, 18-22 October 2004)**

#### Main achievements:

- Audio, efficient coding of spatial audio: four competing technologies evaluated
- Video, Scalable Video Coding Single video. Codec working across a wide range of bit rates without compromising quality based on the Advanced Video Coding (AVC) standard. New technologies: motion-compensated temporal filtering, spatial up-sampling prediction and progressively-refined quantization

#### Open Calls:

- Call for Evidence on 3D AV Technology
  - Free viewpoint video. Allows the viewer to choose an arbitrary viewing position, in the same way as a hologram does, but with all the quality of a normal, full-color video signal and without a hologram's 'see-through' effects
  - Call for "Evidence": If convincing evidence is received that such technologies exist and can be standardized, MPEG would proceed by issuing a Call for Proposals for appropriate technology and start to develop a new standard
- Call for Proposals on Multimedia Middleware (M3W) in order to:
  - Allow applications to execute multimedia functions with a minimum knowledge of the middleware. This goal can be achieved by standardizing the API that the middleware offers
  - Allow applications to trigger updates to the middleware to extend this API. This is more difficult, as it requires mechanisms to manage the middleware API and to ensure that this process functions correctly

### **71<sup>st</sup> MPEG meeting: Hong Kong (China, 17-21 January 2005)**

#### Main achievements:

- Scalable Video Coding (SVC): First working draft produced in Hong Kong.
  - Part of Advanced Video Coding (AVC) standard (ISO/MPEG-4 part 10 | ITU-T/H.264) by the Joint Video Team (JVT)
  - Adds enhanced forms of quality scalability to further enable advanced video coding uses in a wide variety of applications, particularly including highly-heterogeneous environment
- Discrete Cosine Transform (DCT) Addition of a fixed-point version considered as a compatible, but optional, means of improving decoder precision. Draft call for proposals issued for this technology
- Audio: MPEG-4 Lossless Coding of one-bit Oversampled Audio (final ballot phase). One-bit oversampled audio is a high quality consumer format on high-density optical discs, also valuable as an audio archiving format.

#### Open Calls:

- Call for Proposals on Multimedia Middleware (M3W)
- Draft Call for Proposals on Fixed-point 8x8 IDCT and DCT
  - Fixed-point approximation to the ideal IDCT/DCT function
  - Draft Call for Proposals on Multi-View Video Coding
  - Final call with results to be evaluated in October of 2005. Improved coding efficiency over existing MPEG compression tools. Applications: FTV (free-viewpoint television), 3DTV (3D television) and surveillance



- A Call for information a Scalable Coding of Speech and Audio
  - Seeking Information on Scalable Coding of Speech and Audio. Responses due in July 2005

#### **Main MPEG activities in meetings #68- #72 according to the different standards:**

- MPEG-21:
  - Digital Item Declaration (DID is MPEG-21 part 2): 2nd edition completed
  - Digital Item Streaming for multimedia collection creation and delivery
  - Fragment Identification Schema: technology to assist consumers in accessing specific parts of a movie stored in a consumer device
  - MPEG-21 part 17: addresses physical media such as DVDs
- MPEG-7:
  - Simple Metadata Profile (SMP): describes simple multimedia clips such as music, images, etc.
  - User Description Profile (UDP): describes personal preferences and usage patterns of users of multimedia content
  - Core Description Profile (CDP): describes a collection of multimedia contents for images, such as videos, and audio that can be used for distribution, broadcast and educational courseware

#### **72<sup>st</sup> MPEG meeting: Busan (Corea, 17-22 April 2005)**

##### Main achievements:

- Scalable Video Coding (SVC): second working draft produced, after evaluation the results of CE about extended spatial scalability using cropped areas, adaptive Gop structure, quality layer, ...New Core experiments proposed.
- VidWav group: new tools (de-ringing filter, intra-coding, GSTS) integrated in the reference MSRA software. New Exploration Experiments proposed to improve coding efficiency and extend SNR scalability.
- Multi-View Video Coding: test condition, anchors, evaluation procedure fixed.
- Future Video Coding: tutorial and discussion on distributed source coding, 3DAV based on Ray technology, New High Fidelity RGB Coding, ...

##### Open Calls:

- Call for Proposals on Multi-View Video Coding to choose a working software model of an encoder and decoder
- Call for Contribution on Future Video Coding for Nice meeting
- Refined Call for Proposals on Fixed-point 8x8 IDCT and DCT Standard

#### ***Summary of MPEG activity and discussion of potential joint contributions from SCHEMA to MPEG***

When compared with the aim and objective of SCHEMA, it was clear that the timing of open calls in MPEG was unfortunate for potential joint contributions from the NoE. Advanced video coding is not a primary target in the list of activities of SCHEMA, nor is the description related to multimedia Digital Items in terms of Access Rights or Adaptation to Networks and Terminals (DIA deals with adaptation to different devices making it possible to successfully render multimedia content on multifunction devices by taking into account the network and terminal). These two topics were still open for contributions under MPEG-4 and MPEG-21 at the time of the project.

Apart from the specific contributions to the standards by the individual partners, it was proposed that, rather than aiming at a SCHEMA joint contribution through one of the currently open MPEG calls, **SCHEMA could provide feedback to the standards community with respect to usability of MPEG-7 reference software.** The MPEG-7 reference software (XM or eXperimentation Model) has been implemented within the SCHEMA Reference System, in the framework of Work Package 3. The integration of the XM in the SCHEMA Reference System has allowed SCHEMA partners to provide feedback and suggestions for improvement to MPEG about this software tool, which is key to the wide acceptance of the standard.

Besides the MPEG-7 eXperimentation Model, the SCHEMA Network of Excellence has played a key role within MPEG not only by its active participation in open Call of Proposals but as well by running experiments and disseminating results. The next paragraphs summarise the contributions from SCHEMA NoE members to standardisation activities, from input documents to software releases of the MPEG-7 Reference Software (eXperimentation Model) as well as experimental results and participation in other Standard related activities.

