MPEG-2 over ATM

John Koiste

Senior Engineer

NORTEL
NORTHERN TELECOM

INFORM '96

Agenda

- Why digital video and video compression?
- Fundamentals of color video in five minutes
- Video compression tool kit for MPEG-2
- An overview of MPEG-2
- Fitting MPEG-2 on ATM
- NetCalc3 modeling of MPEG-2 and ATM

Why Digital Video?

- Digital format has evolved
  - from telephone network switching and transmission
  - to high fidelity audio (compact disks)
  - to video

- Attraction of digital video seems irresistible:
  - predictable and controllable quality
  - infinite life for motion picture recordings
  - ability to interface with computers
  - unbounded future flexibility for computer-aided features (editing, splicing, combining, masking, filtering, recording,...)

Why Video Compression?

- **Uncompressed digital video needs much more bandwidth and storage capacity than equivalent analog video**
- **Digital video requires compression, since:**
  - **digital storage requirements are huge**
  - **affordable digital storage devices have low access bit rates**
  - **transmission bit-rate requirements are high**
  - **high-definition TV (digital U.S. broadcast version) must fit into existing 6 MHz channels**

High-bandwidth optical fiber does not reduce the need for compression

Agenda

- Why digital video and video compression?
- **Fundamentals of color video in five minutes**
- Video compression tool kit for MPEG-2
- An overview of MPEG-2
- Fitting MPEG-2 on ATM
- NetCalc3 modeling of MPEG-2 and ATM

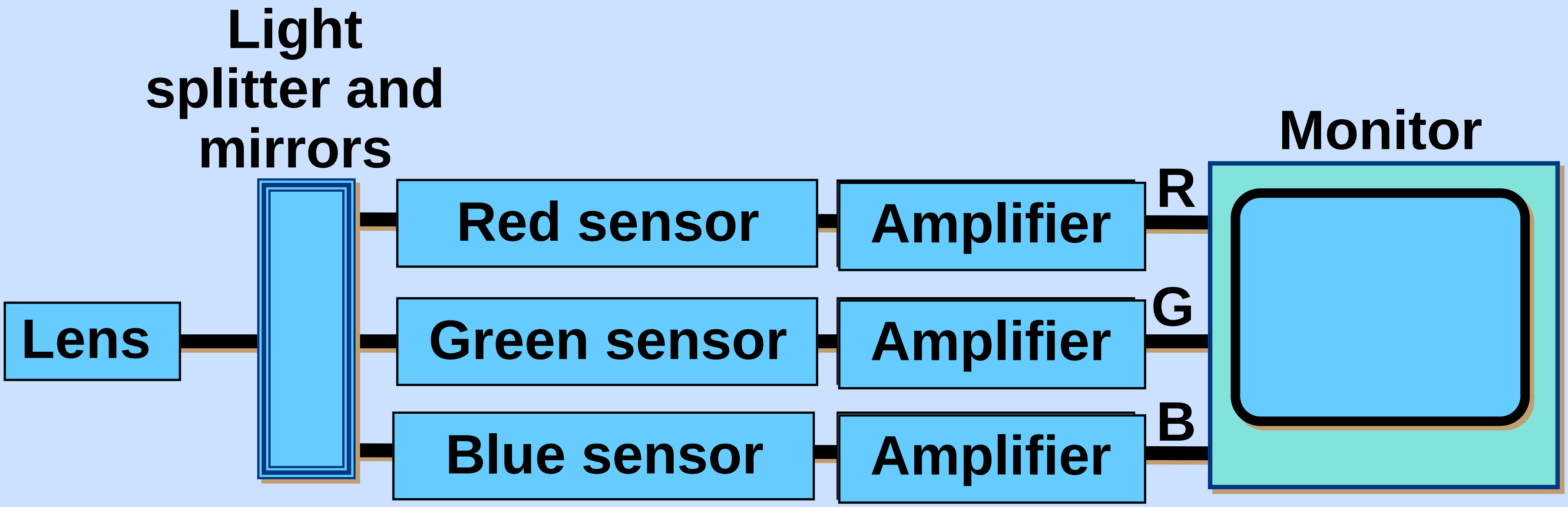
Color Video Basics

- **Screen image**
  - created on a phosphor coating inside the viewing surface of the monitor's picture tube
  - composed of many small momentarily projected dots of light called picture elements (also called pixels or pels)
- **Pixel color is achieved by mixing red, green and blue (RGB) lights using three projection guns**

