Introduction to the MPEG-4 Standard

Rob Koenen

InterTrust Technologies Corp.

Chair MPEG Requirements Group

President MPEG-4 Industry Forum

NAB 2001, Las Vegas

21 April 2001

What to expect

- Overview of MPEG-4 and its design principles

- A bit of its history

- No fine coder details

MPEG: Moving Picture Experts Group
Coding of Moving Video and Audio

- MPEG-1:   CD-i, (VoD trials), ... – 1992

- MPEG-2:   ... + TV, HDTV – 1994

- MPEG-3:   *HDTV → merged into MPEG-2*

- MPEG-4:   Coding of Audiovisual Objects – V.1: 1998; V.2: 1999
  Extensions ongoing

- MPEG-7:   MM Description Interface – Fall 2001
  'Describing' audiovisual material

- MPEG-21: Digital Multimedia Framework – 1st parts early 2002
  'The Big Picture and The Glue'

MPEG: Moving Picture Experts Group
## *Coding of Moving Video and Audio*

- MPEG-1:   CD-i, (VoD trials), ... – 1992

- MPEG-2:   ... + TV, HDTV – 1994

- MPEG-3:   *HDTV → merged into MPEG-2*

- MPEG-4:   Coding of Audiovisual Objects – V.1: 1998; V.2: 1999
            Extensions ongoing

- MPEG-7:   MM Description Interface – Fall 2001
            'Describing' audiovisual material

- MPEG-21: Digital Multimedia Framework – 1st parts early 2002
            'The Big Picture and The Glue'

A bit of history and background on MPEG-4

- 1993: started as 'Very Low Bitrate Audiovisual Coding'

- 1994: goal changed to 'Coding of Audiovisual Objects'

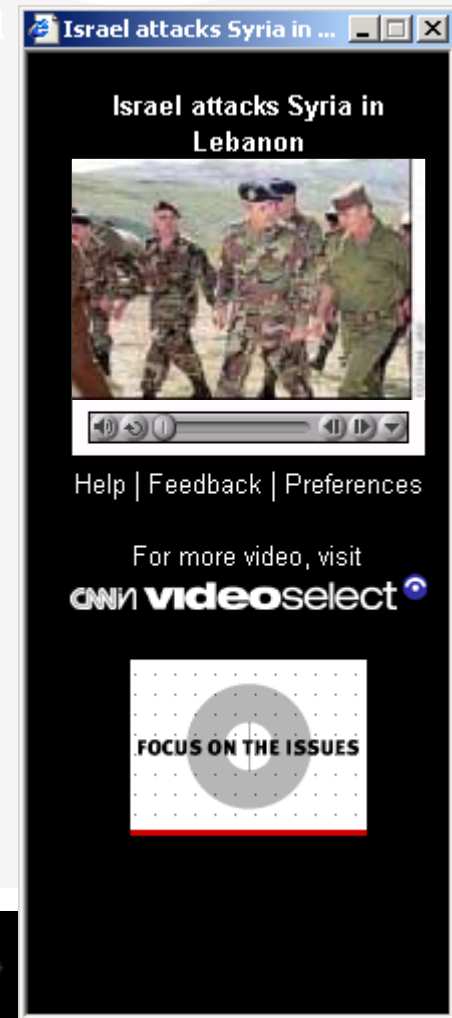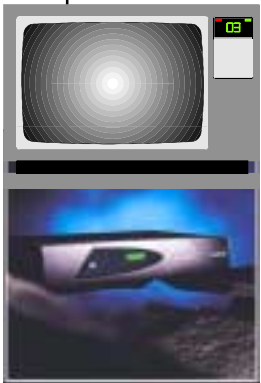- Coding work is now ready; some extensions in Systems