## 2. Contributions to MPEG standards

### *MPEG Input Documents*

Herewith a list of the contributions as input documents to MPEG from SCHEMA members.

- (1) V. Mezaris, R. Medina Beltrán de Otálora, I. Kompatsiaris, S. Herrmann, and J. Casas, "Experimental results for a Search & Retrieval System using MPEG-7 still image descriptors", ISO/IEC JTC1/SC29/WG11, MPEG2004/M10566, presented by TUM at the 68<sup>th</sup> MPEG Meeting, March 2004, Munich, Germany.

This document describes and compares the general architecture and experimental results of two different CBIR (Content Based Image Retrieval) system implementations carried out within the SCHEMA Network of Excellence. These are the first version of the SCHEMA Reference System (using proprietary visual descriptors) and the SchemaXM System (integrating the MPEG-7 eXperimentation Model for descriptor extraction and for supporting search and retrieval functionalities).

As mentioned above, the first version of the reference system uses proprietary region descriptors and simple matching tools based on the Euclidean distance for performing indexing and retrieval. While these are shown to perform satisfactory, it is demanded that superior retrieval accuracy is achieved; this can be accomplished by integrating with the reference system a set of more sophisticated descriptors and matching tools. To this end, it was decided that the MPEG-7 standard are employed. A non-normative part of the Standard realizing the normative descriptors, the MPEG-7 XM (eXperimentation Model), was integrated with the SCHEMA reference system, to endow the latter with advanced descriptors and matching tools and provide the resulting system with full compliance with the MPEG-7 Standard. Under the proposed scheme, images are first segmented using segmentation algorithms integrated with the Schema Reference System indexing descriptors are extracted from the segmented images using a part of the MPEG-7 XM, namely the Extraction Application and stored in binary encoded database files (.mp7 files).

The most important milestone of this input document is an initial feedback to the standards community regarding the usability of the MPEG-7 Standard and its eXperimentation Model (XM). This document additionally underlines the opportunity given to research institutions and organisations to use the Schema Reference System as a test-bed for the evaluation and comparison of different indexing and retrieval modules and interfaces within the context of CBIR.

- (2) V. Mezaris, S. Herrmann, H. Doulaverakis, I. Kompatsiaris, and M. G. Strintzis, "Improving the efficiency of a search and retrieval system using MPEG-7 Visual descriptors", ISO/IEC JTC1/SC29/WG11, MPEG2004/M11342, presented by ITI at the 70<sup>th</sup> MPEG Meeting, October 2004, Palma de Mallorca, Spain.

The latest version of the Schema Reference System is a region-based CBIR (Content Based Image Retrieval) system implementation developed under the auspices of the SCHEMA Network of Excellence, integrating the MPEG-7 eXperimentation Model for descriptor extraction and search and retrieval functionalities. This document describes the introduction to the Schema Reference System of an extension of the XM, realizing an indexing structure for the MPEG-7 visual descriptors. The indexing extension was designed at the Technical University of Munich as part of the EU-IST project BUSMAN and integrated into the SCHEMA Reference System. It is described in detail in this document and its impact on the indexing and retrieval of approximately 33000 key-frames is

evaluated. The use of the XM and its indexing extension in a realistic application scenario is demonstrated. In particular, the presented comparisons clearly demonstrate the improvement in time-efficiency attained by adopting the XM indexing module and its necessity for enabling the use of the XM in a realistic interactive search application scenario involving a large media collection.

- (3) V. Mezaris, S. Papadopoulos, I. Kompatsiaris, and M. G. Strintzis, "On the potential use of machine learning for image categorization (VCE-2)", ISO/IEC JTC1/SC29/WG11, MPEG2004/M11344, presented by ITI at the 70<sup>th</sup> MPEG Meeting, October 2004, Palma de Mallorca, Spain.

This document reports a machine learning approach to be explored for the task of image categorization to predefined categories. The proposed approach makes use of the standardized descriptors that can be extracted using MPEG-7 XM, but does not rely solely on the XM for estimating image similarity; instead, a number of support vector machines and a constrained similarity measure are also used to this end. Instead, it is based on a previously proposed method, where it was used for image retrieval using global image properties under the query-by-example scheme. This method combines support vector machines (SVM) with a constrained similarity measure (CSM).

More specifically, each support vector machine employs a training set (i.e. the model set defined for each category of VCE-2) to learn the boundary separating the two considered classes (images belonging and not belonging to the particular category, respectively). For estimating this boundary, each member of the model set is represented by a vector of low-level MPEG-7 visual descriptor values. Following the boundary estimation, all test-set samples, also represented by a vector of low-level MPEG-7 visual descriptor values each, are classified as relevant or irrelevant, depending on whether they are inside the boundary or not. However, this means that a large part of the test set may be classified to none of the predefined categories, since the intersection of the subspaces inside the boundaries is not expected to be equal to the entire feature space. This, along with a number of other open issues regarding the use of the aforementioned techniques for the application considered by VCE-2 are highlighted in this document.

- (4) V. Mezaris, H. Doulaverakis, I. Kompatsiaris, and M. G. Strintzis, "On the use of MPEG-7 XM-based SCHEMA Reference System for segmentation algorithm evaluation", ISO/IEC JTC1/SC29/WG11, MPEG2004/M11684, presented by ITI at the 71<sup>st</sup> MPEG Meeting, January 2005, Hong Kong, China.

This document describes the current efforts on exploiting the MPEG-7 XM-based Schema Reference System for an innovative task: segmentation algorithm evaluation.

Whilst there exist methods for the objective evaluation of segmentation algorithms, primarily using pre-existing or manually-generated ground truth for the segmentation, these do not provide a clear indication of the suitability of a given segmentation algorithm for the application of content-based image/video indexing and retrieval. This is not only due to the nature of objective evaluation methods, which does not in any way consider the application of content-based retrieval, but also to the need for ground truth, which typically limits evaluation to a few dozen images at most.

