Video Signal Processing

杭學鳴 兼任教授
(Hsueh-Ming Hang, Adjunct Professor)

國立台北科技大學 資訊工程系

Time/Date/Instructor

- **Multimedia Compression**/Data Compression
- Time: Thursday (234; 9:10-12:00AM)
- Classroom: 科研大樓334
- Instructor: Hsueh-Ming Hang, 杭學鳴 hmhang@nctu.edu.tw
- Classnotes: (https://mcube.nctu.edu.tw/wiki/core/pmwiki.php?n=Course.VSP2021)

VSP21

Hsueh-Ming Hang 杭學鳴

- 1984: Ph.D. in EE, RPI (Rensselaer Polytechnic Inst), NY, USA
- 1978,1980: BS and MS, Chiao Tung Univ, Hsinchu
- 1991.11~2021.7: Nat'l Chiao-Tung Univ., Professor
- 2010.2~2014.1/2014~2017: NCTU, ECE College, Associate Dean/Dean
- 2006.8~2009.7: Nat'l Taipei Univ Tech, EECS College, Dean
- 1998.11~2004.7: NCTU, Telecom Research Center, Director
- 1984.6~1991.11: AT & T Bell Labs., USA; MTS
- 2008: IET Fellow
- 2002: IEEE Fellow
- 2000: IEEE Third Millennium Medal
- Associate Editor, *IEEE Trans Image Processing* (1992-1994, 2008-2012), *IEEE Circuits and Systems for Video Technology* (1997-1999)

Text Book and Recommended Readings

- Textbook: (1) **Lecture Notes**

  (2) K. Sayood, *Data Compression*, 5th ed., Morgan Kaufman, 2017. (4th ed; okay)
- Recommended Readings:

1) R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, 4th, Pearson, 2018.
2) J.-R. Ohm, *Multimedia Signal Coding and Transmission*, Springer, 2015.
3) D A. Murat Tekalp, Digital Video Processing, 2nd Ed, Pearson College, 2015.
4) D. Salomon, and G. Motta, Handbook of Data Compression, 5th, Springer, 2010. (on-line)
5) K.R. Rao, J.J. Hwang, and D. N. Kim, *High Efficiency Video Coding and Other Emerging Standards,* River Pub., Sept 2017.

Topics to Be Covered

1. Representation of Digital Images
2. Quantization and Lossless Compression
3. Color and Human Visual System
4. Transform Image Coding (JPEG)
5. Wavelet transform and Coding (JPEG2000)
6. Motion Estimation
7. Video Coding (ITU/MPEG Video)
8. Deep Learning and Image Processing
9. Deep-learning based image/video Coding

* There is no single book that covers all the above subjects in adequate depth.

Applications

- Video, Audio and Speech devices/services are used daily; examples: mobile phone, TV, ...
- "All" captured digital media is *compressed*.
- Huge consumer electronics market.
- Many companies in Taiwan/World:

(1) Devices (IC): MediaTek, Realtek, Qualcomm, ..

(2) Systems: HTC, Foxconn, Apple, GoPro, …

(3) Services: Google, Facebook, Netflix, ...

Grading

- **Computer Assignments**: 45 % -- Two (in C/Matlab): 20%, 25%
- **Examine**: 20% (2 hours, closed book, two pages of A4 notes)
- **Project** or **Paper Study** (simulations) and Report: 35%

Grading Policy

General Requirements: All examines and reports can be in either English or Chinese.

- You may be graded on *method* and *ideas*, and the *clarity* with which you organize them, but accurate numbers are needed for full credits.
- **Examples:** *Principles, formulas* ➔ *plug-in* ➔ *computational tricks* ➔ *accurate results*

Missing projects, homework, …: To claim credits, please retain a copy of your projects, homework,…

Final Project: Experiments/Paper Study

- 20 (?) mins. oral presentation for each person; pick up one topic from the given list.
- Grading of Final Presentation:

-- Written report (contents, organization, clarity, …): ~70%

-- Oral report (contents, organization, clarity, …): ~30%

-- Experiments (computer simulations)

- Your final submission should include: (electronic files)

(1) Slides (2) Report in .doc or .pdf file

(3) Programs (description and codes)

