Source Coding

杭 學 鳴

國 立 交 通 大 學

電 子 工 程 學 系

Tel: (03)5731861

Email: hmhang@mail.nctu.edu.tw

Chapter 1 Introduction

- Why need compression
- How compression possible
- Principles of Compression Techniques
- Multimedia Standards
- Subjects to be discussed

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

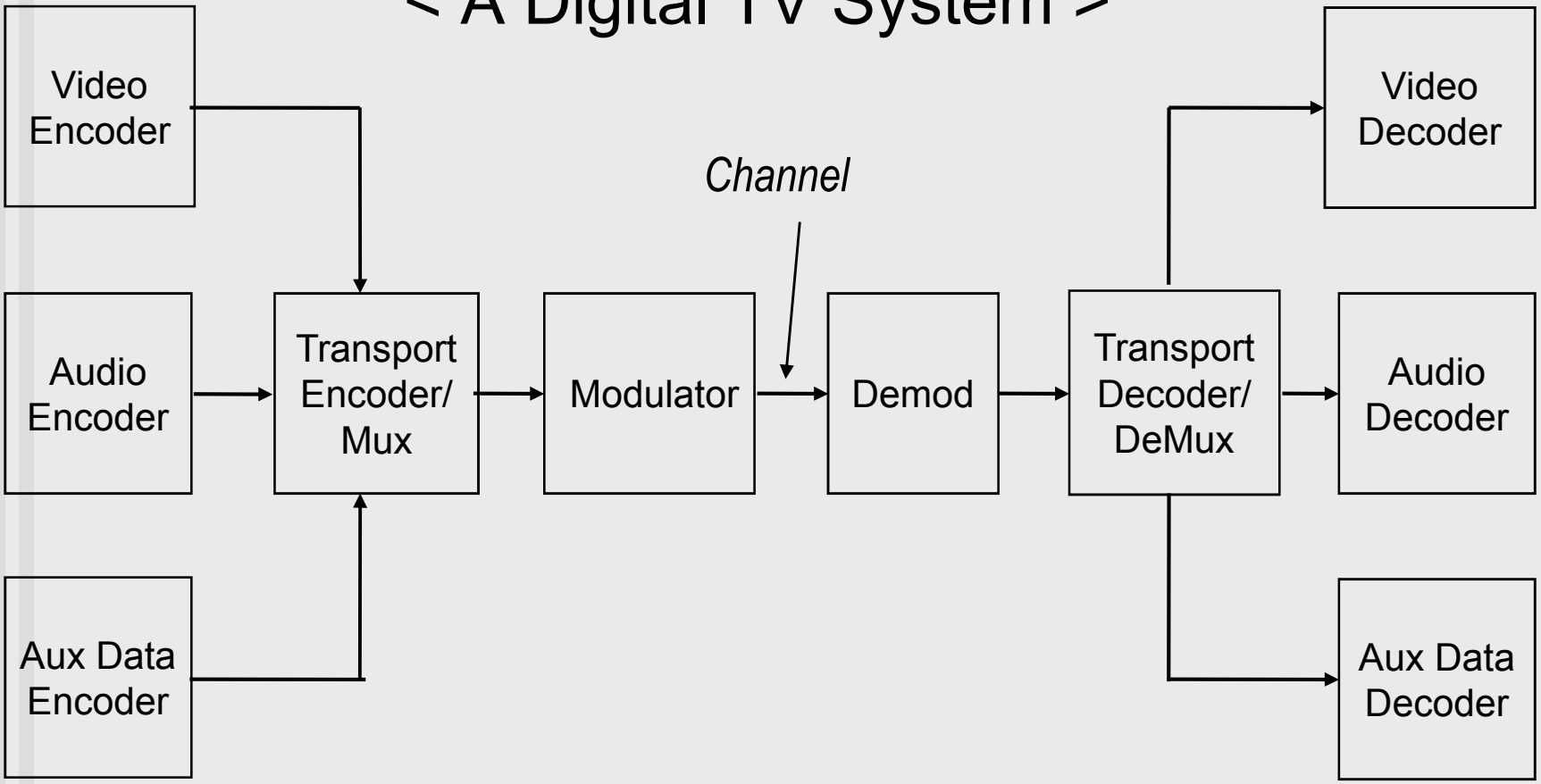
Motion picture: 5.564 Mbits x 29.97 (frames/sec) = 167Mbits/sec

- **Digital HDTV:** (ATSC)

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

An Example of Comm. Systems

< A Digital TV System >



Examples of Data

-- Classified based upon compression techniques

- **General digital data:** various types of computer files
- **Text**
- **Speech:** human conversation
- **Audio:** music
- **Bi-level images:** fax, dithered (two-tone) images
- **Still gray-level pictures:** graphics, natural images
- **Motion pictures:** video conferencing, television.

