MPEG: A Video Compression Standard for Multimedia Applications

Václav Hlaváč

CTU Prague, hlavac@cmp.felk.cvut.cz

Initial material were slides of Didier Le Gall, Worcherster Polytechnic Institute.

Introduction

- 1980's technology made possible full-motion video over networks
  - Television and Computer Video seen moving closer
  - (Today, Sony and Microsoft are squaring off)
- Needed a standard
  - Often, triggers needed volume production
    - Ala *facsimile* (fax)
  - Avoid *de facto* standard by industry
- 1988, Established the Motion Picture Experts Group (MPEG)
  - Worked towards MPEG-1
  - Primarily video but includes audio (MP3)

The Need for Video Compression

- High-Definition Television (HDTV)
  - 1920x1080
  - 30 frames per second (full motion)
  - 8 bits for each three primary colors (RGB)

  →Total 1.5 Gb/sec!
- Cable TV: each cable channel is 6 MHz
  - Max data rate of 19.2 Mb/sec
  - Reduced to 18 Mb/sec w/audio + control …

  →Compression rate must be ~ 80:1!

Compatibility Goals

- 1990: CD-ROM and DAT key storage devices
  - 1-2 Mbits/sec for 1x CD-ROM
- Two types of application videos:
  - Asymmetric (encoded once, decoded many times)
    - Video games, Video on Demand
  - Symmetric (encoded once, decoded once)
    - Video phone, video mail …
- (*How do you think the two types might influence design*?)
- Video at about 1.5 Mbits/sec
- Audio at about 64-192 kbits/channel

Requirements

- Random Access, Reverse, Fast Forward, Search
  - At any point in the stream (within ½ second)
  - Can reduce quality somewhat during this task, if needed
- Audio/Video Synchronization
- Robustness to errors
  - Not catastrophic if some bits are lost
  - Lends itself to Internet streaming
- Coding/Decoding delay under 150 ms
  - For interactive applications
- Ability to Edit
  - Modify/Replace frames

Relevant Standards

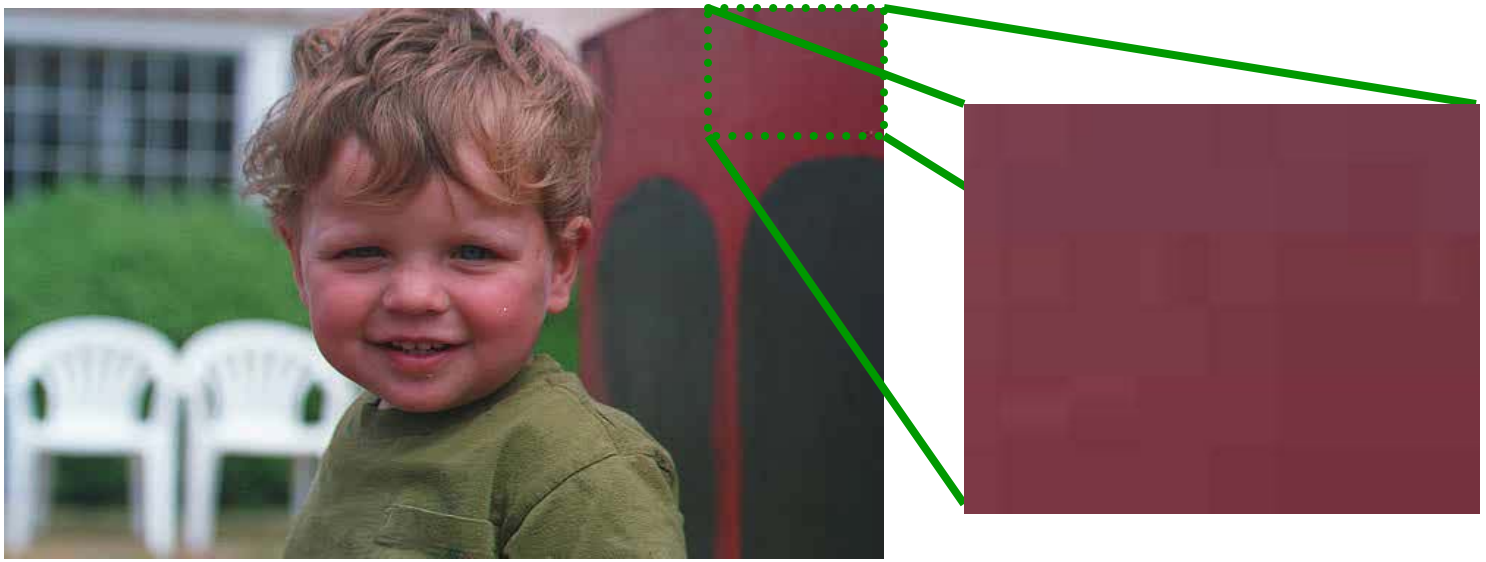
- Joint picture Experts Group (JPEG)
  - Compress still images only
- Expert Group on Visual Telephony (H.261)
  - Compress sequence of images
  - Over ISDN (64 kbits/sec)
  - Low-delay
- Other high-bandwidth “H” standards:
    - H21 (34 Mbits/sec)
    - H22 (45 Mbits/sec)