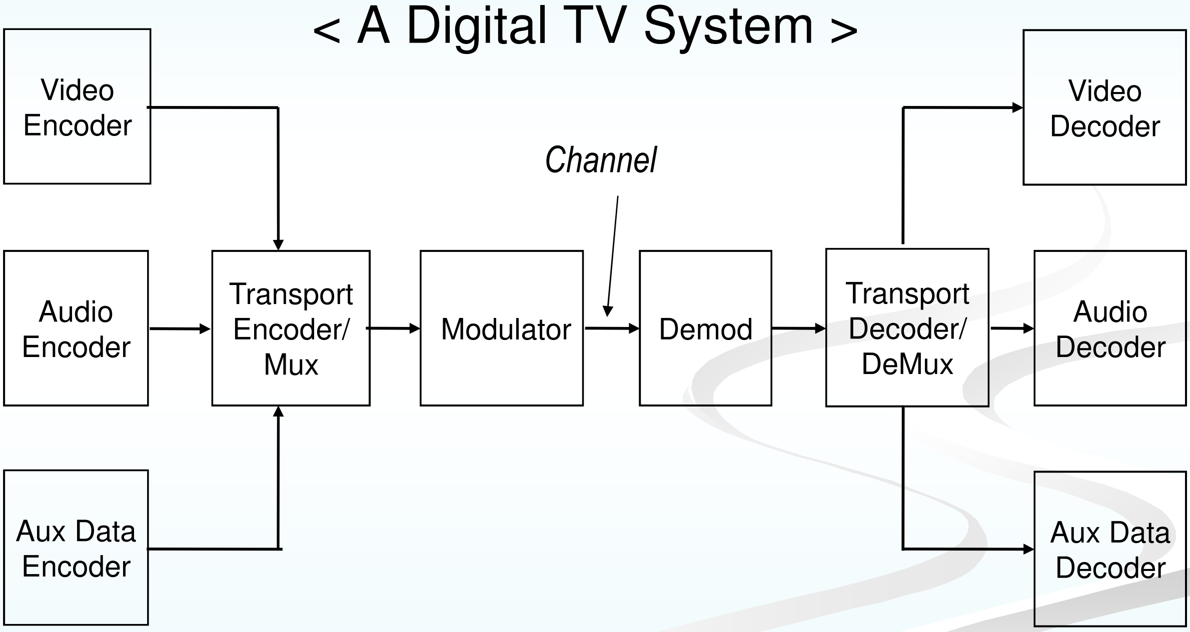
(4) Major references (files)

Chap. 0 Introduction

Digital Media in Daily Life

An Example of Comm. Systems

< A Digital TV System >



Example: ATSC Digital TV

-- *Advanced Television Systems Committee (ATSC) Digital Television Standard*

- ATSC: 1995
- For digital HDTV (terrestrial) broadcasting
- A: Video (MPEG2 Video)
- B: Audio (Dolby AC3)
- C: Transport Systems (MPEG Systems + ...)
- D: RF/Transmission Systems
- (E: Receiver)

Why Compression?

-- Massive data

- **Speech**: 8 bits (per sample) x 8K (samples/sec) = 64Kbits/s
- **CD audio:**

16 bits x 44.1K (samples/sec) x 2 (channels) = 1.411Mbits/sec
(44.1K = 60 (fields) x 245 (lines) x 3 (samples) (J. Watkinson, *The Art of Digital Audio*, p.28, Focal Press, 1989))

- **Digital TV:** (4:2:2, NTSC in CCIR 601)

Still picture: 720 (pels) x 483 (lines) x 2.0 bytes = 5.564 Mbits
Motion picture: 5.564 Mbits x 29.97 (frames/sec) = 167Mbits/sec

- **Digital HDTV:** (ATSC)

1920 (pels) x 1080 (lines) x 1.5 bytes x 30 (frames/sec)=746Mbits

How Compression Possible?

Characteristics of data:

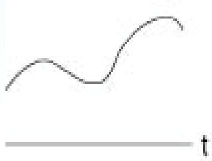
- Stationary statistical model
  -- Shannon information theory
- Non-stationary properties such as local correlation

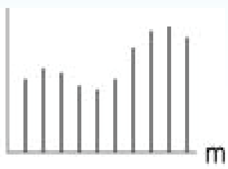
Characteristics of human perception:

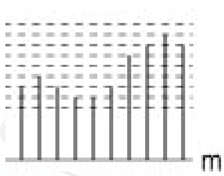
- Finite resolution of hearing and vision
- Auditory masking effect
- Color representation
- Visual masking effect

Signals (Waveforms)

	Time/ Space	Amp.
Analog Signals x(t)	Conti.	Conti.
Discrete-time (discrete-space) (sampled-data) signal x(m)	Discrete	Conti.
Digital signals x(m)	Discrete	Discrete