-- Data characteristics can be used to reduce bit rate.

How Compression Possible?

Characteristics of data:

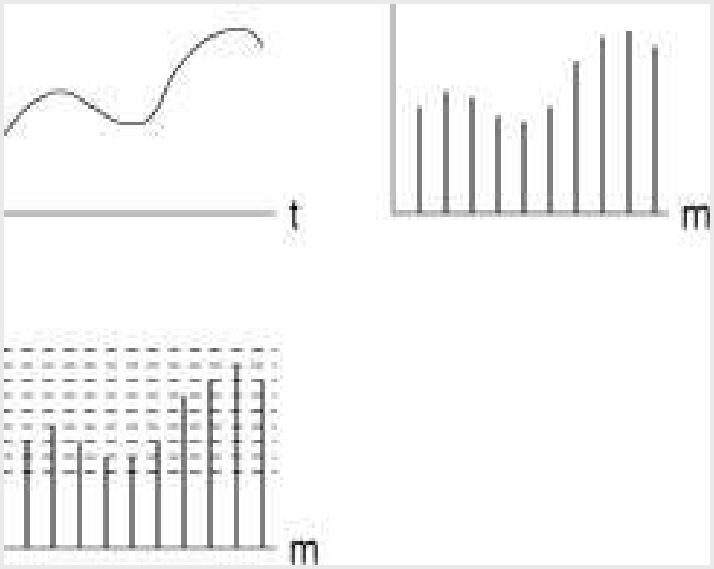
- Stationary statistical model
  -- Shannon information theory
- Non-stationary properties such as local correlation

Characteristics of human perception:

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

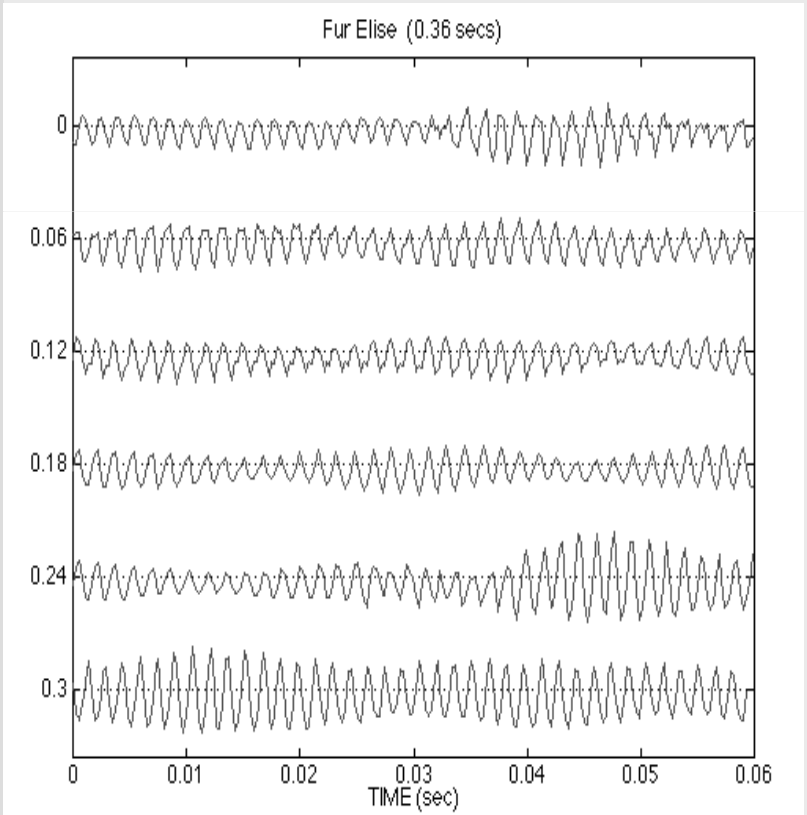
Signals (Waveforms)

