Introduction to Visual Communication Standards and System
Outline

- Introduction
- Standards for Visual Communication Systems
- Compression Technologies
- Implementation Issues
- Optional Enhancements
Applications

- Video Phone
- Video Conference
- Distance Learning
- Broadcasting TV
- Video On Demand
- Digital Library
- Multimedia communication applications
ITU Visual Communication Standards

- H.310 – Broadband ISDN, ATM LAN
- H.320 – Narrowband switched ISDN
- H.321 – Broadband ISDN, ATM LAN
- H.322 – Guaranteed bandwidth packet switched networks
- H.323 – Non-guaranteed bandwidth packet switched network (Ethernet)
- H.324 – GSTN, POTS, the analog phone system
## Components of the Visual Communications

<table>
<thead>
<tr>
<th>Standard</th>
<th>H.310</th>
<th>H.320</th>
<th>H.321</th>
<th>H.322</th>
<th>H.323</th>
<th>H.324</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network</td>
<td>ATM LAN</td>
<td>ISDN</td>
<td>ATM LAN (QoS) LAN</td>
<td>Ethernet</td>
<td>GSTN (POTS)</td>
<td></td>
</tr>
<tr>
<td>Video</td>
<td>MPEG-2 (H.262) H.261</td>
<td>H.261 (H.263)</td>
<td>H.261 (H.263)</td>
<td>H.261 H.263+</td>
<td>H.261 H.263+</td>
<td></td>
</tr>
<tr>
<td>Multiplexing</td>
<td>H.222.0 H.222.1 (MPEG)</td>
<td>H.221</td>
<td>H.221</td>
<td>H.221</td>
<td>H.225.0</td>
<td>H.223</td>
</tr>
<tr>
<td>Control</td>
<td>H.245</td>
<td>H.230 H.242</td>
<td>H.242</td>
<td>H.242</td>
<td>H.245</td>
<td>H.245</td>
</tr>
<tr>
<td>Data</td>
<td>T.120</td>
<td>T.120</td>
<td>T.120</td>
<td>T.120</td>
<td>T.120</td>
<td>T.120</td>
</tr>
</tbody>
</table>
ITU-T DRAFT REVISED H.32

TELECOMMUNICATION STANDARDIZATION SECTOR OF ITU

(Draft revisions dated 1997-09-2)

LINE TRANSMISSION OF NON-TELEPHONE SIGNALS

TERMINAL FOR LOW BITRATE MULTIMEDIA COMMUNICATION

ITU-T Recommendation H.324
Scope of H.324

- “cover the technical requirements for very low bitrate multimedia telephone terminals operating over the General Telephone Network (GSTN).”
- “provide real-time video, audio or data, or any combination”
- “communication may be either 1-way or 2-way.”
- “the multimedia telephone terminals… can be integrated into PCs, or workstation, or be stand-alone.”
- “interworking … on the ISDN and on mobile radio networks are also covered.”
System Overview (H.324)

Scope of recommendation H.324

- Video I/O equipment
- Audio I/O equipment
- User data applications
- System control
- Video codec H.263/H.261
- Audio codec G.723
- Data protocols T.120, v.42
- Control protocol H.245
- Multiplex/Demultiplex H.223
- Modem V.34/V.8
- Modem control V.25ter
- GSTN Network
- MCU

NTUEE DSP/IC Lab

Liang-Gee Chen
Information Streams of H.324

- Video/Audio streams are continuous traffic carrying moving color pictures.
- Data streams may represent still pictures, facsimile, documents, computer files, computer application data, undefined user data, and other data streams.
- Control streams pass control commands and indications between remote counterparts.
Functional Elements covered by H.324---Data Protocols

• The Data Protocols support data applications such as **electronic whiteboards, still image transfer, file exchange, database access, audiographics conferencing, remote device control, network protocols, etc.** Standardized data applications include **T.120** for real-time audiographics conferencing, **T.84** simple point-point still image file transfer, **T.434** simple point-point file transfer, **H.224/H.281** far-end camera control, **T.30** facsimile transfer, **T.140** Text conversation protocol, ISO/IEC **TR9577** network protocols including PPP and IP

• using buffered **V.14** *(asynchronous characters without error control)* or **LAPM/V.42** *(with error control)* **V.42bits** *(with data compression)*, **V.24** *(with frame)*. Other applications and protocols may also be used via **H.245** negotiation.
Functional Elements covered by H.324---Control Protocols