MPEG Compression

- Compression through
  - Spatial
  - Temporal

Spatial Redundancy

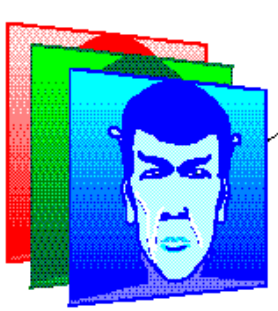
- Take advantage of similarity among most neighboring pixels



Spatial Redundancy Reduction

- RGB to YUV
  - less information required for YUV (humans less sensitive to chrominance)
- Macro Blocks
  - Take groups of pixels (16x16)
- Discrete Cosine Transformation (DCT)
  - Based on Fourier analysis where represent signal as sum of sine's and cosine’s
  - Concentrates on higher-frequency values
  - Represent pixels in blocks with fewer numbers
- Quantization
  - Reduce data required for co-efficients
- Entropy coding
  - Compress

Spatial Redundancy Reduction



RGB

YUV

Blocks

Pixel amplitude values

223	191	159	128	98	72	39	16
223	191	159	128	98	72	39	16
223	191	159	128	98	72	39	16
223	191	159	128	98	72	39	16
223	191	159	128	98	72	39	16
223	191	159	128	98	72	39	16
223	191	159	128	98	72	39	16
223	191	159	128	98	72	39	16

Discrete Cosine Transform

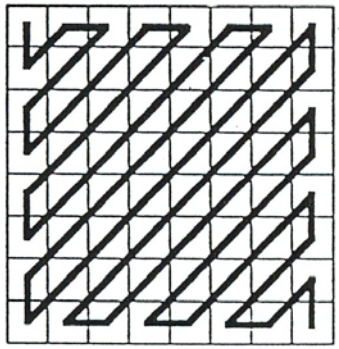
Pixel DCT Values

43.8	-40	0	-4.1	0	-1.1	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

“Intra-Frame Encoded”

Quantization
- major reduction
- controls ‘quality’