From the MPEG-4 Call for Proposals (drafted 1994)
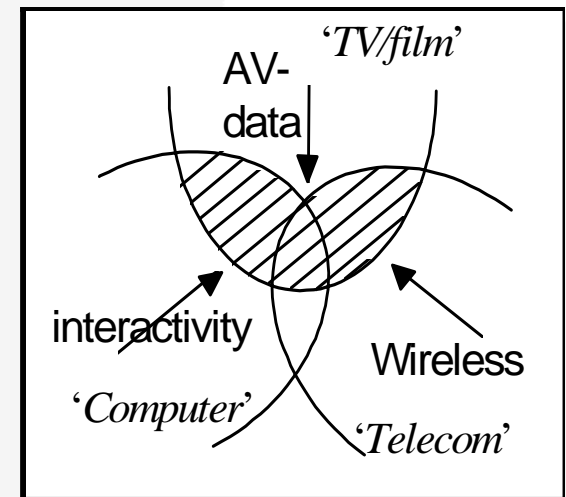
Vision from 1994

- Convergence is a hype. There will not be a single network or terminal. Rather, we will see a proliferation of multimedia services over different (access) networks, terminals.

- Therefore, we need is a common multimedia technology that supports the three main service paradigms:

    – **Broadcast**

    – **Communication**

    – **Retrieval**

Israel attacks Syria in ...

Israel attacks Syria in Lebanon

Help | Feedback | Preferences

For more video, visit

CNN videoselect

FOCUS ON THE ISSUES

The most important objectives

- Common technology for many types of services:

  interactive, broadcast, conversational

- Allowing more & different interactivity - not just stop/play/slow, but interactivity involving elements within the 'scene'

- Integrating natural and synthetic content

- Covering a wide range of access conditions

  – Includes low bitrates, error resilience, scalable coding

- Helping manage and protect Intellectual Property
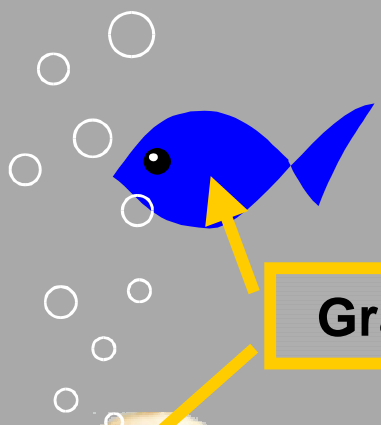
The Nature of MPEG-4

- MPEG-4 is an object based multimedia content representation standard

- Many innovations come from the Systems Layer

- The coders are pretty good, too

- It's got Intellectual Property Management and Protection ('IPMP') hooks deeply built into it

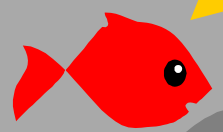# An MPEG-4 Scene (free of IP)