- The Control Protocol (H.245) provides end-to-end signaling for proper operation of the H.324 terminal, and signals all other end-to-end system functions including capabilities exchange, opening and closing of logical channels, mode preference requests, multiplex table entry transmission, flow control messages, and general commands and indications.
Functional Elements covered by H.324---Multiplex Protocols

• The Multiplex Protocol (H.223) multiplexes transmitted video, audio, data and control streams into a single bit stream, and demultiplexes a received bit stream into various multimedia streams. In addition, it performs logical framing, sequence numbering, error detection, and error correction by means of retransmission, as appropriate to each media type.
Functional Elements covered by H.324---Modem

• **The Modem** (V.34) converts the H.223 synchronous multiplexed bit stream into an analog signal that can be transmitted over the GSTN, and converts the received analog signal into a synchronous bit stream that is sent to the Multiplex/Demultiplex protocol unit. **V.25ter** is used to provide control/sensing of the modem/network interface, when the modem with network signaling and **V.8/V.8bis** functional elements is a separate physical item.
Functional Elements covered by H.324---Video/Audio Codec

- **H.261/263**: Video streams are continuous traffic carrying moving color pictures.
- **G.723.1**: Audio streams are real-time and maintain synchronization with the video streams.
## Estimated Processing Requirements

<table>
<thead>
<tr>
<th>Function</th>
<th>Minimum requirement (MIPS)</th>
<th>Maximum requirement (MIPS)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>H.320 video</td>
<td>40</td>
<td>100</td>
<td>Frame rate : 5~30, Frame size : CIF, QCIF, Bitrate : 64K ~ 1.544Mbps</td>
</tr>
<tr>
<td>H.320 Audio</td>
<td>5</td>
<td>60</td>
<td>Bitrate : 64k~16kbps</td>
</tr>
<tr>
<td>H.324 Video</td>
<td>40</td>
<td>1200</td>
<td>Frame rate : 10, Frame size : SQCIF, QCIF, CIF, 4CIF, 16CIF</td>
</tr>
<tr>
<td>Control</td>
<td>10</td>
<td>15</td>
<td>Include H.223, H.245</td>
</tr>
<tr>
<td>Speaker phone</td>
<td>5</td>
<td>50</td>
<td>Not in ITU standard, Echo suppression, Echo cancellation</td>
</tr>
<tr>
<td>I/O and overhead</td>
<td>15</td>
<td>50</td>
<td>Operating system</td>
</tr>
<tr>
<td>Multi-point</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Video Formats

<table>
<thead>
<tr>
<th>Picture format</th>
<th>Luminance pixels</th>
<th>Luminance lines</th>
<th>Uncompressed bitrate (Mbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>10 frames/sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mono</td>
</tr>
<tr>
<td>SQCIF</td>
<td>128</td>
<td>96</td>
<td>1.0</td>
</tr>
<tr>
<td>QCIF</td>
<td>176</td>
<td>144</td>
<td>2.0</td>
</tr>
<tr>
<td>CIF</td>
<td>352</td>
<td>288</td>
<td>8.1</td>
</tr>
<tr>
<td>4CIF</td>
<td>704</td>
<td>576</td>
<td>32.4</td>
</tr>
<tr>
<td>16CIF</td>
<td>1408</td>
<td>1152</td>
<td>129.8</td>
</tr>
</tbody>
</table>
Video Compression

original

Source Encoder

Channel Encoder

01011000...

Channel Decoder

Source Decoder

reconstructed
Video Compression Processing


Video Coder

Video Sequence → Marco Block → Forward DCT → Quantization → Entropy Coding → 0101110...

− Inverse Quantization
− Inverse DCT
− Motion Estimation/Compensation

0101110...
# Picture Formats for Video Terminals

<table>
<thead>
<tr>
<th>PICTURE FORMAT</th>
<th>LUMINANCE PIXELS</th>
<th>ENCODER</th>
<th>DECODER</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQCIF</td>
<td>128 x 96 for H.263³</td>
<td>Optional³</td>
<td>Required¹,²</td>
</tr>
<tr>
<td>QCIF</td>
<td>176 x 144</td>
<td>Required</td>
<td>Required¹,²</td>
</tr>
<tr>
<td>CIF</td>
<td>352 x 288</td>
<td>Optional</td>
<td>Optional</td>
</tr>
<tr>
<td>4CIF</td>
<td>704 x 576</td>
<td>Not defined</td>
<td>Optional</td>
</tr>
<tr>
<td>16CIF</td>
<td>1408 x 1152</td>
<td>Not defined</td>
<td>Optional</td>
</tr>
</tbody>
</table>

