INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC 1/SC 29/WG 11 N2114

March 1998

Source: **Leonardo Chiariglione – Convenor**
Title: **Report of the 43rd WG 11 meeting**
Status:

1. Opening
The 43rd WG11 meeting was held in Tokyo, JP on 98/03/16-20 at the kind invitation of JITSC, the Japanese National Body.

2. Roll call of participants
This was not made for reasons of time

3. Approval of agenda
The agenda was approved (Annex 1)

4. Allocation of contributions
Annex 2 gives the list of documents submitted by members, National Bodies and other organisations.

5. Communications from Convenor
There were no special communications made

6. Report of previous meeting
This could not be approved because one Chair failed to deliver the report of his group.

7. Processing of NB Position Papers
Documents were processed and responses made.

8. MPEG Phase 1
No work was done

9. MPEG Phase 2

9.1 Part 2 (Video)
No work was done

9.2 Part 10 (DSM-CC Conformance)
The FCD was produced

9.3 Verification of MPEG-2

No new results became known

9.4 Amendments

9.4.2 AAC Conformance Testing (part 4 #1)
A Study was produced

9.4.3 Simulation Software (Part 5 #1)
A Study was produced

9.5 Corrigenda
A "Study on MPEG-2 AAC 13818-7 Draft Technical Corrigendum 1" and a Study on DoC on MPEG-2 AAC 13818-7 Draft Technical Corrigendum 1 were produced.

9.6 Error! Bookmark not defined.
This was approved.

10. MPEG Phase 4

10.1 Version 1

10.1.1 ***Error! Bookmark not defined.***
Some statements were collected

10.2.1 Final Committee Draft

10.2.1.1 System
This was produced and approved

10.2.1.2 Video
This was produced and approved

10.2.1.3 Audio
This was produced and approved

10.2.1.4 Reference software
This was produced and approved

10.2.1.5 DMIF
This was produced and approved

10.1.3 Verification Tests
10.1.3.1 Video
Firms plans were made to produce results by the next meeting

10.1.3.2 Audio
Firms plans were made to produce results by the next meeting

10.1.4 Quality of service
Further work was done during the meeting that built on previous work.

10.1.5 Conformance Testing

10.1.5.1 System
Some work was done

10.1.5.2 Video
Some work was done

10.1.5.3 Audio
Some work was done

10.1.5.4 DMIF
Some work was done

10.2 Version 2

10.2.1 Requirements
Requirements were updated

10.2.2 Tools

10.2.2.1 DMIF
New work was done

10.2.2.2 Systems
New work was done

10.2.2.3 Natural Video
New work was done

10.2.2.4 Synthetic Visual
New work was done

10.2.2.5 Natural Audio
New work was done

10.2.2.6 Synthetic Audio
New work was done

10.2.3 Verification Models

10.2.3.1 System
A new document was produced

10.2.3.2 Video
A new document was produced

10.2.3.3 Audio
A new document was produced

10.2.3.4 SNHC
A new document was produced

10.2.5 Working draft

10.2.5.1 Systems
A new document was produced

10.2.5.2 Video
A new document was produced

10.2.5.3 Audio
A new document was produced

10.2.5.4 Simulation software
A new document was produced

10.2.5.5 DMIF
A new document was produced

10.3
This was approved

11. MPEG Phase 7

11.1 Requirements

A new document was produced

11.2 Call for proposals

A new document was produced

11.3
This was approved

12. Overall WG11 workplan

This was approved

13. Explorations

Discussions were made about possible new ways to encode high resolution video in scalable mode.

14. Liaison matters

Documents were considered and responses provided.

15. Administrative matters

15.1 Error! Bookmark not defined.
This was approved

15.2 Promotion of MPEG

New versions of the MPEG-4 and MPEG-7 documents were produced

16. Organisation of this meeting

16.1 Tasks for subgroups

Requirements
- MPEG-4 Profiles
- MPEG-4 Applications
- MPEG-4 ver. 2 Requirements
- MPEG-7 Applications
- MPEG-7 Requirements
- MPEG-7 Proposal Evaluation
- MPEG-2 4:2:2P@HL Requirements
- ALC Requirements

Delivery
- MPEG-2 part 10 DoC and Final Text
- MPEG-4 part 6 DoC and Final Text
- Reference software
- MPEG-4 part 6 ver 2 WD

Systems
- MPEG-4 part 1 DoC and Final Text
- Reference software (Res. 11 San José)
- Conformity of ver. 2 to versioning policy
- MPEG-4 part 1 ver 2 WD containing
- BIFS extensions
- IP Management and Protection
- AAVS
- File Format

Video
- MPEG-4 part 2 DoC and Final Text
- Reference software (Res. 11 San José)
- Conformity of ver. 2 to versioning policy
- MPEG-4 part 2 ver 2 WD
- MPEG-2 4:2:2P@HL Amendment

Audio
- MPEG-2 part 7 Cor DoC and Final Text
- MPEG-4 part 3 DoC and Final Text
- Reference software (Res. 11 San José)
- Conformity of ver. 2 to versioning policy
- MPEG-4 part 3 ver 2 WD
- SNHC
- MPEG-4 part 2 DoC and Final Text
- Reference software (Res. 11 San José)

- Conformity of ver. 2 to versioning policy
- MPEG-4 part 1&2 ver 2 WD

Test
- Audio tests
- Video tests

ISG
- Complexity of all parts of MPEG-4
- Quality of service
- Conf.
- Appoint editor

Liaisons
- Study incoming liaison statements
- Produce outgoing liaison statements
- Produce responses to NB papers

16.2 Finalisation of meeting allocation
The following Joint meetings were held

Group 1	Group 2	Day	Time	Room	About
Req	Vid	Tue	2-3	Req	4:2:2
Req	Sys	Tue	4-5	Sys	Profiles and levels
Req	Aud	Tue	5-6	Aud	Profiles and levels
Aud	Test	Mon	3-4	Aud	NADIB VT
Del	Sys, Vid, Aud	Tue	3-4	Sys	QoS
Aud	ISG	Tue	10-10:30	Aud	Complexity
SNHC	Sys	Tue	11-12	Sys	CD comments
Req	ISG	Mon	4-6	Req	Visual Prof&Lev.
SNHC	ISG	Tue	3-4	SNHC	CGD
Req	Vid	Tue	12-13	Vid	Prof & Lev
Test	Vid	Tue	11-12	Test	Error Rob.
Req	SNHC	Tue	10-11	Req	Profiles
Audio	Test	Wed	5-6	Audio	VT
Audio	Req	Thu	12-13	Audio	P&L
Req	Video	Thu	2-3	Video	P&L
Video	Req	Thu	9-11	Video	MPEG-7
Video	Test	Thu	11-12	Test	VT
Req	Sys	Thu	8:30-10	Sys	P&L
Sys	A, V	Thu	5-6	Sys	M4F
Req	SNHC	Wed	5-6	Req	P&L
Req	ISG	Wed	3-4	Req	P&L

17. Planning of future activities
The following ad hoc groups were established

2185	AHG on MPEG-2 4:2:2@High Level

2132	AHG on MPEG-2 AAC dynamic range control (Schreiner)
2131	AHG on MPEG-2 Audio AAC Conformance and Technical Report (Coleman/Thom/Paley)
2131	AHG on MPEG-2 Audio AAC Conformance and Technical Report (Coleman/Thom/Paley)
2236	AHG on Systems Specifications Editing Eleftheriadis et. al.
2238	AHG on Systems and DMIF Specification Evolution Guido Franceschini et. al.
2244	AHG on IM 1 Software Platform Lifshitz et. al.
2238	AHG on Systems and DMIF Specification Evolution Guido Franceschini et. al.
2139	AHG on MPEG-4 Audio/Systems issues (Teichmann/Herpel)
2183	AHG on MPEG-4 video encoder optimization
2169	AHG on MPEG-4 video verification tests
2137	AHG on MPEG-4 Audio NADIB verification tests (Contin/Dietz)
2136	AHG on MPEG-4 Audio Verification tests (Edler/S-W Kim)
2129	AHG on video decoder Quality of Service
2239	AHG on Systems Conformance Dufourd et. al.
2134	AHG on MPEG-4 Audio Conformance (Spille)
2191	AHG on MPEG-4 normative composition
2190	AHG on MPEG-4 Profiles and Levels
2124	AHG on Mobile Network Connections
2237	AHG on Elementary Streams Management Carsten Herpel et. al.
2240	AHG on Advanced BIFS Signes et. al.
2241	AHG on MPEG-J Specification and Implementation Fernando et. al.
2242	AHG on M4F Dave Pawson
2243	AHG on Intellectual Property Management & Protection within MPEG-4 Niels Rump
2181	AHG on core experiments on coding efficiency in MPEG-4 video
2178	AHG on core experiments on coding of arbitrarily shaped objects in MPEG-4 video
2179	AHG on core experiments on error resilience in MPEG-4 video
2180	AHG on core experiments on multifunctional coding in MPEG-4 video
2184	AHG on MPEG-4 texture coding
2128	AHG on Computational Graceful Degradation
2233	AHG on Face and Body Animation
2234	AHG on 3D Model Coding
2142	AHG on MPEG-4 Audio error resilience for Version 2 (Dietz/Miki)
2138	AHG on MPEG-4 Audio tool complexity (Spille)
2144	AHG on MPEG-4 CELP Optimization(Sugiyama)
2135	AHG on MPEG-4 Structured Audio (Scheirer/Ray)
2141	AHG on MPEG-4 TTSI/FBA convergence (Lee/Ostermann)
2182	AHG on editing the documents of the MPEG-4 video verification model and the MPEG-4 visual working draft
2235	AHG on SNHC VM Editing
2182	AHG on editing the documents of the MPEG-4 video verification model and the MPEG-4 visual working draft
2133	AHG on MPEG-4 Audio FCD and Reference Software FCD progression (Grill/Purnhagen)
2133	AHG on MPEG-4 Audio FCD and Reference Software FCD progression (Grill/Purnhagen)
2192	AHG on MPEG-7 Requirements
2189	AHG on MPEG-7 Test and Evaluation Methods
2143	AHG on MPEG Audio web site (Thom/Purnhagen)

18. Resolutions of this meeting

These were approved

19. A.O.B

There were no other businesses

20. Closing

The meeting was closed on 98/03/20T21:45 with thanks to the hosting organisation

Annex 1

Agenda

1. Opening
2. Roll call of participants
3. Approval of agenda
4. Allocation of contributions
5. Communications from Convenor
6. Report of previous meeting
7. Processing of NB Position Papers
8. MPEG Phase 1
9. MPEG Phase 2

9.1 Part 2 (Video)
9.2 Part 10 (DSM-CC Conformance)
9.3 Verification of MPEG-2
9.4 Amendments
9.4.1 AAC Conformance Testing (part 4 #1)
9.4.2 Simulation Software (Part 5 #1)
9.5 Corrigenda
9.6 **Error! Bookmark not defined.**
10. MPEG Phase 4
10.1 Version 1
10.1.1 **Error! Bookmark not defined.**
10.2.1 Committee Draft
10.2.1.1 System
10.2.1.2 Video
10.2.1.3 Audio
10.2.1.4 Reference software
10.2.1.5 DMIF
10.1.3 Verification Tests
10.1.3.1 Video
10.1.3.2 Audio
10.1.4 Quality of service
10.1.5 Conformance Testing
10.1.5.1 System
10.1.5.2 Video
10.1.5.3 Audio
10.1.5.4 DMIF
10.2 Version 2
10.2.1 Requirements
10.2.2 Tools
10.2.2.1 DMIF
10.2.2.2 Systems
10.2.2.3 Natural Video
10.2.2.4 Synthetic Visual
10.2.2.5 Natural Audio
10.2.2.6 Synthetic Audio
10.2.3 Verification Models

Annex 2

Documents submitted

Source	No.	Title
Pete Schirling	3239	Document Register for 43rd Meeting in Tokyo, Japan
	3240	VACANT
Igor Pandzic, Claudio Lande, Yuval Fisher	3241	FBA revisions to the Systems CD
Gershon Bar-On, Gene Itkis	3242	Basic requirements for Content Management and Protection (CMP)
SC 29 Secretariat	3243	Revised Summary of Voting on ISO/IEC CD 13818-10
SC 29 Secretariat	3244	Revised Summary of Voting on ISO/IEC 13818-7/DCOR 1
Vahe Balabanian	3245	DMIF February 1998 (San Jose) Meeting Report
Caspar Horne	3246	Template for NB comments on visual CD
De Vleeschouwer Christophe, Macq Benoit	3247	Subband Dictionaries for Low Cost Matching Pursuits
Olaf Barheine (within AHG MPEG-4 Profiles and Levels)	3248	MPEG-4 Profiles/Levels Summary
DAVIC via the SC 29 Secretariat	3249	Liaison Statement from DAVIC on 1) Call for Proposals on End to End Digital Audio Visual Systems over IP based Networks, Storage in the Home, and Content Contribution Systems and 2) Call for Replacement of a Technology in the DAVIC Specifications (SC 29 N 2445)
Julien Signes	3250	Review of the Systems CD
SC 29 Secretariat	3251	Summary of Voting on CD 14496-1
SC 29 Secretariat	3252	Summary of Voting on CD 14496-2
SC 29 Secretariat	3253	Summary of Voting on CD 14496-3
ITU-T SG 16	3254	Liaison Statement from ITU-T SG 16 on MPEG-4 CDs
SC 29 Secretariat	3255	Summary of Voting on CD 14496-5
SC 29 Secretariat	3256	Summary of Voting on CD 14496-6
ITU-T SG 16	3257	Liaison Statement from ITU-T SG 16 on General Work Plan for MPEG-4 Verification Tests
ITU-T SG 16	3258	Liaison Statement from ITU-T SG 16 on DMIF Cooperation with H.310/H.323
ITU-T SG 16	3259	Liaison Statement from ITU-T SG 16 on Progress towards GII Project F.3 and F.4
Katsumi Tahara, Yoshihiro Murakami	3260	Requirements for the standardization on 422P@HL
Roberto Pockaj, Eric Petajan	3261	FBA revision to the Visual CD
Italian National Body	3262	URLs in BIFS and ODs/ESDs
Italian National Body	3263	DMIF URL format
Guido Franceschini	3264	The value of DMIF
Peter G. Schreiner III	3265	Report of the Ad Hoc Group on MPEG-2 AAC Dynamic Range Control
D.CURET	3266	AL-PDUs duplication
D.Curet	3267	The SDM and the DTS and CTS timestamps
D.Curet	3268	The SDM and the FlexMultiplexing
D.Curet	3269	OTB recovery with OCRs
T.K. Tan	3270	Report of the Ad Hoc group on core experiments on coding efficiency in MPEG-4 video
Hiroyuki Fukuchi	3271	Brief report of AAC conformance testing (SSR profile)
Niels Rump (as chair of ad-hoc group)	3272	Report of IPR Ad-hoc Group

C. P. Quist, A. Sandvand, G. K. Klungsoyr	3273	H.263 baseline decoding in MPEG-4 Visual
Jean-Claude Dufourd	3274	Corrections and clarifications for the Form node
Jean-Claude Dufourd	3275	Report of AHG on Systems conformance
A. Sandvand, Cor Quist	3276	Proposed syntax changes in the visual CD for H.263 decoding
Anne-Claude Doux, Philips LEP, Jean Gobert, Philips LEP, Andrea Barbieri, PACT, UK	3277	Shape parameter: Conversion Ratio and Alpha_Threshold
Cecile Dufour	3278	Extension of the VLC table for sprite point trajectories encoding
Roberto Pockaj, Loris Ambrosini, Maurizio Costa, Fabio Lavagetto	3279	Results of Core Experiment FA1 on Face Calibration
NNI (Nederlands Normalisatie Instituut)	3280	Dutch NB Votes and Comments to MPEG-4 CD's
Kevin Barron, Laurent Issner	3281	Rendering issues in Normative Composition
Kevin Barron, Graham Thomas, Mike Coleman	3282	Report of Ad-hoc group on Normative Composition
SC 29 Secretariat	3283	Late vote on CD 13818-10
SC 29 Secretariat	3284	Late vote and comments on 13818-7/DCOR 1
James Van Loo	3285	Adaptive Terminal Requirements
James Van Loo	3286	Adaptive Terminal Object Descriptor Protocol Interface
James Van Loo	3287	Adaptive Terminal Resource Allocation
James Van Loo	3288	Extensible Object Descriptor Protocol
Hideaki Kimata, Shigeru Fukunaga	3289	Backward Channel Scheme for Video Error Resilience
Shigeru Fukunaga, Hideaki Kimata	3290	Bitstream Exchange for Backward Channel
Peter Kauff, Klaas Schueuer, Minhua Zhou	3291	Results of mini experiment (SM2) for arbitrarily shaped texture
C. Herpel (Thomson), M. Rei?mann (Sican), J. Vollmer (Bosch)	3292	A First Approach Towards the Definition of Conformance Levels in MPEG-4 Systems
Gael RICHARD, Caroline VENOT	3293	Report of practical complexity evaluation of an optimised HILN decoder
Y.- B. Thomas Kim, S.-H. Park, S.-W. Kim	3294	Improved BSAC description in ISO/IEC CD 14496-3 Subpart 4
Seiji Kameyama	3295	Comments for Structured Audio
Singapore National Body	3296	Singapore NB comments on CD 14496-2 (Visual) and the Study of CD 14496-2 (Visual)
Young-Kwon Lim,, Youngjik Lee,, Jungchul Lee	3297	Consideration for Conformance of the MPEG-4 Audio TTSI
Singapore National Body	3298	Response of the Singapore NB comments to resolution 2.1.5 of the 42nd WG11 meeting
Kai-Kuang Ma, Prabhudev Housr, Lei Huang, Shan Zhu	3299	Status Report of Core Experiment on Fast Block-Matching Motion Estimation
T.K. Tan	3300	On Visual Combination Profiles and Levels
Franco Casalino, Gianluca Di Cagno	3301	IM-1 2D platform: supported BIFS nodes
Dutch National Body	3302	MPEG-4 Audio
Ralf Funken, Werner Oomen, Frans de Bont	3303	AHG report on Quality assessment of the 8kHz sampling mode of the MPEG-4 Audio VM w.r.t. G723.1
Ralf Funken, Werner Oomen, Frans de Bont	3304	Results of an informal asessment of the Quality of the 8kHz sampling mode of the MPEG-4 Audio VM w.r.t. G723.1

