# Multimedia Software Systems CS4551

Video Compression Standards

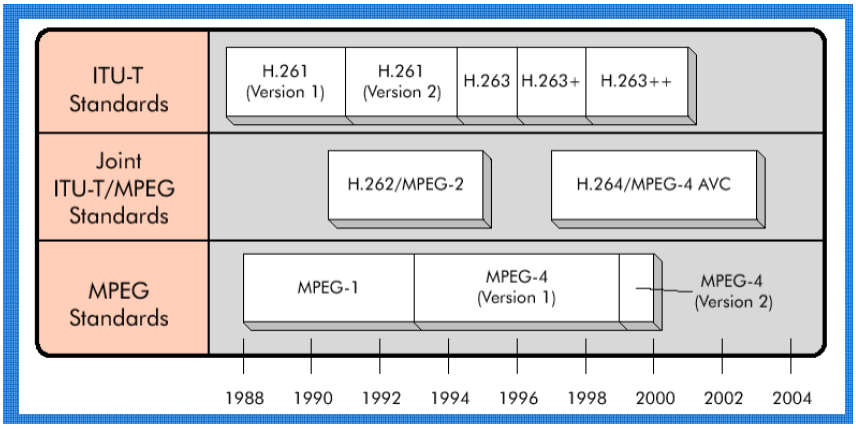


# Video Standards

- Video Standards and Implementation
  - Motion JPEG, Motion JPEG2000
  - ITU-T Standards : H.26x (e.g H.261, H.263, etc)
  - ISO Standards : MPEG-x (e.g MPEG-1, MPEG-2, etc)

- ITU (International Telecommunication Union)
  - ITU-T (ITU Telecommunication Standardization Sector)
  - ITU-T Video Coding Experts Group (VCEG)
- ISO (International Organization for Standardization)
  - MPEG (Motion Picture Experts Group)
- JPEG is a joint working group of the International Standardization Organization (ISO) and the International Electrotechnical Commission (IEC)

# Video Standards (2)



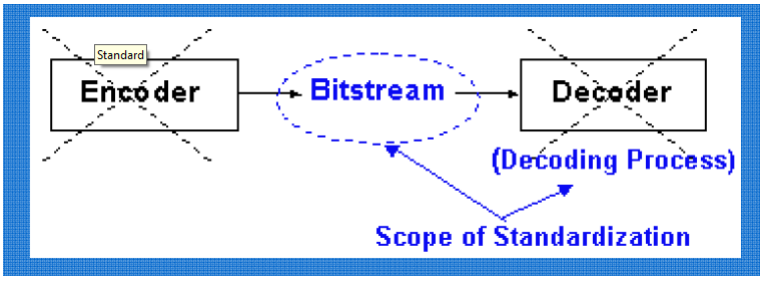



# What Do the Standards Specify?

- A video compression system consists of the following:
  - An encoder
  - Compressed bit-streams
  - A decoder
- What parts of the system do the standards specify?

# What Do the Standards Specify?

- Not the encoder, not the decoder.






# What Do the Standards Specify?

- Just the bit-stream syntax and the decoding process, for example it tells to use IDCT, but not how to implement the IDCT.
- Enables improved encoding and decoding strategies to be employed in a standard-compatible manner.

# Video Standards

| STANDARD | APPLICATION | BIT RATE |
|---|---|---|
| Motion JPEG | Continuous-tone still-image compression | Variable |
| H.261 | Video telephony and teleconferencing over ISDN | p x 64 kb/s |
| MPEG-1 | Video on digital storage media (CD-ROM) | 1.5 Mb/s |
| MPEG-2 | Digital Television | > 2 Mb/s |
| H.263 | Video telephony over PSTN | < 33.6 kb/s |
| MPEG-4 | Object-based coding, synthetic content, interactivity | Variable |
| H.264 | From Low bitrate coding to HD encoding, HD-DVD, Surveillance, Video conferencing. | Variable |



# Motion JPEG

- Motion JPEG
  - A digital video sequence can be represented as a series of JPEG pictures. The advantages are the same as with single still JPEG pictures – flexibility both in terms of quality and compression ratio.
  - The main disadvantage of Motion JPEG (a.k.a. MJPEG) is that since it uses only a series of still pictures it makes no use of video compression techniques. The result is a slightly lower compression ratio for video sequences compared to “real” video compression techniques.

