MPEG-4
A (r)evolution in Multimedia

Rob Koenen
President MPEG-4 Industry Forum

Streaming Media West 2002
25 April 2002 – Los Angeles, USA

What is MPEG-4 Anyway?

- Low bitrate video coding standard?
- Standard for Internet streaming?
- Standard to replace MP3?
- Standard to replace MPEG-2?
- Video coding method that requires extraction of objects?
- A new-found goldmine for patent holders?

What is MPEG-4 Anyway:

- An architecture and coding methods for representing rich multimedia content
- An interoperable yet competitive, cross-platform ecosystem
- An ISO/IEC standard

MPEG-4 is Ready and Available

- MPEG-4 is a revolution in functionality
- MPEG-4 is a an evolution in infrastructure
- MPEG-4 is an open, interoperable standard
- Using MPEG-4 makes business sense

Agenda

- What is MPEG-4 and how does it work?
- Why does it make sense to use MPEG-4?
- How does the MPEG-4 Industry Forum drive adoption?

Agenda

- **What is MPEG-4 and how does it work?**
- Why does it make sense to use MPEG-4?
- How does the MPEG-4 Industry Forum drive adoption?

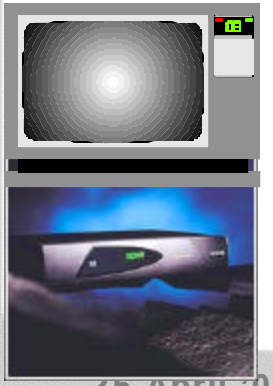
ISO / IEC MPEG: Moving Picture Experts Group

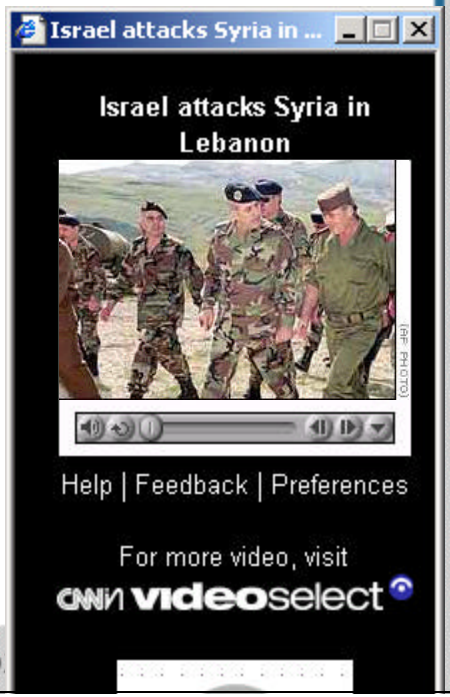
"Coding of Moving Video and Audio"

- MPEG-1: CD-interactive, … – 1992
- MPEG-2: … + TV, HDTV – 1994
- MPEG-3: *HDTV → merged into MPEG-2*
- MPEG-4: Coding of Audiovisual Objects – 1999, ongoing work on extensions
- MPEG-7: Multimedia Description Interface – Fall 2001 'Describing' audiovisual material
- MPEG-21: Digital Multimedia Framework – first parts ready

MPEG-4 Vision

- No convergence, but proliferation of multimedia over different networks, terminals.
- Common multimedia technology necessary that supports:
  - Broadcast
  - Communication
  - Retrieval (online, on-media)





MPEG-4's Objectives

- Common technology for many types of services:
  interactive, broadcast, conversational
- New levels of interactivity
- Integrating natural and synthetic content
- Wide range of access conditions
  - Includes low bitrates, error resilience, scalable coding
- Accommodating content protection