**NOTE 1** - Optional for H.320 interworking adapters.

**NOTE 2** - Mandatory to encode one of the picture formats QCIF and SQCIF; optional to encode both formats.

**NOTE 3** - H.261 SQCIF is any active size less than QCIF, filled out by a black border, coded in QCIF format.
## Differences Between the H.26x Coding Algorithm

<table>
<thead>
<tr>
<th></th>
<th>H.261</th>
<th>H.263</th>
<th>H.263+</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Picture Size</strong></td>
<td>QCIF, CIF</td>
<td>Sub-QCIF, QCIF, CIF, 4CIF, 16CIF</td>
<td>Sub-QCIF, QCIF, CIF, 4CIF, 16CIF, Custom picture size</td>
</tr>
<tr>
<td><strong>Target Bitrate</strong></td>
<td>P x 64 kbps (P=1,2, … 30)</td>
<td>Below 64 kbps</td>
<td></td>
</tr>
<tr>
<td><strong>Frame Format</strong></td>
<td>I, P</td>
<td>I, P, PB</td>
<td>I, P, PB, Improved PB, B, EI, EP</td>
</tr>
<tr>
<td><strong>Frame Rate</strong></td>
<td></td>
<td>30 frames/second</td>
<td>15 to 1800 frames/second</td>
</tr>
<tr>
<td><strong>Composition of Picture</strong></td>
<td>GOB</td>
<td>GOB</td>
<td>GOB, Slice</td>
</tr>
<tr>
<td><strong>Macroblock Size</strong></td>
<td>16x16</td>
<td>16x16</td>
<td>16x16, 32x32</td>
</tr>
<tr>
<td><strong>Block Size</strong></td>
<td>8x8</td>
<td>8x8</td>
<td>8x8, 16x16</td>
</tr>
<tr>
<td><strong>Search Range</strong></td>
<td>-16~15</td>
<td>-16~+15.5, -31.5~+31.5</td>
<td>-16<del>15.5, -31.5</del>31.5, -32<del>31.5, -64</del>63.5, -128<del>127.5, -256</del>255.5</td>
</tr>
<tr>
<td><strong>MC Accuracy</strong></td>
<td>integer-pel accuracy</td>
<td>half-pel accuracy</td>
<td></td>
</tr>
<tr>
<td><strong>Filter Effect</strong></td>
<td>spatial lowpass loop filter</td>
<td>bi-linear interpolation for half-pel MC</td>
<td></td>
</tr>
<tr>
<td><strong>VLC Table</strong></td>
<td>pairs (run, level)</td>
<td>triplets (run, level, cob)</td>
<td></td>
</tr>
<tr>
<td><strong>MV Predictor</strong></td>
<td>MV of previous MB</td>
<td>MV of previous MB</td>
<td>MV of previous MB</td>
</tr>
<tr>
<td><strong>Multi-point</strong></td>
<td>None</td>
<td>None</td>
<td>Up to 4 separate Sub-Bitstreams</td>
</tr>
<tr>
<td><strong>Supplemental Enhancement Information</strong></td>
<td>None</td>
<td>None</td>
<td>Freeze, Resizing, Snapshot, Time Segment, Refinement Segment, Chroma Keying</td>
</tr>
</tbody>
</table>
Discrete Cosine Transform

- **Block size:** 8 x 8
- **Two-dimensional DCT:**

\[
F(u,v) = \frac{2}{N} C(u)C(v) \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} f(x,y) \cos \frac{2\pi(2x+1)u}{4N} \cos \frac{2\pi(2y+1)v}{4N}
\]

\[
C(u), C(v) = \begin{cases} 
1/\sqrt{2}, & u,v = 0 \\
1, & \text{otherwise}
\end{cases}
\]

- **Most large coefficients concentrate on the upper-left corner**
- **Quantization and zig-zag scan**
Discrete Cosine Transform
Motion Compensation/Estimation

- Displacement vector
- Best-matched search-area block
- Current block
- Previous frame
- Present frame
- Time
Motion Compensation/Estimation