To counter these drawbacks we propose the comparative evaluation of segmentation algorithms in terms of their contribution towards improving retrieval accuracy. More specifically, we propose the application of the different segmentation algorithms to be evaluated on an extended common image/video collection, the extraction of a plurality of common indexing features for all resulting

regions, the submission of a large number of queries on each set of regions generated by a different segmentation algorithm, and the comparative evaluation of the segmentations on the basis of the precision-recall curves resulting from the retrieval experiments.

Based on the aforementioned ideas, the suitability of the SCHEMA Reference System and particularly of its underlying indexing and retrieval mechanism, based on the MPEG-7 XM, for the evaluation task, are discussed in this document.

- (5) E. Cooke, N. O'Connor, A. Smolic: "Proposal for Specification of 3D Camera Parameters.", MPEG-4 SNHC, Document ISO/IEC JTC1/SC29/WG11 M11716, Presented by DCU at the 71<sup>st</sup> MPEG Meeting, January 2005, Hong Kong, China.

Recently, due to the SNHC Exploration Experiment detailed in input document M11715 and related discussion on the SNHC reflector, it has become apparent that there is a discrepancy between the camera parameter specification provided by MPEG-4 AFX and those required by 3DAV related technologies. The difference can be summarised by stating that the AFX specification is strongly related to the traditional computer graphics (CG) approach, while the 3DAV requirements are more computer vision (CV) oriented. This proposal attempts to rectify this problem by specifying the 3DCameraNode AFX node description that contains the parameters required for both CG and CV applications.

- (6) E. Cooke: "Multiple Image View Synthesis for Virtual Viewpoint Rendering.", MPEG-4 SNHC, Document ISO/IEC JTC1/SC29/WG11 M11238, presented by DCU at the 70<sup>th</sup> MPEG Meeting, October 2004, Palma de Mallorca, Spain.

The purpose of this contribution is to describe a new view synthesis process suitable for Free Viewpoint Video that can create arbitrary virtual views of a scene captured by an N camera set-up. The Multi-View Synthesis algorithm differs to other view synthesis approaches in that it is camera set-up independent and therefore flexible enough to create virtual views given 1 to N of the available video inputs. Unlike most multiple image view synthesis approaches it does not blend surfaces from the set of available camera views but instead selects the best quality surfaces from the individual reference images for virtual view rendering. This flexibility reduces perceptual errors and ensures that the best quality virtual view reconstruction is produced from the available views. MVS also provides solutions to the view synthesis issues of virtual view surface visibility and hole filling while using multiple images. The approach is designed for image-based rendering systems and assumes that all the cameras have been calibrated and that each reference image has an associated disparity map.

- (7) E. Cooke, A. Smolic, N. O'Connor: "Results on Exploration Experiment: Multi-View Synthesis Reconstruction using DIBR Nodes", MPEG-4 SNHC, Document ISO/IEC JTC1/SC29/WG11 M11715, presented by DCU at the 71<sup>st</sup> MPEG Meeting, January 2005, Hong Kong, China.

The goal of this Exploration Experiment was to determine if the nodes currently defined in the AFX Depth Image-based Representation version 1 or version 2 provide the information required to allow virtual view creation via the Multi-View Synthesis approach presented in input document m11238.

- (8) N. Adami, S. Benini, R. Leonardi and H. Nishikawa , “Report of CE on Metadata Adaptation Integration”, ISO/IECJTC1/SC29/WG11/M9849, presented by UNIBS at the 65<sup>th</sup> MPEG Meeting, July 2003, Trondheim, Norway.

In this document the Metadata Adaptation tool presented at the Pattaya meeting was validated, using a slightly adjusted syntax and semantics. An alternative syntax has also been proposed in order to be combined with the adaptation hint DS. This tool supports adaptation operations such as metadata integration.

The presented scheme includes two descriptors, named AverageValue and InvariantProperties, that can be efficiently used by an integration engine in order to improve the quality of the integrated final description and to speed up the whole integration process. Examples and experiment results were given to support the validation of the proposed tool.

- (9) N. Adami, M. Brescianini, R. Leonardi, A. Signoroni, “Fully embedded entropy coding with arbitrary multiple adaptation capabilities”, ISO/IECJTC1/SC29/WG11/M11378, presented by UNIBS at the 70<sup>th</sup> MPEG Meeting, October 2004, Palma de Mallorca, Spain.

In this contribution, we describe a fully embedded texture coding technique which is suitable for scalable video coding with arbitrary multiple adaptation requirements and operating point selection. This technique is a GOF based extension of the Embedded Morphological Dilation Coding algorithm. This technique has been used in our contribution to SVC-CE1 called STool System-2.

- (10) N. Adami, M. Brescianini, R. Leonardi, A. Signoroni, “SVC CE1: STool - a native spatially scalable approach to SVC”, ISO/IECJTC1/SC29/WG11/M11368, presented by UNIBS at the 70<sup>th</sup> MPEG Meeting, October 2004, Palma de Mallorca, Spain.

This document describes the UNIBS-SCL proposal in response to the MPEG21 SVC CE1. Our scalable video-coding scheme, called STool, is based on a 2D+t+2D structure and is implemented using a modified version of the Microsoft Research Asia (MSRA) reference software plus some modifications and tools which has been used in substitution. The STool architecture has been implemented in two different systems. In System-1 the modules provided in the MSRA software have been used to build the new STool architecture. In System-2 we test a new entropy coder, called GOF-EMDC, which is an extended version of the EMDC coder. At the time GOF-EMDC codec and other parts of System-2 have not been optimized in many aspects, therefore we can expect better performance from our system in the next future. Despite this fact System-2 provides similar coding performances when compared to System-1. In addition, System-2 is much more flexible in many aspects, it guarantees a major number of functionalities and better fulfill the requirements list. Therefore with System-1 we intend to demonstrate the characteristics of the STool architecture, especially with respect to the reference software used, while with System-2 we customize and add functionalities to STool.

