7

Closing Plenary

Ad Hoc Group Meeting on MPEG-7 Test and Evaluation Issues

February 15-19, 1999

7

Special Recognition of Individuals for Extraordinary Support to the AHG

Purpose of this Plenary

- Summaries of Team Evaluations
- Non-normative Tool (NNT) work plan
- AHG report plans
- Comments from Requirements Group
- Final comments and dismissal

Summaries of Evaluations

- 14 Teams
- 132 Evaluators
- 392 Proposals evaluated
- Substantial effort and investment by those interested in success of MPEG-7

D1 - Visual Descriptors

V.V. Vinod

D2 - Visual Descriptors

Sylvie Jeannin

D3 - Visual Descriptors

B.S. Manjunath

D3 - Visual Descriptors Team

B.S. Manjunath
Miroslaw Bober
Sylvain Devillers
Chee Sun Won
Adriana Dumitras
Paulo Villegas
Titus Zaharia
Lilian Labelle
Shunichi Sekiguchi
Leszek Cieplinski
Dragutin Petkovic

Proposal Evaluation Summary

- 42 proposals
- Texture, image structure/signature, video structure, face, video caption and text, edge structure, editing effects, Image maps, shot descriptors, pose, object behaviors, metadata, motion and e-motion!

Promising Technologies

- Image/video structure
- Image texture
- Object behavior
- Face representation

Descriptor Feature Summary

- What features did you evaluate?
- What conclusions did you make about these features?
  - Relevant to highly relevant in general.

Recommendation Overview

List of features that where judged as being relevant or highly relevant	Number of proposals recommended for		
	XM	CE	MP
Emotion conveyed in images/video	-	-	-
Time of the day/location (indoor-outdoor) / lighting conditions	-	-	-
Spatial image intensity distribution	-	P102, P104, P566, P612	-
Characterization of homogeneous texture regions	-	P077, P566, P574	-
Representation of video structure	-	-	P179,P337,P638, P636
Text & captions on images/video	-	-	P180, P183, P639
Object behaviour in video	-	-	-
Labeled image map	-	-	-
Spatial/temporal relations between objects	-	-	-
Faces in images & video	-	-	P181, P551, P650
Human pose in images & video	-	-	-
Hyperlinks for media objects	-	-	P557
Textual metadata associated with images/video	-	-	-
Image and video signature	-	-	-

Other Recommendations

- **Other recommendations:**
  - Merging of proposals (MP)
  - close interactions between Ds and DSs.
  - More descriptor specific Test data sets
  - Related technologies (e.g., segmentation) should be carefully considered.

D4 - Audio Descriptors

Adam Lindsay

Proposal Evaluation Summary

- 34 eclectic Audio Proposals
  - 31 were relevant
  - high-level, relations, perceptual, physical...
- Promise for future development
  - Often the weakest proposals were set towards the most promising areas
- Essential technologies not proposed
  - Library-style meta-data
  - high-level/low-level bridges
  - Existing standards
    - e.g. ISO Language tables

Descriptor Feature Summary

- Links
- Temporal Structure
- High-level speech/music
- Excitation/Resonance Model
- Spectral Features
- Speech summary/content
- Classes
- Low-level building blocks
- Correlation
- Tempo

Recommendation Summary

- **Recommend to XM**
  - **Nothing**
- **Recommend to CE's**
  - **five different features/technologies**
- **Also Recommend…**
  - **Look at existing work/standards**
  - **Develop "essential" features**
  - **Thanks to D4 team for a good week**
    - **Barry, Carola, Jyri, Nakajima, Nick, Richard**

D5 - Audio Descriptors

David Thom & Team

Proposal Evaluation Summary

- We evaluated a total of 34 proposals
- They are all low level technical descriptors
  - Some proposals lacked completeness
- We found 31 which were relevant to MPEG-7
- No descriptor stood out as exceptional but most were sensible and practical

Descriptor Feature Summary

- The features were grouped as follows
  - » Frequency Spectrum
  - » Signal Energy
  - » Room Acoustics
- We identified several missing elements:
  - » Reverberation time
  - » Speech recognition
  - » Impulse response…..