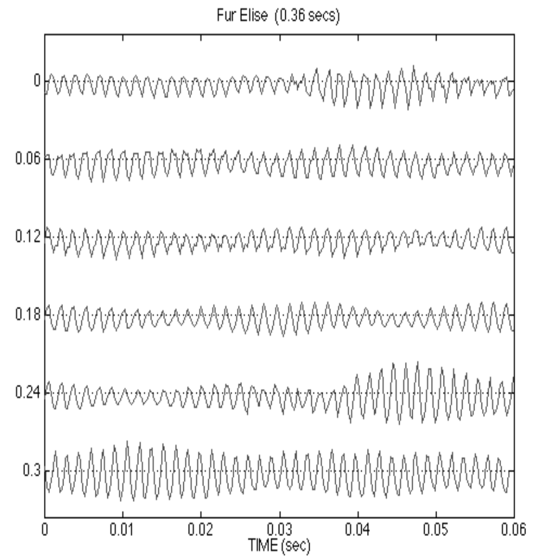


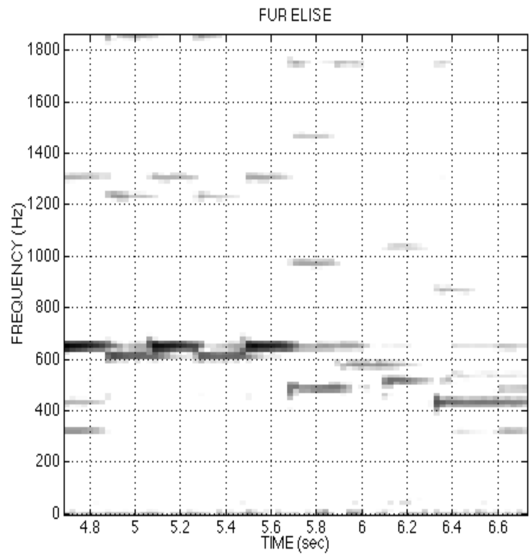




Audio Samples

- Piano (fur Elise) samples and spectrogram (McClellan et al., *DSP First*, Prentice-Hall, 1998)





Data Compression Techniques

Information lossless: Redundancy reduction — The original data can be completely recovered.

- Direct: Huffman codes, arithmetic coding, Ziv-Lempel coding, … (narrow-sense *data compression*)
- Predictive: Run length coding, …

Information lossy: Information (entropy) reduction -- The reproduced data are *approximations* of the original data. This may not be meaningful for a computer file.

- Block coding: vector quantization, transform coding, …
- Sequential: DPCM, tree coding, …
- Multi-resolution (non-block): sub-band, wavelet, ...

Compression Techniques

- **Waveform coding:** Reproduce waveform,

  e.g., DPCM, transform, sub-band, …

  -- Universal but lower efficiency.
- **Content-based coding:** Reproduce *contents*, e.g., (speech) vocoder, (image) contour-texture coding, (video) model-based coding.
- **International standards:** JPEG, MPEG, H.261/3, …
- **Learning-based schemes**: CNN, …

Elements in Waveform Coding

- **Decorrelation:** Reduce spatial and temporal redundancy. *Techniques:* prediction, transform, …
- **Selecting representatives:** Reduce the number of possible signals.

  Techniques: quantization, …
- **Entropy coding:** Equalize the probability distribution of the output symbols.

  Techniques: Huffman codes, Ziv-Lempel coding, ...

Multimedia Coding Standards

- Complete, practical coding algorithms
  -- A balance between (compression) performance and (implementation) complexity (and a compromise among various interest parties)
- Critical for telecommunication products and consumer audio/video media products

Standards Organizations

- **CCITT** – Comité Consultaitif International Télégraphique et Téléphonique (International Telegraph and Telephone Consultative Committee)
- **ITU** – International Telecommunication Union
- **ISO** – International Standardization Organization
- **IEC** – International Electrotechnical Commission
  - **ISO/IEC MPEG, JPEG, AVC, HEVC, VVC**
  - **ITU-T VCEG: H.263, H.264, H.265, H.266**

Image/Video Standards

- **ISO/IEC JTC1 SC29 –** ISO and IEC Joint Technical Committee (on Information Technology) Subcommittee 29 (Coding of audio, picture, multimedia and hypermedia)
  - **Working Group (WG) 1:**
    JBIG (Joint Bi-level Image Group) **–** 1-bit to 4/5-bit still pictures
    JPEG (Joint Photographic Experts Group) **–** 8-bit or more still pictures
- **ISO/IEC JTC1 SC29**
  - **WG 11: MPEG** (Moving Picture Experts Group) **–** Motion pictures
  - **WG 12: MHEG** (Multimedia-Hypermedia Experts Group) **–** Multi/Hyper-media exchange format