# Motion JPEG-2000

- Motion JPEG-2000
  - As with JPEG and Motion JPEG, JPEG 2000 can also be used to represent a video sequence. The advantages are equal to JPEG 2000, i.e., a slightly better compression ratio compared to JPEG but at the price of complexity.
  - The disadvantage reassembles that of Motion JPEG. Since it is a still picture compression technique it doesn't take any advantages of the video sequence compression. This results in a lower compression ration compared to real video compression techniques.



# Main Proposed Digital Video Standards

- ITU standards
  - H.261 (videoconferences over ISDN)
  - H.263 (videoconferencing and video telephony over POTS)
- ISO standards
  - MPEG-1 Part 2(movies on CD-ROM)
  - MPEG-4 Part 2(more versatile distribution)
- Joint Effort
  - H.262/MPEG-2 Part 2(digital television, movies on DVD)
  - H.264/MPEG-4 Part 10 (AVC - Advanced Video Coding, HD-DVD)

# Video Compression – ITU-T Standards



# ITU H.261

- The first member of the H.26x family of video coding standards.
- The ITU H.261 standard was initially designed for Integrated Services Digital Network (ISDN) and was intended to support video conferencing applications, which have relatively small amounts of motion (mainly head and shoulder movements).
- It supports the following features
  - Produces bit-rates of *p* x 64 Kbps (*p* ranges 1 to 30)
  - Requires the video encoders delay to be less than 150ms so that the video can be used for real-time bidirectional teleconferencing
  - Only non-interlaced video
  - Use YCbCr color model and performs 4:2:0 chroma subsampling
  - Only CIF (352x288 for Y) and QCIF (176x144) formats
  - Can encode in *intraframe* and *interframe* mode

# Common Intermediate Format (CIF)

- **QCIF** (Quarter CIF)
- **SQCIF** (Sub Quarter CIF)
- **4CIF** (4× CIF)
- **16CIF** (16× CIF)

| Format | Video Resolution |
|---|---|
| SQCIF | 128 × 96 |
| QCIF | 176 × 144 |
| CIF | 352 × 288 |
| 4CIF | 704 × 576 |
| 16CIF | 1408 × 1152 |

- Format used to standardize the horizontal and vertical resolutions in pixels of YCbCr sequences in video signals, commonly used in video teleconferencing systems. It was first proposed in the H.261 standard.
- The CIF "image sizes" were specifically chosen to be multiples of macroblocks (i.e. 16x16 pixels) for compression.

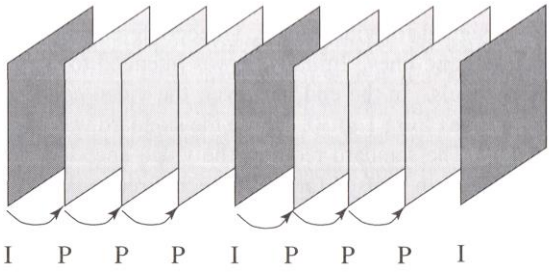


# ITU H.261

Video formats supported by H.261.

| Video format | Luminance image resolution | Chrominance image resolution | Bitrate (Mbps) (if 30 fps and uncompressed ) | H.261 support |
|---|---|---|---|---|
| QCIF | 176 × 144 | 88 × 72 | 9.1 | Required |
| CIF | 352 × 288 | 176 × 144 | 36.5 | Optional |

# ITU H.261



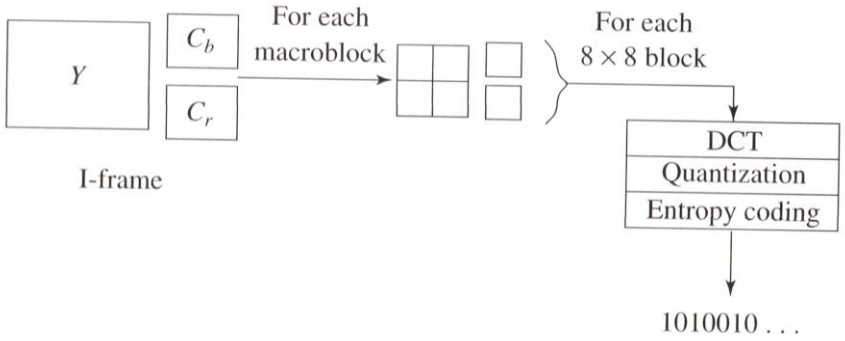

- Interval between I-frames is a variable and is determined by the encoder. Usually a couple of I-frames per second.
- Motion vectors is measured in pixel unit and ranges -15 ≤ vector ≤ 15



# ITU H.261

- ***Intraframe (I-frame)*:**
  - treated as independent image
  - Macroblock : 16 x 16 for Y, 8x8 for Cb|Cr
  - DCT on 8x8 blocks (like JPEG). Y macro block will be divided into 4 8x8 blocks.
  - Quantization
    - Use a constant (not 8x8 quantization table). 8 for DC coefficients, 2x*scale* for AC coefficients where *scale* is a value in the range of [1,31].
    - `QDCT = round (DCT/8)` for DC
    - `QDCT = floor(DCT/(2xscale))` for AC
  - Zigzag code and entropy coding (arithmetic coding)
  - Usually H.261 generates a couple of *I-frames* per second.