One single technology for …

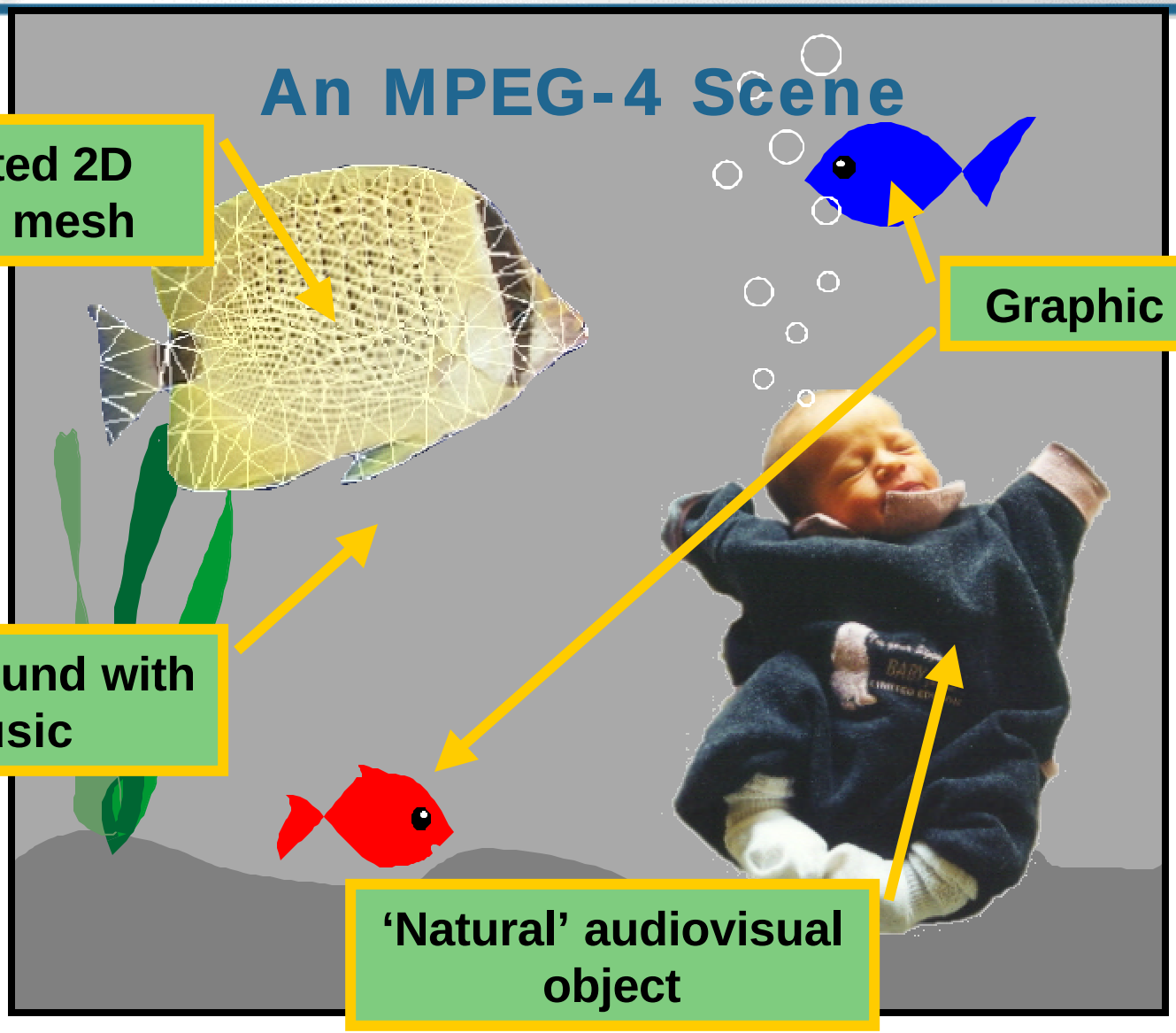
The Nature of MPEG-4

- Object-based multimedia content representation standard
- Revolutionary Systems Layer
- State-of-the art coders, responsibly upgraded
- Profiles and Levels restrict complexity and guarantee interoperability