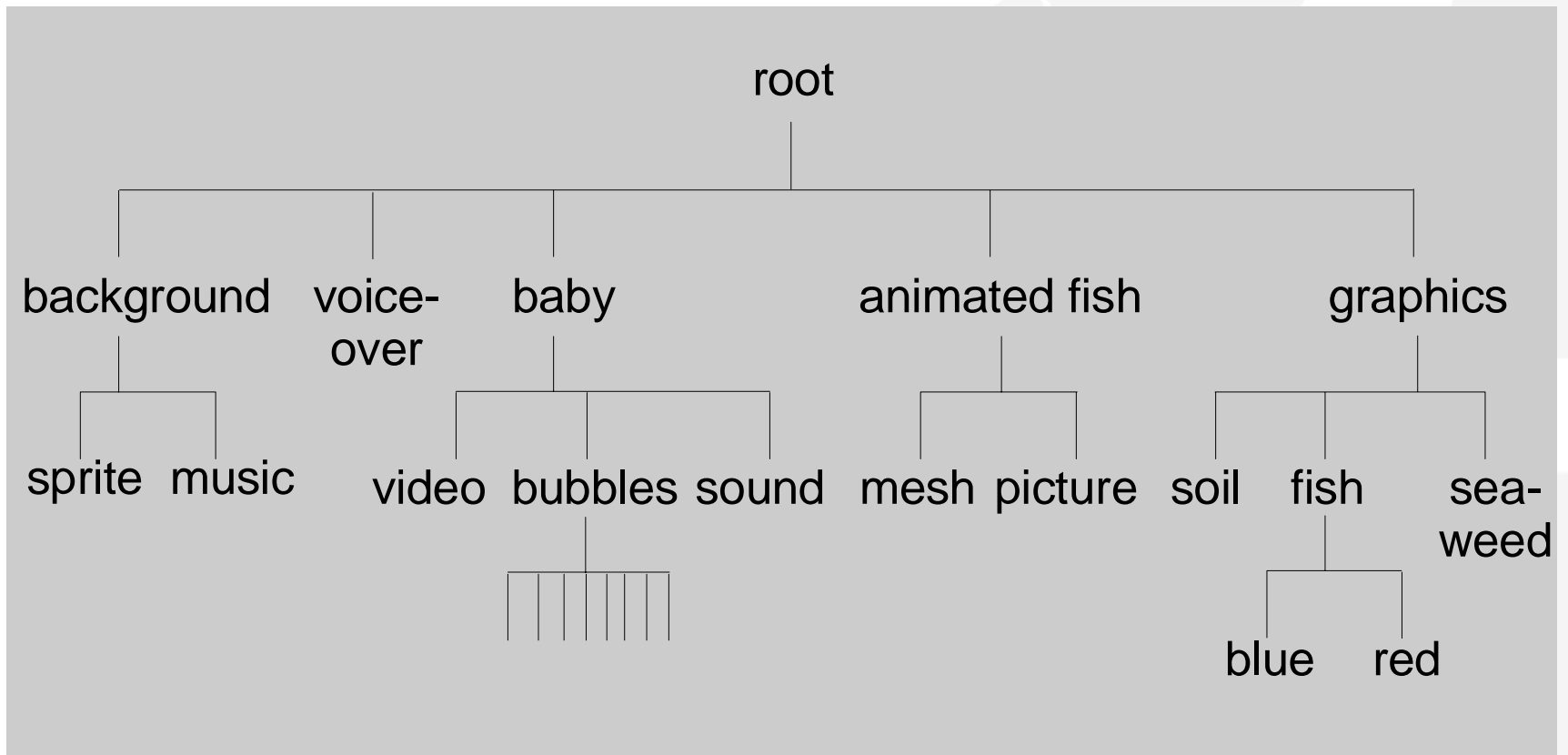**Animated 2D photo / mesh**
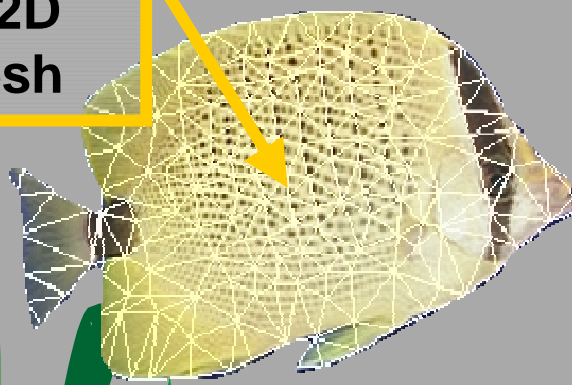
Graphic objects

Background with music

'Natural' audiovisual object

The Scene Tree

MPEG-4: Coding of Audiovisual Objects

- Audiovisual **Scene** is composed of '**Objects**' (A&V)

- '**Compositor**' puts objects in scene (A&V, 2&3D)

- Objects can be of different nature
  - natural or synthetic A&V, text & graphics, animated faces, arbitrary shape or rectangular

- **Coding** scheme can differ for individual objects

- Principle is **independent** of bitrate!
  - from low bitrates to (virtually) lossless quality

The Parts of the MPEG-4 Standard

1. Systems — presentation

2. Visual

3. Audio — decoding

1. Systems — demux & buffer

6. DMIF — transport interface

Transport layer — not in standard (well …)