Grey Scale and Refresh Rates

- **Equal levels of RGB produce familiar grey scale range (from black, dark to light shades of grey, to white)**
- **Pixels glow for only about 40 ms**
- **To prevent picture fading, all pixels must be refreshed by reprojection**
  - **at 25 to 30 times/second for TV screens**
  - **at over 70 times/second for computer monitor screens (which are used in bright rooms)**
- **This refresh requirement also enables a rapid change in images – i.e., simulation of motion**

Color Video Camera and Monitor



- Video camera produces three analog video signals (RGB) to control the three guns in a color picture tube (i.e. in a TV set or computer monitor)
- RGB signals contain all the information necessary to create a color picture

Distribution of RGB Signals

- **RGB signals cannot be widely distributed easily in exact synchronism and relationship**
  - **RGB is converted in the video camera to a composite format better suited to distribution**
- **Composite format:**
  - **is a linear matrix transformation of RGB**
  - **results in one luminance signal (called Y) and two chrominance or color difference signals (called either U and V or Cb and Cr)**
  - **the (composite) color formats are commonly collectively referred to as YUV or YCbCr**

Composite Color Formats

Name	Maximum Pels/Line H	Max. Lines V	Picture Aspect Ratio H:V	Refresh Rate Pict. /s
NTSC*	711	483	4:3	30
PAL*	702	575	4:3	25
SECAM*	702	575	4:3	25
HDTV-1440	1440	1152	16:9	25/30
HDTV(max)	1920	1152	16:9	25/30

*** Per ITU-R-Rec. 601**

Motion Picture Format

- Motion pictures are totally film-based (no pixels)
- Resolution of 35 mm movie frame is not as good as a 35 mm still-camera photo resolution, but better than HDTV
- Shape of movie screen is even slightly wider in proportion than HDTV
- Refresh rate is a rather low 24 frames/ second because the viewing environment is quite dark

Color Video Formats—Computers

- IBM PC
  - VGA produces 640x480 (HxV) pixels per frame
  - Super VGA monitors and others with higher resolutions are common
- Macintosh
  - 832x560 is most common
  - 1152x870 is increasingly used
- Sun Sparc 10
  - 1152x900 is typical
- DEC
  - 1024x864 is typical

Agenda

- Why digital video and video compression?
- Fundamentals of color video in five minutes
- **Video compression tool kit for MPEG-2**
- An overview of MPEG-2
- Fitting MPEG-2 on ATM
- NetCalc3 modeling of MPEG-2 and ATM

Technical Toolkit (for MPEG-2)

- **Discrete cosine transform, forward and inverse (FDCT and IDCT)**
- **Quantization of numbers**
- **Zig-zag ordering and entropy encoding**
- **Motion estimation by block matching**
  - **predicting from past frames**
  - **interpolating between past and future frames**
- **Timing recovery from received data**

Forward DCT Applied to 8x8 Pixels

2	12	22	32	42	54	66	78
5	15	25	35	45	57	69	81
8	18	28	38	48	60	72	84
11	21	31	41	51	63	75	87
14	24	34	44	54	66	78	90
17	27	37	47	57	69	81	93
20	30	40	50	60	72	87	102
23	33	43	53	63	75	90	105

- **Transforms data from image to frequency domain**
- **Transformed data makes compression techniques feasible**