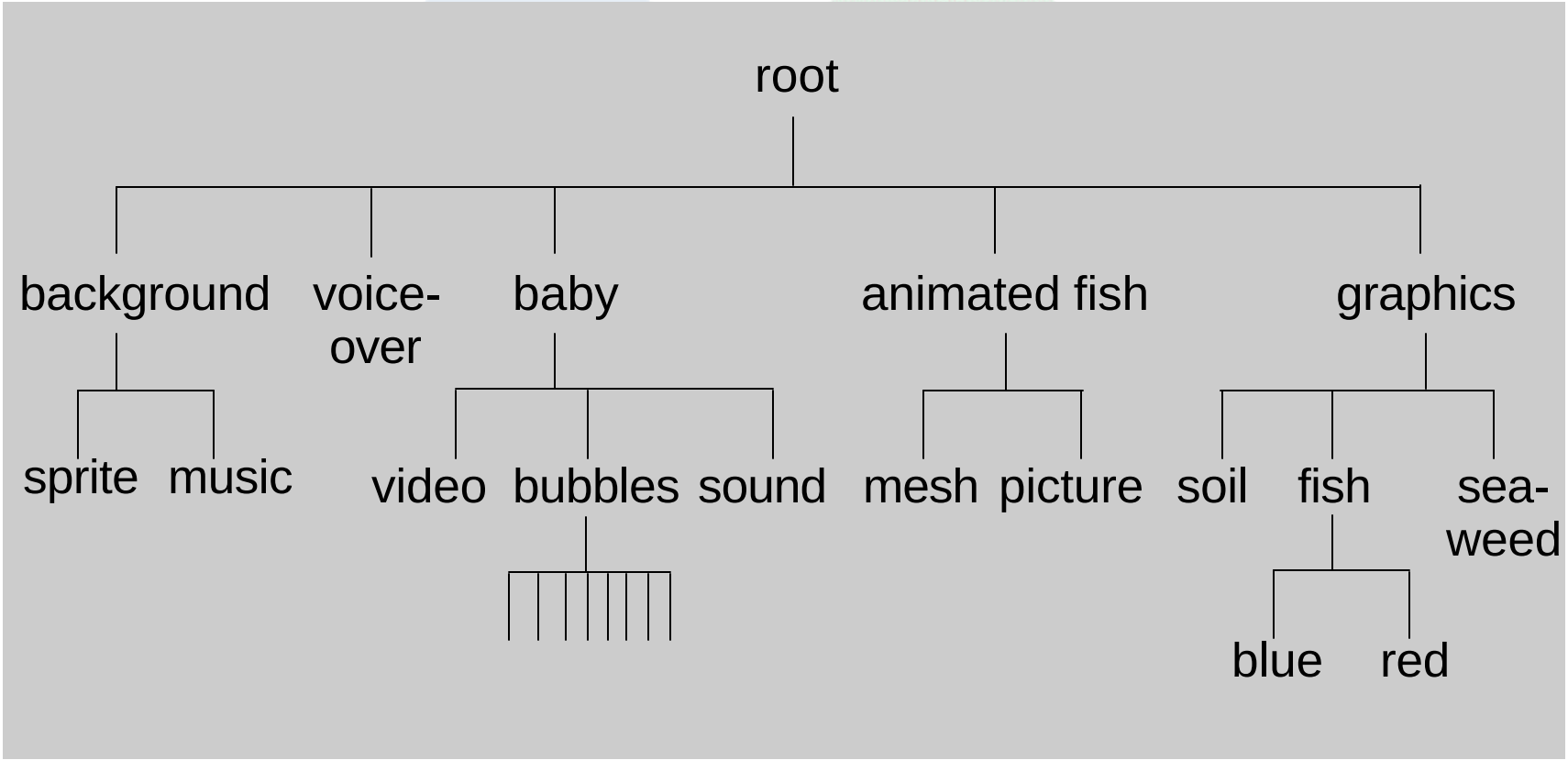
CREATION
Video
3D
Images and Graphics
Audio
Text
Interaction
Layout
Animation
DRM
DISTRIBUTION
MPEG-4 Content
MPEG-4 Server
Packaged Media
Wireless
Satellite
Internet
Broadcast
Retail
CONSUMPTION
Personalization
Interaction
Personal Computer
Television
Game Console
Mobile
DVD
CE Devices
Digital Radio
mp4

MPEG-4: Coding of Audiovisual Objects

- Audiovisual **Scene** is with ‘objects’
- These **Objects** can be of different nature
- **Compositor** puts objects in scene
- Efficient, Real-time Binary **scene description language**
- **Coding** scheme is optimal for object type
- Principle is **independent** of bitrate



The Scene Tree



MPEG-4 Supports a Variety of 'Objects'

- Visual
  - Video
  - Animated face & body;
  - 2D and 3D animated meshes
  - Text and Graphics
- Audio
  - General audio – mono - 5.1 channels
  - Speech
  - Synthetic sounds ('Structured Audio')
  - Tex-to-speech
  - 'Environmental spatialization'