D5 Summary Table

P#	Description	Grade	P#	Description	Grade
P595	Zero Crossing Rate	1	P368	Sharpness	4
P216	Zero Crossing Rate	1	P370	Loudeness	4
			P597	Peak Tracking	2
P596	Fundamental Frequency	4	P276	Forman Time Envelope	3
P376	Fundamental Frequency	4			
P278	Pitch Mark	4	P372	Sound Elevation	3
P399	Melodic	4	P373	Sound Distance	3
P400	Rhythm	4	P375	Sound Direction	3
			P279	Azimuth	3
P199	Intensity Envelope	3	P280	Elevation	3
P606	Short Term Energy	3	P281	Source Presence	3
P378	Energy Of Audio Sequence	3	P282	Warmth & Brilliance	3
P208	Amplitude Sinusoidal	4	P283	Room Presence	3
P209	Amplitude Stoshastic	4	P284	Running Reverberance	3
P204	Average Intensity	3	P285	Envelopment/diffussion	3
			P286	Late Reverberance	3
P198	Frequency contour	1	P287	Heaviness and Liveness	3
P195	Local Spectrum	4			
P197	Global Spectrum	2			
P377	Audio Frequency Content	4			
P96	Frequency Shell	3			

D5 Recommendation Summary

- **No proposals were recommended for inclusion in XM, but several areas are good candidates for the core experiment route if**
  - **Core experiment procedure is defined during Seoul meeting**
- **We need to identify other organizations which are also working on Meta-Data and co-ordinate activity**

D6 - Audio Descriptors

Andrew Collins

Yasuhiro Toguri

Wes Curtis

Pierrick Philippe

Georgio Dimino

Nevenka Dimitrova

Proposal Evaluation Summary

- **How many proposals did you evaluate?**
  - **29 Proposals Evaluated**
  - **Audio Low Level Features**
  - **(Audio) High Level Features**
- **Technologies with most promise for future development**
  - **Fundamental Descriptors**
  - **High Level Descriptors**
- **Any essential technologies not proposed?**
  - **Many, varied (temporal, identification & technical)**

Descriptor Feature Summary

- **What features did you evaluate?**
  - **Audio Noise / Distortion Descriptors**
  - **(Audio) Fundamental Descriptors (Bitrate, Size, Format)**
  - **Audio Performance Descriptors (Vibrato, Tremolo)**
  - **(Audio) Technical Quality Descriptors**
  - **Genre Descriptors**
- **What conclusions did you make about these features?**
  - **Many “obvious” descriptors still unclear**
  - **Much work still required**

Recommendation Summary

- **Summarize your recommendations for XM technologies**
  - **None. Technology still immature or ill defined**
- **Summarize your recommendations for Ces**
  - **Require clarification of features**
  - **Require consolidation of features**
- **Other recommendations**
  - **Mechanisms for linking proposals together**
  - **Require communication with other existing standards / cataloguing bodies**

DS1 - Visual Description Schemes

Mike Zeug

Proposal Evaluation Summary

- Evaluated 29 Visual Description Schemes Primarily targeted at Still Images.
- Most interesting Proposals focused on Description Schemes which combined signal descriptors and semantic descriptors and provided methods for linking between the two.

DS Functionality Summary

- Simple combinations of descriptors
- DSs providing structures combining signal and semantic information.
- Semantic Information, tree/graph, Relationships between D's, and Relevance feedback were deemed important.

Recommendation Summary

- Nothing that we reviewed should be immediately put in the XM.
- The 'Low Level" DSs may be candidates for CEs.
- The concept of tree/graph representations and relationships between Ds needs to be explored further.