Original → Reconstructed

Difference
G.723 General description

- **Bit rates:** 5.3 or 6.3 kbit/s
  - It is possible to switch between the two rates at any 30 ms frame boundary.
- **Linear prediction analysis-by-synthesis coding**
- **Frame size:** 30ms, 240 samples at 8kHz sampling rate.
G.723 Architecture

Draft Recommendation G.723 - Dual Rate Speech Codec for Multimedia Communications Transmitting at 5.3 & 6.3 kbit/s
Parallel Architectures for Media Processor

Multiple Instruction Multiple Data

Single Instruction Multiple Data

Very Long Instruction Word

Current VLIW
## Announced Media DSP Architectures

<table>
<thead>
<tr>
<th>Architecture</th>
<th>TI C80</th>
<th>Chromatics Mpact</th>
<th>Philips Tri-Media</th>
<th>TI C6x</th>
<th>Samsung MSP-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architecture</td>
<td>4 x 64b DSP’s + 32b RISC + cross-bar</td>
<td>VLIW/SIMD 4 ALUs ME engine 792b bus</td>
<td>VLIW 25 exec units + VLD</td>
<td>VLIW 8 instr/clk 2 MACs/clk cond. exec.</td>
<td>32-way SIMD + 32b RICS</td>
</tr>
<tr>
<td>Clock</td>
<td>40 MHz</td>
<td>62 MHz</td>
<td>100 MHz</td>
<td>200 MHz</td>
<td>100 MHz</td>
</tr>
<tr>
<td>Peak Perf.</td>
<td>1.2 Gops</td>
<td>2 Gops</td>
<td>4 Gops</td>
<td>1.6 Gops</td>
<td>6.4 Gops</td>
</tr>
<tr>
<td>Memory</td>
<td>DRAM 400 MB/s</td>
<td>RAMBUS 500 MB/s</td>
<td>SDRAM 400 MB/s</td>
<td>SDRAM 400 MB/s</td>
<td>SDRAM 800 MB/s</td>
</tr>
<tr>
<td>Programming</td>
<td>compiler + assembler</td>
<td>in-house</td>
<td>VLIW compiler</td>
<td>VLIW compiler</td>
<td>compiler + assembler</td>
</tr>
</tbody>
</table>

Peak Perf.:
- TI C80: 1.2 Gops
- Chromatics Mpact: 2 Gops
- Philips Tri-Media: 4 Gops
- TI C6x: 1.6 Gops
- Samsung MSP-1: 6.4 Gops

Memory:
- DRAM 400 MB/s
- RAMBUS 500 MB/s
- SDRAM 400 MB/s

Programming:
- compiler + assembler
- in-house
- VLIW compiler
- VLIW compiler
- compiler + assembler

Clock:
- 40 MHz
- 62 MHz
- 100 MHz
- 200 MHz
- 100 MHz
Performance Issues

- Computation Complexity: HW/SW co-design
- Bitrate: Network Interface
- Quality: Subjective test
- Latency: for two-way communication
- Cost:
- Losslessness: Optional
- Power Consumption:
Prototyping Platform

A snapshot of the video phone system.
MPEG-4 Industries

Computer

Multimedia Computing
Multiplayer Games
Virtual Environment
GUIs

Communication

Wireless Communication
Internet/WWW
Phonelines

Entertainment

Digital TV
HDTV
Film Video, Audio
Production

MPEG-4
Multimedia Communication
Interoperation and Enhancements

- **Interoperation**
  - Speech only terminal
  - H.324 terminal over ISDN (H.324/I)
  - H.324 terminal over mobile radio (H.324/M)

- **Enhancement**
  - Encryption
  - Multipoint Consideration
Where to get more information

- **New ITU-T Standard Drafts**
  - **H.320 suite**: ftp://standard.pictel.com/avc-site/DraftRec
  - **H.324 suite**: ftp://standard.pictel.com/h324-site/
  - **H.263+ video**: ftp://standard.pictel.com/video-site/h263plus

- **ITU-T Standard Documents**
  - Can be purchased from http://www.itu.ch
Prototyping Hardware
H.320 Architecture
H.323 System Structure

- Video I/O
- Audio I/O
- Data App.
- System Control UI
- Video Codec H.261/H.263
- Audio Codec G-series
- Delay
- System Control
  - H.245 Control
  - Call Control
  - RAS Control
- H.225.0 Layer
- LAN Interface