# ITU H.261

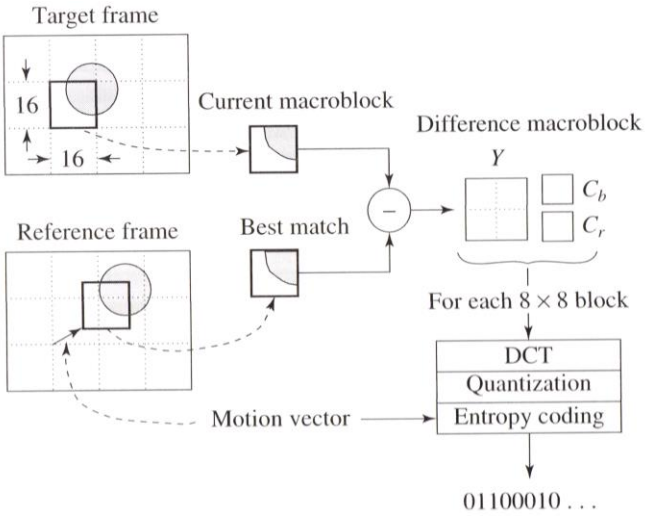





# ITU H.261

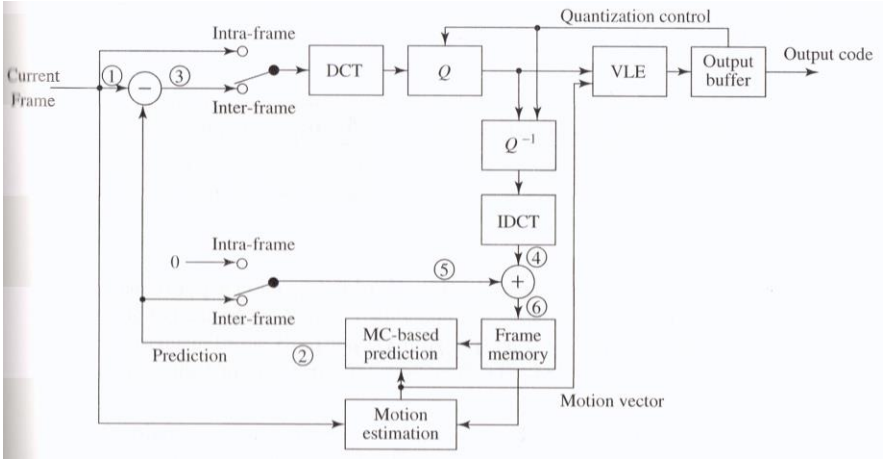
- ***Interframe(P-frame)*:**
  - Not independent. Coded by a forward predictive method which current macroblocks are predicted from the previous *I-frame* or *P-frame*
  - Prediction - Find the best matching block and computes motion vectors on 16x16 macroblocks from the previous frame. This previous frame is called the *reference frame* and it is encoded as *I-frame* or *P-frame.*
  - Encode the difference macroblock (residual image)
  - Encode the **difference motion vectors**. (MVD = $MV_{preceding} - MV_{current}$)
  - If the best matching can not be found (i.e. the error exceeds a certain acceptance level), the macroblock itself is then encoded.
  - Quantization (same as the *I-frame*)
  - Zigzag code and entropy coding