7. MPEG-4 on IP

4. Conformance

5. Reference SW

INTERTRUST®
The MetaTrust Utility™

MPEG
ISO/IEC JTC1/ SC29 WG11
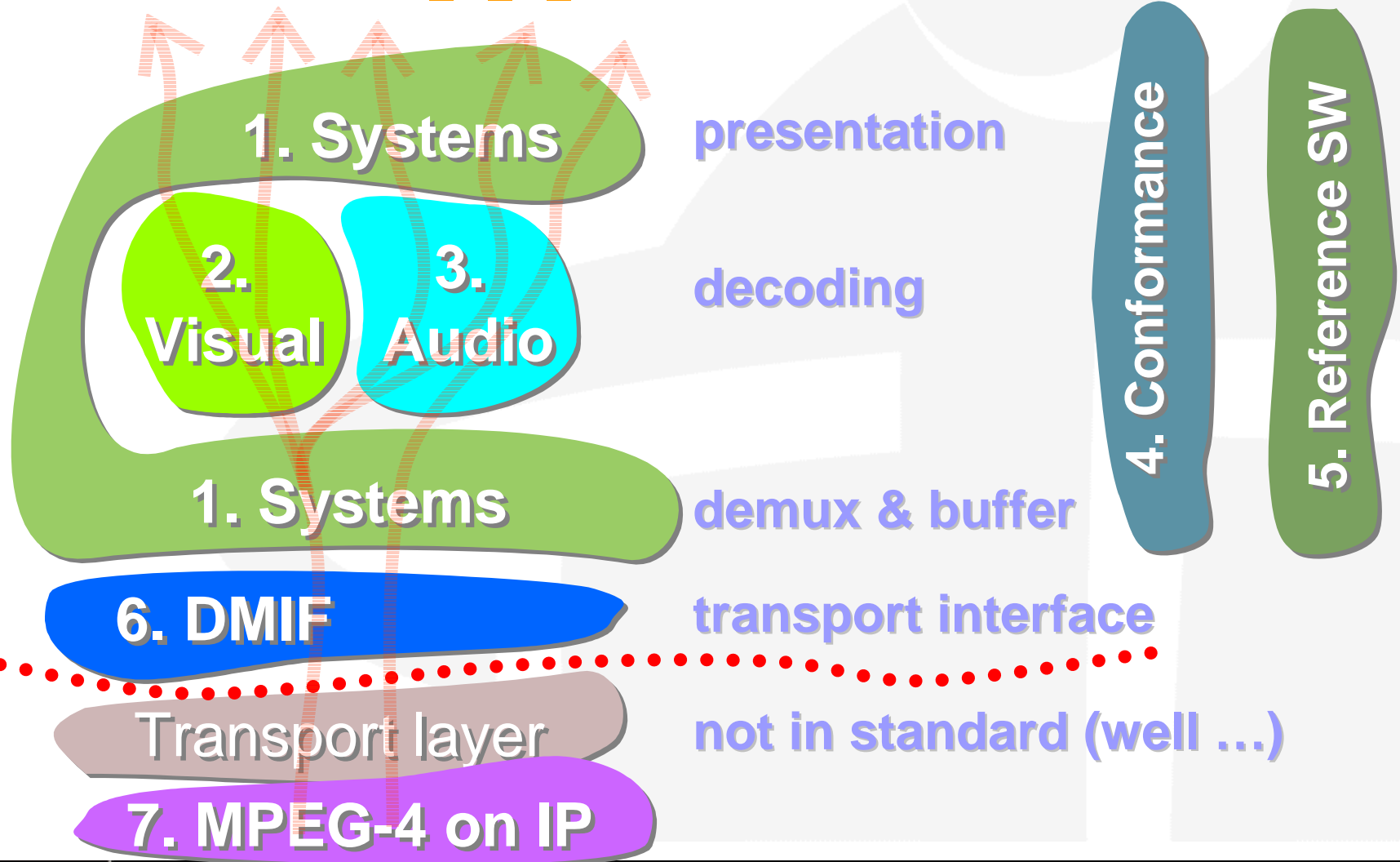
Visual Media Object Types in MPEG-4