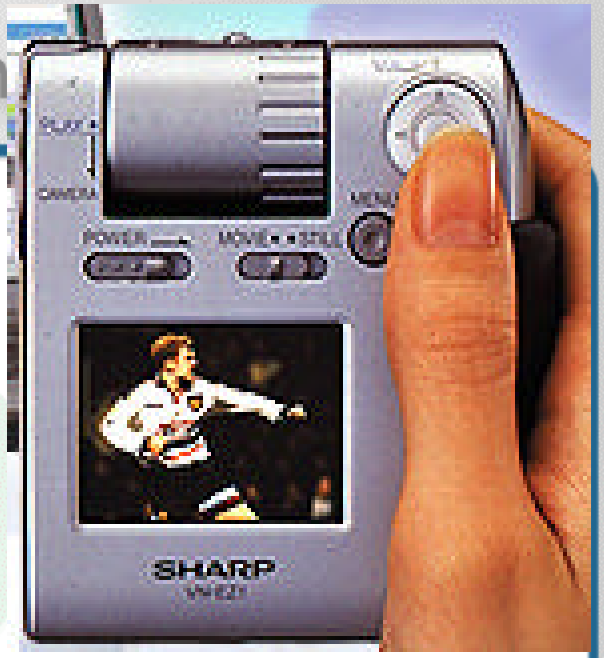
Examples of Interactive Content

- 'We in Music'
  - Develop by Virgin Records and Envivio
  - Windows Media plug-in
- 'TV Lounge'
  - Developed by Universal Music Group and Envivio
  - Real plug-in

- Mixing different objects on the same screen
- They objects have interactive behavior
- The BIFS scene description language describes the scene, the interaction and the behavior

Applications of MPEG-4

- Mobile devices
  - Requires low datarates, error resilience, scalability
- Streaming services
  - Require scalability, low to medium datarates, interactivity
- On-disk distribution (DVD)
  - Interactivity
- Broadcast
  - On-demand services
  - ‘MPEG-4 over MPEG-2’



State of Adoption

- 3G mobile phones
- Internet Streaming Media Alliance
- It is being discussed for next-gen TV services
- RealNetworks announced support yesterday
- QuickTime 6 waiting for licensing to be resolved

Agenda

- What is MPEG-4 and how does it work?
- **Why does it make sense to use MPEG-4?**
- How does the MPEG-4 Industry Forum drive adoption?

Standards Make Sense

- They fuel innovation (GSM, 802.11)
- They survive the test of time (NTSC, MP3)
- They prevent expensive format churn (VHS)
- They help markets mature (DVD, MPEG-2, HTML, HTTP, RTSP, …)
- MPEG-4 is the only cross platform, interactive multimedia standard

MPEG-4 has Clear Benefits

- Code once, use across platforms and players
- Users pick their favorite player
- Content producers pick their favorite tools
- Providers only need to stream in one format
- Competition drives quality up

Increased Operating Profits

- Save money
  - More efficient use of bandwidth
  - Deploy across different platforms
  - Avoid duplication of formats on same platform
- Make money
  - Leverage existing content and infrastructure in new ways
  - Add new dimensions to content use