Video Coding Standards

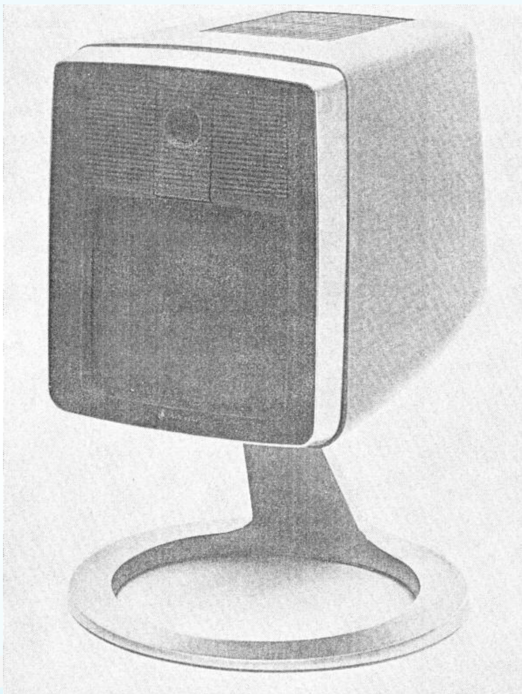
Standards	Typical rates	Applications
ITU-T (CCITT) H.261	128 384k bits/s	Videophone over ISDN
ISO MPEG-1 (11172-2)	1.2 Mbits/s	Video CD
ISO MPEG-2 (13818-2)	4–10 Mbits	Digital TV/HDTV
(ITU-T H.262)	20 Mbits/s	Over air/networks
ITU-T H.263	< 64k bits/s	Videophone
ISO MPEG-4 (14496-2)	Low/high-rates	Object-oriented
ITU-T H.263 v2	< 64k bits/s	PSTN/wireless Videophone
ITU-T/MPEG H.264 (AVC)	< 40k bits/s	Net/wireless Videophone
ITU-T/MPEG H.265 (HEVC)	50% > H.264	High Efficiency Video Coding
ITU-T/MPEG H.266 (VVC)	40% > H.265	Versatile Video Coding

ISDN: Integrated Services Digital Network

AT&T Picturephone@

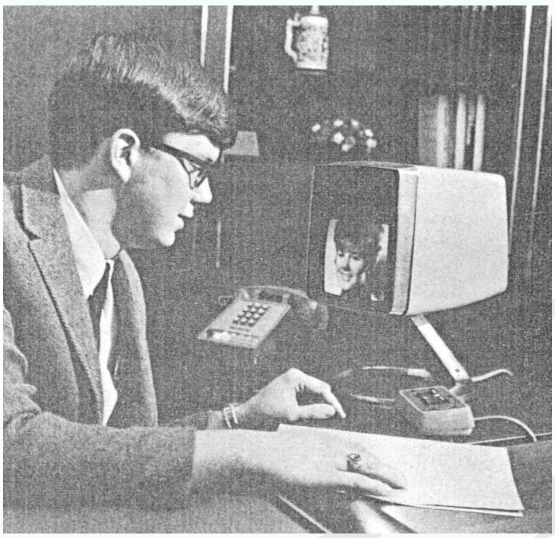
- *1984, Visual Communications Dept. of AT&T Bell Labs,*

  At a corner of a lab. shelf …



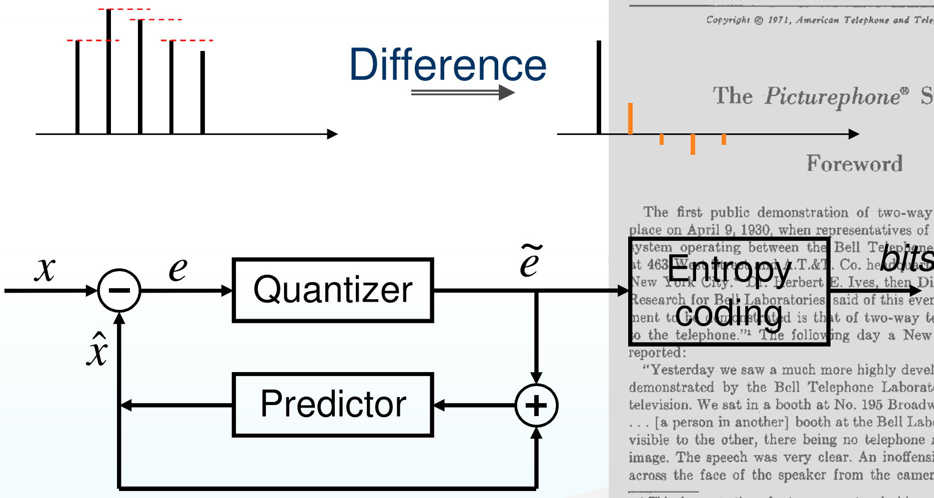
AT&T Picturephone (2)