- Video from 10 – 1,000,000 Kbit/s
  - Multiple rectangular or arbitrary shape objects in the scene
  - Scalable – Interlaced and Progressive

- 'Sprites' (e.g. backgrounds): send once, warp after
  - Matshushita chip for mobile devices uses such functionality

- Computer-generated visual information
  - Face ad body animation,
  - animated 2-D meshes with moving texture,
  - Synchronized graphics & animated text

Audio Media Objects in MPEG-4

- Audio from 4 kbit/s to 64 kbit/s/channel
  - Arbitrary number of objects in the scene
  - AAC extensions (low delay, long term prediction and error resilience,)
  - TWINVQ (Transform-Domain Weighted Interleave Vector Quantization)
  - > 4kbit/s: Harmonic and Individual Lines plus Noise' (HILN)
- Voice from 2 kbit/s to 24 kbit/s
  - 2-4 kbit/s: HVXC (Harmonic Vector eXcitation Coding)
  - 4-24 kbit/s: CELP:
- Large step + Fine Grain Scalability
  - 1 kbit/s steps through Bit-Sliced Arithmetic Coding

MPEG-4's Synthetic Audio Objects

- Structured Audio
  - SAOL (SA Orchestra Language)
  - SASL (SA Score Language)
  - Great 'music' at very low bitrates

- Coded form of MIDI

- Wavetable synthesis for simple decoders
  - Including effects

- Text-To-Speech (interface!)
  - To complement face/body animation

MPEG-4 Systems

- Binary Scene Description
  - VRML concepts + Streaming + Real Time + Efficiency
  - Content in same scene can come from different sources
  - Allows interaction (local/remote)
  - 2D and 3-D
  - Dynamic (continuous) scene updates and scene animation
- XMT: Textual format for BIFS
  - Includes some SMIL harmonization

MPEG-4 Systems

- Predictable behavior of a decoder and decoded content:

  - Tight synchronization of A, V, synthetic, graphic elements

  - Buffer management

- Flextime

  - Spring-like timing model

- IP management and protection

- File format MP4 (based on Quick Time)

MPEG-4 Systems (cntd.)

- MPEG-J
  - API for complicated content behavior,
  - API's to network / terminal / UI resources
  - 'Application Engine'

- Audio Rendering
  - Specify downmix from arbitrary number of channels
  - Environmental spatialization
    - modeling of environment for spatial sound reproduction
    - Physical and perceptual

Applications of MPEG-4

- Mobile devices
  - Requires low datarates, error resilience, some scalability

- Streaming services
  - Require scalability, low to medium datarates, interactivity, IPMP

- Broadcast (later?)
  - Starting with 'individual' bandwidth
  - But not too much later
  - 'MPEG-4 over MPEG-2'
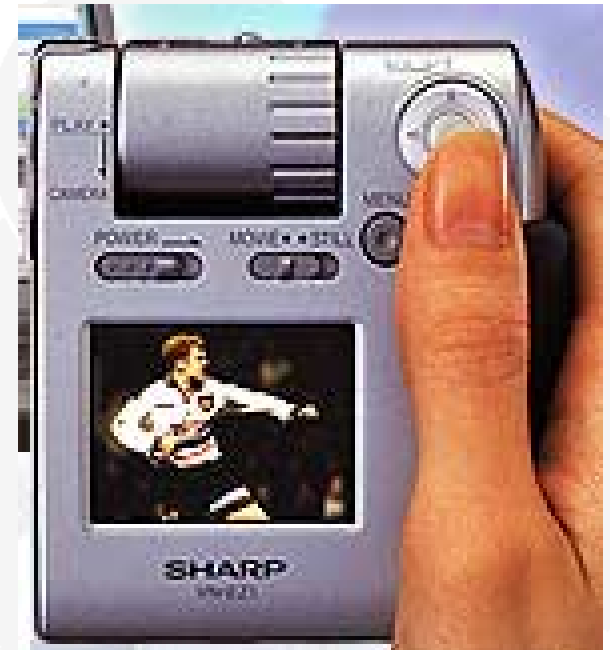
Profiles & Levels