394.25	-282.5	16.0	-33.6	1.1	-7.8	0.7	-3.0
-81.4	3.8	-3.0	2.0	-1.0	0.2	0.2	-0.2
2.9	-2.7	2.2	-1.4	0.7	-0.1	-0.2	0.2
-9.5	1.3	-1.1	0.7	-0.3	0.10	0.10	-0.10
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
-1.5	-0.9	0.7	-0.5	0.2	0.0	-0.10	0.10
-1.2	1.1	-0.9	0.6	-0.3	0.10	0.10	-0.10
0.2	-0.8	0.6	-0.4	0.2	0.0	0.0	0.0

Quantization of Transformed Data

394.25	-282.5	16.0	-33.6	1.1	-7.8	0.7	-3.0
-81.4	3.8	-3.0	2.0	-1.0	0.2	0.2	-0.2
2.9	-2.7	2.2	-1.4	0.7	-0.1	-0.2	0.2
-9.5	1.3	-1.1	0.7	-0.3	0.10	0.10	-0.10
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
-1.5	-0.9	0.7	-0.5	0.2	0.0	-0.10	0.10
-1.2	1.1	-0.9	0.6	-0.3	0.10	0.10	-0.10
0.2	-0.8	0.6	-0.4	0.2	0.0	0.0	0.0

390.00	-280.0	20.0	-30.0	0.0	-10.0	0.0	0.0
-80.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
-10.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

- **Eliminates visually insignificant data**
- **Improves coding efficiency of data that is left**

Zig-Zag Ordering and Entropy Encoding (Zero Run-Length and Huffman Coding)

390.00	-280.0	20.0	-30.0	0.0	-10.0	0.0	0.0
-80.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
-10.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