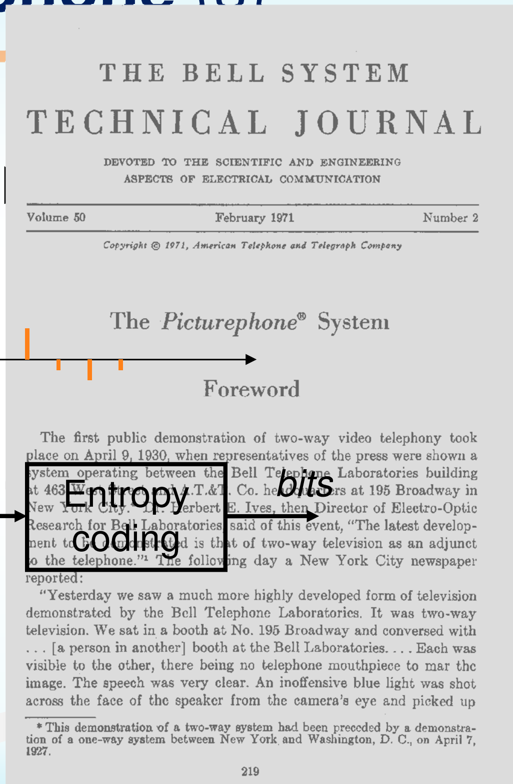
- "Mod II" was developed and field-tested by AT&T Bell Labs around 1966-1969. Commercial service July 1, 1970. (*BSTJ*, Feb. 71)
- Digital: 275 pels x 250 lines; 6.312 Mb/s (T2 line)



AT&T Picturephone (3)

- Image Compression Tech
(Differential Pulse Coded





THE BELL SYSTEM
TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS OF ELECTRICAL COMMUNICATION

Volume 50 February 1971 Number 2

Copyright © 1971, American Telephone and Telegraph Company

The *Picturephone*® System

Foreword

The first public demonstration of two-way video telephony took place on April 9, 1930, when representatives of the press were shown a system operating between the Bell Telephone Laboratories building at 463 [illegible] A.T.&T. Co. [illegible] at 195 Broadway in New York City.* [illegible] Herbert E. Ives, then Director of Electro-Optic Research for Bell Laboratories, said of this event, "The latest development to [illegible] is that of two-way television as an adjunct to the telephone."[1] The following day a New York City newspaper reported:

"Yesterday we saw a much more highly developed form of television demonstrated by the Bell Telephone Laboratories. It was two-way television. We sat in a booth at No. 195 Broadway and conversed with . . . [a person in another] booth at the Bell Laboratories. . . . Each was visible to the other, there being no telephone mouthpiece to mar the image. The speech was very clear. An inoffensive blue light was shot across the face of the speaker from the camera's eye and picked up

* This demonstration of a two-way system had been preceded by a demonstration of a one-way system between New York and Washington, D. C., on April 7, 1927.

219

MPEG Committee

- Convener: Leonardo Chiariglione (resigned 2020.6)
- Standards:

-- MPEG-1: done

-- MPEG-2: done

-- MPEG-4: done

-- MPEG-7: done

-- MPEG-21: done

-- MPEG A,B,C,D,E: on-going

MPEG-2: *1996 Emmy for Technical Excellence*

AVC: 2008 *ATAS Primetime Emmy Engineering Award*

2009 *Paired NATAS Tech & Eng Emmy Award*

ISO/IEC 11172 MPEG-1

MPEG-1 1992 Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s

Part 1 – MPEG-1 Systems
Part 2 – MPEG-1 Video for CD
Part 3 – MPEG-1 Audio (Layers I, II, and III)
Part 4 – Conformance
Part 5 – Software