Issues and Holes

- The ability of MPEG-7 to support links to Ds or DSs outside of MPEG-7 should be considered.
- No comprehensive shape DS proposed
- No DSs which encompass 2D as a subset of 3D properties

DS2 - Visual Description Schemes

Munchurl Kim

DS3 - Audio Description Schemes

Juergen Herre

Proposal Evaluation Summary

- 10 proposals evaluated
- Technologies with most promise for future development:
  - Audio segmentation, indexing, and representation
  - Description of music structure
  - Music classification
- No obvious obmission of essential technologies noticed

DS Functionality Summary

- **Main functionalities evaluated:**
  - **Segmentation, structured representation, and indexing of audio streams**
  - **Linking mechanisms**
  - **Description of Music Structure**
  - **Parametric Sound Description**
  - **Music Classification**
- **Essentially all of these were deemed useful for MPEG-7**

Recommendation Summary

- **Recommendations for XM:**
  No DS candidates for immediate incorporation into XM
- **Collaborative work required on some basic functionalities (e.g. Definition of Audio "Segment")**

  => Combination of structural elements from several proposals

Recommendation Summary (2)

- **3 Candidate Technologies for CEs (some after reconciliation) :**
  - **Audio Segmentation / Indexing**
  - **Music Structure Description**
  - **Parametric Sound Description**
- **Further investigations needed on some proposals (including AhG)**

DS3++ - Multi-modal Description Schemes

Juergen Herre

Proposal Evaluation Summary

- 15 proposals evaluated
- Technologies with most promise for future development
  - Multimedia segmentation, indexing, and representation
  - Linking to outside resources
  - Client-Adapted Delivery & Presentation
- No obvious obmission of essential technologies noticed

DS Functionality Summary

- Main functionalities evaluated:
  - Segmentation, structured representation, and indexing of audio streams
  - Linking mechanisms
  - TV Content Description & Programm ID
  - Multi-Level Digest
  - Support for Multimedia Production
  - ...
- Essentially all of these were deemed useful for MPEG-7

Recommendation Summary

- Recommendations for XM:
  No DS candidates for immediate incorporation into XM
- Collaborative work required on some basic functionalities (e.g. Definition of AV/MM “Segment”)
  => Combination of structural elements from several proposals

Recommendation Summary (2)

- 4 Candidates for CEs (some after reconciliation):
  - MM Segmentation / Indexing
  - DS for Broadcasting Delivery
  - Linking to Outside Resources
  - Multi-Level Digest Scheme
- Further investigations needed on some proposals (including AhG)

DS4 - Application Description Schemes

Ed McDermid

DDL - Description Definition Language

Frank Nack

CS - Coding Schemes

Fernando Pereira

Team Members

- Fernando Pereira (TL) — Instituto Superior Técnico
- Mark Buxton (DL) — Intel Corporation
- Koichi Emura — Matsushita [illegible] Co., Ltd.
- Joerg Heuer — [illegible]
- Gu[illegible] — [illegible] Laboratories
- Kim[illegible] — [illegible] Ltd.
- Hae[illegible] — [illegible] Electronics, Co
- Mas[illegible] — [illegible] University
- Fun[illegible] — Information Broadcasting Lab. Inc.
- Jim[illegible] — Sony BPE
- Li [illegible] — NTTDoCoMo
- Noel O'Connor — Teltec
- Charilaos Christopoulos — Ericsson

CS Evaluation Criteria

- Compression efficiency - Complexity of the description encoding and decoding processes.
- Lossless compression - Ability to losslessly compress descriptions instantiations.
- Streaming capability - It is possible to multiplex and stream the coded description.
- Error resilience - The coded description shall be robust against transmission errors.
- Universality - Ability to be applied to a wide range of descriptions.

Proposal Evaluation Summary

- Four (4) proposals were evaluated.
  - Two (2) proposals were under-specified and ‘could not be evaluated’.
  - The other two (2) proposals brought interesting technology but still some work needs to be done before consideration for the XM.

Missing Pieces

- The requirements more closely related to this team, notably efficient representation/compression, error robustness and streamability were not sufficiently addressed by the proposals.
- After the selection of the descriptors and description schemes, companies should be strongly encouraged to bring to MPEG the required coding technology to fill this gap.