Portuguese National Body (IPQ)	3305	Portuguese Votes to MPEG-4 CD
R. Taori, R. Funken, W. Oomen	3306	Proposal for an optimized VQ for the MPEG-4 CELP
Shun-ichi Sekiguchi, Shin-ichi Kuroda, Yoshimi Isu and Kohtaro Asai	3307	Semantics of vop_rate_code
Di Cagno Gianluca, Casalino Franco	3308	IM1-2D platform version 2.7
French National Body	3309	Comment on OCI part of MPEG4 CD
AFNOR French National Body	3310	Letter ballot and comments on CD14496-1/3
Dominique Curet, Christine Guillemot, Patrice Soyer, Emmanuel Gouleau, Carlos Islas	3311	Some questions on MPEG4 over RTP
Gabriel Abrantes, Fernando Pereira	3312	Implementation and Evaluation of MPEG-4 Facial Animation Tools
Pascal Faudemay, Marie-Jo Caraty, Claude Montacie	3313	Benchmarking methods for audio
Karen Hsing, Chilsung Seo	3314	Conformance Abstract Test Cases for testing DSMCC-UN SDB-CCP protocol and Pass-thru protocol
Karen Hsing, Chilsung Seo	3315	A subset of the Conformance Abstract Test Cases for testing DSMCC-UN Session Protocol
Schuyler Quackenbush, Jim Johnston	3316	Proposal for AAC Dynamic Range Control
Yuval Fisher, Wei Wu, Homer Chen	3317	Binary Encoding of ECMAscript for the Script Node
Yuval Fisher, Hai Tao, Jorgen Ahlberg, Roberto Pokaj, Homer Chen	3318	Revised Description of FIT Node
David Singer, Michael Speer (Editors)	3319	Proposed Revised Intermedia Format (M4F) VM text
Giuseppe Russo, Stefania Colonnese	3320	User interaction modes in semi-automatic segmentation: development of a flexible graphical user interface in Java
Tom White (MMA)	3321	MMA Proposed Enhancements to SASBF
Michael Frater	3322	Australian National Body Comments on CD 14496-2
Michael Frater, Wee Sun Lee	3323	Proposal and Results for Error Resilient Coding of Still Texture
Jiankun Li, C.-C. Jay Kuo, Homer Chen	3324	Compression of Mesh Connectivity by Dual Graph Approach (M1)
Jiankun Li, C.-C. Jay Kuo, Homer Chen	3325	Embedded Coding of Mesh Geometry (M2)
Jiankun Li, C.-C. Jay Kuo, Homer Chen	3326	Progressive Mesh Coding by Vertex Split (M3)
C.S. Boon (1), M. Wollborn (2)	3327	Report of ad-hoc group on content-based coding pre-screening activity
C.S. Boon	3328	Extension of fixed length codes for DCT coefficients
Dan Tamir	3329	Summary of the AHG on Decoder QoS
Itaru Kaneko	3330	Comments on MPEG-4 content IPR and compatibility with Version 2
Francoise Preteux, Gerard Mozelle, Jose Paumard	3331	Report on Core Experiment M1, M2 and M3 on 3D Model Coding
Francoise Preteux,, Patrick Horain,, Hocine Ouhaddi,, Marius Preda	3332	Report on Core Experiment 3 on Hand BAPs interpretation

Francoise Preteux	3333	Liaison to CEN on Metadata for Multimedia Information on MPEG-7
Frank Bossen	3334	Results of Core Experiments on 3D model coding
Belgian National Body	3335	Belgian NB Comments on Computational Graceful Degradation
Jan Bormans, Marco Mattavelli	3336	Report of the ad-hoc group on Computational Graceful Degradation
Jan Bormans, Marco Mattavelli	3337	Report of the ad-hoc group on Computational Graceful Degradation
Michael Frater	3338	Report of ad-hoc group on core experiments on error resilience aspects in MPEG-4 video
Mike Coleman, David Thom, Chuck Leuck	3339	Report of Ad-hoc group on AAC conformance and Technical Report
Mike Coleman, Eric Scheirer	3340	Report of Ad-hoc group on MPEG-4 Audio/Systems Reference Model
Frank Nack, V V Vinod	3341	MPEG-7 Requirements Document V.5
Frank Nack, V V Vinod	3342	MPEG-7: Context and Objectives V.7
Seokwon Han, Sang Hoon Lee, Sung Ryul Cho, Jin Hun Kim	3343	Interlaced texture coding for arbitrary shape object
Seokwon Han, Sang Hoon Lee, Sung Ryul Cho, Jin Hun Kim	3344	Result of S12 (Interlaced video coding for arbitrary shape object)
Munchurk Kim,, Myoung Ho Lee,, Chieteuk Ahn	3345	Workplan for the combined algorithm of automatic segmentation schemes by ETRI, FUB and University of Hannover
Sang-Wook Kim, Bernd Edler	3346	Report of the Ad Hoc Group on MPEG-4 Audio Verification Tests
Iole Moccagatta, Osama Alshaykh	3347	Proposal for a Core Experiment on Still Texture Error Resilience using a Packet Approach
Ibrahim Sezan and Frank Nack (Chairmen)	3348	Report of AHG MPEG-7 Requirements
Jae Gark Choi,, Munchurl Kim,, Jinsuk Kwak,, Myoung Ho Lee,, Chieteuk Ahn	3349	User-assisted Video Object Segmentation by Multiple Object Tracking
AHG on MPEG-7 Requirements	3350	MPEG-7 Draft Proposal Package Description (PPD)
Fan Ling, Hongqiao Sun	3351	Results of Core Experiment on Bit-Plane Coding (T-14d)
Weiping Li	3352	Summary on Bit-Plane Coding Core Experiment
Hiroyuki Katata	3353	Definition of levels for new visual combination profiles
Tomoko Aono, Norio Ito, Hiroyuki Katata	3354	Proposal of Scalable Texture Coding with Fine Granularity
Norio Ito, Shin-ya Hasegawa, Masahiro Shioi, Hiroyuki Katata	3355	Proposal of Tiling function for Scalable Texture Coding
Toshio Miki	3356	Report of the Ad Hoc Group on MPEG-4 Video Verification Tests on Error Resilience
Toshiyuki Nomura, Masahiro Iwadare, Naoya Tanaka	3357	Listening test results and quality improvement of the MPEG-4 Narrow-Band CELP coder for multi languages
Yoichi Yagasaki, Teruhiko Suzuki	3358	Request for New Profile for Studio & Post-production application
Yuji Maeda, Masayuki Nishiguchi	3359	Subjective test results of codeword reordered HVXC
Yoichi Yagasaki, Teruhiko Suzuki	3360	Additional Requirements for Profile used in Broadcast application
Yo-Sung Ho,, Daehee Kim,, Minho Kim	3361	Implementation of the Spatio-temporal Segmentation Algorithm proposed by ETRI

The National Body of Japan	3362	Comment on the parametric speech coder part of the MPEG-4 Audio
Shinichi Sakaida, Michihiro Uehara, Tadashi Isobe, Yoshiaki Shishikui, Yuji Nojiri	3363	ISDB services as broadcasting applications of MPEG-4
Shinichi Sakaida, Yuko Yamanouchi, Masaki Hayashi, Yoshiaki Shishikui, Yuji Nojiri	3364	MPEG-4 studio production application and requirements
Gerard Fernando	3365	Report of AAVS AHG
Takefumi Nagumo,Teruhiko Suzuki,Yoichi Yagasaki	3366	The results of arbitrary shaped scalability in core experiment B1.2
Zhixiong Wu, Toshifumi Kanamaru, Yoshihiro Ueda, Yoichi Yamada	3367	Core Experiment Results of T-14d
Sang Gyu Park, Young-Kwon Lim, Chieteuk Ahn	3368	Proposal for use of same shape information in video coding
Teruhiko Suzuki, Yoichi Yagasaki	3369	Identification of the enhancement layer bitstreams
T. Ebrahimi, E. Jang, C. Horne, A. Puri, Y. Nakaya, J. Ostermann	3370	Report of Ad hoc group on editing the documents of the MPEG-4 video verification model and the MPEG-4 visual working draft
T. Ebrahimi, P. Doenges	3371	Report of ad hoc group on 3D model coding
Teruhiko Suzuki, Yoichi Yagasaki	3372	The Preliminary Results on Advanced Layered Coding
Ad hoc group members on 3D model coding	3373	Description of core experiments on 3D model coding
Iole Moccagatta, Osama Alshaykh	3374	Report on the investigation for the shortcomings of MPEG4: the "singer" case
Lionel Bouchard, Francoise Caillaud, Philippe Gentric, Jean Gobert, Laurent Herrmann, Beatrice Nezot,, Laurent Tricaud	3375	MPEG-4 2/3D players
Graham Pereboom	3376	Clarification of the NNI comments on CD 13818-10
Osama Alshaykh, Iole Moccagatta, Homer Chen, Houng-Jyh Wang, Yi-Liang Bao, C.-C. Jay Kuo	3377	Status of the core experiment on improving compression efficiency of still texture coding - adaptive scanning of wavelet subbands
Naoki Iwakami, Takehiro Moriya, Akio Jin, Kazunaga Ikeda, Takeshi Mori, Satoshi Miki	3378	Proposal of a revised text for the T/F part of the 14496-3 CD
Kazunaga Ikeda, Takeshi Mori, Takehiro Moriya, Naoki Iwakami	3379	Reports on the prescreening tests of the revised UEP tools
Se-Hoon Son, Sung-Jin Kim, Jae-Seob Shin	3380	Results of CE S12 (Interlaced Shape & Texture Coding)
Dae-Sung Cho, Euee S. Jang, Jae-Seob Shin	3381	Results of CE B 1.1.A: Combined scalable shape & texture
Dae-Sung Cho, Jae-Seob Shin	3382	Results of lossy shape using the CR and Alpha_TH parameter

Akio Jin, Takehiro Moriya, Naoki Iwakami, Takeshi Norimatsu, Mineo Tsushima, Tomokazu Ishikawa	3383	Modification of scaleable TwinVQ for harmonizing with AAC
Jae-Seob Shin(As a chairman of ad-hoc group)	3384	Report of the Ad-Hoc group on MPEG-4 Version Management
Graham Thomas	3385	Results of mini-experiment on composition methods for greyscale shape
Frederic Gilbert	3386	Liaison to WG11
Korean National Body	3387	On the efficient utilization of resources for the success of MPEG-4 Version 1
Korean National Body	3388	Comments on Study of CD 14496-2 (Visual)
Korean National Body	3389	Comments on CD 14496-2 Visual
Korean National Body	3390	Comments on the rate control of CD 14496-2 (Visual)
Korean National Body	3391	On efficient shape coding of CD 14496-2 Visual
Korean National Body	3392	Comments on Interlaced Texture Coding for Arbitrarily Shaped Object
Korean National Body	3393	A Study on the CD 14496-3 Audio
Korean National Body	3394	Profile and Audio/System issue of TTS in the CD14496-3 (Audio)
Korean National Body	3395	On BSAC of ISO/IEC CD 14496-3 (Audio)
Korean National Body	3396	Comments on ISO/IEC 13818-7/DCOR1
Martin Dietz, Ralph Sperschneider	3397	VM integration of HCR
Martin Dietz, Thomas Schäfer, Ralph Sperschneider	3398	Evaluation for the common EP tool
Stephen PALM, Taketoshi MORI, Tomomasa SATO	3399	Visually Based Control: another Application for MPEG-7
Adam Lindsay, Frank Nack, Ed Hartley	3400	An example Description for MPEG-7
Joaquim Esmerado, Tolga K. Capin	3401	EPFL Results of the Body Animation CEs
SNHC FBA group	3402	Proposal for update to body animation section of SNHC VM
Jin Soo Choi, Young-Kwon Lim, Myoung Ho Lee, Chieteuk Ahn	3403	Results of core experiments M2/M3: Geometry coding using PRVQ
Jinwoong Kim,, Munchurl Kim,, Kyu-Won Lee,, Hankyu Lee,, Jae-Gon Kim,, Donghyun Kwon	3404	Suggestions on the improvements of Draft of MPEG-7 Requirements
Jens Spille	3405	Report of Ad Hoc Group on MPEG-4 Audio Conformance
Sung-Moon Chun, Jong-Deuk Kim, Cheol-Soo Park, Jae-won Chung, Dongkyoo Shin, Hae-Kwang Kim and Joo-hee Moon	3406	The Results of CE S12: Interlaced Shape and Texture Coding
Jens Spille	3407	Report of Ad-Hoc Group on MPEG-4 Audio Tools Complexity
Jong-Deuk Kim, Choel-Soo Park, Jae-won Chung, Dongkyoo Shin, Joo-hee Moon	3408	Improvement on interlaced texture coding of arbitrarily shaped object with minimal syntax changes in the CD
Jae-won Chung, Dongkyoo Shin, Joo-hee Moon	3409	Technical Description of Vertex-based Binary Shape Coding

Ji-heon Kweon, Dongkyoo Shin, Hae-Kwang Kim, Joo-hee Moon	3410	Report on Bitstream Exchange for Greyscale Alpha Shape
C.P. Quist	3411	Test results H.263 decoding in MPEG-4 Visual
Ji-heon Kweon, Hae-Kwang Kim, Joo-hee Moon	3412	Results of Mini-Experiment SM2
Escher Marc, Pandzic Igor-Sunday	3413	Result of core experiment F1
German National Body	3414	Efficiency of MPEG-4 Video in Version 1
Zvi Lifshitz	3415	Systems Software Implementation AHG Report
Gauthier Lafruit, Jan Bormans	3416	Assessment of video spatial scalability complexity
Zvi Lifshitz	3417	MPEG-4 Off-line Multiplex Software Release 1.2
Zvi Lifshitz	3418	BIFS/OD Encoder Software Release 1.2
Zvi Lifshitz	3419	MPEG-4 Player Core Code Release 1.2
Zvi Lifshitz	3420	Test Application for Systems Core Code Release 1.2
Zvi Lifshitz	3421	Decoder Development Kit for IM1
Toshiyuki Nomura, Masahiro Iwadare, Naoya Tanaka	3422	Listening test results of the MPEG-4 Wide-Band CELP coder with the Multi-Pulse and the LSP-VQ tools
Jorgen Ahlberg	3423	Result of Core Experiment F2
Peter Westerink, Lurng-Kuo Liu	3424	AAVS API to the BIFS scene graph
Peter Westerink, Lurng-Kuo Liu	3425	AAVS API's for set-top box functionality
M.Zhou (HHI), E.S.Jang(SAIT) and S.Sakaida(NHK)	3426	Support of Multiple Alpha Channels in MPEG-4
Carsten Herpel	3427	Report of AHG on Systems Elementary Stream Management
Anthony Vetro, Huifang Sun	3428	Results of core experiment T14d: Bit-plane coding
Adam Lindsay, Ibrahim Sezan, Masahiro Shibata, Thomas Sikora, Toby Walker	3429	Program of MPEG-7 Seminar in Tokyo
Martin Dietz, Toshio Miki	3430	Report of the Ad-hoc group on MPEG-4 Audio error resilience for Version 2
Martin Dietz, Sang-Wook Kim	3431	List of selected items for the MPEG-4 Audio NADIB verification test
Ajay Luthra, Yoichi Yagasaki	3432	Report on MPEG2 4:2:2P@HL Ad-Hoc Group's Work
Gauthier Lafruit, Jan Bormans	3433	Computational Graceful Degradation and backward compatibility issues
SMPTE	3434	4:2:2 Profile @ HL - Requirements
SMPTE	3435	4:2:2 Profile @ HL - Table Entries
Liam Ward, Franco Casalino, Eddie Cooke, Karl Anders Oygard, Zvi Lifshitz	3436	Status of Systems Software
Bernhard Grill, Heiko Purnhagen	3437	Report of the adhoc group on MPEG-4 Audio CD and Reference software progression
Bernd Edler, Heiko Purnhagen	3438	Scalability for Parametric Audio by Means of Simple Transcoding
Bodo Teichmann	3439	Report of the Ad-hoc group on Audio/Systems issues
Riitta Vaananen, Jyri Huopaniemi	3440	Advanced sound parameters in MPEG-4 Version 2 BIFS

Sylvie Jeannin, Adam Lindsay	3441	Report of the Ad Hoc Group on MPEG-7 evaluation methods
Sylvie Jeannin, MPEG-7 evaluation Ad Hoc Group	3442	MPEG-7 Evaluation Process Document : draft of version 2
M. Wollborn	3443	Report of AHG on Random Access Video Coding
U. Benzler	3444	Results of bitstream exchange for quarter-pel motion compensation
Regis J. Crinon, Sam Narasimhan, Tim Addington	3445	Encapsulation of MPEG-4 streams in MPEG-2 DSM-CC Data Carousel
M. Wollborn, U. Pestel	3446	Results of University Hannover for core experiment B1.2 part B ("Real" temporal B-VOPs in spatial scalability)
Eric Scheirer	3447	Report of the AHG on Structured Audio
M. Wollborn, R. Mech	3448	Refined procedure for objective evaluation of VOP generation algorithms
Albert Hsueh, Anurag Bist, Wei Wu	3449	Intelligent Pre-Quantization in Motion Compensated Video Coding
Krit Panusopone, Xuemin Chen	3450	Comparison results between bit-plane coding and intra 3-D VLC for interlaced video
Frederic Dufaux	3451	More results on dynamic sprites and GMC
A. Puri	3452	Report of Ad hoc group on Multifunctional Coding in MPEG-4 Video
Dave Singer (Apple Computer), Michael Speer (Sun Microsystems), et al.	3453	MP4 File Format VM0.0
B. G. Haskell	3454	Support for Simple shape in conjunction with simple profile of MPEG-4
A. Eleftheriadis, A. Basso, H. Kalva, A. Puri, R. L. Schmidt	3455	Input for development of MP4 File Format
A. Puri, R. L. Schmidt, B. G. Haskell	3456	Strawman framework for MPEG-4 AAVS API
Peter K. Doenges	3457	Report of Ad Hoc Group on SNHC FAQs Editing
Viswanathan Swaminathan (Sun Microsystems), Gerard Fernando (Sun Microsystems)	3458	AAVS Architecture and Application Scenarios
Viswanathan Swaminathan and Gerard Fernando (Sun)	3459	Contributions for AAVS VM
Gary Sullivan	3460	On MPEG-4 Video Compatibility with H.263
Vahe Balabanian	3461	Mapping DMIF into IETF RSVP
Vahe Balabanian	3462	MPEG-4 End-to-End QoS Management
Vahe Balabanian-Nortel, Andrea Basso-AT&T	3463	MPEG-4 Media Control with IETF RTSP
Touradj Ebrahimi (editor)	3464	Video verification model 10.1
Vahe Balabanian	3465	Role of DMIF with RTP MPEG-4 Payloads
Cassandra Swain, Joern Ostermann, Qian Huang, Barry Haskell	3466	Requirements for Descriptors for 3D Human Body Model
Kevin O'Connell, Karlheinz Brandenburg	3467	Report of the MPEG-4 Profiles and Levels Ad-Hoc Group
Kevin O'Connell, Manish Singhal, James Brailean	3468	OBMC Complexity Evaluation Results
Gabriel Taubin (editor)	3469	SNHC Verification Model 7.0
Gabriel Taubin (editor)	3470	Report of AHG on SNHC Verification Model 7.0 editing
Gabriel Taubin	3471	Description of Core Experiments on 3D Model Coding
Gabriel Taubin	3472	Report of AHG on Core Experiments on 3D Model Coding

Eric Petajan, Tolga Capin	3473	Report of adhoc group on Face and Body Animation
Gabriel Taubin, Bill Horn, Andre Gueziec, Claudio Silva	3474	Report on Results of Core Experiments on 3D Model Coding (M1-M2-M4)
Gabriel Taubin, Bill Horn, Andre Gueziec, Claudio Silva	3475	Report on Results of Core Experiments on 3D Model Coding (M3)
Tihao Chiang (Sarnoff Corporation) and Huifang Sun (Mitsubishi Electric Information Technology Center America)	3476	Report of Ad Hoc Group on Encoder Optimization
Adam Lindsay	3477	MPEG-7 Applications Document
Iraj Sodagar	3478	Report of MPEG-4 Visual Texture Coding Ad-Hoc
Heiko Purnhagen, Bernd Edler	3479	Scalability for Parametric Audio Coding by Means of Simple Transcoding
Iraj Sodagar, Hung Ju Lee, Paul Hatrack, Bing-Bing Chai, Shipeng Li	3480	A Proposal for Improving Coding Efficiency of Visual Texture Coding
Iraj Sodagar, Hung Ju Lee, Paul Hatrack, Bing-Bing Chai, Shipeng Li	3481	Result of Core Experiment T1: Improvement on Visual Texture Coding Using 1-D Adaptive Wavelet
Euee S. Jang, Visual CD Editors	3482	MPEG-4 Version 2 Visual Working Draft Rev 2.1
David Thom, Heiko Purnhagen	3483	Report of the Ad-Hoc group on Audio web page
William Miller	3484	Proposed Draft Amendment to ISO/IEC 13818-2 (MPEG-2 Video)
William Miller	3485	Proposed Draft Amendment to ISO/IEC 13818-2 (MPEG-2 Video)

Annex 3

Report of Requirements meeting

Source: Rob Koenen, Chair

The Requirements meeting was very fruitful and took place in a cooperative atmosphere. Thanks go to all the participants who made this possible, as the discussions were important and difficult.