Revolution in Functionality, but Evolution in Infrastructure

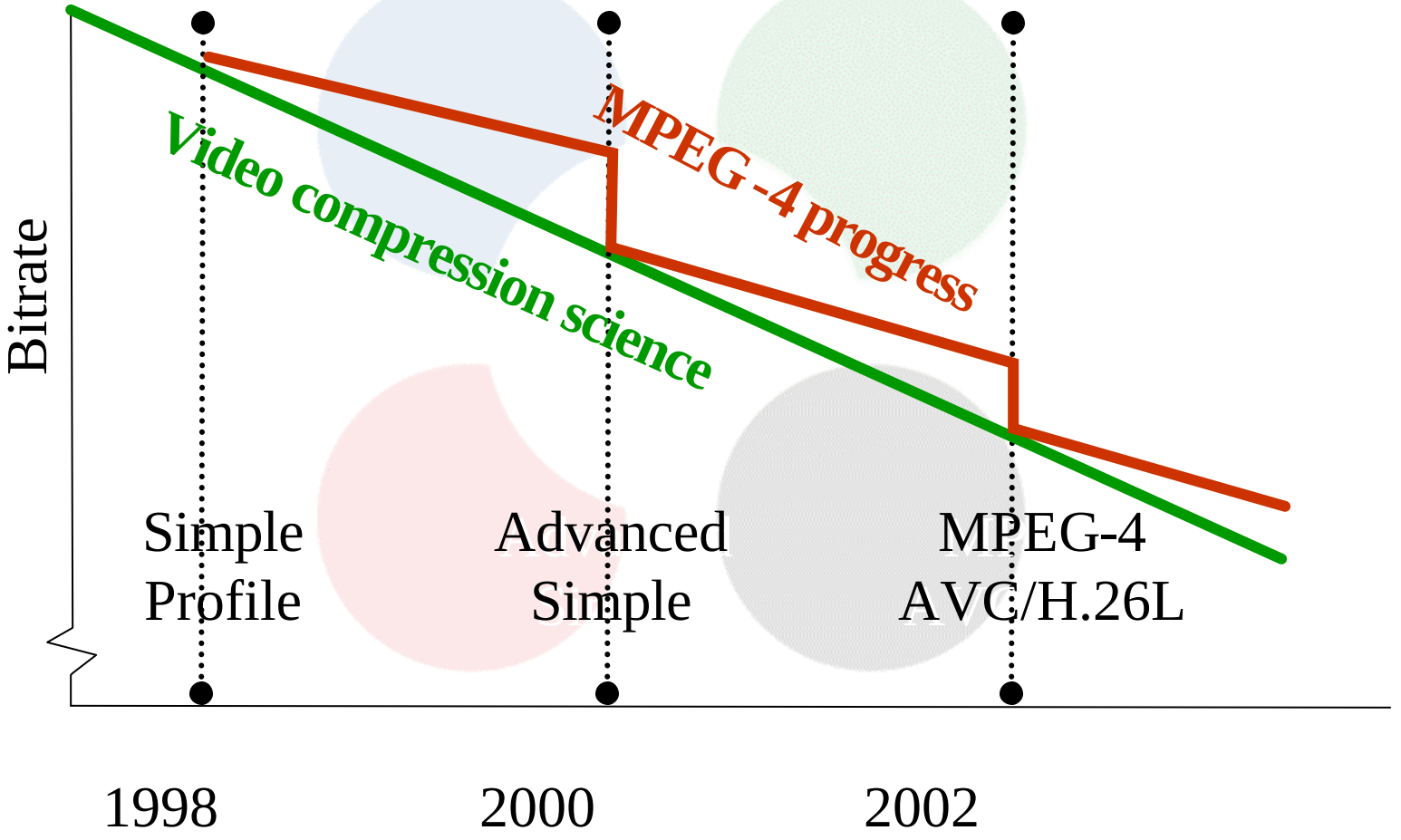
- Agnostic to transport and allows all types of 'casting'
- Use on all types of networks without need for new networks or transport formats
- Leverage existing MPEG-2 infrastructures

Write once - Play everywhere

Quality improves through competition

- 'Industry-standard MPEG-4 quality' does not exist
- MPEG standardizes the minimum: only ***de***coders are standardized
- Competition on encoding and post processing
- MPEG-2 bitrates for same quality have reduced by 50% since 1996 -- ***after*** the standard was frozen and ***without*** upgrading decoders

Responsible Upgrade Strategy



Risks of Proprietary Technology

- Third party business and pricing models
- Proprietary, confidential 3rd party technology roadmaps
- Potential channel conflicts
  - supplier also competitor
  - license includes more “bundled features" than required
- Single sourcing problems

The Insurance of an Open Standard

- Open standard creates an interoperable ecosystem of tools and equipment
- Different providers make what they excel in
- No vendor lock-in
- Market-based price control mechanisms
- Second sourcing of equipment
- No monopolies
- Avoid channel conflicts
- No single vendor controls the format

Agenda

- What is MPEG-4 and how does it work?
- Why does it make sense to use MPEG-4?
- **How does the MPEG-4 Industry Forum drive adoption?**

MPEG-4 Industry Forum

- Picking up where ISO MPEG stops
- Interoperability testing
- Logos for conformant Products
- Marketing
- Information clearinghouse
- Bootstrapping licensing pools
  - But real work is independent of M4IF
- Continuing to drive MPEG-4 adoption