Recommendation Summary

- NO Coding Schemes recommendations for XM.
- NO Coding Schemes recommendations for CEs.
- Coding requirements need reviewing, clarification, and re-classification.

Good work !

See you in Seoul !

ST - System Tools

Jens-Rainer Ohm

Work done

- 4 regular proposals summarized
- 3 additional proposals reviewed
- Summary contains statement about “MPEG-7 relevance”
- Evaluation criteria for Systems Tools set up along requirements

Proposal Summary

- Systems technology will have to be developed for push and pull applications
- Basic solutions in both of these areas were proposed
- Further development must be harmonized with DDL and DSs

Proposal Summary

- Basic systems aspects :
  - Flexible access to descriptions (also partial) and multiplexing
  - Synchronization of descriptions and of descriptions with content
  - Transport and file format
  - Robustness, Quality of Service and IP management

Proposal Summary

- More focus to be set on normative elements
- Open issues :
  - Distributed systems
  - APIs

To be done

- Discuss ST evaluation criteria on reflector
- Review and evaluate ST tools with Systems subgroup in Seoul
- Recommendations for XM ?

NNT - Non-Normative Tools

Philippe Salembier

Non-normative proposals

9	Descriptor extraction / computation
5	Descriptor similarity, search
1	General histogram distance
3	Still image segmentation
4	Video segmentation, spatial, motion
1	Text segmentation in video
1	Motion estimation
1	Audio segmentation
1	AV Scene segmentation
5	Shot detection and Key-frame extrac.
31	Total

Work to be done

- Proposals will also be received for non-normative MPEG-7 tools that will not be formally evaluated at the AHG meeting in Lancaster. There will be no presentations or demonstrations associated with proposals submitted for non-normative tools. However, a group of MPEG-7 experts will meet via teleconference and/or electronic mail following the Lancaster meeting to review and summarize the proposals received and prepare a final summary as a contribution to the Seoul MPEG meeting.
  - W2569 (evaluation guide)

Work organization

- 3 Teams:
  - Related to descriptors (NNT1)
  - Spatial segmentation (NNT2)
  - Temporal segmentation (NNT3)

1	Creation of 3 working teams	Feb. 24
2	Assignment of proposals to review	Feb. 25
3	Individual review done	Mar. 3
4	Consolidation of report	Mar. 5

Non-normative components of normative proposals

- Goal: create a summary of non-normative elements that may be used for XM development
- Some of these elements may be in proposals evaluated this week!
- Team leaders: identify proposals describing:
  - Still image segmentation
  - Video segmentation, spatial, motion
  - Text segmentation in video
  - Motion estimation
  - Audio segmentation
  - AV Scene segmentation
  - Shot detection and Key-frame extraction
  - Any other major non-normative elements not directly related to the normative proposals

AHG Final Report

- Contents
  - » Summaries of Team recommendations
  - » ST and NNT summaries
  - » Documentation of evaluation process
  - » Appendices
    - – Team summary forms
    - – Proposal evaluation forms
    - – AHG participants
- Schedule
  - » Discussion of draft on reflector (25 Feb-5 Mar)
  - » Approval of draft (before 8 March)
  - » Upload to MPEG FTP site (by 8 March)

Additional AHG Issues

- Proposals will be automatically moved to the MPEG Seoul FTP site as contributions from this AHG unless authors request that they NOT be uploaded
- Presenters are asked to send us PowerPoint copies of their presentations to be included with their proposals
- Contact: philippe@gps.tsc.upc.es

No AHG meeting planned for Seoul

MPEG-7 - Where to from here?
+ Q&A

by

MPEG Requirements Group Chair

Final Comments
and
Dismissal

Final Comments and Dismissal

- **Amount and quality of participation**
- **Significance of the results for MPEG-7**
- **Significance of the process for future MPEG technology evaluations**
- **Importance of the collaborative phase and participation in Seoul**
- **Thank you and have safe travel back to your homes!**