ISO/IEC 13818 MPEG-2

MPEG-2 1994 Generic coding of moving pictures and associated audio information

1996 Emmy for technical excellence

Part 1 Systems
Part 2 Video
Part 3 Audio
Part 4 Conformance
Part 5 Technical Report
Part 6 DSM CC - Digital Storage Media Cmd & Cntl
Part 7 AAC - Advanced Audio Coding
Part 9 RTI - Real Time Interface
Part 10 Conformance Extensions
Part 11 IPMP on MPEG-2 Systems

ISO/IEC 14496 MPEG-4

MPEG-4 1998 Coding of audio-visual objects

Part 1 Systems
Part 2 Visual
Part 3 Audio
Part 4 Conformance
Part 5 Reference Software
Part 6 Delivery Multimedia Integration Framework (DMIF)
Part 7 Optimized Software
Part 8 MPEG 4 on IP
Part 9 Reference Hardware
Part 10 Advanced Video Coding (AVC) (JVT, H.264)
Part 11 Scene Description and Application Engine
Part 12 ISO Base Media File Format
Part 13 IPMP Extensions
Part 14 MP4 File Format
Part 15 AVC File Format
Part 16 Multimedia Animation Framework eXtension (AFX)
Part 17 Streaming Text Format
Part 18 Font Compression and Streaming
Part 19 Synthesized Streams
Part 20 Lightweight Application Scene Representation
Part 21 MPEG-J Extension for rendering
Parts 22 -- 31

ISO/IEC 15938 MPEG-7

MPEG-7 2001 Multimedia content description interface

Part 1 Systems
Part 2 DDL - Description definition language
Part 3 Visual Part 4 Audio
Part 5 Multimedia description schemes
Part 6 Reference software
Part 7 Conformance testing
Part 8 Extraction and use of description
Part 9 MPEG-7 Profiles
Part 10 Schema Definition
Part 11 Profile Schemas Part 12 Query format
Part 13 **Compact Descriptors for Visual Search (CDVS)**
Part 14 Reference Software for CDVS

ISO/IEC 21000 MPEG-21

Part 1 Vision, Technologies and Strategy
Part 2 Digital Item Declaration (DID)
Part 3 Digital Item Identification (DII)
Part 4 Intellectual Property Management and Protection (IPMP)
Part 5 Rights Expression Language (REL)
Part 6 Rights Data Dictionary (RDD)
Part 7 Digital Item Adaptation (DIA)
Part 8 Reference Software
Part 9 File Format
Part 10 Digital Item Processing
Part 11 Persistent Association
Part 12 Multimedia Test Bed Resource Delivery
Part 15 Event Reporting
Part 16 Binary Format
Part 17 Fragment Identification for MPEG Media Types
Part 18 Digital Item Streaming
Part 19 Media Value Chain Ontology
Part 20 Contract Expression Language
Part 21 Media Contract Ontology

MPEG-A,B,C,…

- **MPEG-A** (ISO/IEC 23000) Multimedia Application Formats
- **MPEG-B** (ISO/IEC 23001) MPEG Systems Technologies
- **MPEG-C** (ISO/IEC 23002) MPEG Video Technologies
- **MPEG-D** (ISO/IEC 23003) MPEG Audio Technologies
- **MPEG-E** (ISO/IEC 23004) Multimedia Middleware
- **MPEG-G** (ISO/IEC 23092) Genomic Information Representation
- **MPEG-H** (ISO/IEC 23008) **High Efficiency Video Coding**
- **MPEG-I** (ISO/IEC 23090) Part 3 - **Versatile Video Coding**
- **MPEG-M** (ISO/IEC 23006) MPEG Extensible Middleware
- **MPEG-U** (ISO/IEC 23007) Rich-Media User Interface
- **MPEG-V** (ISO/IEC 23005) Media Context and Control
- **MPEG-DASH** (ISO/IEC 23009) Dynamic adaptive streaming over HTTP

MPEG Meetings

- 4 (3) meetings a year; 5(→10) days
- ~300 participants (→ 500!)
- Over 200 companies
- Meetings are divided into groups (~2010)

ITU-T VCEG Committee

- **Gary J. Sullivan** (Microsoft) Chair/co-chair: VCEG, JVT, and JCT-VC
- **H.263, H.264, H.265, H.266**

◆ MPEG/H.264 AVC received ATAS Primetime Emmy Engineering Award (Aug. 2008) & Paired NATAS Tech & Eng Emmy Awards (Jan 2009)

100th MPEG Meeting

- *Brief History*
  - 1st meeting: **May 1988** – Hiroshi Yasuda & Leonardo Chiariglione (Ottawa, Canada)
  - …25 years …
  - 100th meeting: **April 2012** -- Leonardo Chiariglione (Geneva, Switzerland)
- *Proc. IEEE*, April 2012

INVITED PAPER

Multimedia Standards: Interfaces to Innovation

A history of the Motion Picture Experts Group is provided and its probable future activities are discussed, including understanding 3-D audio–video, machine design, and creating best practices and models.