Members MPEG-4 Industry Forum

Airia
Amnis Systems Inc
Agfa Monotype Corp.
Amphion Semiconductor
AOL Time Warner
Apple Computer, Inc.
ARM Ltd
AT&T Ltd
Avipix
Blaxxun
Canon Inc.
Celvibe
ClearBand, LLC
Cirrus Logic, Inc.
Cisco Systems
Comlink Group, Inc.
Coding Technologies GmbH
Comverse network systems
DiamondBack Vision, Inc.
Dicas Digital Image Coding GmbH
DivXNetworks, Inc.
Dolby Laboratories Inc.
Dynapel Systems
Edge Networks Corporation
Emblaze Research
Enformatica Limited
Enquad
Envivio
ETRI
France Telecom
Fraunhofer Institute IIS-A
Fujitsu Limited
GMV Network
Hantro Products Oy
Hitachi, Ltd
Hypnotizer
IBM Ltd
Improv Systems
Indigo Vision Ltd
Intel Corp.
Interactive Sapience Corporation
Interra Digital Video Technologies
Intertrust Technologies Corp.
iVAST Inc.
JiGami Corporation
Latham & Watkins
Lumic Electronics Inc.
Luxxon Corp.
Matsushita Electric Industrial Co.,Ltd.
Media Excel, Inc.
Microsoft Corporation
Motorola
Mitsubishi Electric Corporation
mp4cast
MPEG LA, LLC
nCube
NEC Corporation
Neomagic Corporation
Nippon Telegraph and Telephone Corporation
Nokia
Nogatech Ltd.
Novatek Microelectronics Corp
Oki Electric Industry Co., Ltd.
OnTimeTek
Optibase
Packetvideo
PACE Soft Silicon Limited
Panasonic Disc Services Corporation
ParallelGraphics
Philips Electronics
Philips Semiconductors
Protocom Corporation
Reakosys Inc.
RealNetworks
Samsung Electronics Co., Ltd.
Scientific-Atlanta, Inc.
Serome Technology, Inc.
Sharp Corporation
Siemens AG
Sigma Designs
SolidStreaming, Inc.
Sony Corporation
Sorenson Media
Soundball, Inc.
SPaSE BV
Sun Microsystems
Symah Vision
Tandberg Television ASA
TDK Corporation
Tektronix Cambridge Limited
Telecom Italia Lab
THOMSON multimedia
Toshiba
Tvia, Inc.
VCS Video Communication Systems AG
Vianet Technologies
VideoSpheres Inc.
Vusix SA.
WebCast Technologies, Inc.
Wiral
Xilinx Inc.
Zoran
ZTE Corporation
DRM Digital Radio Mondiale
Ecole Nationale Superieure des Telecommunications
Multimedia Technology Research Center, HKUST, Hong Kong
ViDe

MPEG-4 is Interoperable Today

- One year of interop work on Audio, Video and MP4 files
- The program has 30 companies and is growing:
  - Amphion, Apple, Avipix, Cirrus Logic, Comverse, Diamondback Vision, Emblaze, ENST, Envivio, Fraunhofer, Hantro, IBM, iVast, Luxxon, MTREC, NeoMagic, OnTimetek, Philips, PacketVideo, Reakosys, Samsung, Serome, Sorenson, Thomson Multimedia, VideoSpheres, WebCast Technologies
  - Recently joined: Dicas, DivX Networks, RealNetworks
- Next steps: streaming and certification to get the logo

Licensing?

- MPEG-4 AAC licensing announcement by 5 MPEG-4 Audio licensors well received by market
- Pre-release of Visual licensing controversial, but will hopefully be resolved soon

Thank you

Further information:

www.m4if.org (website M4IF)

mpeg.telecomitalialab.com (website MPEG)