# ITU H.261

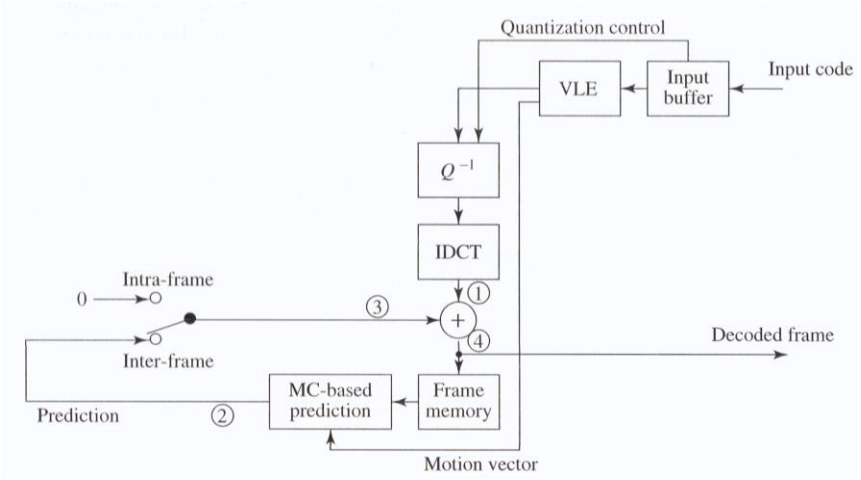





# ITU H.261



# ITU H.261






# ITU H.263

- Improved video coding standard for video conferencing and other audio-visual services transmitted on POTS (Plain Old Telephone System) or PSTN (Public Switched Telephone Networks)
- H.263 is part of the H.324 standard for communication over POTS with a modem with a maximum available rate of 33.6 Kb/s and a normal available bit-rate of 26-28 Kb/s
- Supports a wider range of picture formats, including 4CIF (704x576 for Y) and 16CIF (1408x1152) in addition to QCIF and CIF.

# ITU H.263

Video formats supported by H.263.

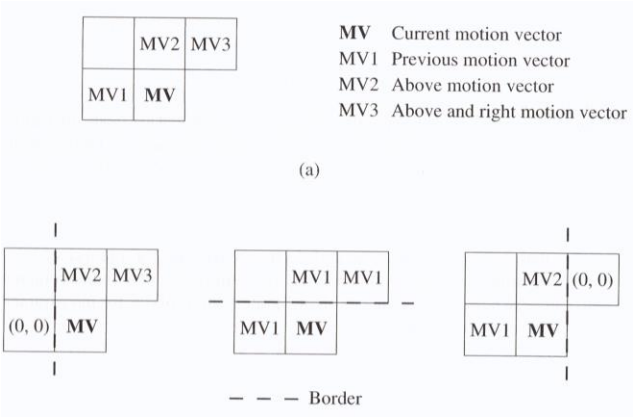
| Video format | Luminance image resolution | Chrominance image resolution | Bitrate (Mbps) (if 30 fps and uncompressed) | Bitrate (kbps) BPPmaxKb (compressed) |
|---|---|---|---|---|
| Sub-QCIF | 128 × 96 | 64 × 48 | 4.4 | 64 |
| QCIF | 176 × 144 | 88 × 72 | 9.1 | 64 |
| CIF | 352 × 288 | 176 × 144 | 36.5 | 256 |
| 4CIF | 704 × 576 | 352 × 288 | 146.0 | 512 |
| 16CIF | 1408 × 1152 | 704 × 576 | 583.9 | 1024 |



# ITU H.263

- Based on the same DCT, motion compensation and variable-length coding (using a Run-Level variable-length code) technique used in H.261.
- Incremental improvements are
  - Coding error vector instead of the motion vector itself
    - The prediction starts from taking median values of the motion vectors of "previous", "above", and "above-right" macroblocks.
    - Motion vector for current macroblock : MV(u,v)
    - Motion vector for previous macroblock : MV1(u1, v1)
    - Motion vector for above macroblock : MV2(u2,v2)
    - Motion vector for above-right macroblock : MV3(u3, v3)
    - up = median(u1, u2, u3)
    - vp = median(v1, v2, v3)
    - Instead of coding MV(u,v), compute the error vector (eu, ev) = (u-up, v-vp) and encode

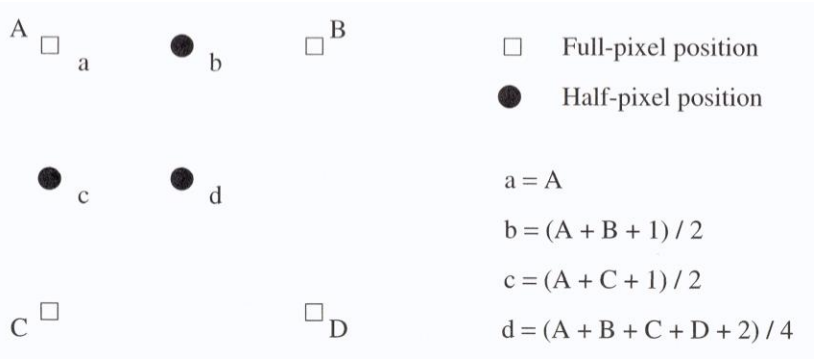
# ITU H.263



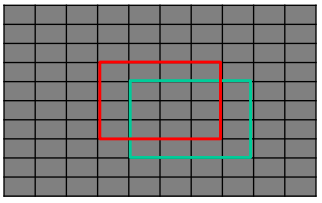



# ITU H.263

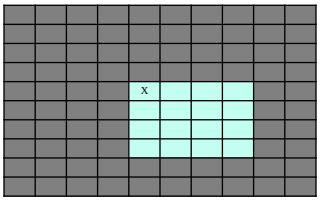
- Incremental improvements are
  - Half-pixel motion compensation to reduce the prediction error