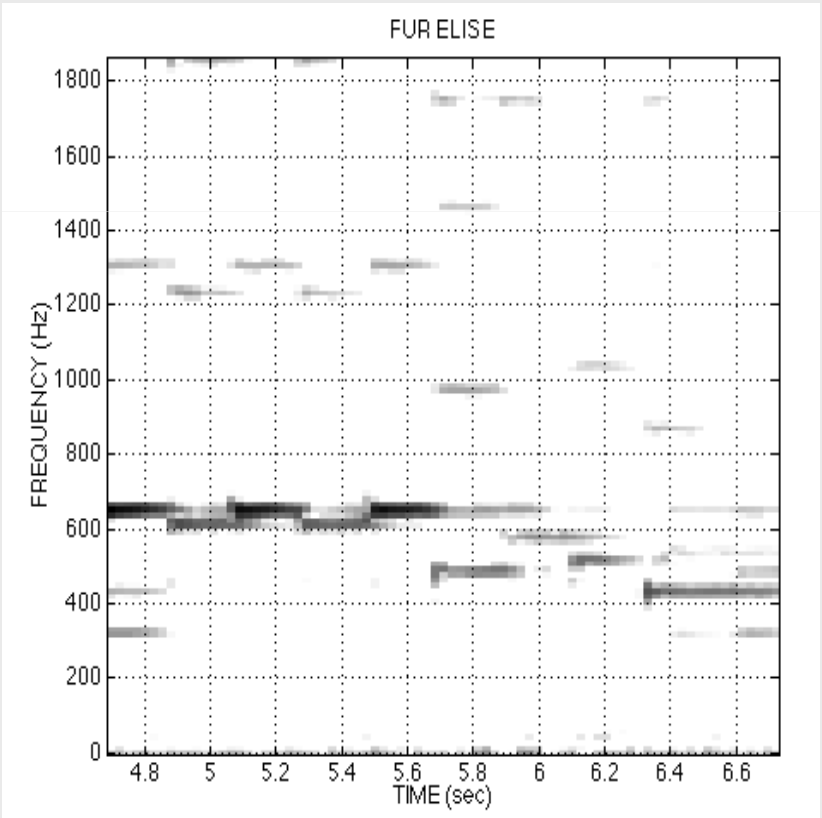
	Time/ Space	Amp.
Analog Signals x(t)	Conti.	Conti.
Discrete-time (discrete-space) (sampled-data) signal x(m)	Discrete	Conti.
Digital signals x(m)	Discrete	Discrete



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, …

Compression Algorithm Development

- **Phase 1: Modeling.** Describe redundancy in the form of a model. E.g., prob. distribution, Markov model, …

  --*Techniques* are then invented to remove redundancy; e.g.,: predictor, transform, …

- **Phase 2: Coding**. Representations of "model" and "residual" (unpredictable) information.

  --*Techniques:* e.g., quantization and entropy coding of the prediction errors, …

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
- **ITU** – International Telecommunication Union
- **ISO** – International Standardization Organization
- **IEC** – International Electrotechnical Commission

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

Speech Coding Standards

Standards	Typical rates (Year)	Quality: MOS (1-5)
PCM	64 kbits/s (1972)	4.4 (PSTN)
G.721 ADPCM	32 kbits/s (1984)	4.1 (PSTN)
GSM	13 kbits/s (1991)	3.6 (Cellular)
G.728 (low delay)	16 kbits/s (1992)	4.0
IS-96A (CDMA)	0.8-8.55 kbits/s (1993)	~3.4 (Cellular)
G.729	8 kbits/s (1995)	~4.2 (VoIP)
G.723.1	5.3, 6.3 kbits/s (1995)	~4.0
Half-rate GSM	5.6 kbits/s (1995)	~3.4 (Cellular)
AMR	5.15-12.2 kbits/s (1999)	~3.9 (3GPP)

MOS: Mean Opinion Score -- 5=excellent, 4=good, 3=fair, 2=poor, 1=bad

MPEG Audio Standards

MPEG-1 Layer 1: 1992	(good: 256k /2ch)	1-2 chs
MPEG-1 Layer 2: 1992	(good: 192k /2ch)	1-2 chs
MPEG-1 Layer 3: 1993 (MP3)	(good: 128k /2ch)	1-2 chs
MPEG-2 Layers 1,2,3: 1994		1-5.1 chs
MPEG-2 AAC: 1997 (Advanced Video Coding)	(good: 96k /2ch)	1-96 chs
MPEG-4 (v1) subpart 3 General Audio Coding, AAC: 1999	(new tools: PNS, LTP, TwinVQ)	1-96 chs

MPEG-4 Amd 1: (2003) Bandwidth extension (SBR -- Spectral Band Replication)	HE-AAC, AAC+ (good: 48k?)
MPEG-4 Amd 2: (2004) Parametric Audio extension → MPEG surround	(good: 24k?)