By Leonardo Chiariglione

ABSTRACT | Standardization is concerned with interfaces; industry is concerned with systems. This papers brings the evidence brought by the MPEG standardization group to show how, through the proper management of interface evolution, the constituent industries have been able to achieve product and service interoperability, room for differentiation, and opportunities for innovation in the context of the tectonics shift also known as convergence.

KEYWORDS | Compression; convergence; digital media; standards

same "standard" if C is ever to be able to have bolts that screw into the nut. In the case of nuts and bolts, the standard is the "thread" which represents the interface between the nut and the bolt.

Standards play a fundamental role in enabling a diversified industry. Once a standard has been published—and the industry has adopted it—independent manufacturers can build products conforming to it that can immediately reach a potentially global market. Users can choose products that are more convenient for their needs from different suppliers.

There is also a prevailing view that standards choke

The MPEG Process

1. **Exploration**
   Search for new technology
2. **Requirements**
   Establish work scope
   Call for Proposals
3. **Competitive phase**
   Do Homework
   Response to CfP
   Initial technology selection
4. **Collaborative phase**
   Core Experiments
   Working Drafts
5. **Standardization**
   Ballots
   National Body Comments
6. **Amendment**
   Adding new technology
7. **Corrigenda**
   Corrective actions
8. **New subdivisions**
   Add new non-compatible technology

Stages of Standardization

WD	Working Draft
CD	Committee Draft
FCD	Final Committee Draft
FDIS©	Final Draft International Standard
IS©	International Standard
PDAM	Proposed Draft Amendment
FPDAM	Final Proposed Draft Amendment
FDAM©	Final Draft Amendment
AMD©	Amendment

How I Got Involved?

- 1984: Joined AT&T Bell Labs – Visual Comm. Dept.
  → H.261 video standard started
- 1988.1: MPEG started
- 1991.12: I joined NCTU ← discontinued standard activities
- 1999.9: NCTU formed a small group to participate in the MPEG activities

NCTU MPEG Activity

- Tihao Chiang (蔣迪豪), C.J. Tsai (蔡淳仁), Wen Peng (彭文孝), H.-M. Hang (杭學鳴) and Chris Lee (李國君), NCKU
- Tihao Chiang : Co-editor, MPEG-4 Part 7 **Optimised Reference Software** (Done)
- C.J. Tsai : Co-editor, MPEG-21 Part 12 **Multimedia Test Bed for Resource Delivery** (Done)
- 100+ contributions (input and output documents) in the past 8 years. [Dr. Y.-S. Tung (童怡新); **ITRI**]
- ***Examples:*** **1) Call for Proposal on Scalable Video Coding** (2004.2) – 2 out of 14 proposals

  2) Call for Proposal on HEVC (2010.2) – one out of 27

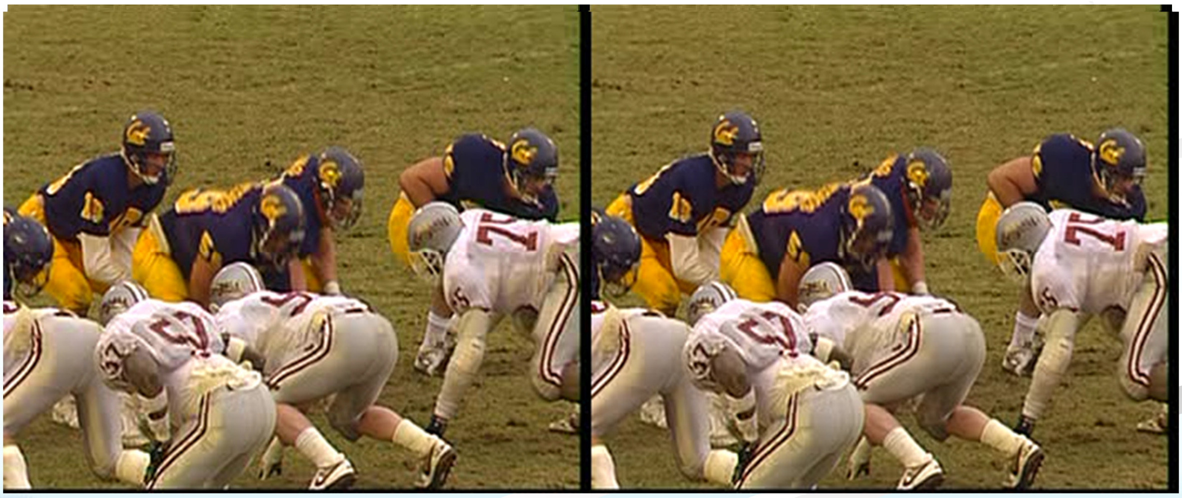
  3) Call for Proposal on SCC (2014.4) – one out of 7