Zig-Zag Scan, Run-length coding



Question

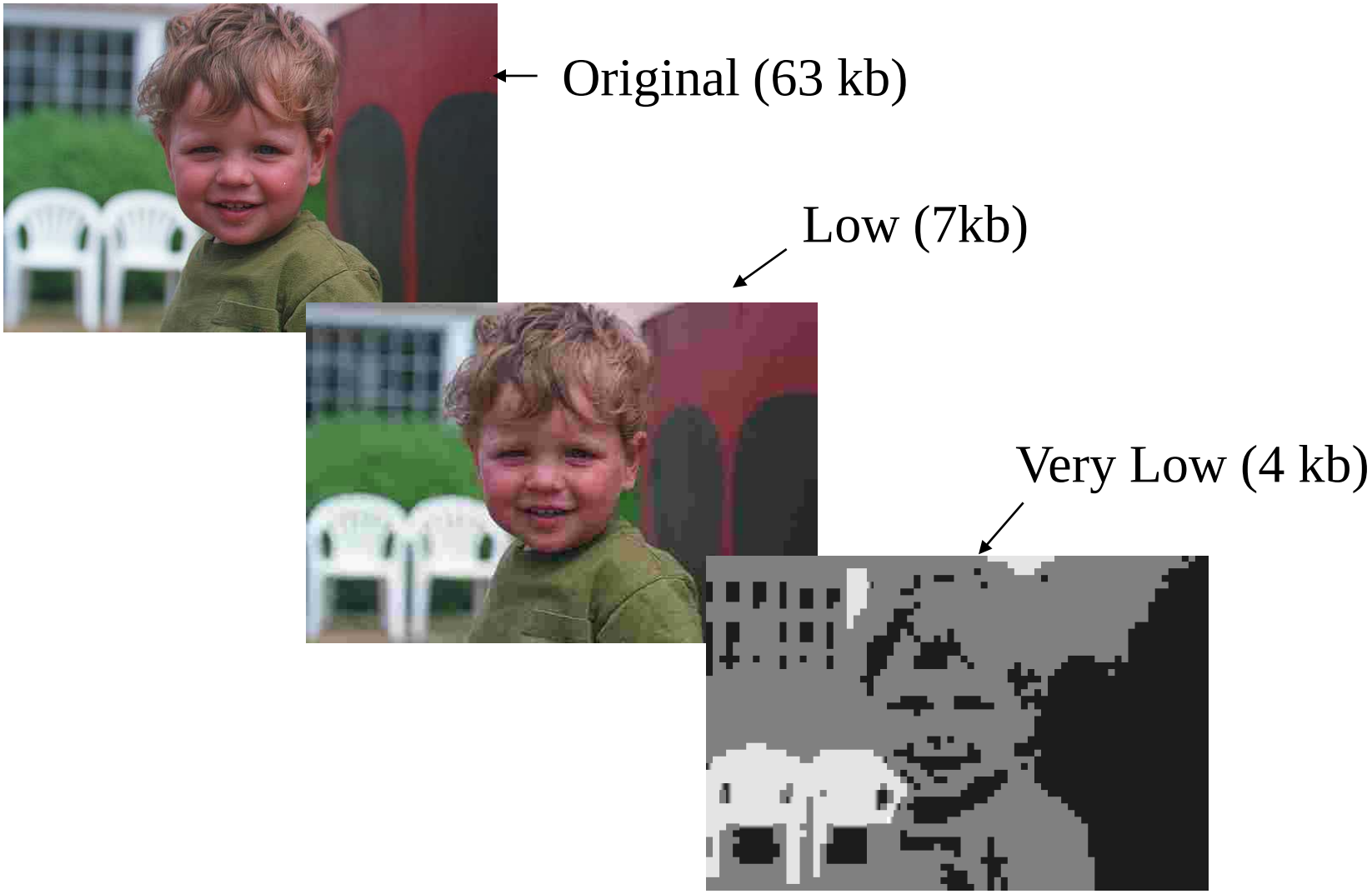
- *When may spatial redundancy reduction be ineffective?*

- What kinds of images/movies?

Answer

- *When may spatial redundancy elimination be ineffective?*
  - High-resolution images and displays
    - May appear ‘coarse’
- What kinds of images/movies?
  - A varied image or ‘busy’ scene
    - Many colors, few adjacent