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
ISO MPEG-7 (15938)	Database	Content description
ITU-T H.263 v2	< 64k bits/s	PSTN/wireless Videophone
ITU-T H.264 (JVT,AVC)	< 40k bits/s	Net/wireless Videophone
ITU-T H.264 ext (SVC)	Multi-layer	Net/wireless streaming

ISDN: Integrated Services Digital Network

Progress of Image/Video Coding

H.261 (CCITT/ITU;1984, 88, 90) – *video (videoconf.)*

→ **JPEG** (1986, 89, 92) – *image (**Digital Camera**)*

→ MPEG-1 (1988 – 92) – *video (VCD)*

→ **MPEG-2** (1990 – 94) – *video (**DVD, DTV**)*

→ MPEG-4 part 2 (1992 – 99) – *video (Internet, WL)*

→ H.263 (1993 – 95; ver.3: 2000) – *video (WL)*

→ JPEG2000 (1996 – 2001) – *image*

→ **H.264** (**MPEG-4 part 10**) AVC (1998 – 03) – *video (**WL, HD-DVD**)*

→ AVC Amd.1 (2003 –) – ***scalable video coding** (?)*

Demo: AVC vs MPEG-2

- Test sequence: 61.9 secs (1857 frames)
  - Resolution: 432 x 240
  - Frame rate: 30 frm/sec

Demo: AVC vs MPEG-2

- Test sequence: 61.9 secs (1857 frames)
  - Resolution: 432 x 240
  - Frame rate: 30 frm/sec

- Left: AVC (JM2.1) – 606 kbps
  - SNR: 45 dB (Y)
- Right: MPEG2 – 2,000 kpbs
  - SNR: 44.94 dB (Y)

AT&T Picturephone@

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

  At a corner of a lab. shelf …

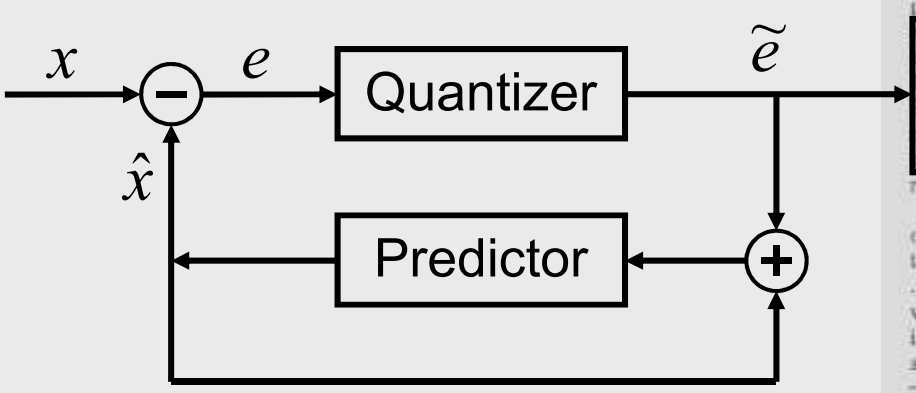
AT&T Picturephone (2)

- "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 Technique: **DPCM** (Differential Pulse Coded Modulation)

Difference



Entropy coding → *bits*

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ... 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 ... T&T. Co. ... at 195 Broadway in New York City. ... Herbert E. Ives, then Director of Electro-Optic Research for Bell Laboratories, said of this event, "The latest development ... is that of two-way television as an adjunct to the telephone."* 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

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

MPEG Committee

- Convener: Leonardo Chiariglione

- Standards:
  - -- MPEG-1: done
  - -- MPEG-2: done
  - -- MPEG-4: very much done
  - -- MPEG-7: very much done
  - -- MPEG-21: on-going

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
Part 6 – Specification for IDCT implementation

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 and 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
Part 22 Open Font Format Specification