Special thanks for support during the meeting go to Jim Brailean, for editing all the work related to the Video FCD; to Niels Rump and Ibrahim Sezan, for chairing break-out sessions; to Fernando Pereira, for chairing a part of the meeting, to Olaf Barheine and Olle Franceschi, for editing the documents.

MPEG-2

Advanced Layered Coding

Advanced Layered Coding: (ALC) there are letters from various companies stating they would be interested in using ALC for their products in the area of high quality audiovisual services. An 'ALC Requirements and Applications Document' (N2193) has been drafted, starting from the Preliminary ALC Requirements Document (N1842, issued at the 41st MPEG meeting, in October 1997, in Fribourg.)
Work will continue in collecting evidence, so that a decision can be taken. It will then also be decided what the best context is: MPEG-2 or MPEG-4. It was felt that the evidence is not yet conclusive. Currently, the Requirements Document only mentions Progressive scan. Some members think also Interlaced Video should be covered, but there are no people present that intend to carry out that work, and hence the requirement has been left out of the document.

A High Level for the MPEG-2 4:2:2 Profile

An new Level for 4:2:2 Profile was drafted. This is a working draft that will become a Proposed Draft Amendment (PDAM). Currently, the Profile lists a number of lines and a number of pixels per line, but the suggestion has been made to replace this by just the number of pixels. It is thought that MPEG should allow this kind of flexibility, and MPEG members would like to make use of it, for supporting broader aspect ratios. (a similar approach is being taken in MPEG-4) This has been worded in resolution 1.1.3, as follows:

> *[The Requirements Group recommends] to ask NB to comment on the PDAM N2125 (4:2:2 Profile @ High Level) with respect to the desirability to replace the maximum horizontal and vertical picture size with a maximum number of pixels, given that the maximum number of pixels per second is already specified.*

Input from NB's is asked on the issue of splice parameters as well.
There are also thoughts about a Very High Level.

MPEG-4 Version 1

Profiling - General

3305 *Portuguese National Body (IPQ)* - Portuguese Votes to MPEG-4 CD

Guidelines for profiling were accepted, as they were expressed in the votes from the Portuguese NB and the Dutch NB comments. The basis of these guidelines is that care should be exercised in defining new Object Profiles (OP's), Combination Profiles (CP's) and Levels. A good reason should

exist from an application point of view. One guideline was added to those: to create A, V and S Combination Profiles that can be used together. For example: when Systems has a 2D-only Profile because from an implementation point of view it is sensible to separate 2D and 3D, then in Video there should also exist a Combination Profile that only requires 2D composition. The guidelines can be found in document N2199. An overview of all the Profiles in the different parts of the MPEG-4 standard can be found in N2226. Olaf Barheine of FZI compiled this overview during the work of the Profiles and Levels AHG, were it proved very useful.

Systems

During the joint meeting with Systems, a major step was taken towards a cleaner profiling approach. It was decided to separate Scene Description form Graphics, the latter being a media issue like Audio and Video profiling. Profiles were drafted in both areas. Both the Requirements and Systems Groups were very happy with this decision, which solves an issue that has been bothering us for a long time now.

This gives us the following Profiles for Scene Description (Called *Scene Profiles*):

- Audio-only
- Simple
- 2D
- VRML (includes the previously separate 3D)
- Complete

It is believed that 2D and Simple can probably be merged into one single Scene Profile now that media nodes are no longer part of the Profile.

The following Graphics Profiles have been defined for the moment:

- 2D
- 3D

It is likely that more Graphics Profiles are needed, but they will only be defined following an apparent need.

Level information is not included for either Scene or Graphics Profiles at the moment. There is, however, a list of parameters that could be used to define levels for Scene Profiles.

Audio

The situation in Audio Profiling did not change too much in comparison with the last meeting, and this is a good thing. There are quite a number of Object Profiles, but a limited set of Combination Profiles.

The object Profiles are: (also see N2226, the Overview of Profiles in MPEG-4):

- AAC Main, AAC LC, AAC SSR, T/F, T/F Main scalable, T/F LC scalable, TwinVQ core, CELP, VXC, HILN, TTSI, Main Synthetic, Wavetable Synthesis.

The combination Profiles are:

- Main (4 levels)
- Scalable (4 levels)
- Speech (one default level)
- Low Rate Synthesis (2 levels)

A new method of defining levels was developed for ‘Main’, using the concepts of Processor Complexity Units” (PCU, specified in integer numbers of MOPS) and “RAM Complexity Units” (RCU, RAM usage is given as an integer numbers of kWords (1000 words).

Specifying Structured Audio in these terms still needs some work. In Structured Audio it is possible to send complex instructions using only very few bits.

Visual

Scalability in General

Because there was much discussion about conformance points for scalability, it is good to note that scalability can be done for two different reasons:
1) *for decoder complexity reasons*. A decoder that only understands the base layer is significantly less complex than the decoder that understands all layers. In this case we need a conformance point at base layer and the base layer decoder needs not be able to decode enhancement layer.
2) *for service scalability reasons*: The reason for doing scalability is that you don't always have the bandwidth to send the full quality. In that case you still want a meaningful picture. The same decoder can, however, decode both base and base + enhancement layer(s), depending on the information it receives.
In this case, there is no need for conformance point at base layer; the decoder is required to decode correctly:

a) base layer, and
b) base + enhancement layer(s)

Object Profiles and Tools

The Requirements Group decided to endorse all the changes with respect to the selection of tools for the Natural Video Object Profiles that are proposed by the study on Visual CD. This included (but was not limited to):

- Adding tools for Grey Scale and Interlace to what should have been called ‘Core OP’ in the Visual CD and to rename it to Main’;
- Adding a new OP, below what is called ‘Simple’ in the Visual CD
- To call this new OP ‘Simple’, to rename what is called ‘Simple OP’ in the CD to Core OP’,

Subsequent discussions about tools in Object Profiles started from this situation. An overview of the decisions:

- Following a request from the US National Body received during the 42nd meeting in San Jose, the ISG has performed an evaluation of the complexity of Overlapped Block Motion Compensation (OBMC). In the absence of consensus within Requirements an ISG, the conclusion was that there was no consensus for making this change to the Visual part of the standard. There were some concerns about whether having it only in Main makes sense at all, as Main is expected to be used at bitrates where OBMC does not have much effect. Some experts would like to try it on very difficult sequences at several Megabits/s – maybe it could still be somewhat helpful. The cost was pinned down at 25% (addition in complexity when decoding already ‘difficult’ cases . It was not so simple to quantify benefit.
- It was confirmed that the tools ‘4 Motion Vectors’ and ‘Unrestricted Motion Vectors’ are in all Natural Video Profiles.
- Considering the fact that Core was no longer the simplest Object Profile, B-VOPs were included in Core because it was thought to increase the quality.
- A scalable version of Simple was created, that includes B-VOP scalability.
- Other scalable versions of the Natural Video OP were not created, because spatial scalability does not yet support arbitrary shape.
- New Object Profiles were created in the area of Mesh and Scalable Texture, according to NB requests.

- An arbitrary-shape version of the scalable Wavelet Object was deferred until V.2, when scalable shape should be available.
- A very simple wavelet Object (Simple Scalable Texture) was defined, e.g. for use in audiographics terminals. A liaison was sent to JPEG on this issue.
- An Object Profile was defined that combines Wavelet with Scalable texture and shape; this OP can decode deformable non-video objects.

Chroma key shape.

3454 *B. G. Haskell* - Support for Simple shape in conjunction with simple profile of MPEG-4

M3454 made a request for adding syntactic elements to Systems that would allow Chroma Key to be used for the (de)coding of shape information shape in the Simple Video Object Profile. The question arose because since we added the new Simple Profile, there is an Object Profile that does not have arbitrary shape. Shape coding for such low complexity situations was not considered yet by the Video group (nor had it been requested). The Requirements Group thought that while the Simple OP was created for Mobile applications, and its proponents did not much care about arbitrary shape, Simple OP would also enable other applications, which *do* care about shape. The example of arbitrary shape objects moving on a Web page was given. With respect to the proposal, the Reqs Group felt that:

- if there is were really 'cheap' way of adding Shape to the Simple Video Object, that would be useful.
- Any proposed technique would have to be evaluated against
  - Its quality
  - Its cost in terms of complexity
- It would not be acceptable to leave completely non-normative the way that the shape was extracted at the decoder. It then does not matter so much where the extraction took place. As the Systems Group feels that this is not a Systems issue, the Video Group was asked to further process the proposal, based on the considerations of the Requirements Group.

Some additional observations:

- If it is in the Video syntax, and we want to keep hierarchy, then we have a duplication of tools at the higher OP's. This was not considered a major problem since the functionality is provided at clearly different quality levels.
- If it is in the Systems syntax, then it will be applicable to any 2D Video Object, whatever the Object Profile. On the other hand, the technology would require MPEG-4 Systems to be used.

Combination Profiles

Following requests from several NB's, nine Combination Profiles were defined. Not all requested CP's were defined, because of a number of reasons:

- we tried to keep the amount of different CP's low;
- For some of the CP's there was just not enough support;
- it was decided to postpone definition of scalable OP's to Version 2, and hence the corresponding scalable CP's were not defined.

The issue of having Facial Animation added to Main and Core CP's was not resolved during the meeting. There are two views :

- **View A:** "FA is very simple to implement, so we should add it to Core and Main, thus limiting the number of possible CP's"
- **View B: "**FA is simple to implement because in its easiest case, all you have to do is move around pixels on the screen. Such unpredictable behavior does not well suit the philosophy behind Main

and Core, which is getting good and predictable output."

NB's were asked in a resolution to say something about these views in their votes in the FCD.

Levels

Level definitions were taken from the Study of Visual CD, following a number of NB Votes. Also following these votes, the lowest bitrates for *Simple CP @ Level 1* and *Facial Animation CP @ Level 1* were when compared to the Study.

Also, the Requirements group decided not to specify normatively the Visual Session Size, but rather to use commonly used sizes (QCIF, CIF, BT.601) as indications. The user of the standard is free to choose other screen sizes and aspect ratios, but the total amount of 'object surface' is always defined by the Level definition.

There was agreement on the Levels and the approach, but due to the late stage in the week when levels could start being addressed, not all input contributions could be taken into account.

The integration of the Level information to be included in the Visual FCD still needs to take place in the AHG on Profiles and Levels.

MPEG-4 Version 2

Support for Studio and Broadcast Applications

3358 *Yoichi Yagasaki, Teruhiko Suzuki* – Request for New Profile for Studio & Post-production application
3360 *Yoichi Yagasaki, Teruhiko Suzuki* – Additional Requirements for Profile used in Broadcast application
3363 *Shinichi Sakaida, e a.* – ISDB services as broadcasting applications of MPEG-4
3364 *Shinichi Sakaida, e.a.* – MPEG-4 studio production application and requirements
3500 Ralf Schäfer – MPEG-4 Video Studio Object Profile *(Late submission)*

Broadcast and Studio applications were discussed. Requirements for supporting these have been included in the MPEG-4 Requirements Document (N2194). A remark is made, however, that the requirements have been tentatively included, because enough support is needed for carrying out the work, as this is a completely new area for MPEG-4 with much higher quality than supported so far. Example Applications were included in the Applications Document (N2195).

MPEG-7

A break-out Group on MPEG-7 met separately for the full week, reporting back to the Requirements group on Tuesday afternoon and Thursday afternoon. MPEG-7 Seminar was held on Wednesday.

The main accomplishments are:

- Terminology, requirements and scope of MPEG-7 is further refined and developed; A new diagram illustrating the concepts of MPEG-7 is introduced into the documents
- A Call for Test and Evaluation Material has been issued
- Further details were added to evaluation document regarding evaluation criteria and procedure
- An Advance Call for Proposals has been issued

AHG Reports

3348 I. Sezan, and, F. Nack - Report of AHG on MPEG-7 Requirements
3441 S. Jeannin and A. Lindsay - Report of AHG on MPEG-7 evaluation methods

Document m3348 reports the results of the discussions in the AHG reflector. Main issues that were discussed and agreed upon were (i) the importance of and the need for standardizing a set of descriptors and description schemes; (ii) semantic bootstrapping via DDL; (iii) expectations from the DDL. A major recommendation of the AHG was to improve the evaluation procedures and synchronise the PPD with the evaluation procedures document.

In Document m3441, the AHG discussions on (i) generation of criteria for evaluation; and (ii) improvement of the evaluation procedures document are summarised, with major recommendation being the improvement and harmonisation of the evaluation procedures document with the other documents.

Liaison

3333 F. Preteux - Liaison to CEN on Metadata for Multimedia Information on MPEG-7

This document reported that MPEG-7 activities were presented at the CEN/ISSS-MMI Workshop, held in Brussels, 24-25 February 1998. The CEN group greatly appreciates the MPEG-7 work. They approved the following resolutions:

i) To register the main output documents of CEN/ISSS-MMI in the 44th MPEG Meeting in Dublin, and
ii) To report the progress work and investigation of CEN/ISSS-MMI in the MPEG working group.

Applications

3477 A. Lindsay - MPEG-7 Applications Document
3399 S. Palm et al. - Visually based control: Another application for MPEG-7

In document m3399, a potential application of MPEG-7 to robotics, based on visually based control, is proposed. In vision based control, instead of using text based approaches for control programming, images and image sequences are used to specify the control behavior and are an integral part of the control loop (e.g. visual servoing).

This application is now added to the section of the Applications Document, called Specialised Professional Applications. There were no other changes in the Applications Document.

Requirements

3341 F. Nack and V. V. Vinod - MPEG-7 Requirements Document V.5
3342 F. Nack and V. V. Vinod - MPEG-7: Context and Objectives V.7
3404 J. Kim et al. - Suggestions on the improvements of Draft of MPEG-7 requirements
3400 A. Lindsay et al. - An example description for MPEG-7

Document m3342 was reviewed and approved. This document was revised at the end of the week by incorporating the new diagram illustrating the main concepts of MPEG-7.

The Requirements document m3341 was reviewed and approved. During the review, we checked whether the requirements do reflect our expectations from Descriptors, Description Schemes, and Description Definition Language.

In the revised output Requirements document, we have added “a scheme for coding the description” to the list of What MPEG-7 will standardise, in addition to D, DS and the DDL. We added Fig. 1,

graphically describing the interrelation of the concepts of MPEG-7. This diagram also indicates the normative and non-normative parts of MPEG-7. This figure replaced the previous Fig. 1.

Contribution m3404 proposed addition of a table to the Requirements Document showing feature types, example features and their descriptors. This proposal was viewed favourably and we decided to refine the table during our subsequent reflector discussions for inclusion in the document.

PPD Document

3350 MPEG-7 Requirements AHG - MPEG-7 Draft Proposal Package Description

The PPD document was revised to reflect the changes in the requirements document as well as the Evaluation Procedures Document..

Evaluation Process Document

3442 S. Jeannin - MPEG-7 Evaluation Process Document: Draft Version 2.0
3313 P. Faudemay et al. - Benchmarking methods for audio

The group has heard the presentation of m3313 that discussed benchmarking issues for audio. We have spent most of the time on improving the evaluation process document. The discussions were centred around criteria for evaluation and the evaluation process itself.

The evaluation criteria for D, DS, DDL, and coded descriptions were reviewed, improved and made consistent with the associated requirements. It is still planned that there will be "common criteria" and "application-specific criteria", although we did not have the chance to discuss evaluation criteria and process for different application domains. We were not able to reach consensus on how to define various application domains. It was agreed however that this may be a necessity sooner or later. We decided that we will work on such a classification as we classify the test and evaluation content.

There was a long discussion on how we would evaluate the initial proposals and start building an Experimentation Model. First, there seemed to have support for the idea of submitting proposals along with a search engine, where the proposers would disclose their search engines. Those who did not have a search engine would be provided with one. Then, the evaluation of each submission would be performed using their own search engine. Then the Experimentation model would be built using the best search engines that provided the best results.

The problems with this approach, namely not comparing competing technologies on a common platform (e.g., search engine) and impracticality of providing people with search engines, led us to make a different decision regarding the evaluation process, which can be expressed as follows:

Evaluations will be both objective and subjective and will be carried out by a panel of experts.

- The evaluation group will perform:
    1. Test over common requirements which are common across applications.
    2. Test for other specific requirements associated with selected applications.

  To this aim, the evaluation group will define some application based tests environments (defined from available content and applications of the application document).

- For these tests, proposals should include a write-up and may include search/filtering tools (i.e. provide its executable code plus its paper description) . The proposals are encouraged to

include experimental results over the MPEG-7 test data set. (Experimental conditions to be defined)

- A first level screening of all the proposals will be done based on the submitted experimental results and the associated documentation included with the proposal. Each proposal will be considered according to the evaluation criteria, determining which criteria are satisfied.
- An eXperimental Model (XM) will be built from the screened proposals. This XM is a "common framework" on which relative evaluations and improvements of tools will be done. The XM may contain several tools for the same functionality and will have components for evaluating and improving the DDL, DS and Ds.
- The relative evaluation and improvements will start with a first set of Core Experiments which will be defined.

MPEG-7 Seminar

3429 A. Lindsay et al. - Program of MPEG-7 seminar in Tokyo

The seminar was well attended by some 80 to 100 people, and proved both very interesting as well as extremely useful for the development of the upcoming MPEG-7 standard. The document m3153 contains the program of the seminar, abstracts as well as speakers' bio.

There were a total of 5 presentations. Three presentations addressed the video indexing problem, where human interface issues and professional production applications were also addressed. The remaining presentations were on audio and music. A beat tracking system was presented as an example of musical feature extraction. The other presentation addressed the description of non-speech and non-music audio, namely synthetic audio.

As usual a conclusions and discussion period followed the presentations. The presentations on audio and music was particularly useful for the MPEG-7 participants since most of the participants have video expertise.

Call for Content for Testing and Evaluation

A document was issued calling for content that will be used in evaluation of MPEG-7 proposals as well as developing the MPEG-7 standard. The document calls for content in a number of categories ranging from feature, short and documentary movies to consumer home-movies. A call for associated information, if any, is also included in the document. It is clearly specified in the Call that the material provided will be used only for the purposes of test and evaluation of responses to the MPEG-7 call for proposals and the development of the standard. 5n particular it will be categorised into application domains and used to form test database(s).

The materials will be submitted to Dr. V. V. Vinod or Dr. Remi Ronfard by **12th of June 1998.** We also invite those providing material to participate in the next MPEG meeting at Dublin during July 6-12 when the test material for MPEG-7 would be collected and categorised.

The group will consider the content in AHG mtg. in Dublin (July 3-5). It is also planned that the content will be given to HoDs during October Meeting for distribution.

Meetings with the Video Group and the Test Chairperson

We had a joint meeting with the video group where we have reviewed the scope and the (video) requirements. It is expected that we will hold similar joint meetings in the upcoming MPEG meetings. We also had a lunch meeting with Laura to introduce her to testing and evaluation issues that we will be dealing with.

Decisions on AHG Meetings before the Dublin Meeting

We have decided that we will hold an Test and Evaluation AHG meeting in June 3-5 in Paris (hosted by INA) provided that we have enough participants. The goal of this meeting is to further improve the Test and Evaluation Procedures document and start discussing the classification into different application domains. It was decided that an email poll will be taken by April 10th to determine the attendance to this possible meeting and make a decision.