Loss of Resolution



Temporal Redundancy

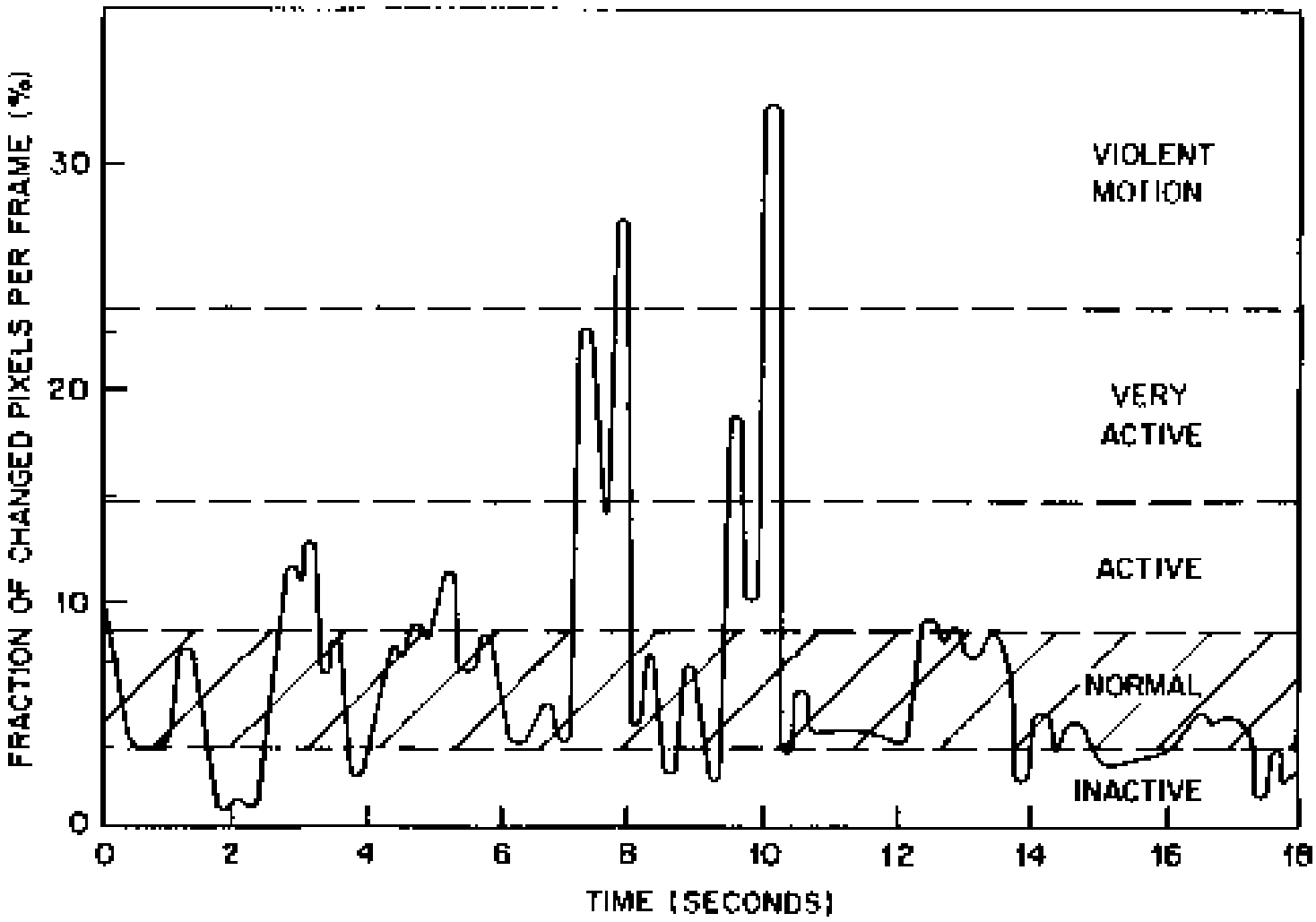
- Take advantage of similarity between successive frames