- MPEG-4's Conformance points are Profiles@Levels

  – A bit like in MPEG-2

- Profiles determine tool set

  – E.g. B frames, ¼ pel Motion Compensation

- Levels limit complexity

  – E.g. MBs/sec, max # objects, Complexity Units (Audio)

- Profiles will be convergence point for Industry Standards built on MPEG-4

  – They will be the vehicle for market decisions and uptake

Profile Dimensions:

- Visual (natural, synthetic, natural + synthetic)

- Audio (natural, synthetic, natural + synthetic)

- Graphics

  Media Profiles

- Scene Description (Scene Graph)

  – Tools to describe an manipulate scene

- MPEG-J (Main and Personal)

- Object Descriptor (Synch and Buffers)

- MPEG does not prescribe how to combine these

Some MPEG-4 technology providers

- Philips
- IBM
- (Microsoft)
- Fraunhofer
- InterTrust
- Envivio
- Toshiba
- Matsushita
- WebCast
- SUN
- (Apple)

- PacketVideo
- Robert Bosch
- iVAST
- NEC
- Sony
- Optibase
- Samsung

- Not nearly exclusive.

- 'MPEG-4 will explode'

MPEG-4 Industry Forum

3D pipeline
Adherent Systems Ltd.
America Online
Anystream, Inc.
Apple Computer, Inc.
ARM Ltd
AT&T
Blaxxun Interactive AG
Canon Inc.
celvibe
Cirrus Logic Inc.
Cisco Systems
Clear band, LLC
ComLink Group, Inc.
comverse network systems
CSELT S.p.A.
DiamondBack Vision, Inc
Digital Innovation Ltd.
DIRECTV
DivXNetworks
Dolby Laboratories
Dynapel
Edge Networks
Enformatica
Enquad
Envivio
ETRI
France Telecom
Fraunhofer Institute IIS-A
Fujitsu Limited

Generationxcellent LLC (GMV)
Geocast Network Systems
Gordon & Glickson LLC
Hantro Products Oy
Hitachi, Ltd. Central Research Lab.
Hyundai Electronics Industries Co.
Indigo Vision
Intel Corp.
IBM
Intertrust Technologies International
LightSurf Technologies, Inc.
Lumic Electronics Inc.
Luxxon
Matsushita Elec. Industrial Co.,Ltd.
Media Excel
MedioStream, Inc.
Microsoft Corporation
Mitsubishi Electric Corporation
Motorola/General Instrument
mp4cast
MPEG LA
nCUBE
NEC Corporation
Neomagic Corporation
Net2Wireless
NTT Corporation
Nogatech Ltd.
Nokia
Oki
Optibase

Optivision, Inc.
Packetvideo
ParallelGraphics
Philips Electronics
Philips Semiconductors
samsung Electronics Co., Ltd.
Scientific-Atlanta, Inc.
Serome Technology
SHARP CORPORATION
Siemens AG
SolidStreaming, Inc.
Sony Corporation
SPaSE BV
Sun Microsystems
TANDBERG Television ASA
TDK Corporation
THOMSON multimedia
TOSHIBA
Tvia, Inc.
Vianet Technologies
VideoSpheres Inc.
Vusix SA.
WebCast Technologies Inc.
Wiral
Xilinx Inc.
Zapex Research Ltd.
Zoran

MPEG-4 Industry Forum

- Doing the things that ISO cannot do

- Bootstrapping licensing pools

  – Studying alternative licensing schemes

- Interop testing

- Marketing

  – Trade shows, tutorials, papers

- Logos for conformant Products

Thank you

Further info:

www.cselt.it/mpeg

www.m4if.org

MPEG
ISO/IEC JTC1/ SC29 WG11

INTERTRUST
The MetaTrust Utility™