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

ISO/IEC 21000 MPEG-21

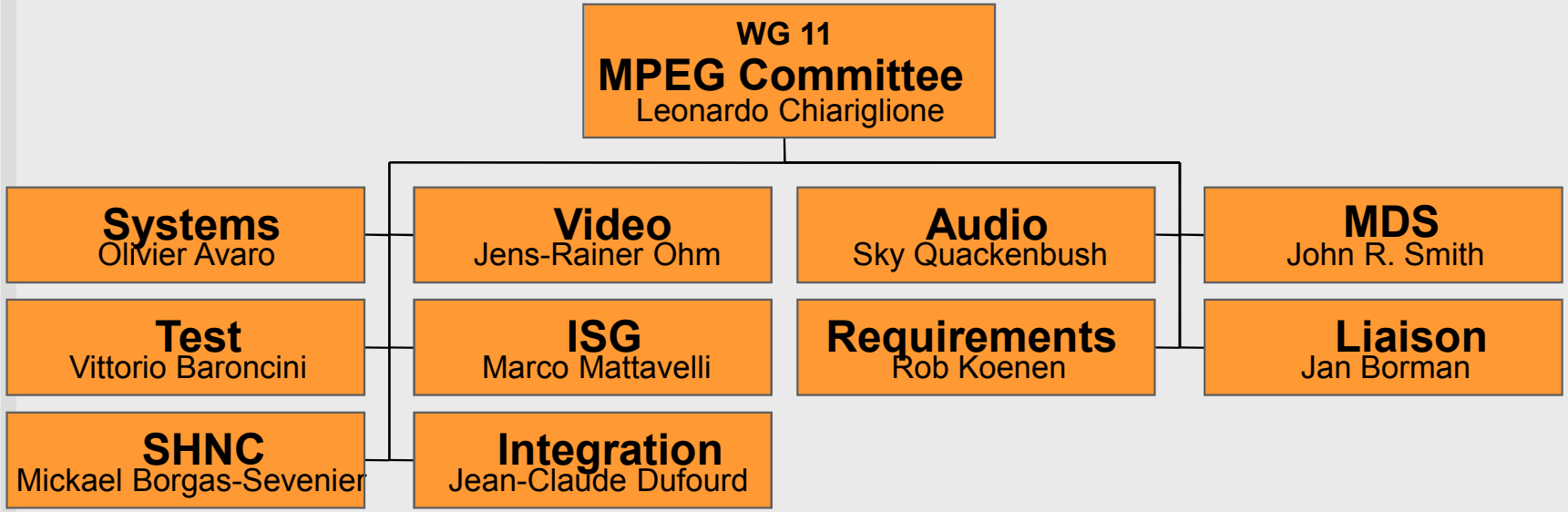
Part 1 Vision, Technologies and Strategy
Part 2 Digital Item Declaration (DID)
Part 3 Digital Item Identification (DII)
Part 4 Intellectual Property Management and Protection (IPMP)
Part 5 Rights Expression Language (REL)
Part 6 Rights Data Dictionary (RDD)
Part 7 Digital Item Adaptation (DIA)
Part 8 Reference Software
Part 9 File Format
Part 10 Digital Item Processing
Part 11 Persistent Association
Part 12 Multimedia Test Bed Resource Delivery
Part 14 Conformance Testing
Part 16 Binary Format
Part 17 Fragment Identification for MPEG Media Types
Part 18 MPEG -21 Schema files

MPEG Chair Dr. Chiariglione at NCTU (2003.12)

http://www.chiariglione.org

MPEG Meetings

- 4 meetings a year; 5+ days per meeting
- ~300 participants
- Over 200 companies
- Meetings are divided into groups



NCTU MPEG Activity

- Tihao Chiang (蔣迪豪), C.J. Tsai (蔡淳仁), and H.-M. Hang (杭學鳴) attend MPEG meetings constantly
- 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)
- 95 contributions (input and output documents) in the past 4 years (2002 -- 2005). [Dr. Y.-S. Tung, NTU]
- ***Example:*** **Call for Proposal on Scalable Video Coding** (Feb. 2004) – 2 out of 14 proposals

Subjects To Be Discussed

- **Basic theory:** Shannon theory (outline)
- **Lossless coding algorithms:** Huffman, arithmetic, Ziv-Lempel, run-length.
- **Lossy coding algorithms:** (vector) quantization, differential coding, Transform coding, subband coding, wavelet-coding, motion estimation/compensation.
- **Standards**: G.723.1, MPEG-1/2 audio, JPEG, JPEG2000, H.261/3 basic, MPEG-1/2/4 video.