950

951

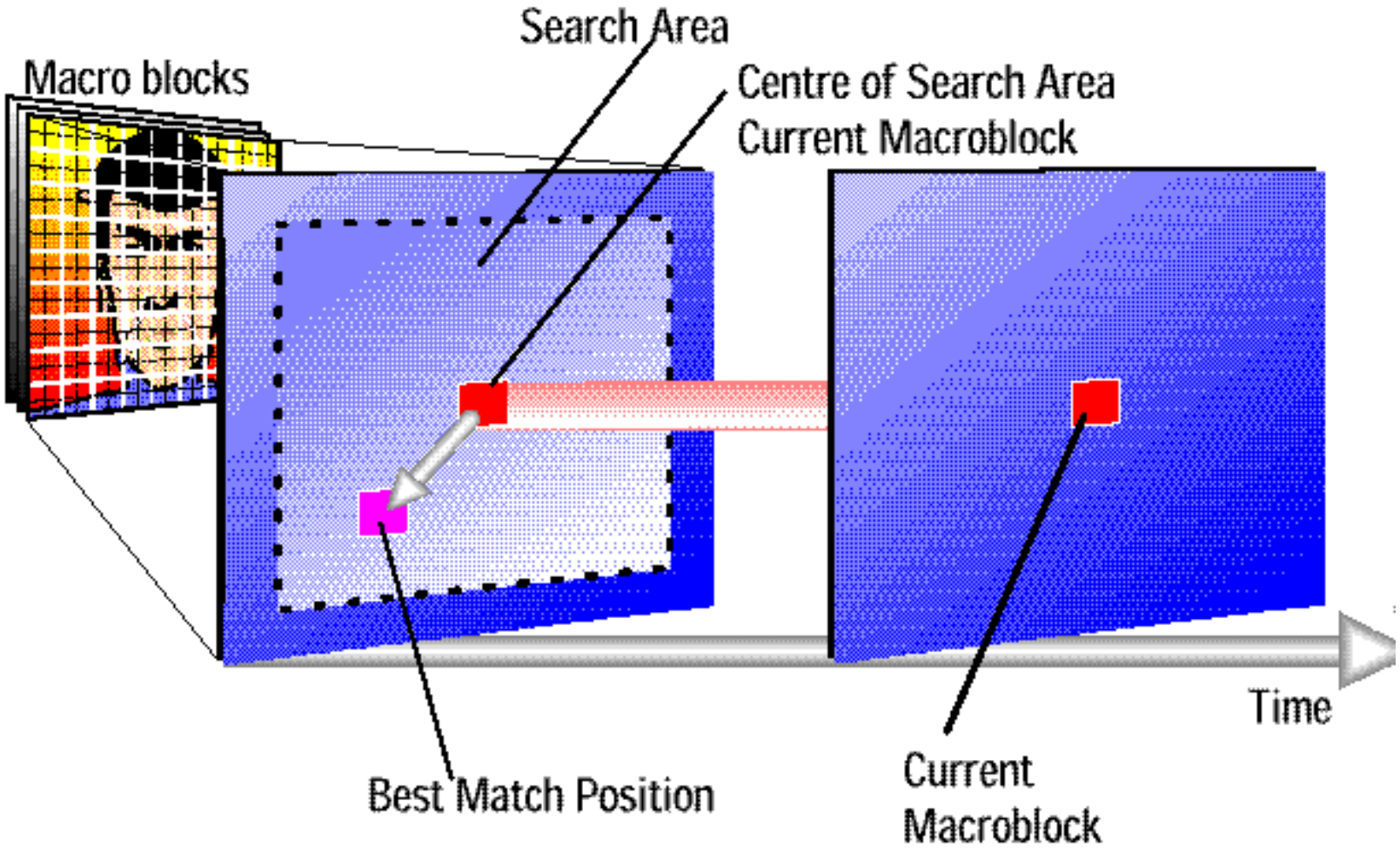
952

Temporal Activity