Another Test and Evaluation AHG meeting was scheduled for 3-5 July, immediately before the Dublin meeting to discuss the responses to Call for Content and classify them into different application domains.

When	What	Where
Monday		
9.00 - ??.??	*opening plenary meeting*	
Until 1 hour after closing of plenary (estimate: 13.30–14.30)	lunch	
14.30–15.30	**Meeting goals Requirements, Approval of agenda, Assignments (incl. IPR & MPEG-7)**	Reqs
15.30–16.00	**Profiling Policy**	Reqs
	3248 *Olaf Barheine* (within AHG) – MPEG-4 Profiles/Levels Summary	√
	3280 *NNI* – Dutch NB Votes and Comments to MPEG-4 CD's	√
	3305 *Portuguese National Body (IPQ)* – Portuguese Votes to MPEG-4 CD	√
	3467 *Kevin O'Connell, Karlheinz Brandenburg* – Report of the MPEG-4 Profiles and Levels Ad-Hoc Group	√
16.00–20.00	**Visual Profiling (together with ISG)**	Reqs
	3252 *SC 29 Secretariat* – Summary of Voting on CD 14496-2 (SC 29 N 2482)	√
	3280 NNI – Dutch NB Votes and Comments to MPEG-4 CD's	√
	3298 *Singapore National Body* – Response of the Singapore NB comments to resolution 2.1.5 of the 42nd WG11 meeting	(√)
	3300 *T.K. Tan* - On Visual Combination Profiles and Levels	√
	3322 *Michael Frater* – Australian National Body Comments on CD 14496-2	√
	3353 *Hiroyuki Katata* – Definition of levels for new visual combination profiles	√
	3454 *B. G. Haskell* – Support for Simple shape in conjunction with simple profile of MPEG-4	√
	3467 *Kevin O'Connell, Karlheinz Brandenburg* – Report of the MPEG-4 Profiles and Levels Ad-Hoc Group	√
	3468 *Kevin O'Connell, Singhal, Brailean* – OBMC Complexity Evaluation Results	√
Tuesday		
8.30–9.00	**ALC Applications and Requirements**	√
9.00–10.00	**Version Management**	
	3384 *Jae-Seob Shin* – Report of the Ad-Hoc group on MPEG-4 Version Management	√
10.00–11.30	**Levels for Synthetic Visual elements in Combination Profiles**	Reqs
11.00–12.00	**t.b.d.**	
12.00–13.00	**Joint with Video about Visual Profiles and Levels**	Video
13.00–14.00	Lunch	
14.00–14.45	**Joint with Video on MPEG-2 4:2:2@HL**	Reqs
	3432 *Ajay Luthra, Yoichi Yagasaki* – Report on MPEG2 4:2:2P@HL Ad-Hoc Group's Work	√
	3434 *SMPTE* - 4:2:2 Profile @ HL – Requirements	√
	3435 *SMPTE* - 4:2:2 Profile @ HL – Table Entries	√
14.45–15.45	**Broadcast and Studio Applications and** Requirements	Reqs
	3358 *Yoichi Yagasaki, Teruhiko Suzuki* – Request for New Profile for Studio & Post-production application	√
	3360 *Yoichi Yagasaki, Teruhiko Suzuki* – Additional Requirements for Profile used in Broadcast application	√
	3363 *Shinichi Sakaida, et. al.* – ISDB services as broadcasting applications of MPEG-4	√
	3364 *Shinichi Sakaida, et. al.* – MPEG-4 studio production application and requirements	√
	3500 Ralf Schäfer – MPEG-4 Video Studio Object Profile *(Late submission)*	√
16.00–17.00	**Joint meeting with Systems about Systems Profiles and Levels**	Systems
	Also: how to signal Object Profiles and Combination Profiles @ Levels in Systems syntax	
	3292 *C. Herpel, M. Reißmann, J. Vollmer* - A First Approach Towards the Definition of Conformance Levels in MPEG-4 Systems	√

	3251 *SC 29 Secretariat* – Summary of Voting on CD 14496-1 (SC 29 N 2481)	√
17.00–18.00	**Joint with Audio (Profiles and Levels)**	Aud
18.00–18.50	**Progress Report from IPR people**	Reqs
18.55–19.30	**Progress report from MPEG-7 break out**	Reqs
19.30–20.30	**Continued discussion from 14.45 (Broadcast + Applications)**	Reqs
Wednesday		
9.00–11.30	**Plenary meeting**	Video
11.30–12.00	**MPEG-7 Seminar**	
12.00–13.00	Lunch	
13.00–17.00	Visual Profiles and Levels continued (Joint with ISG from 15.00) (for contributions see Monday agenda)	
17.00–18.00	**Joint meeting with SNHC (Profiles and Levels)**	Reqs
	Social Event	
Thursday		
8.30–10.00	**Joint with Systems (Profiles, Levels, Normative Composition)** 3282 *Barron, Thomas, Coleman* –AHG on Normative Composition 9.00–11.00 **Joint with Video (MPEG-7)** Intended for MPEG-7 people (Video room)	Reqs
10.00–12.00	**Visual Profiling**	
12.00–13.00	**Joint with Audio (Profiles and Levels)**	Audio
13.00–14.00	Lunch	
14.00–15.00	**Joint with Video (Profiles and Levels)**	Video
15.00–16.00	**Levels**	Reqs
16.00–17.00	**Report back from IPR break out group**	
17.00–18.00	**Report back from MPEG-7 break out group**	
18.00 - ??.??	**Final editing tasks**	
Friday		
9.00– 9.30	**Approving ALC Requirements Document; 4:2:2 @ HL document**	
9.30–10.00	**Approving CMP Documents**	
10.00–11.00	**Approving MPEG-4 documents (inclusing DoC's and FCD deltas)**	
11.00–11.45	**Approving MPEG-7 Documents**	
11.45–12.30	**Remaining AHG's, resolutions**	
12.30–14.00	Lunch	
14.00–22.22	plenary meeting	

Annex 4

Report of Delivery meeting

Source: Vahe Balabanian, Chair

1. Disposition of comments on DMIF V1 CD

All the comments were dealt with successfully. The results are in N2122

2. DMIF V1 FCD

Changes agreed to the CD document. Final text N2206 will be uploaded prior to May 15.

3. End-to-End MPEG-4 QoS Management

End-to-end QoS management is added with a minimum set of Media-based QoS metrics.

4. Reviewed the progress of the DMIF Ref. Source Code

The work progress is on schedule for integration of DMIF with systems according to the plan in N2058

5. DMIF V1 conformance has been updated N2210

6. AHG is recommended on Mobile Network Connections N2124

7. Working with IETF (For Info)

A small group of experts will look at how MPEG-4 can make use of IETF RTSP with intent to create a technical proposal to IETF.

Annex 5

Report of Systems Meeting

Source: Olivier Avaro, Chair

Overview

The MPEG-4 Systems Sub-group delivered in Tokyo two key documents : Systems Version 1 Final Committee Draft and Systems Version 2 Verification Models.

The Systems Version 1 Final Committee Draft is the result of the processing of the numerous comments made by National Bodies on the Committee Draft previously released in Fribourg and already commented on in San Jose. The document will be released on the 15th of May and will provide a substantial improvement of the current specification. In parallel to these efforts, activities for Systems Version 2 took place.

Ongoing Systems Version 2 activities will augment the set of tools delivered in Systems Version 1. Systems has delivered the Version 2 VM and the WD. They consist of enhanced specifications of the following activities :

- In Advanced BIFS, a first set of new BIFS nodes have been identified. They include functionality existing in VRML and not yet supported by MPEG (e.g. : PROTO, scripts, …) and also new MPEG-4 specific nodes (e.g. advanced spatialization).
- The MPEG Committee continued the MPEG-J (previously AAVS) activity to facilitate the integration of features for applications such as set-top box, interactive games and mobile AV terminals in MPEG-4. These features will enable a high level of interaction for both local and remote terminal control. It will also provide platform independence by using Java technology. MPEG-J experts have defined a precise framework within which work will progress on architecture, enhanced specifications of APIs, and implementation environment.
- Following the Call for Proposals (CFP) issued in October and the starting point decided in San Jose, the “Intermedia Format” now called "M4F" is moving fast and the VM has achieved a quality which allows for a first implementation of the specification. MPEG invites experts in content production and delivery to join its experts in completing the development of this specification, scheduled for December 1998.
- Finally, the Intellectual Property Management and Protection group has begun to build technical specifications according to the requirements that have been previously defined within the context of the MPEG-4 Requirements group.

The report of the Plenary for Systems can be find in an attached presentation(sys.ppt).

Detailed Report

Version 1 Specification and Implementation

Scene Description(Liam)

The report of this activity can be found in an attached presentation (bifs.ppt).

Elementary Streams Management (Carsten)

The report of this activity can be found in an attached presentation (esm.ppt).

Implementation framework (Zvi)

The report of this activity can be found in an attached presentation (im1.ppt).

Conformance

Fruitful discussions with the Requirement group lead to clarification of the definition of the profiles in MPEG-4 and particularly Systems. Systems profiles have been refined. An input proposal to Systems gave first elements that will be useful to the definition of Systems levels. This work on complexity evaluation of a BIFS scene will continue in a Systems AHG.
Normative composition activity is taking place. First items have been resolved for version 1 (semantic for BIFS transparency values in linear spaces). More issues (composition and rendering of media objects, audio composition), will continue in Version 2.

Methodology for evolution from FCD to DIS

We have produced in Tokyo the Final Systems CD. This document has an official value and will be submitted for ballot.

As soon as this document will be released 15th of May, the Systems sub-group will :

- Begin to work on a study of FCD from the 15th of May to Dublin in order to provide for clean normative text for the bugs that have been identified.
- Take care that all the comments that we foresee are taken into account by at least one national body, that will report them at the second round of comment.

In Princeton, we will produce the DIS, and no more technical changes (only editorial) will then be possible).

Version 2 Tools and Implementation

Advanced BIFS

The following items have been treated in advanced BIFS :

- The new audio nodes have been enhanced.
- The user interaction nodes have been enhanced.
- The coding of Ecmast scripts has been enhanced.
- The nodes related to multi-user world have been suppressed because of lack of activity in that domain.

MPEG-J (Gerard)

The report of this activity can be found in an attached presentation (mpj.ppt).

MPEG-4 Intermedia Format (David)

The report of this activity can be found in an attached presentation (m4f.ppt)

Intellectual Property Management & Protection (Niels)

The work on Intellectual Property Management and Protection (IPMP) was kicked off by the Requirements Group and the IPMP break-out group met throughout the entire week. During the week several joint meetings were held with Requirements, Systems and Audio.

Since the future work for MPEG-4 IPMP is more a system’s than a requirement’s issue, it was agreed to move the IPMP activities from the Requirements to the System Group.

General Documentation

The overview documents for MPEG-4 Version 1 and Version 2 have been reviewed and the relevant parts of both documents have been updated.

The MPEG-4 Requirements document has been reviewed but no changes have been made since the document is, from the IPMP point of view, up-to-date.

Intellectual Property Management & Protection Framework

The Framework, which has been drafted in the previous meeting in San José, has been revised in Tokyo. During this process, the input documents to the IPMP Break-out (M 3330 and M 3242) were presented/taken into account.
A first version of an educational document, called »The Why and How of the Intellectual Property Protection Framework« (N2198) was drafted.

During the Tokyo meeting it was possible to finalise the high-level view of the IPMP Framework. Technical details of the IPMP framework will be addressed due to the work plan as follows:

until the 43rd meeting	Revise Educational Document Technical Discussion Finalise Technical Specification 2 ad-hoc meetings: May 6-8 in the New York area July 5, Dublin
at the 43rd meeting	Finalise Educational Document Provide Text for Systems Standard for MPEG-4 Version 2 MPEG-4 IPMP Educational Plenary

Additional information on this activity can be found in an attached presentation (ipmp.ppt)

Methodology for evolution in Version 2

The methodology for evolution of the VM is described on the MPEG home page and follows the Core experiment process.
Syntax and semantic specified in the VM will go in the WD only when validated (implemented, tested and exercised).
The first implementation of Version 2 activities are expected in Dublin.

`

Annex 6

Report of the Video Meeting

Source: Thomas Sikora, Chair

The video group addressed in its meeting issues related to MPEG-4 (Version 1 and Version 2) and MPEG-2. The major effort in the group was dedicated towards the release of the Final Committee Draft (FCD) of the Visual part of the MPEG-4 standard.

MPEG-2/4
Advanced Layered Coding

Proposals for advanced layered coding of video at high bit rates were reviewed. Based on a number of discussions and results generated between the San Jose and the Tokyo meeting some of the remaining questions on the techniques used in the proposed schemes were clarified. At the meeting the attempt was to separate pre-processing techniques used in the proposals from encoder issues required to be specified for standardization. A plan to verify the merits of the proposals in comparison to the established MPEG-2 and MPEG-4 scalable coding approaches was issued.
It was decided to discuss the proposals on Advance Layered Coding in the context of MPEG-4 rather than for an ammendment MPEG-2.

High Quality Video at High Resolution

At the meeting the parameters for coding video at MPEG-2 4:2:2@HL were specified. A PDAM document was issued.
MPEG-4 (Version 1)
New Tools

The Korean and German National Bodies requested the support for the following new tools to be provided for MPEG-4 Version 1 (NB Comments)
Global Motion Compensation (GMC)
¼ pel motion compensation (1/4 MC)

- Shape Adaptive DCT (SA-DCT)
- Boundary Block Merging (BBM)

The video group evaluated the maturity of these tools and the possibility to integrate these tools into the official software platforms. Although the video group considered all tools to be mature, the critical issue remains the integration of these tools into the software platforms in due time and the sufficient bitstream testing of the tools. Furthermore, the time between the Tokyo and Dublin MPEG meetings will be required to test the tools already adopted for Version 1.

The decision was taken to adopt the GMC, ¼ MC and BBM methods at the Dublin meeting, if the already specified and integrated tools for Version 1 are fully tested by then. In that case the time between the July Dublin meeting and the October USA meeting could be used to integrate and bitstream verify the new tools. Regarding the SA-DCT algorithm the video group identified that the main application area for this method would be at high

bit rates not yet specified for any profile in Version 1. Since for Version 2 a Studio Profile is under discussion it was felt that this would be an appropriate Profile for the SA-DCT algorithm and that the method should be adopted should such a Profile be established.

Backward Compatibility with H.263

Various National Bodies requested to support backward compatibility of Version 1 video with H.263. The video group identified an appropriate syntax extension to accommodate backward compatibility of MPEG-4 video with baseline H.263. This syntax extension was adopted for Version 1.

Software Integration/Bitstream Exchange

The software intergration and bitstream testing continued as one of the most important activities between the meetings. At the Tokyo meeting the activity was reviewed and a new schedule revised. It was decided to release software based on MOMUSYS and MICROSOFT platforms two weeks after the meeting - a cleaned-up version in app. 4 weeks after the meeting.

Final Committee Draft

The FCD document for the visual standard was issued taken into account the Comments on the CD provided by National Bodies through the formal voting process.

Verification Tests

The planning for the Error Resilience tests progressed well. At the meeting new source material was reviewed.

The video group also progressed well towards the Content-Based Coding tests. Here a number of shortcomings identified at the San Jose meeting were clarified and most of the problems were solved. A new activity was started with the aim to resolve remaining problems and to prepare the Error Resilience Tests.

MPEG-4 (Version 2)

The discussions related to MPEG-4 Version 2 covered the following aspects:

Coding Efficiency
Shape Coding
Multifunctional Coding
Error Resilience
Encoder Optimization
Random Access

In general Core Experiments were reviewed and the VM document was updated. No new WD

document was issued since no time at the meeting could be allocated to review the Version 2 tools and appropriate text in detail. At the July MPEG meeting the tools under consideration for Version 2 will be reviewed again.

Annex 7

Report of the Audio meeting

Source: P. Schreiner, Chairman Audio Subgroup
D. Meares, Secretary Audio Subgroup

Opening of the meeting

The MPEG/Audio Subgroup meeting was held during the 43rd meeting of WG11 in Tokyo, Japan on 16 to 20 March 1998. The list of participants is given in Annex A-1. The Chairman welcomed the delegates to the meeting and outlined the work for the five days

Administrative matters

Approval of agenda

The agenda as presented in

Annex A-II was discussed and approved.

San Jose meeting report

The Audio Subgroup portion of the San Jose meeting report, February 1998, had been previously distributed by email. Comments had been returned demonstrating the need for clarification of what was written. These changes were made and entered into the record of that meeting. The resultant report was therefore approved.

Allocation of contributions

All contributions were listed (see Annex A-VI) and allocated to the agenda. All contributions directly related to the Subgroup were presented in task group discussions, or in Audio plenary. Several relevant documents from Test, Systems and Requirements were brought to the attention of the group.

Communications from the Chair

The Chairman summarised the detailed allocations and questions raised at the Chairman's meeting held on the evening before the main meeting started. The majority of these, by design, were already in the agenda.

- NB comments on MPEG-4 reviewed and distributed to the Subgroups.
- MPEG-4 conformance editors need to be identified by the Wednesday Plenary.
- Reference software needs to be reviewed
- Definition of the syntax of elementary streams for Audio back channel applications needed
- Complexity estimates of Structured Audio part needed with ISG coordination
- Joint meetings identified

Joint meetings

Joint meetings were scheduled with Test, Systems, Requirements, and ISG.

Report of ad hoc group activities

All of the ad-hoc group reports had been presented in the opening MPEG Plenary. Additionally, some extra details were presented to the Audio Subgroup and their contents were taken into account at the appropriate point of the meeting.

Received National Body Comments

The National Body comments were reviewed in the context of specific agenda items as reported below. Responses were drafted.

Liaison matters

Mr. Schreiner discussed the content of the ITU-T liaison, M3254. Essentially, they are encouraging MPEG-4 to use the existing standards rather than create new ones which are no better. The matter was discussed and a response was prepared.
MMA are requesting that WG11 adopt their current and proposed concepts. A response was prepared indicating the value to WG11 of the MMA integration work and the continuing effort to further the joint work of the MMA members and the MPEG experts in developing the Structured Audio part of MPEG-4.

Temporary task group formation

To accomplish the large number of tasks to be performed by the Audio Subgroup, 13 task groups were formed as indicated in Annex A-V. The results of each of the task groups were presented to and discussed by the entire Audio Subgroup, including iterations as necessary. The conclusions of

the task groups are presented elsewhere in this report and are included in the output documents.

MPEG-2

IS 13818-7 AAC

Conformance 13818-4 /DAM 1 (Jul 98) and Technical Report 13818-5 /DAM 1 (Jul 98)

Mr. Paley's task group continued the work from the last meeting on Conformance and Technical Report. There was some confusion over whether or not one can make changes now as the matter is out for ballot at present. However, as the work is completing what was requested by NBs then we need to do it.
Mr. Fukuchi presented M3271 on AAC SSR conformance testing. This shows that 10 SSR bitstreams were delivered by Sony and decoded by Nippon Steel at each of 12 different sampling rates. In each case, errors less than 1LSB were achieved.
Mr. Paley co-ordinated the review of the Conformance and Technical Report including input by email from Mr. Lueck. Parts were identified as missing , namely coupling channel and lfe in encoder, and volunteers to supply the code were identified (N2152). Mr. Herre prepared a "how to do an LFE and Coupling channel" guide to support the completion of the MPEG-2 AAC TR encoder software. (FhG has volunteered to provide the LFE part of the work.)

The record of the updates is given in documents WG11/N2151 and WG11/N2152.