Yiannis Kompatsiaris (ITI) attended the 72nd MPEG meeting in Korea, where he gave presentations in relation to the following two input documents:

- (11) V. Mezaris, H. Doulaverakis, I. Kompatsiaris, and M. G. Strintzis, "Results of the use of MPEG-7 XM-based SCHEMA Reference System for segmentation algorithm evaluation" ISO/IEC JTC1/SC29/WG11, MPEG2005/M12034, presented at the 72nd MPEG Meeting, April 2005, Busan, Korea.
- (12) V. Mezaris, H. Doulaverakis, I. Kompatsiaris, M.G. Strintzis, "Application of the MPEG-7 XM-based SCHEMA Reference System to art image retrieval" ISO/IEC JTC1/SC29/WG11, MPEG2005/M12035, presented at the 72nd MPEG Meeting, April 2005, Busan, Korea.

The first input document outlines the results of the use of the MPEG-7 XM-based SCHEMA Reference System for segmentation algorithm evaluation, being the first publication to report results of this activity undertaken by SCHEMA with the participation of most of its partners. The reported results involve those acquired using the MPEG-7 VCE-2 dataset, as agreed at the time the MPEG-7 VCE-2 dataset was made available to SCHEMA for testing purposes. There was considerable interest of the MPEG community for the testing procedure that was followed and the results that were reported.

The second input document outlines the application of the SCHEMA Reference System, and therefore of MPEG-7 XM as well, to art image retrieval. This is an exciting and demanding application with significant exploitation possibilities, therefore being of importance to the MPEG community. The input document and the corresponding presentation given during the meeting provided valuable feedback on the usability of the MPEG-7 eXperimentation Model for this task.

In addition to the presentations that reported standards-related work within SCHEMA, discussions on MPEG's Multimedia Application Formats (MAF) took place during the meeting. Multimedia Application Formats provide the framework for integration of elements from several MPEG standards into a single specification that is suitable for specific, but widely usable applications. Possible MAF application scenarios under consideration are the delivery of music and the delivery of pictures or home videos over a network. SCHEMA, being a real-world application that successfully integrated MPEG technologies in a system suitable for the retrieval and delivery of retrieved image/video content over the Internet, was represented in these discussions. As a result of this interaction with the MPEG community, the SCHEMA Reference System will possibly be used as a platform for testing MAF-related technologies, strengthening the interaction between SCHEMA and MPEG. Furthermore, it was discussed that the SCHEMA Reference System could serve in one of the two Core Experiments needed for adopting these technologies by MPEG.

- (13) E. Cooke, A. Smolic, N. O'Connor "Results of Exploration Experiment: Multi-View Synthesis Reconstruction using DIBR Nodes" ISO/IEC JTC1/SC29/WG11, MPEG2005/M11973, presented at the 72nd MPEG Meeting, April 2005, Busan, Korea.

The 72nd MPEG meeting was held in Busan in South Korea between the 17-22/04/05. DCU attended this meeting to participate in both the MPEG-7 group and the Video and Synthetic Natural Hybrid Coding (SNHC) subgroup 3D Audio-Visual (3DAV) meetings. The importance of attending this meeting was twofold: Firstly to keep abreast with the latest work being completed within the MPEG-7 community, since we have both implemented and extended numerous MPEG-7 descriptors within the SCHEMA reference system; and to present and discuss our ideas on extending research beyond using just 2D descriptors for content-based information retrieval by developing 3D descriptors within the 3DAV group.

One of the main objectives of SCHEMA was the creation of a Reference System, which is a general architecture for content-based analysis, representation, indexing and retrieval. The Reference System

utilised the MPEG-7 XM (eXperimentation Model) for descriptor extraction and in order to support search and retrieval functionalities. At the MPEG-7 group meeting a number of potentially interesting Visual Core Experiments (VCE) were presented and discussed which had various levels of similarity to the work we completed in SCHEMA. These VCEs were based on new tools for image libraries: (1) Image categorisation information classes, (2) Applications of face recognition to image and video libraries, (3) Situation view based photo clustering. Also an ad-hoc group on the maintenance of MPEG-7 Visual related documents and reference software was set-up. Here, there is potential for SCHEMA members to continue the good work we completed during the project in correcting bugs and extending the MPEG-7 descriptors. A highlight for the SCHEMA group at this MPEG-7 meeting was the presentations by Ioannis Kompatsiaris of ITI. He presented two input documents written by Vasileios Mezaris, Haralambos Doulaverakis, Ioannis Kompatsiaris and Michael G. Strintzis entitled "Results of the use of MPEG-7 XM-based SCHEMA Reference System for segmentation algorithm evaluation" and "Application of the MPEG-7 XM-based SCHEMA Reference System to art image retrieval". These were both well received by the MPEG-7 group.

At the 3D Audio-Visual group, which is currently an ad-hoc group with members from both the Video and SNHC communities, DCU presented input document M11973 entitled "Results of Exploration Experiment: Multi-View Synthesis Reconstruction using DIBR Nodes". In this document we presented the results of work we are doing in an Exploration Experiment in the area of 3D video-based processing. Our goal here is to extend the normally 2D centric content-based information retrieval paradigm with the use of 3D descriptors. Such descriptors are intended to support queries such as "show me this object/item from a different user defined angle" and "show me similar views of similar objects/items". In order to define these descriptors we are examining leveraging camera calibration, 3D depth information, internal and external camera matrix parameters, and view orientation information. These 3D descriptors will be useful for improving and extending tasks such as the TRECVID 2004 collaboration which SCHEMA was involved in called SchemaTREC. The 3DAV group is currently identifying ways of standardising such 3D information within video streams. We had some discussions on such camera specific details and how storage of such information might be implemented particularly in a multiple camera environment. Work also completed within this group was the revision of the requirements for Multi-View Video Coding and a creation of a draft document on the exploration perspectives of Multi-View Video Coding.