# Pixel-accuracy Motion Compensation



The Reference frame

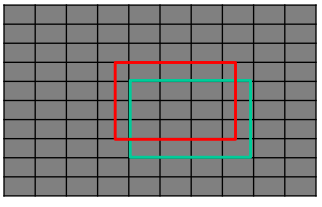



A target macro block at (4,4)
at the current frame

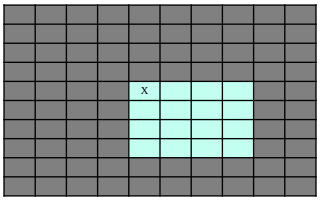
Matching with a candidate/reference block at d=(-1,-1)



# Half-pixel Motion Compensation



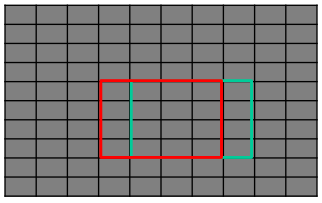
The Reference frame



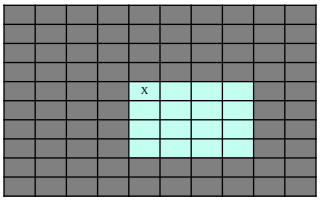

A target macro block at (4,4)
at the current frame

Matching with a candidate/reference block at d=(-0.5,-1)

# Pixel-accuracy Motion Compensation



The Reference frame

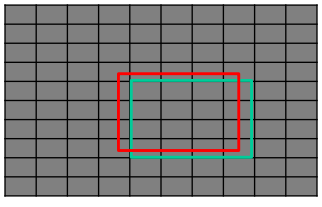



A target macro block at (4,4)
at the current frame

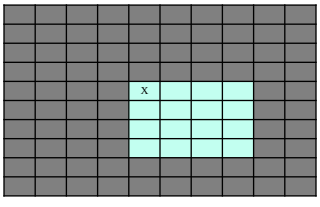
Matching with a candidate/reference block at d=(-1, 0)



# Half-pixel Motion Compensation



The Reference frame

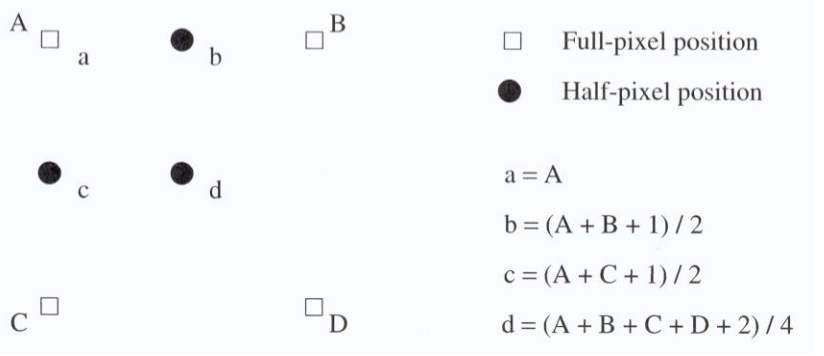



A target macro block at (4,4)
at the current frame

Matching with a candidate/reference block at d=(-0.5,-0.5)

# Generating Half-pixel Motion Block

- Depending on the motion, generate pixels using a through d.






# ITU H.263

- Optional coding mode of H.263
  - Unrestricted motion vector mode. Range is [-31.5~31.5]
  - Syntax-based arithmetic coding mode
  - Advanced motion prediction mode
    - Overlapped block motion estimation (In luminance image, 16x16 block is divided into four 8x8 and produce 4 motion vectors to give precise description of motions -> better compression)
  - ***B-frame mode*** (version 2) –frame encoded by ***Bi-directional motion estimation***

# H.263+ and H.263++

- Fully backward compatible with H.263
- Developed after MPEG-1 and 2 so it adopts many aspects of the MPEG such as SNR scalability and better bidirectional prediction.



# H.264 & H.265

- Will be covered after MPEG compression standards.