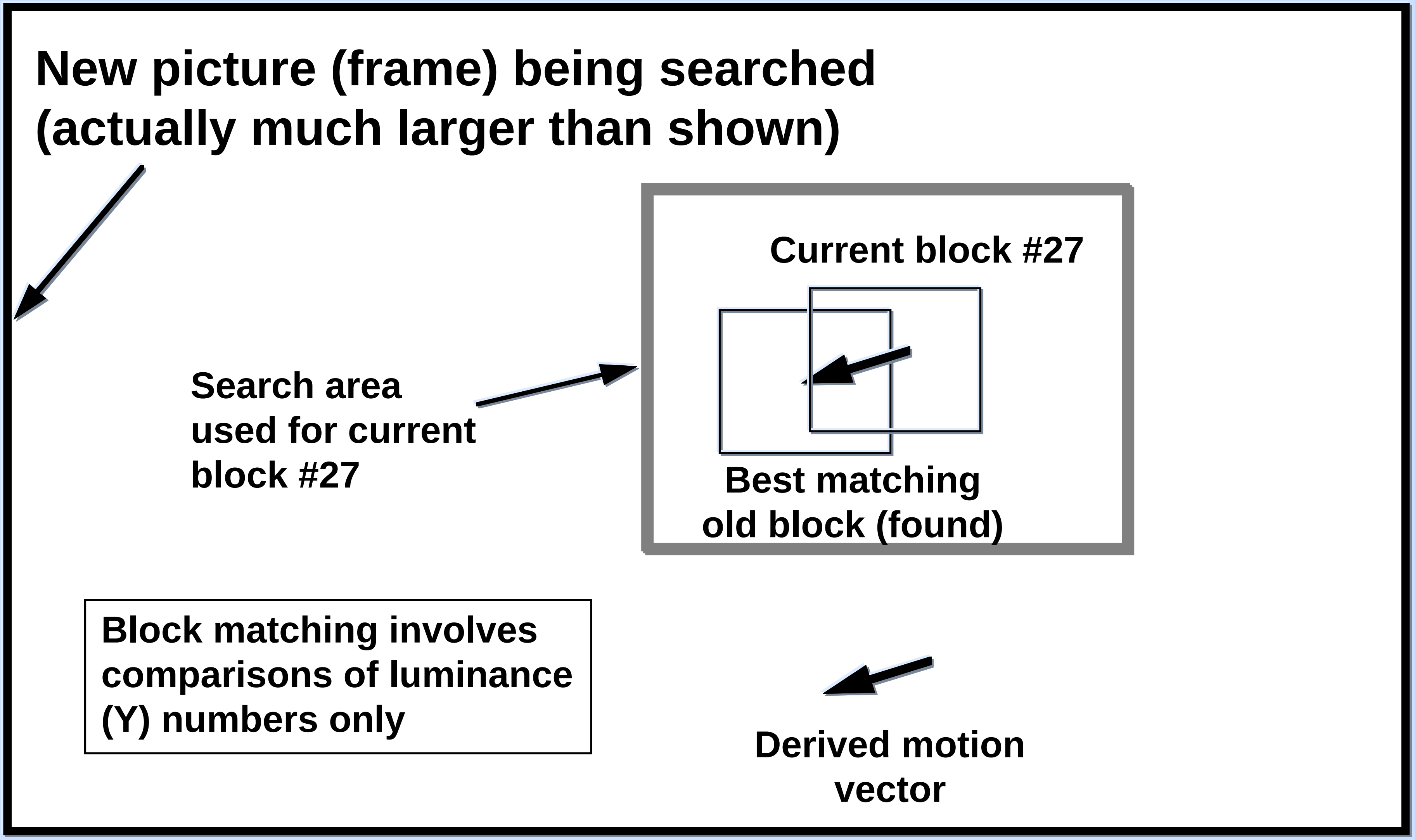
End

- **Values near upper left are encoded most precisely because they are most important visually**
- **Zeroes are most likely to be bunched up toward end of string and are not important visually**
- **Consecutive zero runs are coded very compactly**
- **Only 10 numbers instead of 64 need to be stored or transmitted in this example of one 8x8 pixel block**

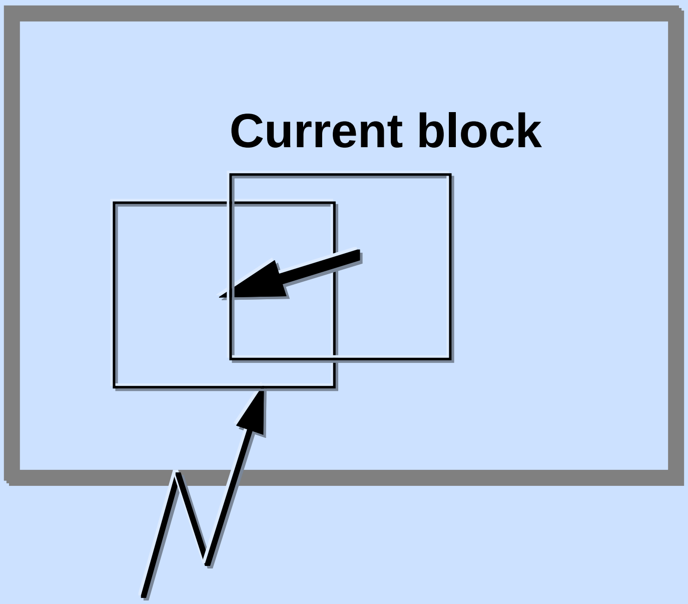
Intraframe Coding: “I-Picture” Summary

- **Previous pages have described a coding method for any picture without relying on other pictures**
  - **pictures coded this way are called Intraframes or I-pictures in MPEG-2**
- **The coding method typically yields a compression of about 15:1 for video frames**
  - **pictures without much detail compress more while detailed ones compress less**
- **The still-picture coding called JPEG and its video compression variant (Motion JPEG) just use intraframe coding**
- **But MPEG-2 goes further as described next**

Motion Estimation by Block Matching



FDCT of YUV Differences



Previously stored YUV numbers for the most similar old block, which are the estimate for the current block's YUV

- The motion vector points to the best old block for estimating the luminance and chrominance (YUV) numbers for current block
- The video encoder can also calculate the actual YUV numbers for the current block
- It then calculates the difference between the old block and actual current block YUV numbers
- It finally calculates the FDCT of the differences

Effectiveness of Motion Estimation

- **Effective because**
  - **motion vector (sent once for the search area) plus**
  - **FDCT of difference between estimated and actual current block YUV numbers**

  usually requires significantly fewer bytes than full FDCT of current blocks

- **The decoder recreates "actual" current block YUVs from the motion vector and FDCT of the differences**

- **If YUV estimate is perfect for a current block, only the vector needs to be sent or stored for the search area**

What if the Estimated YUVs Turn Out to be Totally Unsimilar to the Actual YUVs?