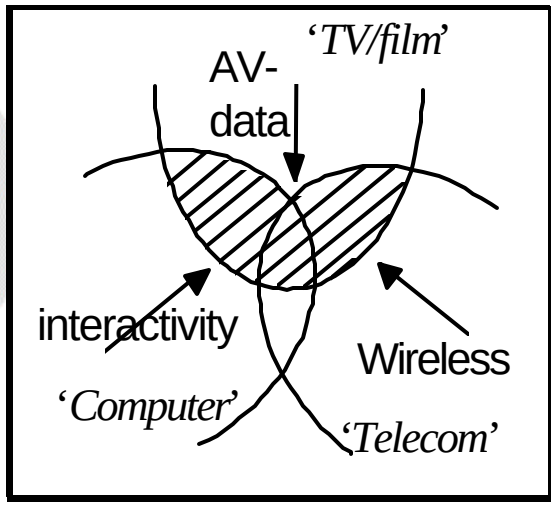
Additional Information

The following slides provide more detail.
They were not used during the keynote

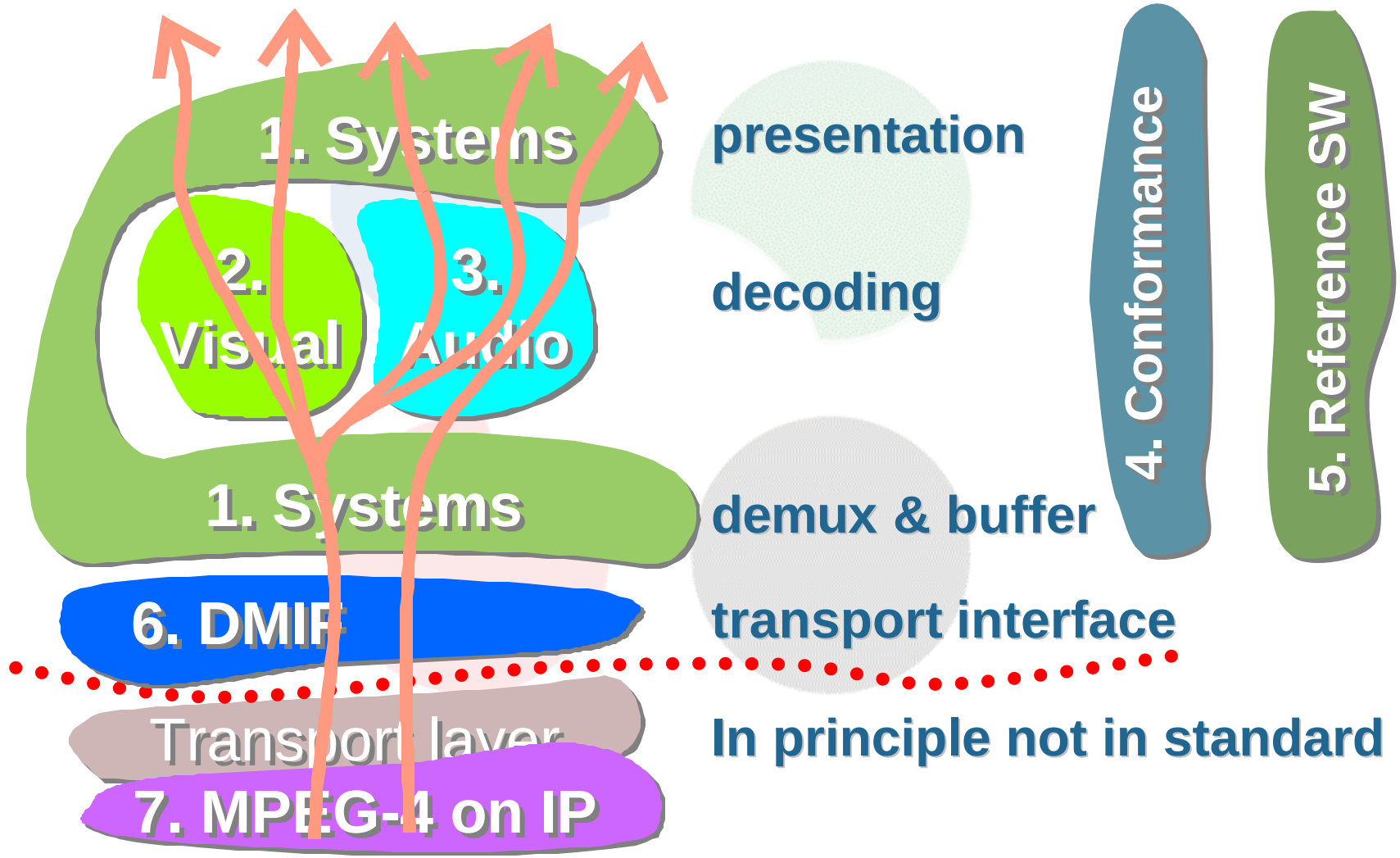
A bit of history and background on MPEG-4

- 1993: started as 'Very Low Bitrate Audiovisual Coding'
- 1994: goal changed to 'Coding of Audiovisual Objects'
- Ready since 1999, being extended in backward compatible ways
- Now working on MPEG-4 part 10, the 'JVT Codec', Audio extensions and some Systems refinements

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



The Parts of the MPEG-4 Standard



Recent additions

- Part 8: optimized video encoder
- Part 9: VHDL description of MPEG-4 (tentative)
- Part 10: JVT Coder
  - New video coder, built jointly with ITU-T
  - H.26L was starting point
  - Ready end 2002

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, then warp
  - Matshushita chip for mobile devices uses this
- Computer-generated visual information
  - Face and 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
  - Extended AAC, TWINVQ
  - > 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

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