### **Activities related to MPEG-7 XM**

The list of contributions to MPEG-7 related to the Reference Software includes the development of the following modules:

**Module combining MPEG-7 Extraction Tools for still image descriptors:** This module provides a simple interface to extract all still image descriptors with a few function calls. It also encapsulates the memory management for descriptor data and the individual extraction modules. The main benefit of this module is the improved re-usability of the visual descriptors in other system environments.

**Module combining Binary Coding Schemes of MPEG-7 visual descriptors:** Complementary to the extraction tools, this module allows encoding and decoding MPEG-7 visual descriptors to and from a file to simplify the interface to a few function calls. The purpose is also to increase the re-usability. This module includes memory management of coding schemes of visual descriptors and an abstract buffer model class. Besides the abstract buffer model, a real record based buffer model was designed.

**Module combining MPEG-7 Matching Tools for still image descriptors:** This module allows to call all matching functions of MPEG-7 still image descriptors with a few function calls. On one hand, this module increases the re-usability of the matching functions. On the other hand, this module is a platform to research on how the distance metrics of different descriptors can be combined by providing a complete simulation infrastructure. Research on this item is important because it is a central obstacle on pattern recognition. Solving an open problem in this domain reduces the gap between low level features and high-level semantics.

**Indexing module for the record based buffer class:** This module enables to use existing index structures to optimise the search strategies in descriptor databases. Previously, only a linear search was implemented. The target of this module is to test different indexing strategies, which are not part of this module.

**Library model for MPEG-7 descriptors:** This work groups the different modules of the MPEG-7 Reference Software (XM) thus they can be reused in a more selective way. Now there are libraries for:

- Basic components (e.g., media data structures, media decoders,...),
- Visual components
- Audio components
- Multi Media Description Schemes (MMDS) components
- Common components (components used by two of the parts: Visual, Audio, and MMDS)
- Application components (that call the Visual, Audio, and MMDS components within the MPEG-7 XM System)

**Module replacing the MPEG-7 XM Application module for visual low level descriptors:** The purpose of this module is to provide another interface to call the extraction and the matching (search & retrieval) application than the implemented MPEG-7 XM Application module. In the XM, the output interface (of the search & retrieval application) is the stdout of the process. In order to use the retrieval results in a different system (e.g., in the QIMERA system) than the XM, access to the retrieval results is required.

Other contributions to MPEG-7 related to the Reference Software include

- Draft of MPEG-7 Part 6 (Reference Software) and Amendment 1
- Bug fixes in the MPEG-7 XM
- Bug fixes in the Segment DS (central to MPEG-7)
- Update of software implementation templates
- W6188 FDAM-1 15938-6 (MPEG-7 Part 6, Reference Software)
- W6251 FPDAM-1 15938-6 (MPEG-7 Part 6, Reference Software)
- W6511 FPDAM-1 15938-6 (MPEG-7 Part 6, Reference Software) (editorial fixes)

- XM Software Release 6.1 (No documentation)

### **Participation to VCE-2 & the SCHEMA segmentation evaluation activity**

The MPEG-7 Visual Core Experiment 2 (VCE-2) is concerned with the grouping of images into semantic classes. It is one of a number of Core Experiments aiming at investigating the suitability of the MPEG-7 tools for advanced applications, such as semantic image categorization. ITI closely follows this ongoing activity and has already contributed to the generation of the necessary ground truth data, by manually categorizing a number of images into the 7 image classes currently defined under VCE-2. This collaboratively generated ground truth is to be used by all VCE-2 participants.

In addition to the VCE-2 experiments, this image data set and the corresponding ground truth will be used, after discussions in the relevant MPEG Ad-Hoc group, as one of the image databases of the ongoing SCHEMA segmentation evaluation activity. The motivation behind this activity is as follows: whilst there exist methods for the objective evaluation of segmentation algorithms, primarily using pre-existing or manually-generated ground truth for the segmentation, these do not provide a clear indication of the suitability of a given segmentation algorithm for the application of content-based image/video indexing and retrieval. This is not only due to the nature of objective evaluation methods, which does not in any way consider the specific application of content-based retrieval, but also to the need for ground truth, which typically limits evaluation to a few dozen images at most. To counter these drawbacks we propose the comparative evaluation of segmentation algorithms in terms of their contribution towards improving retrieval accuracy. More specifically, we proposed the application of the different segmentation algorithms to be evaluated (those already integrated with the SCHEMA Reference System) on an extended common image/video collection (already part of the SCHEMA Reference System). Besides, the extraction of common indexing features for all resulting regions, the submission of a large number of queries on each set of regions generated by a different segmentation algorithm, and the comparative evaluation of the segmentations on the basis of the precision-recall curves resulting from the retrieval experiments.

All these functionalities are currently supported by the SCHEMA Reference System. A significant number of partners and affiliated members are involved in performing the experiments required for the evaluation activity. The results of it are expected to be reported to MPEG, to highlight the suitability of a general-purpose image retrieval system, the SCHEMA Reference System, for differentiating between images belonging to the different semantic categories, and to indicate the effect of different segmentation algorithms on its performance.

### **Other Standards-related activities**

- (14) V. Mezaris, H. Doulaverakis, S. Herrmann, B. Lehane, N. O'Connor, I. Kompatsiaris, and M. G. Strintzis, "Combining Textual and Visual Information Processing for Interactive Video Retrieval: SCHEMA's Participation in TRECVID 2004", presented by ITI at TRECVID 2004 Workshop, November 2004, Gaithersburg, MD, USA.

In this paper, the two different applications based on the Schema Reference System that were developed by the SCHEMA NoE for participation to the search task of TRECVID 2004 are illustrated. The first application, named "Schema-Text", is an interactive retrieval application that employs only

textual information while the second one, named "Schema-XM", is an extension of the former, employing algorithms and methods for combining textual, visual and higher level information. Two runs for each application were submitted, I\_A\_2\_SCHEMA-Text\_3, I\_A\_2\_SCHEMA-Text\_4 for Schema-Text and I\_A\_2\_SCHEMA-XM\_1, I\_A\_2\_SCHEMA-XM\_2 for Schema-XM. The comparison of these two applications in terms of retrieval efficiency revealed that the combination of information from different data sources can provide higher efficiency for retrieval systems. Experimental testing additionally revealed that initially performing a text-based query and subsequently proceeding with visual similarity search using one of the returned relevant keyframes as an example image is a good scheme for combining visual and textual information.