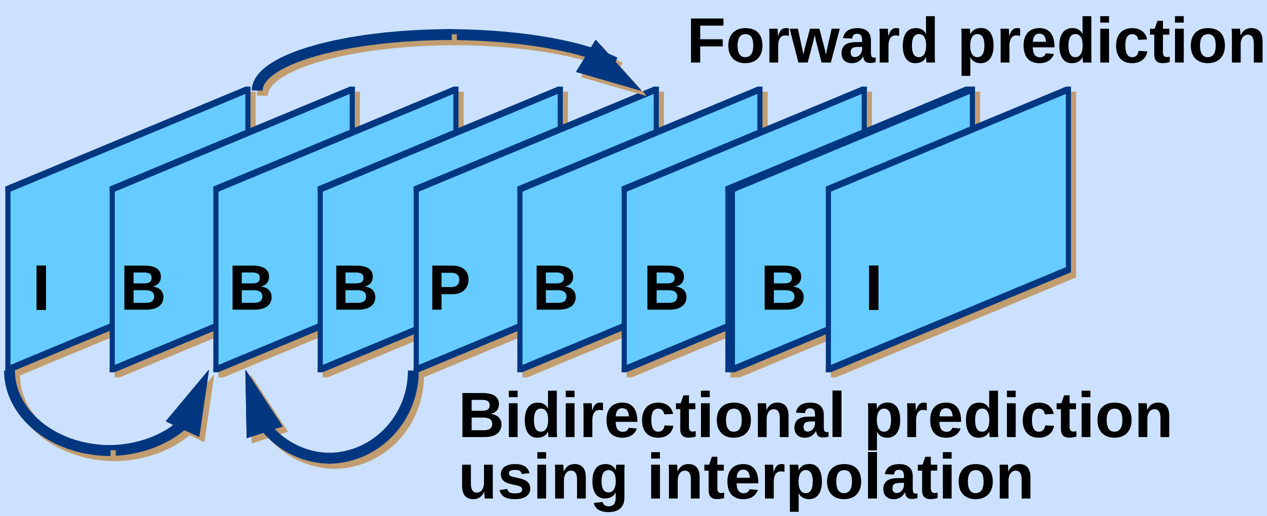
- **This will happen many times in any long video sequence because of scene changes and camera panning**
  - **but the frequency of occurrence is a tiny % of all blocks**
- **Here the FDCT of the YUV differences is of the same complexity as the actual current block FDCT (a minor penalty at worst)**
- **And the motion vector representation will add a few percent to the transmission and storage requirements (another minor penalty)**
- **The main penalty is because of the encoder processing in creating the motion vector and YUV differences—and that processing resulting in slightly worse compression**
  - **but those functions in the encoder could not be used for other tasks anyway**

- **Estimation backfires for some blocks**
- **But pays dividends overall**

Forward Predictive Motion Estimation: "P-Picture" Summary

- Previous pages have described a method for more compactly coding the YUV numbers for blocks in a current frame based on
  - finding the most similar earlier nearby block
  - derivation of a motion vector
  - having the YUV numbers for the earlier block
  - using those YUVs as predicted numbers for the current block
  - information in the encoder of the actual new block YUVs
  - calculating the FDCT of the differences in YUVs
- The resulting predicted frames are called P-pictures
- P-pictures can be predicted to advantage several frames into the "future" (i.e. 4 or 7)
- Decoded P-pictures are just as accurate as I-pictures if there is no total scene change in between