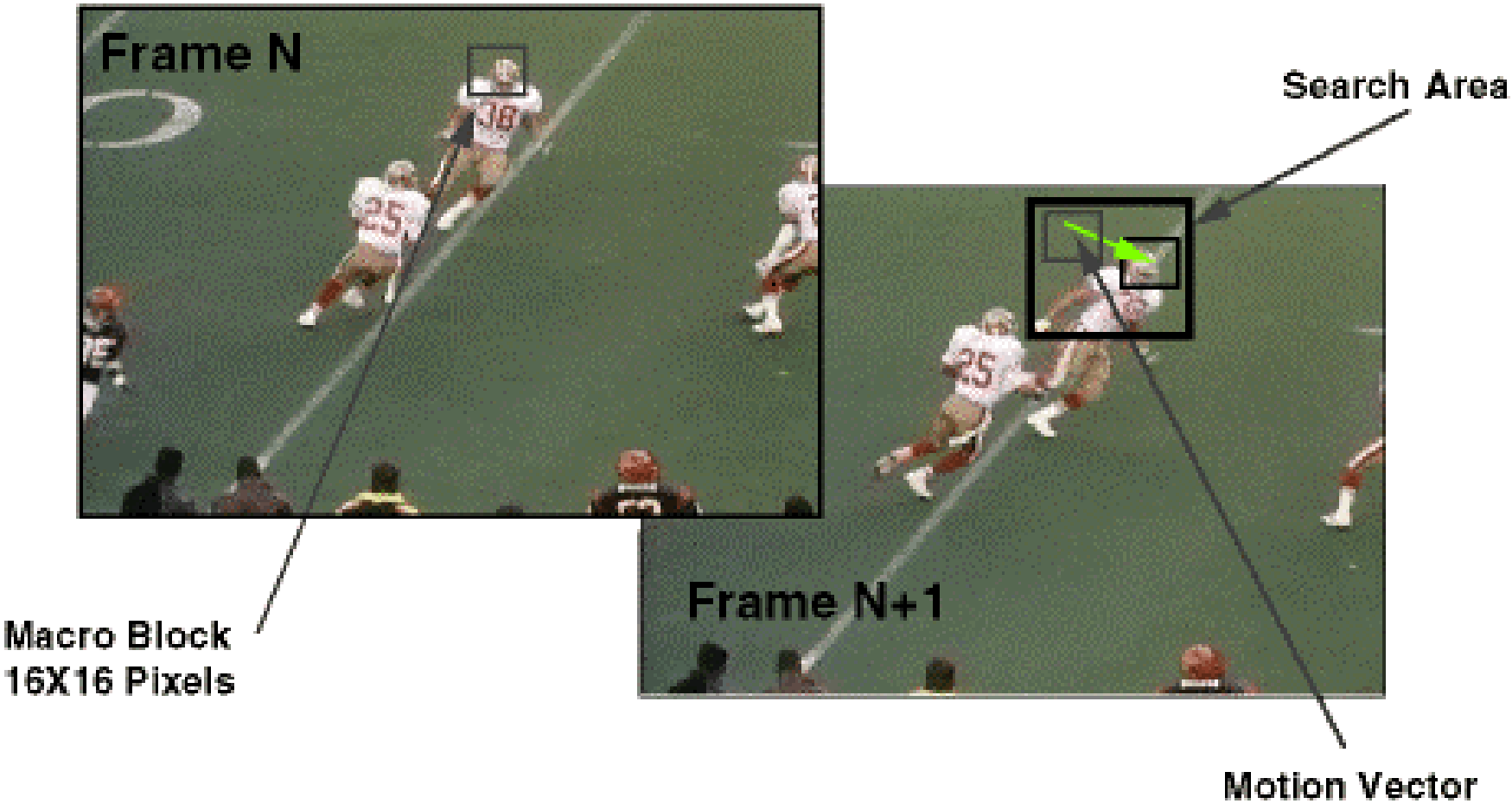


"Talking Head"