The work carried out for the participation of Schema to TRECVID 2004, reported in the above paper, and the corresponding oral presentation and demo of the system given at the TRECVID conference by DCU, resulted in Schema being invited by the organizer (US National Institute of Standards and Technology - NIST) to give a presentation at the Joint MPEG/JPSearch sessions held during the 71<sup>st</sup> MPEG Meeting in Hong Kong, China. This presentation was given by ITI, which attended the 71<sup>st</sup> MPEG Meeting and the Joint MPEG/JPSearch sessions and had the opportunity to start following closely the new JPSearch standardization activity.

- (15) Video, "Description of Core Experiments in SVC", ISO/IECJTC1/SC29/WG11/N6373, 68<sup>th</sup> MPEG meeting, Munich, Germany, March 2004.

At this meeting, UNIBS joined the MPEG-21 Scalable Video Coding group. A preliminary core experiment, see doc N6373, was set up with the intent to test performance and functionalities of the available scalable video architecture and algorithms.

- (16) Video, "Description of Core Experiments in SVC", ISO/IECJTC1/SC29/WG11/N6521, 69<sup>th</sup> MPEG meeting, Redmond, USA, July 2004.

On the basis of the evaluation of core experiments defined at the 68th meeting two main scalable video coding approaches were identified: one based on scalable extension of MPEG-4 AVC and the other based on wavelet technology. A competitive phase was then started with the intent to select the class, MPEG-4 AVC or Wavelet, and the tool to use as reference model in the standardization process. UNIBS decided to attend the competition proposing a Wavelet based encoder named Stool (see document N6521).

- (17) Hyoung Joong Kim, Yong Soo Choi, Yongju Cho: "On the Symbolic Music Representation of Traditional Korean Music"

KNU (Kangwon National University), an associate member of SCHEMA, contributed SMR (Symbolic Music Representation) to extend existing electronic music notation standardization by adding traditional Korean music notations which are quite different from western music notations. SMR is a new and essential part of MPEG-4 audio standardization

- (18) "Workplan for the Evaluation of Responses to the CfP on Symbolic Music Representation"

This document, even if not supported by all the information and test material required by the CfP, as

been reviewed by the Audio subgroup and regarded as a very interesting proposal that may be easily integrated into an available core technology.

### 3. Contributions to JPEG 2000

#### *OpenJPEG*

The OpenJPEG library is an open-source JPEG 2000 codec written in C language. It has been developed in order to promote the use of JPEG 2000, the new still-image compression standard from the Joint Photographic Experts Group (JPEG). In addition to the basic codec, various other features are under development, among them the JP2 and MJ2 file formats, an indexing tool useful for the JPIP protocol, JPWL-tools for error-resilience, a Java-viewer for j2k-images, etc.

So far, until the redaction of this document the status of JPEG200 Parts 1 to 6 have been completed :

- Part 1: JPEG 2000 Core technologies
- Part 2: JPEG 2000 extensions
- Part 3: Motion JPEG 2000
- Part 4: Compliance Testing
- Part 5: Reference Software
- Part 6: Mixed Raster Coding

In addition new parts of JPEG 2000 are listed:

- Part 8: Security (JPSEC) WD v3.6, FDIS scheduled for 02/2005
- Part 9: APIs and Interactive Protocol (JPIP) Final FCD (v2.0), FDIS scheduled for 07/2004 (almost completed)
- Part 10: 3D JPEG 2000 and floating point data (JP3D)WD v3.0, FDIS scheduled for 02/2005
- Part 11: Wireless JPEG 2000 (JPWL)WD, FDIS scheduled for 02/2005
- Part 12: ISO Media File Format (superset of MP4 & MJ2)
- Common with MPEG (ISO/IEC 14496-12)

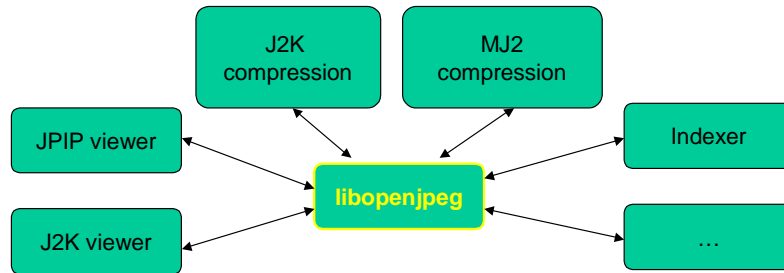
UCL is developing a JPEG 2000 coder and decoder (codec) called OpenJPEG. This codec is totally open source and distributed under a BSD license on the internet at the address: <http://www.openjpeg.org>. The codec respects the part 1 of the JPEG 2000 standard.

This codec is based on the library of J2000 (<http://www.j2000.org>). Many improvements have been done on it by the UCL and his collaboration with the I3S.

The coder is compliant with the profil-0 defines in the part 1 of JPEG 2000 standard. Besides all these options, the coder has been also modified to support the coding of mega images. These coder and decoder are developed in the optic to make them completely compliant with all the part 1 of the standard JPEG 2000.

#### **Application Demonstrator**

A mega-image navigator based on the JPEG 2000 technology is currently being developed by UCL. The targeted applications are medical imagery, remote sensing and on high-res satellite imagery. It studies the user behavior to anticipate its actions and improve the visual smoothness of the navigation.



**OpenJPEG Structure**

New tools based on OpenJPEG, the open-source JPEG 2000 codec of the lab have also been developed during the first year of the WCAM project:

- A JPEG 2000 image viewer based on a Java interface,
- A video viewer for MJ2 files transmitted on a network interfaced with a codestream transcoder.