Bidirectional Motion Estimation: "B-Picture" Summary



- **Frames between I and P-pictures (or between two P-pictures) can be bidirectionally predicted**
  - such interpolated frames are called B-pictures
  - result is even greater savings in the coded bytes required for storage or transmission of B-pictures compared to P-pictures

Timing Recovery from Received Data

- The decoder video output must be kept essentially perfectly synchronized with
  - the encoder video input
  - the associated audio
- This becomes an issue when sending video/audio streams over jitter-inducing networks such as ATM
- A solution is offered by at least three approaches
  - for ATM CBR service, using network-specific encoders and decoders with FIFO buffers for jitter smoothing
  - for CBR and VBR service, using ATM adaptation layer assistance in the form of time stamps
  - for CBR and VBR, using a common network clock, if available end-to-end

Agenda

- Why digital video and video compression?
- Fundamentals of color video in five minutes
- Video compression tool kit for MPEG-2
- **An overview of MPEG-2**
- Fitting MPEG-2 on ATM
- NetCalc3 modeling of MPEG-2 and ATM

MPEG-2 (ISO/IEC DIS 13818)

- **A coding and compression method for moving pictures and associated audio for digital storage media, television broadcasting and communication**
  - MPEG-2 is a specification from the Moving Pictures Expert Group (MPEG) of ISO/IEC
  - MPEG-2 was created by MPEG and the Experts Group for ATM Video Coding in the ITU-T SG15
  - the specification is currently a draft international standard, formally called ISO/IEC 13818
  - MPEG-2 essentially replaces MPEG-1 (ISO/IEC 11172) for new applications

MPEG-2 versus MPEG-1

- MPEG-1 focused heavily on storage media
- MPEG-2 is a much broader standard which is compatible with MPEG-1
- MPEG-2 includes additions in several areas:
  - harmonization with HDTV, NTSC, PAL, SECAM
  - two coding structures: field and frame
  - multiple profiles and levels (picture formats, resolutions, scalability of picture quality)
  - more general color handling
  - multi-channel audio
  - ability to include private data
  - two data stream formats, including one suitable for ATM VBR: the transport stream

Basics of MPEG-2

- I, P and B-pictures may be used in a group
  - typical picture group sequence is IBBPBBPBB
  - block matching is done for 16x16 “macroblocks”
- The encoding process is as described in the previous section
  - encoder outputs must comply with 13818
  - significant flexibility is given for encoder design
- The decoding process, defined in 13818 in detail, follows encoding steps in reverse order and with inverse functions
- Timing recovery may use any of the three approaches covered earlier, or others