IS 13818-7 Corrigendum 1 (Jul 98)

Dynamic Range Control

Mr. Schreiner reported on the work of the ad-hoc group on DRC and the results of the ad-hoc meeting, m3265. That meeting confirmed the need for DRC, decided that full dynamic audio be coded within AAC, recommended that Reference Level and Gain Control data be added to the bitstream using the data fill_element feature and further recommended various features relating to the generation of the DRC meta data. This was discussed in the Audio Subgroup and the conclusions are recorded in document WG11/N2155.

Preparation of output review

Mr. Herre's task group reviewed all the comments that had been received on the Technical Corrigendum including the DRC proposals. Ms. Bosi sent in an e-mail indicating that ISO had recommended making changes to the place of recording the Normative references. The study on the Technical Corrigendum is given in document WG11/N2153 whilst the DoC is given in document WG11/N2154.

MPEG-4

Technical issues

Compression techniques

Mr. Nomura presented M3357 and M3502 showing further test results for the NB-CELP coder and the supporting changes that were made to the Reference Software NB-CELP coder. The results from NEC show no statistical improvement on any items over G.723.1 and worse results on a few items. The various newly proposed changes offer 5% reduction in bitrate and marginally lower delay for no significant change in performance. NEC proposed adding new modes based on the indications of these tests.

Mr. Nomura presented M3422 on WB-CELP improvements. Improvements are there (~0.2[1] of a grade overall, due improved handling of music signals) but they come at the expense of complexity and additional look-up tables. The question was asked, "Are the additional changes really worthwhile?"

Mr. Oomen reviewed the papers from Philips on informal tests on NB-CELP in relationship to G.723.1 both at 8 kHz sampling and about 6 kb/s each. The test details are given in M3304. They used non-MPEG-4 speech items in English, German, Japanese, French and Dutch languages, some with added noise. With noise, there was little difference between coders, but without noise, G.723.1 was preferred by nearly one grade. Additional comments from Philips members indicated acceptable quality for the proposed application most of the time, with some annoying noise artefacts on some items.

Mr. Oomen also presented results, only just available, after NEC had implemented some changes to their NB-CELP coder. These showed similar trends to the results of M3304, with some items showing improvements and some showing loss. These results were added to the input document.

Mr. Oomen made the observation that the Netherlands NB had tabled a paper encouraging MPEG not to expend continuing effort on changes for their own sake but should concentrate on ensuring that what is already there really works. Mr. Fielder asked why there was a mindset for the CELP coder work against trying to stabilise the standard at this stage. The meeting had a general and extensive discussion on the large number of matrix combinations, such as the various modes of CELP, that have never been subject to any form of proof testing. The fear is that there may be combinations of functionalities that are intrinsically incompatible with one another. The need was to try to identify modes which were wanted and had been tested.

Ultimately, it was decided that both NB- and WB- CELP will remain unchanged from what is currently in the FCD unless dramatic improvements in performance can be demonstrated. There are, however, no specific developments or core-experiments planned. The considered position is that we have a good scaleable coder, in CELP, which gives almost as good a result as a non-scaleable non-MPEG coder (G.723.1).

At the request of NEC members, the debate was revisited in the presence of the Convenor. Mr. Schreiner summarised the position that both NB- and WB- CELP modifications had only demonstrated marginal improvement for non-broken functionalities, so why the dissatisfaction in the decision. The Convenor had been led to believe that some combinations of encoding tools are being excluded - if so why? Mr. Grill reiterated that we have pre-screening results for NB- that showed insufficient improvement for European languages. So we should wait until the improvement is worthwhile. For WB-CELP, as well, improvements are marginal only. No exclusion of existing functionalities had taken place. The Convenor wishes to ensure that we do not constrain the choice of tools in the encoder. However, we should not keep changing the normative components of the decoder. As the NEC proposal changes the normative look-up tables, it changes the decoder tool. Thus the previous decision was endorsed.

The issue of postfiltering needed for CELP was causing concern because if it was now 'Informative' then the performance of MPEG-4 without it was unknown. This would also cause difficulties when it comes to Conformance, unless conformance tests are performed in advance of the postfilter. The consensus was in favour of an 'Informative' postfilter being specified with the conformance point before it. It was observed that this also facilitates the applicability of CELP coding to scaleable coding. For completeness, it was noted that the postfilter is comprised of a) pitch filter and b) spectral filter.

If further developments of CELP coding do take place as a result of ad-hoc activities, Philips has volunteered to be an independent test site for any further CELP coder development tests. An ad-hoc group was set up to resolve the possibilities for improvements to CELP. The procedure for

[1] quarter of 'slightly better' according to proponents own test results.

evaluating the value of any change will be based on the core experiment methodology and will be by Subgroup consensus at the Dublin meeting.

TTSI

On the topic of text-to-speech-interface, there was a proposal from the US to add Mark-up Language to the TTSI. This is noted as being a new proposal at a point in time when everything should now be finalised and proven. The Mark-up Language is not in this category and is now far too late for Version 1. The group unanimously noted that this is new work. The proposal will be considered under the constraints of Version 2, i.e. WD 14496-3 Amd/1.
Mr. Ostermann suggested that TTSI and facial animation need to be drawn together. He wishes to use the phoneme data to drive the facial animation algorithm. There is a total of 68 facial animation parameters plus proprietary face models invoked in the FBA activity. Mr Ostermann said there was a need to add visemes to the TTSI to drive the facial expressions and head movements better than at present. This will give, specifically, better animation timing. Mr Ostermann gave a presentation of the proposal and the syntax changes to the text stream. The requirement brings with it the need to change the API. There is a need to ensure that this change does not break the existing TTSI. The verification will have to take place before the next meeting. The mandate document WG11/N2141, of the ad-hoc group was written accordingly.

Structured Audio

Mr. Ray chaired a task group reviewing the SA issues, particularly the NB comments. One interesting proposal was to use AAC to convey the SA samples. This is being considered. The US paper supporting the MMA proposal as part of the SA structure was noted. The intent was to achieve maximum inter-operability between MPEG and non-MPEG SA authoring tools. Member of the Japanese MMA group, AMEI, assisted in this matter and had submitted an input document M3295.
The task group prepared the DoC for SA issues. They additionally studied the issue of complexity and worked with ISG on the issue.

Complexity

The issue of complexity was reviewed in a joint meeting with ISG. Input document M3407 presents the current estimates of complexity in terms of RAM, ROM, MIPS, MOPS etc. for each component of the tools. It was observed that in Video the complexity was not bounded without constraints on number of objects and tools and manipulations. With Audio it was more bounded by just selecting tools and number of objects. Mr. Mattevelli recommended an approach to complexity estimation based on using a specific platform and attempt to run each of the tools on that. Mr. Schreiner indicated the difficulty with this overly conservative approach was that the algorithms would ultimately be performed with application specific silicon implementations. Thus, use of a common platform would not be indicative of the final performance of the tools. However, the common platform approach would give a feel for the problem.
Mr. Richard presented document M3293 giving the complexity evaluation of an optimised HILN parametric speech decoder. This gives values for a Reference Software decoder against an optimised decoder in terms of number of instructions, MIPS and execution times. Much of the savings is shown to be reduced complexity of the Individual Line Synthesis.
An ad-hoc group will continue filling in the missing complexity details for the whole of the MPEG-4 toolset.

Verification demonstrator

Mr. Coleman reported the work of his ad-hoc group, M3340. Much of the work had been done but, in doing it, real problems with the current software had been encountered. These difficulties had

been resolved and resultant changes to the software had been communicated to the Reference Software editors. Work on the demonstrator will continue up to the Dublin meeting.

Systems issues

Mr. Coleman worked with Systems during the week to help resolve some of the problems with current strategies. He reported that the composition buffer definition is 'broken' as far as audio applications are concerned. The framing task of the buffers is not well defined: the last information is overwritten or held until replaced. Instead it should be held until its timestamp matures - it should not be randomly displaced. Particularly with SA, the voices need to be held until the compositor has finished with them, not just until a new voice is received. This is a matter of concern which Systems seem to be unable to grasp.

Mr Coleman presented document WG11/N2160 summarising the Audio/Systems issues that had arisen in discussion during the week. The main detail related to audio composition. The document was approved.

Audio FCD (Mar 98)

Mr. Grill and his task group reviewed the all comments on the CD. Most were straight forward and had in the main been reviewed previously. The detailed set of responses was presented to the Audio Subgroup. Mr. Vannanen asked that his objection to the removal of details relating to low delay filter bank and low complexity backwards adaptive predictor, as per the decision at San Jose, be recorded. The issue of multichannel functionalities in MPEG-4 also caused some concern. The initial proposal of mapping 13818-7 software into MPEG-4 and treating them as one set of multichannel software was rejected as subsequent changes to the software would be likely to cause conflict with 13818-7 devices. The alternative of importing 13818-7 into MPEG-4 but treating this separately from the 13818-7 stabilised software was accepted. The issue of MMA proposed changes to SA was also discussed . It was felt that the proposed changes could not be adequately considered during this meeting but would be reviewed and prepared for integration as scheduled in an ad-hoc activity until the Dublin meeting.

The interim release of the text of the FCD is given in document WG11/N2247 and the approved DoC is in document WG11/N2120.

Conformance Testing WD

Mr. Lee presented his paper, M3297, on the conformance testing of TTSI. The paper notes that there are a number of specific interfaces all of which need to be tested. Mr Lee proposes 6 bitstreams to evaluate different groupings of functionality. This proposal was accepted and is presented in document WG11/N2140.

Due to lack of time and effort, these matters could not be amalgamated into revision of the Conformance document during the meeting. However, this will be dealt with in an ad-hoc group.

Reference Software FCD (Mar 98)

Mr. Grill reviewed the work of the ad-hoc group in M3437. Concern had been expressed in that group over the nominal 'Normative' status of the Reference Software. Does this imply that 'only' the version of the software given should be used and nothing else can be claimed to be compliant? The ad-hoc group had reviewed the NB comments and made a preliminary draft of the necessary DoC that will be required from this meeting. Additionally, there were still some open issues which were carried through to the Task Group during this meeting. For instance, it is proposed that the CELP post-filter be moved to the informative part of the standard, as discussed in section 0. Also, there are multiple proposals for the SA components of the software.

Mr. Paley co-ordinated the review of the Reference Software including input by email from Mr. Lueck. Parts were identified as still missing from the reference software, e.g. coupling channel and

lfe in encoder. Mr. Herre started preparing the missing parts, but he would not be able to complete them during this meeting. An end date for the provision of data for the FCD was set at 8th May. Mr. Grill and Mr. Coleman reported on the DoC preparation. The FCD is given in document WG11/N2205 and the DoC is given in document WG11/N2121.

Requirements

Profiles

Mr. Brandenburg chaired a task group on the review of profiles and levels. A proposal was made for Main Audio Combination Profile to concentrate on a concept where what was specified was so many "decoding units," so much computational power and so much RAM etc. This method of addressing the complexity of a level defined for a profile would bound the complexity for the application of the natural audio elements. The issue of SA complexity bounding is still under consideration. A summary of the currently proposed MPEG-4 Audio profiles and levels with complexity bounds is presented in the Requirements output document WG11/N2226.

Testing

Verification tests

NADIB tests

In a joint meeting with Test, the issues relating to the NADIB tests were reviewed. The use of the ITU-T G.723.1 codec as an anchor had been approved by ITU-T and thus will be used. It was noted that Berkom had unfortunately had to withdraw from these tests and that NEC would not be able to use their version of G.723.1 software for these tests. Alternative volunteers were identified. The schedule of the tests was revised to allow more time for the testing phase.
Over 140 items had been offered for the NADIB tests. Mr. Deitz explained the way in which these had been pre-selected down to a more manageable number: this is given in M3431. Details relating to sampling rate, language in its own right and relative to the natural language at the test site were taken into account.
The range of codecs had been extended to include Twin VQ to ensure completeness of the tests.
The issue of different bandwidths of the codecs relative to the reference was discussed as this will give rise to audible differences - this will be taken into account during the interpretation phase. Care had been taken to eliminate biases from the methodology, e.g. four different randomisations.
The details were reviewed in the Audio Subgroup and are presented in document WG11/N2157.

Speech codec tests

Mr. Kim's task group also reviewed the need for and proposed details of the speech codec tests. The proposals were discussed within the Audio Subgroup. A total of 18 codecs and 8 MNRUs were to be evaluated covering two sampling rates and bitrates from 2 to 12 kb/s. The group discussed the choice of test material: old MPEG-4 material or new NADIB test material. The new material and material with noise was felt to be more useful. Problems were identified due to the very large number of combinations of test item and codecs which would lead to long duration tests. The Subgroup decided to deal with the matter in the ad-hoc and also to allow proponents to be test sites as long as there were multiple sites. The details of the tests are given in document WG11/N2158.

Music on Internet tests

Mr. Dietz proposed specific tests for music on Internet based on the proposals from San Jose. These were reviewed in task group. The preliminary description of the tests is recorded in document WG11/N2159: much of the detail will have to be supplied in the ad-hoc group leading to Dublin.

Version 2 matters

Error resilience

Mr. Dietz reviewed the work of the ad-hoc group. Several input documents had been submitted to the ad-hoc group meeting as follows

#Doc.	Submitter	Title
m3359	Sony	Subjective test results of codeword reordered HVXC
m3379	NTT	Reports on the pre-screening tests of the revised UEP tools
m3397	FhG	VM integration of HCR
m3398	FhG	Evaluation of the common EP tool
m3430	Ad-hoc chairs	Ad-hoc Report
	Sony	UEP Implementation for HVXC
	DoCoMo	Description of the revised EP tool
	DoCoMo	Using the revised EP tool for TWIN-VQ

The ad-hoc group had concluded that the EP tool was very useful and should be integrated into Version 2 by the time of the Dublin meeting.
Huffman code reordering and its progress to integration was reviewed. In this process, a new bitstream element had been found to be needed (length_of_spectral_data) and will be integrated into Version 1.
Mr. Sperschneider reviewed the error resilience workplan update, outlining the modifications from the San Jose version. This was approved by the subgroup as document WG11/N2161. The WD for MPEG-4 Audio proposed amendment was also updated and approved, document WG11/N2213.

IPR and content protection

Watermarking

Mr. N. Johnson presented to the Audio Subgroup the concepts of watermaking as applied to Audio. This is the proposed means of identification of content ownership. The watermaking system is outside the MPEG-4 scenario, but it needs to be sufficiently robust to travel through the MPEG-4 compression. The principles of the ICE watermarking is that a small code word is buried regularly in the audio itself. Inserted data rate is about 20 b/s. If the audio is incapable of hiding the code at any point in time it is not inserted. The end product is an inaudible watermark according to informal tests. The code is inserted at a relatively high level and thus it survives most of the time in 6:1 compression, e.g. Atrac, MPEG-2 Layer II systems. Mr. Johnson agreed to support trials of robustness of the CRL watermarking system if someone could arrange to encode and decode watermarked audio passages.

Intellectual Property Management and Protection (IPMP)

Mr. Niels Rump provided a marketing presentation of a proposal for MPEG-4 version 2 copyright protection management. Concern was expressed that this proposal was not backward compatible with version 1 in that non-protected version 1 content was prohibited from distribution without independent authorization in the version 2 copy protection scheme. This would appear to be in direct contradiction to the requirements for version 2 work items.
The potential increase to the required overhead bitrate for the proposed copyright management scheme was an additional cause for concern for many of the MPEG-4 audio tools that are intended to provide very low bitrate coding as a primary attribute.

Other developments

Ms. Vaananen presented her papers M3440 and M3501 on aural spatialisation and the use of SA BIFS. The nodes were noted to be AcousticScene, AcousticMaterial, DirectiveSound. The first two describe the interaction of the environment whilst the third describes the source. The Group felt that this was a good proposal and much needed.

MPEG-7 Audio (15938)

Mr. Lindsay gave the Audio Subgroup a rundown on MPEG-7 work item and its progress to date, based on M3341. Main units are descriptors and description language. Terms have been identified to allow the objects to be defined. The intention is to enable searches on such things as object features, production features, composition. For Audio, these things might be timbre, frequency profile, spatial structure etc. Mr. Wyse also reported on ideas for text descriptions for audio. Descriptors could be based on source, sound type as well as source independent descriptors. But ultimately, MPEG-7 will have to be based on a mix of textural descriptors, sounds themselves (or sound models) etc. In discussion, it was pointed out that enquiries such as "Find me something that sounds like this piece" or "something with this combination of sound components" would arise and should be supported.

Mr. Lindsay also asked what file formats should be used for MPEG-7? The Subgroup recommended that all formats currently supported by MPEG and linear pcm should be supported.

Promotion of MPEG Audio

FAQ

Mr. Thom chaired a task group on the FAQ and generated document WG11/N2162 as a publicly available document.

Audio Web site and content

Mr. Thom also chaired a task group on the web site content. Various additions and links were approved. Mr. Meares reported that the EBU had given permission to place some of the SQAM copyright cleared items on the web site, for which they were thanked. The tasks of the ad-hoc group through to Dublin were identified. Document WG11/N2163 covers these matters.

Discussion of unallocated contributions

There were no unallocated contributions.

Meeting deliverables

Press statement

Mr. Meares prepared the Audio part of the press statement which was approved.

Dispositions of Comments

The DoC matters are referred to above.

Responses to NB comments

Responses to the National Body comments from Netherlands and Japan were prepared by Mr. Oomen and were approved.

Liaison statements

A liaison statement to ITU-T on speech codec developments was prepared and approved. A liaison response to the MMA on the collaborative work on MPEG-4 SA was prepared and approved.

Recommendations for final plenary

A list of recommendations was prepared for approval at the final MPEG plenary meeting. One

document was approved for public release, see Annex A-VI.

Establishment of new Ad-hoc Groups

The following ad-hoc groups were established:

	Mandate	Meeting
Ad-hoc Group on MPEG-2 Audio AAC Conformance and Technical Report (Coleman/Thom/Paley)	N2131	
Ad-hoc group on MPEG-2 AAC dynamic range control (Schreiner)	N2132	9:00-15:00 Sun
Ad-Hoc Group on MPEG-4 Audio FCD and Reference Software FCD progression (Grill/Purnhagen)	N2133	Sun p.m.
Ad-hoc group on MPEG-4 Audio Conformance (Spille)	N2134	
Ad-Hoc Group on MPEG-4 Structured Audio (Scheirer/Ray)	N2135	
Ad-Hoc Group on MPEG-4 Audio Verification tests (Edler/S-W Kim)	N2136	Sat p.m.
Ad-hoc group on MPEG-4 Audio narrowband audio broadcasting verification tests (Contin/Dietz)	N2137	Sun a.m.
Ad-hoc group on MPEG-4 Audio tool complexity (Spille)	N2138	
Ad-hoc group on MPEG-4 Audio/Systems issues (Teichmann/Herpel)	N2139	Sun p.m.
Ad-hoc group on MPEG-4 TTSI/FBA convergence (Lee/Ostermann)	N2141	
Ad-hoc group on MPEG-4 Audio error resilience for Version 2 (Dietz/Miki)	N2142	Sun p.m.
Ad-Hoc group on MPEG Audio web site (Thom/Purnhagen)	N2143	
Ad-hoc group on MPEG-4 Audio CELP optimisation (Sugiyama)	N2144	

Approval of output documents

All output documents were presented to Audio plenary and were approved.

Future activities

Schedule of future meetings

The dates of the next MPEG meeting in Dublin were confirmed. Dates for the ad-hoc group meetings were decided. The ad-hoc group meetings will all occur on the Saturday or Sunday immediately preceding the Dublin MPEG meeting.

Agenda for next meeting

The agenda for the MPEG Audio Subgroup meeting in July 1998 in Dublin, Eire was discussed and approved (Annex III).