### **MJ2 Module**

UCL has also developed the MJ2 module. Its goal is to handle Motion JPEG 2000 files. A first version of this module was designed to be linked with OpenJPEG, the JPEG 2000 open-source codec developed by UCL. In a second phase, the module was re-designed to be connected to any other JPEG 2000 codec. In the WCAM project, it is connected to the codec developed by TCF.

This capability of being connected to any codec is interesting for two main reasons:

1. Like OpenJPEG, the MJ2 module is an open-source software. Since many hardware and software implementations of JPEG 2000 exist, it was important to provide this modular open-source code, so that it could work with those existing implementations.
2. In the WCAM project, a more optimised implementation of the JPEG 2000 algorithm than OpenJPEG is used (the J2K codec developed by TCF). It was very important for this codec to be able to deal with MJ2 files.

### **Visualisation and streaming tool**

UCL has also developed a visualisation and streaming tool for MJ2 files and transcoded MJ2 streams. It enables a good validation tool for the transcoding and MJ2 modules. The streaming is realized using the RTP on UDP protocol. When the server detects that a client wishes to receive a video stream, it sends the complete stream to the transcoding module, that can be located on the same machine, or on another one. The transcoder detects the user resolution, bandwidth, and modifies the scalable stream if necessary, as explained above. If more than one user wish to receive the video, the stream is sent only once from the server to the transcoder, and the transcoder multiplies the stream if necessary, according to each client's characteristics.

This scalable transmission scheme has been adapted to a multicast environment. The main goal of multicast is to avoid the repetition of a same stream if multiple users close one from another wish to receive it. Each node of the multicast network would be able to transcode the stream if necessary, after multiplying it. A multicast protocol combined with multiple layer compression standards, like scalable video, offers a great opportunity for future multimedia transmission techniques.

### Other Standards-related activities – JPSearch

JPSearch is a new standardization activity, aiming at standardizing not only the low-level descriptors that are necessary for the retrieval of visual information, like MPEG-7, but also a generic architecture for extracting, storing and using the metadata, addressing, among others, issues such as the synchronization of the metadata with the visual content and digital rights management. This is an activity very relevant to SCHEMA. As a result, both isolated SCHEMA partners and SCHEMA as a whole have been actively involved in the development of the JPSearch standard by making contributions such as the following:

- (19) Serkan Kiranyaz: "Compression and Scaling Effects on Content-Based Image Indexing and Retrieval" Document ISO/IEC JTC1/SC29/WG1 N3503 Presented by TUT at the 71st MPEG Meeting, January 2005, Hong Kong, China.

In order to gain processing speed and save memory, a study was conducted to evaluate the effects of digital compression and downscaling on content-based multimedia retrieval using color and texture attributes. The idea is to query compressed/downscaled image databases instead of the original uncompressed ones, thus saving time and memory. Subjective evaluation tests are applied on digital image databases using different compression and visual feature extraction techniques have been performed. The results are particularly relevant to applications in which a mobile device is involved in a multimedia retrieval system. No related results are reported in the literature so far. Therefore, with this ongoing study we aim to cover the gap between two ongoing but independent activities in the digital multimedia world, namely lossy compression and content based indexing and retrieval. It is a significant factor to know the effects of compression (both in size and bits) on the performance of retrieval since ideally the content does not change with the parameters of the digital media, such as bit-rate, frame size, encoding scheme, etc. but the descriptors and hence the retrieval performance are immensely effected by those parameters as the experimental results confirm.

- (20) Serkan Kiranyaz: "MUVIS: A content-Based Multimedia Indexing and Retrieval Framework", Document ISO/IEC JTC1/SC29/WG1 N3502 Presented by TUT at the 71st MPEG Meeting, January 2005, Hong Kong, China.

The current MUVIS system has been developed as a framework to bring a unified and global solution to content-based multimedia indexing and retrieval. Due to variations and factors in today's digital multimedia world such as codec types, file formats, capture and encoding parameters, such factors significantly affect the indexing and retrieval scheme. Therefore, covering a wide-range of multimedia family and especially the last generation multimedia codecs, MUVIS is developed to provide an efficient framework structure upon which robust algorithms can be implemented, tested, configured and compared against each other. Furthermore, it supports browsing, hierarchic video representation and video summarization and most important of all, MUVIS framework supports integration of the aural and visual feature extraction algorithms explicitly. This brings a significant advantage for third parties to develop and test several feature extraction modules that are independent from the MUVIS applications development. In this context we believe that MUVIS can be conveniently used as a development and test-bed platform for the advanced feature extraction techniques.

Supported by the core MUVIS API, M-MUVIS system is an extension of contemporary MUVIS into a realm of embedded devices. Since the existing information delivery structure of World Wide Web is ill-suited for content distribution and reusing the information for diverse range of platforms including mobile phones and PDAs so we have proposed a system which has the capability to cope up with information management across heterogeneous networks and wide range of devices. The proposed

system has been tested with an image database. We have obtained quite encouraging results based on content-based retrieval. However, we are facing limiting factors such as hardware/software features in mobile devices and unpredictable network responses. With 3G round the corner and available high-end devices such limitations and drawbacks can be overcome. M-MUVIS holds a promising future in information content management and retrieval as it can be tuned with different combinations and weights of low-level features to meet requirements of certain enterprise or private usage.

- (21) JPSearch ad-hoc group: "Image Search system components and standardization scope recommendations" Proposal from ad-hoc group, Document ISO/IEC JTC1/SC29/WG1 N3506, with participation of TUT on behalf of SCHEMA at the 71st MPEG Meeting, January 2005, Hong Kong, China.

This document describes a complete image search system. For each component or step in the search process, we have made an initial determination of whether WG1 should attempt to standardize that step, either informatively or normatively.