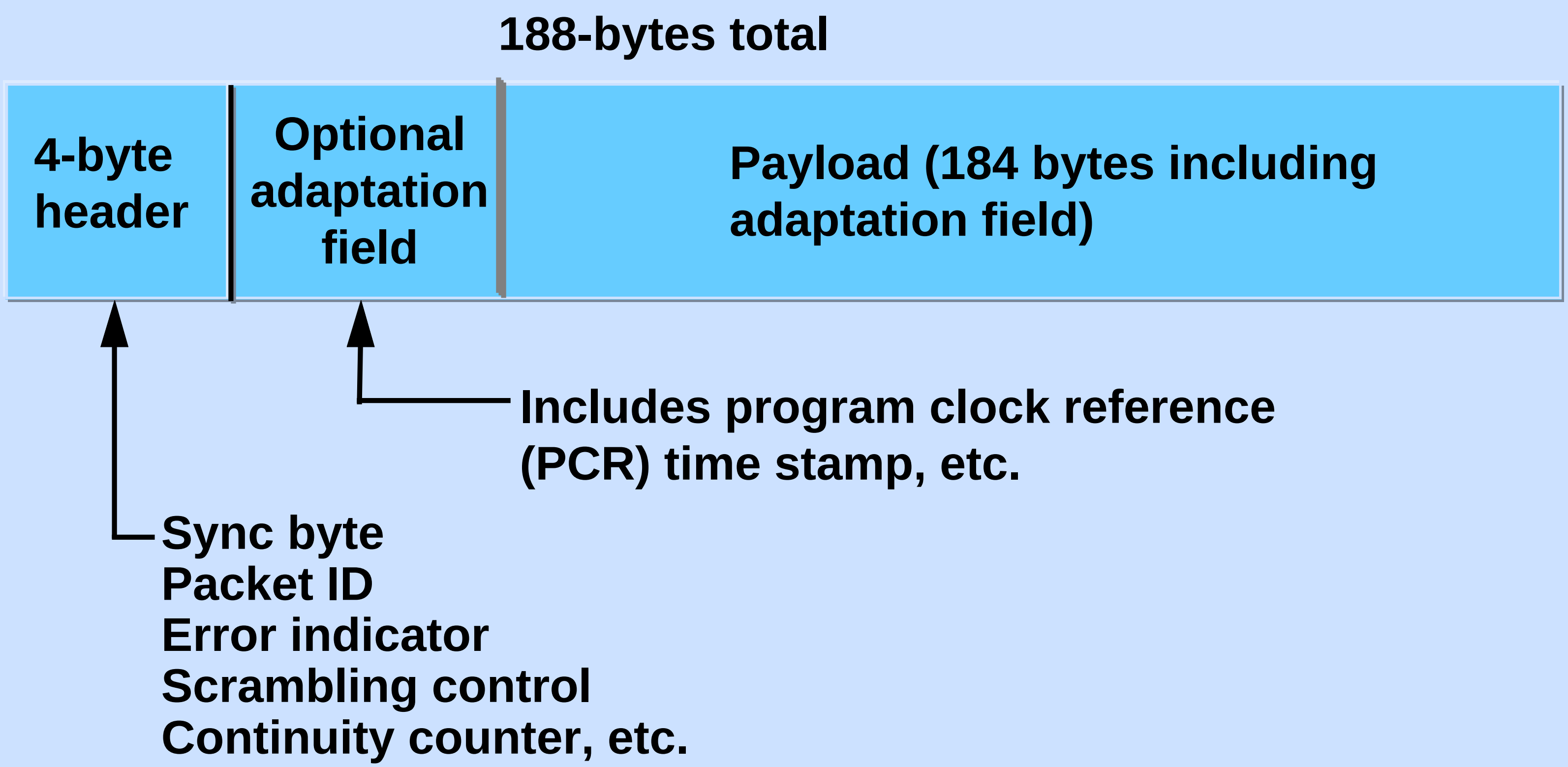
MPEG-2 Profile and Level Structure

Profile → / Level ↑	Simple No B-pict NS 4:2:0	Main With B-pict NS 4:2:0	SNR 2-Layer S 4:2:0	Spatial 3-Layer S 4:2:0	High 3-Layer S 4:2:0/4:2:2
High		1920*1152 60 fps 62.7 Msamp/s Up to 80 Mbps			1920*1152 60 fps 62.7/83.7 Msamp/s Up to 100 Mbps
High 1440		1440*1152 60 fps 47 Msamp/s Up to 60 Mbps		1440*1152 60 fps 47 Msamp/s Up to 60 Mbps	1440*1152 60 fps 47.0/62.7 Msamp/s Up to 80 Mbps
Main	720*576 30 fps 10.4 Msamp/s Up to 15 Mbps	720*576 30 fps 10.4 Msamp/s Up to 15 Mbps	720*576 30 fps 10.4 Msamp/s Up to 15 Mbps		720*576 30 fps 11/14.7 Msamp/s Up to 20 Mbps
Low		352*288 30 fps 3.04 Msamp/s Up to 4 Mbps	352*288 30 fps 3.04 Msamp/s Up to 4 Mbps		