A.O.B.

There was no A.O.B.

Closing of the meeting

Mr. Schreiner thanked the participants for all their hard work in preparation for and during this meeting. With that, he declared the Audio Subgroup meeting closed and wished members a safe return journey.

Annex A-II: Agenda for Tokyo Audio Subgroup Meeting

1.	Opening of the meeting	
2.	Administrative matters	
	2.1. Approval of agenda	
	2.2. San Jose meeting report	
	2.3. Allocation of contributions	
	2.4. Communications from the Chair	
	2.4.1. Joint meetings	
	2.5. Report of ad hoc group activities	*3265, 3272, 3303, 3339, 3340, 3346, 3348, 3405, 3407, 3430, 3437, 3439, 3441, 3447, 3467, 3483,*
	2.6. Received National Body Comments and Liaison matters	3254, *3280, 3302, 3305, 3310, 3362, 3393, 3394, 3395, 3396,*
	2.7. Temporary task group formation	
3.	MPEG-2	
	3.1. IS 13818-3 BC	
	3.1.1. IS 13818-5/DAM 1 (Mar 98)	
	3.2. IS 13818-7 AAC	
	3.2.1. Conformance 13818-4 /DAM 1 (Mar 98)	3271, 3339,
	3.2.2. Technical report 13818-5 /DAM 1 (Mar 98)	3339,
	3.2.3. Systems IS 13818-1 /Amd 5 (Apr 98) - Systems-related table entries for AAC	
	3.2.4. 13818-7 Corrigendum (Jul 98)	3244, 3284, 3396,
	3.2.4.1. Dynamic Range Control	3265, 3316,
4.	MPEG-4	
	4.1. Audio IS 14496-3 FCD (Mar 98)	3253, 3254, 3280, 3294, 3302, 3305, 3310, 3362, 3378, 3393, 3394, 3395,
	4.2. Conformance Testing IS 14496-4 WD (CD Dec. 98)	3297, 3405, 3437,
	4.3. Reference Software IS 14496-5 FCD (Mar 98)	3255, 3437,
	4.4. Requirements	
	4.4.1. Profiles	(3248), 3387, 3467,
	4.5. Overview	
	4.6. Testing	
	4.6.1. Core experiments	3303, 3304, 3357, 3422, 3502
	4.6.2. Verification tests	3346,
	4.6.2.1. NADIB tests	3431,
	4.7. Technical issues	
	4.7.1. Complexity	3293, 3407,
	4.7.2. Other matters	3306, 3383,
	4.8. Systems issues	(3251), 3439,
	4.8.1. Verification demonstrator	3340, 3447,
	4.8.2. Systems audio transport	
	4.8.3. Scene description: composition	3295, 3321,
	4.8.4. IPR and content protection	3272, 3330,

	4.8.4.1. Watermarking	
	4.9. Version 2 matters	3501,
	4.9.1. Error resilience	3379, 3430,
	4.9.2. Other developments	3440, 3501
5.	MPEG-7 Audio (IS 15938)	(3313), (3341), (3342), (3348), (3441), (3442), (3477),
6.	Promotion of MPEG Audio	
	6.1. FAQ	
	6.2. Audio Web site and content	3483,
7.	Discussion of unallocated Contributions	
8.	Meeting deliverables	
	8.1. Press statement	
	8.2. Dispositions of Comments	
	8.3. Responses to NB comments	
	8.4. Liaison statements	
	8.5. Recommendations for final plenary	
	8.6. Establishment of new Ad-hoc Groups	
	8.7. Approval of output documents	
9.	Future activities	
	9.1. Schedule of future meetings	
	9.2. Agenda for next meeting	
10.	A.O.B.	
11.	Closing of the meeting	

1

Annex A-III: Agenda for the Dublin Audio Subgroup Meeting

1. Opening of the meeting
2. Administrative matters
   - 2.1. Approval of agenda
   - 2.2. Tokyo meeting report
   - 2.3. Allocation of contributions
   - 2.4. Communications from the Chair
     - 2.4.1. Joint meetings
   - 2.5. Report of ad hoc group activities
   - 2.6. Received National Body Comments and Liaison matters
   - 2.7. Temporary task group formation
3. MPEG-2
   - 3.1. IS 13818-3 BC
     - 3.1.1. IS 13818-5/Amd 1 (May 98)
   - 3.2. IS 13818-7 AAC
     - 3.2.1. Conformance 13818-4 /Amd 1 (May 98)
     - 3.2.2. Technical report 13818-5 /Amd 1 (May 98)
     - 3.2.3. Systems 13818-1 /Amd 5 (Apr 98)
     - 3.2.4. 13818-7 Corrigendum (Jul 98)
4. MPEG-4
   - 4.1. Technical issues
     - 4.1.1. Compression techniques
     - 4.1.2. TTSI
     - 4.1.3. Structured audio
     - 4.1.4. Complexity
     - 4.1.5. Verification demonstrator
     - 4.1.6. Systems issues
     - 4.1.7. Other matters
   - 4.2. Audio 14496-3 FDIS (Oct 98)
   - 4.3. Conformance Testing 14496-4 WD (CD Dec. 98)
   - 4.4. Reference Software 14496-5 FDIS (Oct 98)
   - 4.5. Requirements
     - 4.5.1. Profiles & levels
   - 4.6. Testing
     - 4.6.1. Verification tests
       - 4.6.1.1. NADIB tests
       - 4.6.1.2. speech codec tests
       - 4.6.1.3. Internet radio tests
   - 4.7. Version 2 matters
     - 4.7.1. Audio 14496-3 /Amd 1 WD (Oct 98)
     - 4.7.2. Reference Software 14496-5 /Amd 1 WD (Oct 98)
     - 4.7.3. IPR and content protection
       - 4.7.3.1. Watermarking
     - 4.7.4. Error resilience
     - 4.7.5. Low delay
     - 4.7.6. Environmental spatialisation
     - 4.7.7. Other developments

5. MPEG-7 Audio (15938 CfP Oct 98))
6. Promotion of MPEG Audio
   - 6.1. FAQ
   - 6.2. Audio Web site and content
7. Discussion of unallocated Contributions
8. Meeting deliverables
   - 8.1. Press statement
   - 8.2. Dispositions of Comments
   - 8.3. Responses to NB comments
   - 8.4. Liaison statements
   - 8.5. Recommendations for final plenary
   - 8.6. Establishment of new Ad-hoc Groups
   - 8.7. Approval of output documents
9. Future activities
   - 9.1. Schedule of future meetings
   - 9.2. Agenda for next meeting
10. A.O.B.
11. Closing of the meeting

2

Annex A-V: Audio Task Groups

1. MPEG Audio FAQ/Web Page - Thom
2. 13818-7 AAC Corrigendum - Herre
3. MPEG Audio - Preparation of press statement - Meares
4. MPEG-4 Audio/Systems Issues - Dietz
5. MPEG-2 AAC Conformance - Paley/Coleman
6. MPEG-2 AAC Technical Report - Coleman
7. MPEG-4 Verification Tests - Kim/Contin/Dietz
8. MPEG-4 CD editing - Edler, Grill, Lee, Nishiguchi, Ray, Vaananen
9. MPEG-4 Reference Software - Edler, Grill, Lee, Nishiguchi, Ray, Vaananen
10. MPEG-4 Error resilience - Dietz
11. MPEG-4 Profiles and levels - Brandenburg
12. MPEG-4 CELP Experiments - Oomen
13. MPEG-4 TTSI/FBA interaction - Lee

Annex A-VI: Input/Output Documentation

Contributed documents

The following documents were contributed to the Audio Subgroup and were considered during this meeting:

Number	Source	Title
3239	Pete Schirling	Document Register for 43rd Meeting in Tokyo, Japan
3244	SC 29 Secretariat	Revised Summary of Voting on ISO/IEC 13818-7/DCOR 1(SC 29 N 2467)
3248	Olaf Barheine (within AHG MPEG-4 Profiles and Levels)	MPEG-4 Profiles/Levels Summary
3253	SC 29 Secretariat	Summary of Voting on CD 14496-3 (SC 29 N 2483)
3255	SC 29 Secretariat	Summary of Voting on CD 14496-5 (SC 29 N 2485)
3256	SC 29 Secretariat	Summary of Voting on CD 14496-6 (SC 29 N 2486)
3265	Peter G. Schreiner III	Report of the Ad Hoc Group on MPEG-2 AAC Dynamic Range Control
3271	Hiroyuki Fukuchi	Brief report of AAC conformance testing (SSR profile)
3272	Niels Rump (as chair of ad-hoc group)	Report of IPR Ad-hoc Group
3280	NNI (Nederlands Normalisatie Instituut)	Dutch NB Votes and Comments to MPEG-4 CD's
3283	SC 29 Secretariat	Late vote on CD 13818-10 “SC 29 N 2473”
3284	SC 29 Secretariat	Late vote and comments on 13818-7/DCOR 1 “SC 29 N 2474”
3293	Gael RICHARD, Caroline VENOT	Report of practical complexity evaluation of an optimised HILN decoder
3294	Y.- B. Thomas Kim, S.-H. Park, S.-W. Kim	Improved BSAC description in ISO/IEC CD 14496-3 Subpart 4
3295	Seiji Kameyama	Comments for Structured Audio
3297	Young-Kwon Lim,, Youngjik Lee,, Jungchul Lee	Consideration for Conformance of the MPEG-4 Audio TTSI
3302	Dutch National Body	MPEG-4 Audio
3303	Ralf Funken, Werner Oomen, Frans de Bont	AHG report on Quality assessment of the 8kHz sampling mode of the MPEG-4 Audio VM w.r.t. G723.1
3304	Ralf Funken, Werner Oomen, Frans de Bont	Results of an informal asessment of the Quality of the 8kHz sampling mode of the MPEG-4 Audio VM w.r.t. G723.1
3305	Portuguese National Body (IPQ)	Portuguese Votes to MPEG-4 CD
3306	R. Taori, R. Funken, W. Oomen	Proposal for an optimized VQ for the MPEG-4 CELP
3310	AFNOR French National Body	Letter ballot and comments on CD14496-1/3
3313	Pascal Faudemay, Marie-Jo Caraty, Claude Montacie	Benchmarking methods for audio
3316	Schuyler Quackenbush, Jim Johnston	Proposal for AAC Dynamic Range Control
3321	Tom White (MMA)	MMA Proposed Enhancements to SASBF
3330	Itaru Kaneko	Comments on MPEG-4 content IPR and compatibility with Version 2
3335	Belgian National Body	Belgian NB Comments on Computational Graceful Degradation
3339	Mike Coleman, David Thom, Chuck Leuck	Report of Ad-hoc group on AAC conformance and Technical Report
3340	Mike Coleman, Eric Scheirer	Report of Ad-hoc group on MPEG-4 Audio/Systems Reference Model
3341	Frank Nack, V V Vinod	MPEG-7 Requirements Document V.5
3342	Frank Nack, V V Vinod	MPEG-7: Context and Objectives V.7
3346	Sang-Wook Kim, Bernd Edler	Report of the Ad Hoc Group on MPEG-4 Audio Verification Tests
3357	Toshiyuki Nomura, Masahiro Iwadare, Naoya Tanaka	Listening test results and quality improvement of the MPEG-4 Narow-Band CELP coder for multi languages
3359	Yuji Maeda, Masayuki Nishiguchi	Subjective test results of codeword reordered HVXC

3362	The National Body of Japan	Comment on the parametric speech coder part of the MPEG-4 Audio
3378	Naoki Iwakami, Takehiro Moriya, Akio Jin, Kazunaga Ikeda, Takeshi Mori, Satoshi Miki	Proposal of a revised text for the T/F part of the 14496-3 CD
3379	Kazunaga Ikeda, Takeshi Mori, Takehiro Moriya, Naoki Iwakami	Reports on the prescreening tests of the revised UEP tools
3383	Akio Jin, Takehiro Moriya, Naoki Iwakami, Takeshi Norimatsu, Mineo Tsushima, Tomokazu Ishikawa	Modification of scaleable TwinVQ for harmonizing with AAC
3387	Korean National Body	On the efficient utilization of resources for the success of MPEG-4 Version1
3393	Korean National Body	A Study on the CD14496-3 Audio
3394	Korean National Body	Profile and Audio/System issue of TTS in the CD14496-3 (Audio)
3395	Korean National Body	On BSAC of ISO/IEC CD 14496-3 (Audio)
3396	Korean National Body	Comments on ISO/IEC 13818-7/DCOR1
3397	Martin Dietz, Ralph Sperschneider	VM integration of HCR
3398	Martin Dietz, Thomas Sch?fer, Ralph Sperschneider	Evaluation fo the common EP tool
3405	Jens Spille	Report of Ad Hoc Group on MPEG-4 Audio Conformance
3407	Jens Spille	Report of Ad-Hoc Group on MPEG-4 Audio Tools Complexity
3422	Toshiyuki Nomura, Masahiro Iwadare, Naoya Tanaka	Listening test results of the MPEG-4 Wide-Band CELP coder with the Multi-Pulse and the LSP-VQ tools
3429	Adam Lindsay, Ibrahim Sezan, Masahiro Shibata, Thomas Sikora, Toby Walker	Program of MPEG-7 Seminar in Tokyo
3430	Martin Dietz, Toshio Miki	Report of the Ad-hoc group on MPEG-4 Audio error resilience for Version 2
3431	Martin Dietz, Sang-Wook Kim	List of selected items for the MPEG-4 Audio NADIB verification test
3437	Bernhard Grill, Heiko Purnhagen	Report of the adhoc group on MPEG-4 Audio CD and Reference software progression
3438	Bernd Edler, Heiko Purnhagen	Scalability for Parametric Audio by Means of Simple Transcoding
3439	Bodo Teichmann	Report of the Ad-hoc group on Audio/Systems issues
3447	Eric Scheirer	Report of the AHG on Structured Audio
3466	Cassandra Swain, Joern Ostermann, Qian Huang, Barry Haskell	Requirements for Descriptors for 3D Human Body Model
3467	Kevin O'Connell, Karlheinz Brandenburg	Report of the MPEG-4 Profiles and Levels Ad-Hoc Group
3477	Adam Lindsay	MPEG-7 Applications Document
3479	Heiko Purnhagen, Bernd Edler	Scalability for Parametric Audio Coding by Means of Simple Transcoding
3483	David Thom, Heiko Purnhagen	Report of the Ad-Hoc group on Audio web page

Output Documents

The following output documents were produced in whole or part by the Audio Subgroup. Those shown in Italics were approved for public release.

Title	number
Study on Conformance 13818-4/DAM 1(covering 13818-7 AAC)	N2151
Study on Technical Report 13818-5/DAM 1 (covering both 13818-3 Second Edition and 13818-7 AAC)	N2152
Study on MPEG-2 AAC 13818-7 Draft Technical Corrigendum 1	N2153
Study on DoC on MPEG-2 AAC 13818-7 Draft Technical Corrigendum 1	N2154
MPEG-2 AAC 13818-7 Dynamic Range Control Proposal and Workplan	N2155

Interim release of ISO/IEC 14496-3 FCD	N2247
DoC on MPEG-4 Audio Committee Draft 14496-3	N2120
Consideration for Conformance of the MPEG-4 Audio TTSI	N2140
MPEG-4 Reference Software Final Committee Draft 14496-5	N2205
DoC on MPEG-4 Reference Software Committee Draft 14496-5	N2121
MPEG-4 Audio AAC/TwinVQ convergence workplan	N2156
MPEG-4 Audio verification test specification: narrowband audio broadcasting	N2157
Proposal for MPEG-4 Audio verification tests: speech codecs	N2158
Proposal for MPEG-4 Audio verification tests: music on Internet	N2159
Information on MPEG-4 Audio systems issues	N2160
MPEG-4 Audio Proposed Amendment WD	N2213
MPEG-4 Audio error resilience workplan update	N2161
MPEG Audio FAQs version 7	*N2162*
Proposals for the MPEG Audio web site content	N2163
Bibliography of MPEG Audio documentation	N2164

Annex 8

Report of SNHC meeting

Source: Peter Doenfes, Chair

Meeting Objectives

The objectives of the meeting for SNHC were to:

1. Generate MPEG-4 SNHC Part 1 Disposition of Comments on the CD and final text for FCD with traceability through National Body comments to CD, referenced Study of CD, and contributions.
2. Coordinate with Systems on DoC and FCD related to BIFS node updates and animation header.
3. Coordinate SNHC Reference Software developments for IM1 contributions and Version 2.
4. Refine SNHC Version 2 contents and work plan for VM and WD supported by work progress and CE objectives and methodology, particularly for new 3D Model Coding and Body Animation work.
5. Confirm SNHC Version 2 technology conformance with N2073 promotion/process methodology.
6. Work closely with Requirements on profiling to achieve the simplest structure possible, and to select for profiling only those tools whose utilization and maturity justifies profiling in Version 1.
7. Continue dialog with ISG about CGD and metrics for 3D synthetic graphical objects, and agree on strategy for closure in Version 2

Output Document Editors

Editors or coordinators responsible for SNHC elements of output documents were:

SNHC Disposition of CD Comments & FCD Text	
• Face animation coding	Eric Petajan, Igor Pandžić
• 2D mesh coding	Ebrahim Sezan
• Still image texture coding	Iraj Sodagar
• View-Dependent Scalable texture	Homer Chen, Touradj Ebrahimi
Coordination with Video on DoC Integration	Touradj Ebrahimi
Coordination with Systems on DoC & FCD Integration	Liam Ward, Julien Signes, Ganesh Rajan
SNHC Software Work Plan	Jiankun Li
SNHC Verification Model 8.0	Gabriel Taubin, Tolga Capin
SNHC Core Experiments	Tolga Capin, Frank Bossen
Press Release	Pete Doenges

SNHC Reports & Working Groups

Summary of results of AHGs & recommendations were given as shown:

3457	MPEG-4	SNHC	Peter K. Doenges	Report of Ad Hoc Group on SNHC FAQs Editing
3470	MPEG-4	SNHC	Gabriel Taubin (editor)	Report of AHG on SNHC Verification Model 7.0 editing
3371	MPEG-4	Video	T. Ebrahimi, P. Doenges	Report of ad hoc group on 3D model coding

3473	MPEG-4	SNHC	Eric Petajan, Tolga Capin	Report of Ad Hoc Group on Face and Body Animation
3478	MPEG-4	Video	Iraj Sodagar	Report of MPEG-4 Visual Texture Coding Ad-Hoc

SNHC/Video Contributions

A number of contributions were presented on recommended changes to the CD, on results and tools associated with core experiments for Version 2 VM work, and on the refinement or introduction of new technology for 3D Model Coding and Body Animation:

SNHC VM & FAQs				
3457	MPEG-4	SNHC	Peter K. Doenges	Report of Ad Hoc Group on SNHC FAQs Editing
3469	MPEG-4	SNHC	Gabriel Taubin (editor)	SNHC Verification Model 7.0
3470	MPEG-4	SNHC	Gabriel Taubin (editor)	Report of AHG on SNHC Verification Model 7.0 editing
Face Body Animation				
3241	MPEG-4	SNHC	Igor Pandzic, Claudio Lande, Yuval Fisher	FBA revisions to the Systems CD
3261	MPEG-4	SNHC	Roberto Pockaj, Eric Petajan	FBA revision to the Visual CD
3279	MPEG-4	SNHC	Roberto Pockaj, Loris Ambrosini, Maurizio Costa, Fabio Lavagetto	Results of Core Experiment FA1 on Face Calibration
3312	MPEG-4	SNHC	Gabriel Abrantes, Fernando Pereira	Implementation and Evaluation of MPEG-4 Facial Animation Tools
3318	MPEG-4	SNHC	Yuval Fisher, Hai Tao, Jorgen Ahlberg, Roberto Pockaj, Homer Chen	Revised Description of FIT Node
3332	MPEG-4	SNHC	Francoise Preteux, Patrick Horain, Hocine Ouhaddi, Marius Preda	Report on Core Experiment 3 on Hand BAPs interpretation
3401	MPEG-4	SNHC	Joaquim Esmerado, Tolga K. Capin	EPFL Results of the Body Animation CEs
3402	MPEG-4	SNHC	SNHC FBA group	Proposal for update to body animation section of SNHC VM
3413	MPEG-4	SNHC	Escher Marc, Pandzic Igor-Sunday	Result of core experiment F1
3423	MPEG-4	SNHC	Jorgen Ahlberg	Result of Core Experiment F2
3473	MPEG-4	SNHC	Eric Petajan, Tolga Capin	Report of ad hoc group on Face and Body Animation
3D Model Coding				
3324	MPEG-4	SNHC	Jiankun Li, C.-C. Jay Kuo, Homer Chen	Compression of Mesh Connectivity by Dual Graph Approach(M1)
3325	MPEG-4	SNHC	Jiankun Li, C.-C. Jay Kuo, Homer Chen	Embedded Coding of Mesh Geometry (M2)
3326	MPEG-4	SNHC	Jiankun Li, C.-C. Jay Kuo, Homer Chen	Progressive Mesh Coding by Vertex Split (M3)
3331	MPEG-4	SNHC	Francoise Preteux, Gerard Mozelle, Jose Paumard	Report on Core Experiment M1, M2 and M3 on 3D Model Coding
3334	MPEG-4	SNHC	Frank Bossen	Results of Core Experiments on 3D model coding
3371	MPEG-4	Video	T. Ebrahimi, P. Doenges	Report of ad hoc group on 3D model coding
3373	MPEG-4	Video	Ad hoc group members on 3D model coding	Description of core experiments on 3D model coding
3403	MPEG-4	SNHC	Jin Soo Choi, Young-Kwon Lim, Myoung Ho Lee, Chieteuk Ahn	Results of core experiments M2/M3: Geometry coding using PRVQ
3474	MPEG-4	SNHC	Gabriel Taubin, Bill Horn, Andre Gueziec, Claudio Silva	Report on Results of Core Experiments on 3D Model Coding (M1-M2-M4)

3475	MPEG-4	SNHC	Gabriel Taubin, Bill Horn, Andre Gueziec, Claudio Silva	Report on Results of Core Experiments on 3D Model Coding (M3)
Still Image Texture Coding				
3323	MPEG-4	Video	Michael Frater, Wee Sun Lee	Proposal and Results for Error Resilient Coding of Still Texture
3347	MPEG-4	Video	Iole Moccagatta, Osama Alshaykh	Proposal for a Core Experiment on Still Texture Error Resilience using a Packet Approach
3354	MPEG-4	Video	Tomoko Aono, Norio Ito, Hiroyuki Katata	Proposal of Scalable Texture Coding with Fine Granularity
3355	MPEG-4	Video	Norio Ito, Shin-ya Hasegawa, Masahiro Shioi, Hiroyuki Katata	Proposal of Tiling function for Scalable Texture Coding
3377	MPEG-4	Video	Osama Alshaykh, Iole Moccagatta, Homer Chen, Houng-Jyh Wang, Yi-Liang Bao, C.-C. Jay Kuo	Status of the core experiment on improving compression efficiency of still texture coding - adaptive scanning of wavelet subbands
3478	MPEG-4	Video	Iraj Sodagar	Report of MPEG-4 Visual Texture Coding Ad-Hoc
3480	MPEG-4	Video	Iraj Sodagar, Hung Ju Lee, Paul Hatrack, Bing-Bing Chai, Shipeng Li	A Proposal for Improving Coding Efficiency of Visual Texture Coding
3481	MPEG-4	Video	Iraj Sodagar, Hung Ju Lee, Paul Hatrack, Bing-Bing Chai, Shipeng Li	Result of Core Experiment T1: Improvement on Visual Texture Coding Using 1-D Adaptive Wavelet

FBA Results

Extensive work was reviewed in AHG meetings on FBA and 3D Model Coding on Sunday and during WG11 meeting, drawing upon contributions above. Much discussion of FCD recommendations driven by National Body comment was conducted, and DoC and FCD text was generated or drawn from references. More detail is given on these efforts in a separate section later.

FBA reviewed face animation contributions, notably on face calibration under CE FA1, the non-normative aspects of how to use the FA system to make facial models increasingly specific to source data, tolerance for error with different FAP bitrates and quantization scale factors, bitrate vs. frame rate, tradeoffs in using frame-based vs. DCT-based coding, and expression control. The calibration work was very useful in revealing what to expect in the fidelity of general speech, expressions and visemes that can be reconstructed. This work also showed that the face calibration process is sufficiently complex and subject to varied encoder and modeling techniques, such that the Predictable and Calibration profiles conceived in Fribourg needed more experimenting before defining object profiles. A revised proposal for the FIT node was input for Systems to consider along with some examples of how to interpret the semantics of the node. Much of this work was used later in the meetings with Systems on BIFS nodes and with Requirements on profiling. Independent implementations of the FAP coder now included the common animation header re: BIFS. Significant effort has gone into IM1 integration, and modules of FA including FIT were nearing completion in IM1.

Initial results on CE FA2 were shown based on use of Facial Action Basis Functions (FBFs) that are intended to exploit correlation in FAPs to achieve higher compression with low lag. The technique provides a mapping of 14 FBFs into 68 FAPs using a transform matrix that is reduced in correspondence with masked and unused FAPs, and that otherwise leaves the specific fitting method to the encoder. The FBFs were also coded using predictive coding to exploit temporal correlation. While FBFs show small integer gains in bitrate efficiency for the same PSNR, they are not yet sufficiently differentiated to justify standardization without significant further improvement targeted in a subsequent CE. The predictive arithmetic coding in the CD for FAPs shows favorably relative to FBFs when left-side/right-side symmetry is used. A number of concepts for FBF enhancement are contemplated by proponents, but this must be tested by the next meeting with N2073 requirements in mind if FBFs are to be considered further. Hopefully the pre-Dublin FBA AHG meeting will deal with this. At this point a very large step in efficiency, perhaps in trade for small fidelity compromises,

needs to be demonstrated to continue.

FBA looked at body animation contributions, including BA3 on hand signing based on procedural generators, BA2 on BAP compression, a constrained examination of quantization effects on the fidelity of body movement (not full conditions of BA6), and a proposal for a significant update to the body animation section of the SNHC VM. BAPs are intended to be model-independent, and video results were targeted to show whether similar postures, behaviors, and hand sign dynamics can be achieved with different models. One set of results from EPFL was shown in a time sequence of renderings with one hand counting while connected to a 3D body model. Video results and corresponding test data for BAP coding under BA2 allowed assessment of waving with one arm, a curious sequence moving both arms, an attentive sequence moving upper body joints, and walking behaviors that move most joints of a body model. Tests were run over a range of global quantization factors while holding the quantization step size fixed. These conditions allowed the viewing of tradeoffs between bitrate, global quantization, and quality of the rendered body animations at a fixed frame rate. Initial results possible by Tokyo showed that 8Kbits/sec was needed for highest quality with unity quantization, while 4Kbit/sec was reachable under a quantization value of 16 with good results. More careful work on quantization is needed.

Hand signing clearly requires a wide range of test data to ultimately represent the alphabet/numerals and higher-level hand signing symbols (some context-sensitive), but test data going into the Tokyo meeting was limited. Another contribution generously offered by SIM/INT showed a procedural system with graphical user interface for generating hand signs. A video analysis system may be used eventually to generate some hand signing. The procedural system provides hand signing in two layers. An inside layer provides the kinematic graph and associated articulation and movement of these morphologically scalable kinematic structures. A second layer corresponds to the visual model and its attributes (standard 3D graphics shapes, with a more general 3D mesh model of the hand under current integration). Time was not available before Tokyo to generate a complete set of alphabet or higher-level signs that should eventually be used in more complete testing of BAPs/BDPs. Generating a complete, natural-looking set of hand signs without unnatural movements or collisions between hands and arms could be a major task. Questions remain about what kind of subjective testing with competent subjects may be needed later.

Efforts toward BA6 on BAP Quantization Step Size and BA7 on BDP Interpretation were tabled for future work. While these latter topics deal more with how to use body animation coding (non-normative), it is still crucial to cover these areas of experimentation over the next year with sufficient cross-checking before profiling can be accomplished properly in Version 2. The lessons learned so far in face calibration may be compounded when attempting to calibrate body models. It is most important before the Dublin meeting to achieve independent testing of BAP coding and presumably bitstream exchange, so there is much work yet to complete that, as always, depends on adequate participation.

3D Model Coding Results

3D Model Coding continued the assessment of PRVQ for geometry coding including progressivity, as well as scalar and successive quantization from IBM and USC/Rockwell for static geometry. Competing techniques were evaluated in connectivity coding based upon dual graph and topological surface partitioning schemes. Progressive connectivity coding based on forest split or independent vertex split would provide for detailing of baseline 3D models for bitstream scalability or for level-of-detail control after a resolution hierarchy is downloaded to a terminal.

Based on the process requirements and testing results available in Tokyo, the SNHC VM was upgraded to use the topological surgery method for topological compression, retaining PRVQ as before for geometry coding. The group agreed that the new round of CEs targeted for completion by the interim 3D Model Coding AHG meeting in New York in May 1998 will be used to qualify other potential technology promotion into the VM baseline per N2073. More information on competing technologies for the next CE round is given in a later report section. The group also agreed to carry descriptions of the competing technologies in the CE document, which has grown large, to allow independent work to proceed quickly among participants as decisions are

made.

USC/Rockwell proposed a different topological coding scheme in San Jose, relative to the baseline spanning tree technology then in the SNHC VM. This technique represents an advancing-front coding strategy that starts from a seed point. The dual graph of the original face/edge/vertex structure is coded, using a set of symbols representing branches in the dual graph, and another set of symbols for end nodes in mergers of the dual graph. Using path length within the dual graph as a probabilistic clue to whether a link is a branch or merge (shorter for branch, longer for merge), an adaptive context-sensitive arithmetic coder was then used. This was explained more thoroughly. Experimental results were presented that involved testing of the method with a large set of big and small models from EPFL (a few vertices to many 10,000s) where the meshes were triangulated before compression. The CE M1 showed 3.6 BPT maximum, .3 BPT minimum, and 1.6 average, representing useful results that should be explored further.

Attempts by other experimenters to repeat these and other related techniques were complicated by changes in the details before the Tokyo meeting. More communication of accurate details on this work was needed to give others a decent chance to implement, and the CE detail on syntax was sometimes simply insufficient for participants. Significant meeting time was dedicated in Tokyo to share different techniques in detail among participants to ensure that subsequent independent work can move quickly. Significant effort was made in an attempt to identify core algorithm elements that are common between the topological surgery and dual graph, hoping to consolidate underlying tools into the smallest combined set with the most power, although to no avail. This discussion did however reconfirm the key importance of comparing total (topology + geometry) performance in new CEs. It was recognized that topological coding for the BPT levels of topological surgery and dual graph represents a small fraction of the total bitstream (geometry being much larger). The final determination of which topological coding technology to use could consider whether the chosen method fosters competition in the encoder.

The topological surgery method from IBM, well known as the top candidate for VRML Compressed Binary Format (CBF) while that was being considered, continued to show well in CE M1 results. Testing with extensive data sets for 8-, 10-, and 12-bit bounding cube quantization gave impressive compression results for M1 as before and also in M2 on Geometry Coding and M4 on Attribute Coding. The full collection of tools used in the proposed VRML CBF was applied to topology, vertices, normals, color, texture coordinates, etc. across these results. EPFL also showed in CE M1 a significant improvement in coding efficiency by combining the spanning tree decomposition of IBM with arithmetic coding for gains between 20% and 85%, with an average of roughly 40% across 300 models tested. EPFL work on M2 using two ancestor vertices for prediction showed BPV averages of roughly 15-25 for bounding cube quantization of 8, 10, and 12 bits, but distortion metrics are needed to make comparisons.

PRVQ continued to test well for M2 using relatively large models with high spatial correlation between the positions of successive vertices in the meshes tested. PRVQ provided compression gains over Gzip of 3.5 to 4.5 times, while constraining losses and consequent distortion (compared with lossless Gzip) to almost imperceptible levels. These results corresponded with 6-27 BPV and distortions of .54-.003 where it is increasing clear that all experimental results must be given in these terms with agreed metrics and visualization to make comparisons possible and meaningful.

USC/Rockwell evaluated an embedded context-sensitive arithmetic coder to code geometric data. Studies of prediction ancestry in clouds of vertices organized in local regions showed that the prediction residues are statistically somewhat insensitive to prediction length (number of ancestors used). So a first-order prediction was used with the predecessor vertex. The residues of successive (x, y, z) coordinates are normalized by means and variances so that each of three residue sequences has zero mean and unit variance. Then the residue vector is transformed into spherical-like coordinates. The r is analyzed with respect to its histogram to fit a Gaussian distribution and scaled for minimum fitting error, after which r is then successively quantized with binary symbols based on how rescaled r fits into layered quantization bins. The (ϕ, φ) of the transformed vertex residue lies within one of six sphere regions defined by +/- 45° into which the angular residues are successively partitioned noting that these regions exhibit symmetry. Binary subdivision of the region in step

with the angular residues produces a string of binary symbols that codes the angular residues. Context-sensitive arithmetic coding then finally codes these symbols. The results of using this method on a large model, like that used in PRVQ testing, showed BPV in the 9-30 range with mean distortion statistics roughly between 1.35-.02 which needs more verification work.

PRVQ was retained in the SNHC VM as the preferred baseline for geometry coding (static and progressive). The next round of experimentation with better agreed metrics could change these conclusions. PRVQ also offers a natural lead-in to progressive coding of geometry analogous to progressive image coding. The issue remained about the most meaningful, agreed metric of geometric distortion for the lossy aspects of 3D model compression with tools available so uniform test data can be used by all participants and so that more objective decisions can be made in down-selecting technologies.

CE M3 for progressive topology was evaluated with two technologies. The progressive forest split from the IBM submission in San Jose was evaluated, and a form of progressive vertex split/edge collapse from USC/Rockwell was also tested, in terms of lossless coding of the expansion and contraction of mesh connectivity according to the strategy of each technique. In both techniques, multiple layers of expanding detail are generated by distributing vertex additions over a base 3D mesh in such a way that local vertex additions are independent of neighboring splits. Forest split opens coherent holes in the base mesh, and then adds vertices in groups. The vertex split/edge collapse method adds single vertices that are just distant enough in the mesh to decouple their effects and coding, and to permit vertex additions in any order. Split flags and patterns are entropy coded with arithmetic coding. The techniques can use the M1 method for coding of the static topology of the base mesh. Since the base mesh corresponds to the simplest level of detail in typical computer graphics terms, then the smaller the base mesh - the less efficient in BPT will be the likely coding of the base mesh.

The USC/Rockwell work explored 15 layers of mesh enhancement with a split rate of 20-25% per layer, BPT/layer from roughly 4 down towards 2 as triangles were added, and cumulative compression achieved at about 3 BPT on average. The IBM results evaluated a 5-level enhancement of the base mesh, under similar test conditions used by them in M1, M2, and M4 (as pertains to the geometry and base mesh coding). The IBM results showed connectivity coding of roughly 4-7 BPT and provided integrated results with geometry coding included. These efforts provide an initial glimpse into the results that can be expected. However, since the context of the specific vertices chosen to accomplish topological mesh enhancement can effect the efficiency of geometry coding, further experiments should look consistently at the integrated sum of geometry and topology compression for static and progressive cases to draw properly informed conclusions about what is the best. Part of the challenge for the next CEs is to design the conditions to make testing as comparable as possible in spite of algorithm differences.

Still Image Texture Coding

DoC and FCD work as well as Version 2 enhancements (including error resilience and improving compression efficiency and/or picture quality using adaptive wavelets of different kinds) are being conducted by the Video Subgroup. SNHC appreciates this work and monitors the results. See the Video meeting report for more details.

DoC and FCD Text

A highly detailed effort was undertaken to make an inventory of NB comments on SNHC-related aspects of the CD to prepare DoCs and FCD recommendations. Considerable time was spent in reviewing NB-referenced meeting contributions on recommendations for repairing defects in the CD and in tracing through NB comments to Study of CD and CD references. The most extensive changes were made in face animation, and other attention was given to 2D dynamic mesh coding, view-dependent scalable texture, and scalable still texture coding. All NB references on the CD (not listed here) were surveyed.

By agreement with Video, the DoC on SNHC items was prepared as an Appendix A, Disposition of SNHC-

related Comments on MPEG-4 CD Visual 14496-2, for incorporation in N2119, Disposition of National Body Comments on CD Ballot on ISO/IEC 14496-2 (Visual). There were approximately 200 SNHC edits or comments requiring action, with about half in the face animation area, and the rest over 2D mesh, VDS, and still texture. Appendix A is ordered by country with cross references between NB input or references and the affected specification section, and with a summary of the problem and the cure. Output document N2227 collected the actual SNHC updates of the FCD text for use by the FCD editors before the publication deadline. In several cases where the overall response should consider the results, for example, of profiling coming from the Requirements group or BIFS node changes governed by Systems, appropriate reference was made to their respective contributions to the DoC and FCD text.

Profiling with Requirements

Extensive joint meetings with Requirements competed with their precious time in attempts to simplify and consolidate object and combination profiling in a realistic manner. The results of this big effort for SNHC and across all the subgroups should be reviewed in the Requirements reformulation of profiles. In the case of SNHC, the object profiles were generally simplified or deferred to Version 2. Specifically, object profiles were defined satisfactorily for simple face animation, 2D mesh, and scalable still texture.

Much work was accomplished in San Jose on quantifying difficult aspects of conformance points in the downloading and reconstruction performance of the more advanced, BIFS-based aspects of tailoring facial models. However the calibration and predictable object profiles of face animation were delayed to Version 2 on the basis that more exploitation work in facial calibration and more industrial input is needed. Due to a lack of adequate industrial input at this time on VDS texture, there is no object profile on VDS, while the tools based on earlier bitstream exchange and verification remain in the specification.

A very useful modularization of the scene description and graphics profiling related to Systems was accomplished that offers choice in VRML-like 3D, 2D, and combined 2D/3D primitives, in a manner somewhat independent of the way in which other visual decoding tools are selected. The current partitioning of Systems and Visual profiling, and the mix-and-match possibilities supported by this structure, should allow good choices of decoder implementations in hardware and software with the minimum footprints in tools for logical application classes.

Harmonizing with BIFS and Systems

DoC and FCD Text

3.1 Extensive joint discussions resolved the final aspects of a common bitstream header for BIFS-Anim and the elementary streams of face and 2D mesh animation. A significant overhaul to BIFS nodes related to FAP, FDP, and FIT was accomplished, also driven by NB comments.

IM1 Support

In review, SNHC contributors committed IM1 integration support for the Tokyo meeting. This committed individuals to contribute encoder software and fast/optimized decoder software to the IM1 Systems implementation software and to integrate the work into IM1 for SNHC functionalities:

2D Mesh	U. of Rochester	P. v. Beek
FAP coder	Rockwell	H. Chen
Wireface	Lucent Bell Labs	E. Petajan
Miraface	MIRALab/EPFL	I. Pandž i æ
FAT	AT&T	J. Ostermann
FIT	U. of Illinois	H. Tao

This work was drawing very close to conclusion with the fine efforts of Karl Oygard of Telenor. Without

going into details, the integration of 2D animated mesh into IM1 encountered some difficulties with what appeared to be 2D node problems in Systems BIFS. The subsequent concerns about completing IM1 integration were such that the 2D mesh software contributor, Peter van Beek, stayed home from Tokyo to concentrate on IM1 software work. Only a little more time after Tokyo will tell how all the IM1 integration work progressed and whether optimization of decoder software is needed.