In addition to the above contributions, a presentation on the SCHEMA Reference System entitled "Region-based image retrieval using MPEG-7 XM and high-level features: The SCHEMA Reference System" was given by ITI at the 35th ISO/IEC JTC 1/SC 29/WG 1 meeting held in Lisbon, Portugal, in March 2005. This presentation familiarized the JPEG community with the SCHEMA Reference System, the application of it that was built for participation of SCHEMA to the TRECVID 2004 activity and its results. The JPEG community showed interest in the outlined architecture and functionality of the system and characterized it as being very relevant to the scopes of JPSearch. There was also interest in the use of the MPEG-7 features in the SCHEMA Reference System, since the JPSearch standard is expected to adopt a similar architecture for utilizing MPEG-7 low-level features as its low-level descriptors.



## 4. Conclusion

The MPEG and JPEG standards are strongly related with the research objectives of the SCHEMA Network of Excellence. Work Package 4 focuses on the exchange of information between the relevant standardization bodies and the partners of SCHEMA.

The scope of this document was to bring the main picture to the reader about the contributions that the SCHMEA Network of Excellence has done in the past 30 months of project. In this report, contributions to standards, experiments as well as other standards-related activities from the SCHEMA partners are reported. Current developments related to standardization activities have been described in section 1 and specific contributions from SCHEMA partners to the MPEG and JPEG2000 standards have been reported in sections 2 and 3.

## 5. References, MPEG standards & parts and table of acronyms

### References

1. Open JPEG codec: <http://www.openjpeg.org>
2. J2000 Library: <http://www.j2000.org>
3. [http://www.chiariglione.org/mpeg/working\\_documents.htm](http://www.chiariglione.org/mpeg/working_documents.htm) for a list of working documents related to MPEG standardization process

### MPEG standards and parts

1. MPEG-1: ISO/IEC 11172 MPEG – 1 (1992) "Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s"  
Part 1 - MPEG-1 Systems - Program Stream  
Part 2 - MPEG-1 Video for CD -I  
Part 3 - MPEG-1 audio including Layers I, II, and III(a.k.a. mp3)  
Part 4 - Conformance  
Part 5 - Technical Report
2. MPEG-2: ISO/IEC 13818 MPEG – 2 (1994) "Generic coding of moving pictures and associated audio information"  
Part -1 Systems - joint with ITU - 2nd edition Dec 2000  
Part -2 Video - joint with ITU - 2nd edition Dec 2000  
Part -3 Audio- forward/backward compatible to MPEG-1(1998)  
Part -4 Conformance (1998)  
Part -5 Technical Report (1998)  
Part -6 DSM CC - Digital Storage Media Cmd and Cntl (1998)  
Part -7 AAC - Advanced Audio Coding (1997)  
Part -9 RTI - Real Time Interface (1996)  
Part -10 Conformance Extensions for DSM-CC (1999)  
Part -11 IPMP on MPEG-2 Systems
3. MPEG-4: ISO/IEC 14496 MPEG – 4 (1998) "Coding of audio-visual objects"  
Part 1 Systems (2001)  
Part 2 Visual (2001)  
Part 3 Audio (2001)  
Part 4 Conformance (2002)  
Part 5 Reference Software (2002)  
Part 6 DMIF - Delivery Multimedia Integration Framework (2000)  
Part 7 Optimized Software (2002)  
Part 8 MPEG 4 on IP (2002)  
Part 9 Reference Hardware (2003)  
Part 10 Advanced Video Coding (AVC) joint with ITU-T (2003)  
Part 11 Scene Description and Application Engine  
Part 12 ISO Base Media File Format  
Part 13 IPMP Extensions  
Part 14 MP4 File Format  
Part 15 Advanced Video Coding File Format  
Part 16 Animation Framework eXtension (AFX)
4. MPEG-7: ISO/IEC 15938 MPEG – 7 (2001) "Multimedia content description interface"  
Part 1 Systems

- Part 2 DDL - Description definition language
- Part 3 Visual
- Part 4 Audio
- Part 5 Multimedia description schemes
- Part 6 Reference software
- Part 7 Conformance testing
- Part 8 Extraction and use of description
- 5. MPEG-21: ISO/IEC 21000 MPEG- 21 (2004?) "Multimedia framework"
  - Part -1 Technical Report (2001)
  - Part -2 Digital item declaration (2002)
  - Part -3 Digital item identification (2002)
  - Part -4 IPMP
  - Part -5 Rights Expression Language
  - Part -6 Rights Data Dictionary
  - Part -7 Digital Item Adaptation
  - Part -8 Reference Software
  - Part -9 File Format
  - Part -10 Digital Item Processing
  - Part -11 Evaluation Methods for Persistent Association
  - Part -12 Test Bed for MPEG-21 Resource Delivery

### **MPEG acronyms**

<b>Acronym</b>	<b>Name</b>	<b>Description</b>
WD	Working Draft	Any MPEG document starts out in life as a working draft
CD	Committee Draft	Final MPEG Committee document (proposed standard specification) to be balloted with comments by ISO/IEC SC29/JTC1 National Bodies
DoC	Disposition of Comments	Comments processed and approved at MPEG meetings
FCD	Final Committee Draft	Proposed standard specification after voting with comments from ISO National Bodies (3 months)
FDIS <sup>©</sup>	Final Draft International Standard	Specification after voting (yes/no, without comments) from National Bodies (4 months)
IS <sup>©</sup>	International Standard	Standard Document (2 months)
PDAM	Proposed Draft Amendment	Amendment intended to add new technology to the specification
FPDAM	Final Proposed Draft Amendment	Same as FCD for amendments
FDAM <sup>©</sup>	Final Draft Amendment	Same as FDIS <sup>©</sup> for amendments
AMD <sup>©</sup>	Amendment	Same as IS <sup>©</sup> for amendments
DCOR	Draft Corrigenda	Corrective actions to one element of the specification
COR <sup>©</sup>	Corrigenda	Same as IS <sup>©</sup> for corrigenda

<sup>©</sup> indicates ISO copyright