MPEG-2 Systems Specification

- Systems specification, 13818-1, defines
  - coding, video/audio synchronization, multiplexing, clock recovery, buffer behavior, program time identification, and the two data stream types (PS and TS)
- Program Stream (PS)
  - carries a single program
    - one MPEG-2 “program” is a set of audio-visual streams (i.e. a set of TV programs) with a common time base
  - uses variable length packets typically 1 or 2 kbytes in size with a maximum of 64 kbytes
- Transport Stream (TS)
  - may carry multiple programs (with different time bases)
  - uses fixed length 188-byte packets
  - more suitable for ATM than PS

Transport Stream Packet Structure



Agenda

- Why digital video and video compression?
- Fundamentals of color video in five minutes
- Video compression tool kit for MPEG-2
- An overview of MPEG-2
- **Fitting MPEG-2 on ATM**
- NetCalc3 modeling of MPEG-2 and ATM

TS Packets to ATM for CBR (AAL-1)

AAL-1 protocol data unit (PDU) structure

SN-SNP (1 byte)	Payload (47 bytes)

- **TS packets are 188 bytes long**
- **Four AAL-1 cells (payloads 4x47 bytes) can carry one TS packet**
- **SN-SNP provide clock rate recovery**
  - SN is cell sequence number
  - SNP provides sequence number protection

TS Packets to ATM for VBR (AAL-5)

AAL-5 PDU structure suitable for MPEG-2 TS

Header (8 bytes)	Payload (8x47 or 376 bytes)

- **TS packets are 188 bytes long**
- **One AAL-5 PDU (payload 376 bytes) can carry two TS packets**
- **The AAL-5 PDU is carried by eight ATM cells**

Where to do Adaptation to ATM?

- **At the network interface**
  - this implies MPEG-2 packet format transmission to the network
    - › not a problem with private networks
    - › for public networks, requires access rates in an inconvenient speed range (4 Mbit/s to 45 Mbit/s)
- **Using customer equipment**
  - requires a local MPEG-2/ATM access device
  - network access is ATM UNI
- **Using mixture, for example:**
  - a local MPEG-2/ATM access device for coder end
  - 4 Mbit/s to 9 Mbit/s access at decoder end

Network traffic calculations are the same for all above cases

Agenda

- Why digital video and video compression?
- Fundamentals of color video in five minutes
- Video compression tool kit for MPEG-2
- An overview of MPEG-2
- Fitting MPEG-2 on ATM
- **NetCalc3 modeling of MPEG-2 and ATM**

Overview of NetCalc3 (for PC, Mac, Sun)

- A program used for “production” wide-area network design for data, voice and video
- Primarily used inside Nortel but will be available to Magellan customers at no charge under non-disclosure
- Finds best routes through any topology
- Does call-level voice engineering, calculates trunk delays and utilizations, path delays, delay variations and trunk costs
  - delay calculations include propagation, nodal, transmission and queuing for each of n(n-1) paths in a n-node network
  - does tariff calculations for several countries
- New in NetCalc3: MPEG-2 video offered traffic preprocessing and a few other features

NetCalc3 Modeling of MPEG-2 on ATM

- User selects
  - MPEG-2 video profile and level
  - CBR or VBR service
    - If CBR, the data rate in Mbit/s
- NetCalc3
  - provides video quality advice on CBR data rate choice
  - calculates CBR or VBR offered traffic estimate
  - incorporates the resulting cell rates into the entire network traffic calculation (along with voice and data)
  - produces a network performance and cost file for voice, data, and MPEG-2 video

Closing Comments

- Nortel's Magellan portfolio is designed to carry data, voice and video traffic, including MPEG-2
- Nortel has the experience and tools to assist customers in designing networks that carry MPEG-2 and other traffic
  - demonstration of NetCalc3 and MPEG-2 traffic entry and network design will be given in whiteboard clinics
  - a precalculated example will be handed out to interested attendees

Appendix A: Presentation References

- **References are listed in the notes**

Appendix B:
MPEG-2 and ATM Terminology

- **Commonly used terms or acronyms are explained in the notes**