CGD with ISG

ISG continued again its fine work on the development of SNHC complexity metrics and the basis for exercising parametric controls in the bitstream to accomplish computational graceful degradation for SNHC media types and compositions in Version 2. Contribution document M3433 provided important details for embedding a placeholder bit for bitstream expansion of CGD metrics in Version 2 that achieves backward compatibility requirements for Version 1 content. The document indicates suggested insertion points for this bit flag, considering face object, mesh object, and still texture object as well as video object. Output document N2126 provides the methodology for experiments to identify SNHC parameters for the implied complexity of decoding (and possibly rendering) of SNHC-related elementary stream types and compositions of them supported by Systems BIFS. Parameters would be drawn from the normative information in Version 1 bitstreams. SNHC contributors are expected to provide software implementations to ISG before Dublin so that experiments can be run to look for correlation between decoder/terminal loading and the intrinsic complexity of the media or its composition. ISG can also provide guidance to SNHC members on running experiments. SNHC members are urged also to contribute media processors or graphics accelerators if this will bracket a more representative sample of (by then) older platforms on which MPEG-4 applications may run.

Version 2 VM/WD Development

The work plan of SNHC Version 2 remains basically the same as that agreed in San Jose. A consolidated Software Work Plan for SNHC was developed in cooperation with various contributors for Version 1 and 2, and is given in output document N2228. 3D Model Coding does not yet have a full set of baseline technologies in the VM for all areas, and FBA needs to decide about FBF coding. The following schedule milestones continue in force with the understanding that 3D Model Coding and FBA will utilize AHG meetings between Tokyo and Dublin to achieve adequate quantitative and subjective differentiation of proposals to finalize the core techniques for the Version 2 VM:

1. Core technologies stabilized in VM – May/June 98, AHG Meetings
2. WD 1.0 with initial bitstream exchanges – July 98, Dublin
3. WD 2.0 with final bitstream exchanges – Oct 98, Atlantic City
4. CD Version 2 – Dec 98, Eilat

Only then can bitstream exchanges commence according to document N2073, and a baseline WD for SNHC Version 2 be achieved in Dublin. There is an acknowledged possibility that CE results in the interim 3D Model Coding AHG meeting may not quite provide all the evidence needed to down-select all the technologies. If this occurs, the plan is either to reduce the scope of 3D Model Coding or preferably to use Dublin for final comparative testing so that bitstream exchanges occur no later than Atlantic City.

Core Experiments

The following Core Experiments were formulated for Dublin:

- **Body animation (N2230)**
  - BA2: BAP Compression
  - BA3: Hand BAP Interpretation

  - BA6: BAP Quantization Step Sizes
  - BA7: BDP Interpretation
- **Face animation (N2231)**
  - FA1: Face Calibration (feature points, mesh texture & shape)
  - FA2: FAP Coding with Facial Action Basis Functions (FBFs)
  - FA3: Error Resilience of FAP Bitstreams
- **3D Model Coding (N2232)**
  - M1: 3D mesh connectivity - lossless compression ratio
    - » Topological Surgery with enhanced arithmetic coding
    - » Dual Graph topology with adaptive context-based arithmetic coding
  - M2: Geometry coding including progressive shape
    - » Predictive Residual Vector Quantization (with Lattice VQ)
    - » Scalar Quantization (IBM VRML-CBF geometry compression)
    - » Successive Quantization (with embedded arithmetic coding)
  - M3: Progressive connectivity coding
    - » Progressive forest split compression
    - » Progressive vertex split/edge collapse
  - M4: Properties coding (normals, color, texture coordinates)
    - » Compressed representation of properties attached to a model, with all bindings defined in VRML (per vertex, per face, per corner)

Other Output Documents

The following output documents were generated:

- **SNHC Software Work Plan (N2228)**
  - Version 1 and 2 covered
- **SNHC VM 8.0 (N2229)**
  - Significant Body Animation update, Topological Surgery, etc.
  - Pre-Dublin AHG meetings may recommend further technology promotion per N2073

Ad Hoc Groups for Dublin

The following groups were established to coordinate core experiments and documents:

- **FBA (N2233)**
  - Combined Face & Body CE coordination, analysis, data gathering, BAP generation
  - Key AHG meeting with VRMLC H-Anim on June 8, 1998 at U. Pennsylvania
- **3D Model Coding (N2234)**
  - CE coordination, analysis, VRML test data gathering
  - Key AHG meeting on May 18-19, 1998 at IBM Research
- **SNHC VM Editing (N2235)**
  - Substantial 3D Model Coding and Body Animation editing

Annex 9

Report of MPEG/Test meeting

Source: **Laura Contin, Chair**

1. Introduction

At the 43rd meeting of WG11, in Tokyo, considerable progresses were made in the specification of the testplan for MPEG-4 verification tests. This was the result of a fruitful cooperation between Test Subgroup and Audio and Video Subgroups.
As a general rule test conditions (e.g. coding tools, coding parameters, kind of audio and video material, etc.) were defined according to Profiles specifications included in the Final CD and having in mind potential applications.

The table below summarises the features of the planned tests.

	Test	Profile	Application	Schedule
Video	Error robustness	Simple	Mobile Communications/	July '98
	Content-based coding	Core	Internet, DVD	October '98
	Scalability1	Simple scaleable	Internet, Simulcast	December '98
	Scalability2	Core	DVD	TBD
Audio	Narrow band broadcasting	Scaleable	broadcasting on AM modulated bands	July '98
	Speech coding	Speech	Speech communication	September '98
	Music on Internet	Main	Music on Internet	September '98

The sections below illustrate the specific goal of each test and provide details about the corresponding experimental conditions.

2. MPEG-4 audio verification tests

Three MPEG-4 verification tests were proposed in Fribourg (doc. N. 1849), related to the following applications:

- Internet Audio applications
- Digital Audio Broadcasting on AM modulated bands
- Speech Communications

2.1 Narrow-band audio broadcasting

This test was proposed by the European consortium NADIB (Narrow Band Digital Broadcasting) and it is aimed at evaluating the performance of digital systems, including MPEG-4, ITU-T G.723.1, MPEG-2 LayerIII, against Perfect AM.
NADIB motivation for having these tests is to get indications about coding efficiency and coding gain of the MPEG-4 Audio, compared under different test conditions, especially in the scaleable versus not scaleable mode.

Being a realistic condition for the NADIB system, this listening test will be done with a bitrate of 6 kbps the core coder and a bitrate of 24 kbit/s for the total bitstreams in monophonic mode. This means that the bitrate of the enhancement layer (or a separate coder working simultaneously) is 18 kbit/s.

Since both music and speech are relevant for the application addressed in this test, both types of items will be used. Considering that among the tested codecs some are suitable for music and others

for speech, the test results will strongly depend on the items. Statistical analysis will investigate the overall performance of the codecs under tests.

The comparisons of interest are:

1. to compare the coding efficiency of different possible core codecs: MPEG-Narrowband CELP as a speech codec and Twin-VQ as a generic audio codec. Furthermore, G723.1 is added as an anchor point. No audio codec anchors are available for narrow-band audio.
2. to evaluate the advantages/disadvantages of a system with two layers versus a system using only one layer. Therefore the unscaled AAC codec is compared to the scaleable versions at the same total bitrate (24 kbps)
3. to compare scaleability against simulcast. In the simulcast mode, as many bit streams as potential decoders are broadcast in parallel. This solution makes the decoding process less complex since only one decoder at a time has to be used in order to get the desired quality. On the other hand, upper layers can not take advantage from lower layers and the coding is expected to be less effective. Therefore it is desirable to compare, at a given bit rate, the audio quality for both solutions.

Details about test design and logistics of this test can be found in document N. 2157.

2.2 Speech coding test

This activity is aimed at evaluating the performance of MPEG-4 speech coding tools at different bandwidths and bit rates. Both coding efficiency and scalability will be evaluated.
Three different tests are planned according to the bandwidth and the range of bit rate. The first one will include bitrates from 2 to 5 kbit/s, the second one bitrates from 6 to 12 kbit/s and finally the third one bitrates from 16 to 56 kbit/s.
Suitable MNRUs an reference codecs will be added to the test conditions
Details about test conditions and logistics can be found in document N.2158

2.3 Music on Internet test

The goal of this test is to evaluate the performance of MPEG-4 audio coding tools at bitrates suitable for analogue modems and ISDN connections.
The previous testplan, proposed in Fribourg, was including only stereo items coded at bitrates ranging from 20 to 56 kbit/s. In Tokyo that testplan was revised and two new tests were added. In these new tests monaural items coded at different ranges of bitrate (6 to 8 kbit/s and 20 to 56 kbit/s respectively) will be evaluated. The bitrates considered in the test on stereo material were also slightly changed.
Details about test conditions and logistics can be found in document N.2159

3. MPEG-4 video verification test

Verification tests were previously planned for the following MPEG-4 video functionalities:

- error robustness
- content-based coding

In Tokyo a further test addressing scalability tools was proposed.

3.1 Error robustness test

The main goal of this test is to evaluate the performance of MPEG-4 video error resilience tools in error-prone environments, like mobile communications.

This test has been planned since July ’97 and most of the coding conditions were specified at the San

Jose meeting (February ’98).

Test Subgroup endorsed the recommendations of the ad hoc group established after the San Jose meeting to investigate test conditions to be used in this test. Coding parameter values were slightly changed, taking into account the results of the TransMux simulation and the new specification of the Simple Profile. Moreover, training conditions and test design were fully specified.

Details about test conditions and logistics can be found in document N.2165

3.2 Content-based coding test

In San Jose the pre-screening highlighted problems in a few conditions. Most of them were fixed by the ad hoc group that was established in San Jose.
It was recognised that in order to test representative conditions and to get more meaningful results more critical sequences should be used and rate control should be included in the coding strategies. Therefore a call for new material (N.2168) was issued and the ad hoc group on coding optimisation was required to provide the ad hoc group for verification test with optimised rate controls to be used in this test.

Details about test conditions and logistics can be found in document N.2166

3.3 Scalability test

The scalability test was proposed for the first time in Tokyo and two possible test configurations were discussed. The former is representative of the Simple Scaleable Profile, while the latter includes tools typical of the Core Profile. In both cases it was proposed that only tools that are specified in MPEG-4 version 1 are used.
In the case of the Simple Scaleable Profile both spatial and temporal scalability are taken into account and the enhancement is frame-based. The performance of these scalable schemes will be evaluated against the Simple Profile (non scalable) and multicast.
In the other case only temporal scalability is used and the enhancement layer is arbitrarily-shaped. The performance of this scaleable scheme will be evaluated against non-scaleable MPEG-4 schemes and MPEG-1.
Details about coding conditions are given in document N. 2167

4. List of output documents

Title	Doc. #
Specification of Formal Verification Tests on Video Error Resilience	2165
Results of the second Pre-evaluation of content-based coding verification test	2166
Preliminary specification of test conditions for video scalability verification test	2167
MPEG-4 audio verification test specification: narrow-band audio broadcastings	2157
Proposal for MPEG-4 audio verification tests: speech coding	2158
Proposal for MPEG-4 audio verification tests: music on Internet	2159
Call for new critical segmented audiovisual sequences	2168

5. Ad hoc groups

The following ad hoc groups were established:

Ad hoc group	Doc. #
Ad hoc Group audio verification tests (B. Edler, S.W. Kim)	2136
Ad-hoc Group on narrowband audio broadcasting verification test (L. Contin, M. Dietz)	2137
Ad-hoc Group on MPEG-4 Video verification tests (T. Miki, M. Wollborn, T. Suzuki)	2169

Annex 10

Report of Implementation Study Group

Source: Marco Mattavelli Chair

Generalities

Five main topics have been the subject of the activity of the Implementation Studies Group (ISG) during the Tokyo meeting:

1) Video decoder complexity analysis for the definition of profiles and levels,
2) Computational Graceful Degradation for SNHC video,
3) Complexity analysis for natural and synthetic audio,
4) Video decoder quality of service,
5) Complexity evaluation of MPEG-4 components.

The last item is a new topic that was discussed for the first time in the ISG during this meeting. Considering the status of the experiments within the various activities, the validation of the acquired results and the new started activity, it is clear the need of new critical video and audio content to be used for complexity evaluations. In particular there is the lack of segmented material containing several video/audio objects. Such content has been requested, in a specific call (W2168), by the ISG jointly with the Test Group to the MPEG community.

1) Video decoder complexity analysis for the definition of profile and levels

The group has reviewed the results of the Video Complexity ad-hoc group concerning the Overlapped Block Motion Compensation (OBMC) tool (M3468, M3488). This activity has been requested by the requirement group to answer NB comments asking the insertion of OBMC in the core profile. The results based on profiling measures were not conclusive and convincing. It was concluded that more care should be taken in the definition of the framework of profiling experiments in order to achieve meaningful results. The group could not reach an agreement to give a final recommendation to the Requirement Group for the inclusion or not of the OBMC tool in the core profile. The controversial results have been reported to the Requirement group that finally decided not to include the OBMC tool in the core profile since no unanimous consensus has been achieved for a change in the CD. The decision seems not consistent from the technical point of view since the OBMC tool is now present only in profiles at high bit-rates, while its usage is desirable and its efficiency is higher at low/medium bit-rates.

Another request from the Requirement group was to investigate the additional complexity of Spatial Scalability. Contribution M3416 was reviewed also reconsidering all the contributions on the subject that has been previously presented to the ISG in the last meetings. The additional complexity of adding a spatial scalability tool to two separate layers for N layered multi-resolution system has been estimated and compared to IDCT complexity as reference in terms of the number of addition and multiplication. The ISG group concluded that the inclusion of spatial scalability tools for rectangular objects does not add significant complexity to the existing profiles.

Assistance to the Requirement group has also been given for various complexity issues related to the definition of profiles and levels, in particular for the inclusion of wavelets texture coding tools in simple profiles. No other precise requests of video complexity evaluations have been done and no

ad-hoc group has been established with that purpose.

2) Computational Graceful Degradation for SNHC video and audio.

The ISG reviewed the CGD contributions and ad-hoc activity (M3336, M3433).
The variability of the complexity of SNHC tools is expected to be very high. No useful measures of the SNHC object included in the CD are available now. The common idea of ISG and SNHC group is that it would be useful to have CGD backward compatibility by the insertion of a one bit flag for version 1. The two groups will further discuss the syntax of such insertion, no agreement has been reached in Tokyo.
The two groups recognized the need to verify the existence of at least one parameter for each SNHC object that cannot be extracted from the bitstream and therefore needs transmission in the header in order to evaluate complexity. Experiments aiming at finding a correlation between SNHC object parameters and complexity have been specified and described in output document N2126.
It has been also concluded that CGD for natural audio is not needed, while further investigations need to be carried on for synthetic audio (no formal contributions have been presented but only some preliminary results).
An ad-hoc group has been established to perform SNHC experiments and to investigate CGD for audio in particular synthetic audio.

3) Complexity analysis for natural and synthetic audio

A joint meeting with the audio group discussed the problem of evaluating natural audio complexity. The metrics used for evaluation have been presented and compared to the results of generically optimised decoder implementations. Good agreement of the results has been found. It was concluded that the complexity of natural audio tools is not critical in absolute terms and it is not dependent on the audio content.
Another joint meeting of the ISG with the Synthetic audio task group has analysed the complexity of the synthetic audio profiles. Concerning the two lower profiles, (1. General MIDI and 2. Wave-table synthesis) decoder complexity is bounded for General MIDI profile or can be bounded by specifying the maximum size of downloadable wave-tables. The ISG recommended to the task group to specify functionally meaningful limits in order to be able to implement conformant decoders.
Concerning synthetic audio (SAOL) complexity, the situation is not satisfactory. SAOL defines a language that in principle is not bounded in complexity. The ISG recommended to find suitable techniques to bound the SAOL complexity in order to be able to define and implement conformant decoders. ISG will assist the future SAOL activities aiming at bounding SAOL complexity to meaningful levels and will actively participate to the related ad-hoc activities.

4) Decoder quality of service

No significant results have been reported from the ad-hoc group activity after the San Jose meeting. The group reviewed the description of QoS experiments and decided to simplify and restructure the experiment description. New simplified objectives for the experiments have been proposed:

- to define complexity metrics capable of describing the intrinsic complexity of video bit-streams
- to verify the applicability of the complexity metrics for specifying levels in video profiles
- to verify the applicability of the complexity metrics for bounding the intrinsic complexity of video bit-streams
- to verify the applicability of the complexity metrics for classifying different decoders

In a second step, on the basis of the obtained results, investigations will be carried out in order to assess the QoS of decoders classified by the specified complexity metrics

Output document N2127 reports the description of the revised experiments that are now limited to the definition of complexity metrics for the setting of meaningful video complexity levels. An ad-hoc has been established with the purpose of carrying on the experiments and further advance in the work.

5) Complexity evaluation of MPEG-4 components.

This is a new activity that the ISG started during the Tokyo meeting. The first step was to define the framework of the analysis and the evaluation criteria. Two possible outcomes can be envisaged.
In one case the complexity analysis is performed in terms of global performance of each system component measured on platforms of large diffusion. For instance, the performance of video decoders could be measured in terms of frame/sec using some standard test sequences. The advantage of this approach is to give a clear idea, although not precise, of the overall complexity of MPEG-4 components referring to the performance of well-known platforms. The disadvantage of this approach is that it must rely on optimised versions of decoders for a meaningful result.
The second case is to analyse the complexity in more abstract forms such as number of elementary operations required for some test sequences or other complexity metrics. The disadvantages of this second approach is that, although more precise than the first, it could result less illustrative of the MPEG-4 complexity.
The group considered the status of the reference software, its level of optimisation and the platforms for which the reference software is available. Most of MPEG-4 video audio and SNHC decoders are available on both Unix and PC platforms in non-optimised versions. Optimised versions are not available but performance evaluation results could be made available for the overall complexity evaluations.
The problem of the analysis of MPEG-4 System complexity has not been discussed in the Tokyo meeting. The definition of what is meant by complexity and performance of the MPEG-4 System and what could be measured to assess its performance will be the task for next meeting.
No ad-hoc group has been established for lack of interest of ISG participants in this activity.

Annex 10

Report of Liaison group

Source: Barry Haskell, Chair

The Liaison group considered the following Tokyo input documents

M3258 from ITI-T SG16 on DMIF

M3259 from ITI-T SG16 on GII

M3257 from ITI-T SG16 on Verification Testing

M3254 from ITI-T SG16 on DMIF, audio and video

M3249 from DAVIC on a new work item on IP networks and character coding

SC29/N2362 from ITU-R SG11 on new questions 207-2/11 and 249/11. This begins a study of compression methods for SDTV, EDTV and HDTV and associated computer technology.

M3321 from MMA on MIDI

M3333 on the first meeting of CEN

SC29/N2291 from SC24 on collaborations with VRML

The following output liaison documents were produced:

N2145	Liaison to VRML on modes of collaboration
N2146	Liaison to ITU-T SG16 on MPEG-4 DMIF
N2147	Liaison to ITU-T SG16 on MPEG-4 Audio
N2148	Liaison to ITU-T SG16 on MPEG-4 Video
N2149	Liaison to MMA on MIDI
N2150	Liaison to JPEG on the VCP for texture coding
N2186	Liaison to CEN on MPEG-7 activities
N2187	Liaison to SMPTE on MPEG-7 activities