MPEG Chair Dr. Chiariglione at NCTU (2003.12)

- http://www.chiariglione.org

Demo: AVC vs MPEG-2

- Test sequence: Football 8 secs
  - Resolution: 352 x 240
  - Frame rate: 30 frm/sec



Demo: AVC vs MPEG-2

- Test sequence: Football 8 secs
  - Resolution: 352 x 240
  - Frame rate: 30 frm/sec

- Left: MPEG2 – 2 Mbps
  - PSNR: 34.5 dB (Y)
- Right: AVC – 1 Mpbs (50%)
  - PSNR: 34.1 dB (Y)

Demo: HEVC vs AVC

- Test sequence: BQ Terrace (Class B) -- LD
  - Resolution: 1920x1080p, 60Hz (512x500 shown)

Demo: HEVC vs AVC

- Test sequence: BQ Terrace (Class B) -- LD
  - Resolution: 1920x1080p, 60Hz (512x500 shown)

- Left: AVC JM16 – 4.11 Mbps
  - PSNR: 33.20 dB (Y)
- Right: HEVC HM3.0 – 1.93 Mbps **(-53%)**
  - PSNR: 33.14 dB (Y)

References

(1) K. Sayood, *Intro. to Data Compression*, 5th ed., Morgan Kaufman, 2017.

(2) J.-R. Ohm, Multimedia Signal Coding and Transmission, Springer, 2015.

(3) Y. Q. Shi and H. Sun, *Image and Video Compression for Multimedia Engineering*, 2nd Ed., CRC Press, 2008.

(4) J.W. Woods, *Multidimensional Signal, Image, and Video Processing and Coding*, 2nd Ed., Elsevier, 2011.

(5) A. Puri & T. Chen, ed., *Multimedia Systems, Standards, and Networks,* Marcel Dekker, 2000.

(6) H.-M. Hang and J.W. Woods ed., *Handbook for Visual Communications*, Academic Press, 1995

(7) A.M. Tekalp, *Digital Video Processing*, 2nd ed, Prentice-Hall, 2015.

(8) Y. Wang et al., *Video Processing and Communications*, Prentice-Hall, 2002.

(9) W.B. Pennebaker and J.L. Mitchell, *JPEG: Still Image Data Compression Standard*, Van Nostrand Reinhold, 1993.

(10) D.S. Taubman and M.W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwer Academic, 2002.

(11) B.G. Haskell, A. Puri, and A.N. Netravali, *Digital Video: An Introduction to MPEG-2*, International Thomson, 1997.

(12) F. Pereira & T. Ebrahimi, *The MPEG-4 Book*, Prentice-Hall, 2002.

(13) B.S. Manjunath et al., ed., *Intr. to MPEG-7*, John Wiley & Sons, 2002.

(14) I.E.G. Richardson, *H.264 and MPEG-4 Video Compression*, John Wiley & Sons, 2003; *The H.264 Advanced Video Compression Standard*, 2nd ed., Wiley, 2010.

(15) K.R. Rao and J.J. Hwang, *Techniques and Standards for Image, Video, and Audio Coding*, Prentice-Hall, 1996.

(16) K.R. Rao, J.J. Hwang, and D. N. Kim, *High Efficiency Video Coding and Other Emerging Standards*, River Pub., Sept 2017.

(17) I.S. Burnett, *The MPEG-21 Book*, John Wiley & Sons, 2006.

(18) 戴顯權,資料壓縮,旗標出版社, 2007. (2nd, 紳藍出版社, 2002)

(19) 戴顯權等,多媒體通訊,滄海書局, 2008. (2nd, 紳藍出版社, 2003)

(20) 戴顯權等, JPEG2000, 紳藍出版社, 2002.

(21) 吳炳飛等, JPEG2000影像壓縮技術,全華出版社, 2003