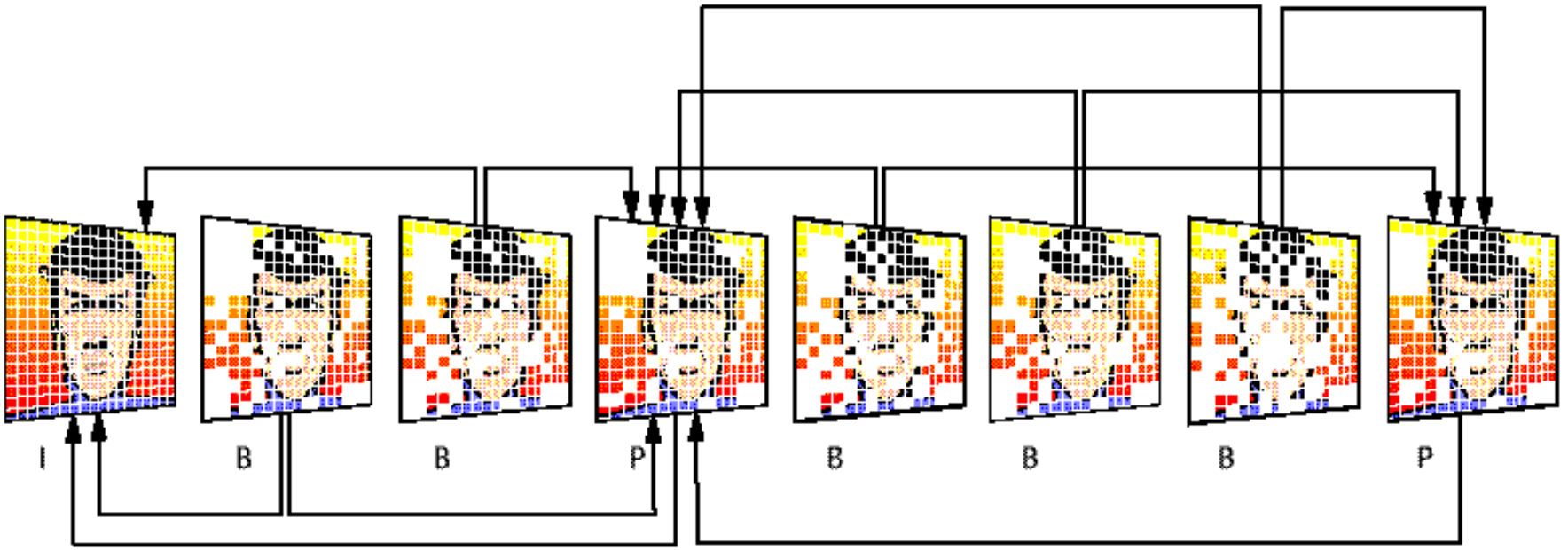
Temporal Redundancy Reduction



Temporal Redundancy Reduction



Temporal Redundancy Reduction



- *I* frames are independently encoded
- *P* frames are based on previous I, P frames
  - Can send motion vector plus changes
- *B* frames are based on previous and following I and P frames
  - In case something is uncovered

Group of Pictures (GOP)

- Starts with an I-frame
- Ends with frame right before next I-frame
- “Open” ends in B-frame, “Closed” in P-frame
  - (What is the difference?)
- MPEG Encoding a parameter, but ‘typical’:
  - I B B P B B P B B I
  - I B B P B B P B B P B B I
- *Why not have all P and B frames after initial I?*

Question

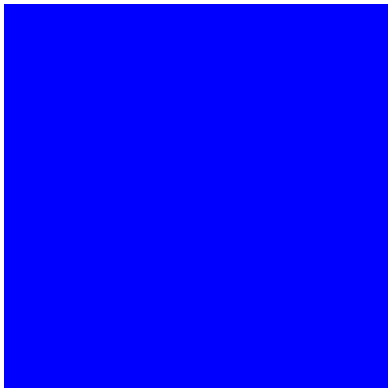
- When may temporal redundancy reduction be ineffective?

Answer

- *When may temporal redundancy reduction be ineffective?*
  - Many scene changes
  - High motion

Non-Temporal Redundancy

- Many scene changes vs. few scene changes



Non-Temporal Redundancy

- Sometimes high motion

Typical MPEG Parameters

Parameters	Value
Image resolution	384x288
Quantization factor	8
Frames between I pictures	5
Frames between P pictures	2
Frames sequence as to be displayed	...IBBPBBI...
Rate control	None

Typical Compress. Performance

Type	Size	Compression
I	18 KB	7:1
P	6 KB	20:1
B	2.5 KB	50:1
Avg	4.8 KB	27:1

Note, results are Variable Bit Rate, even if frame rate is constant

MPEG Today

- MPEG video compression widely used
  - digital television set-top boxes
- HDTV decoders
  - DVD players
  - video conferencing
  - Internet video
  - ...

MPEG Today

- MPEG-2
  - Super-set of MPEG-1
  - Rates up to 10 Mbps (720x486)
  - Can do HDTV (no MPEG-3)
- MPEG-4
  - Around *Objects*, not *Frames*
  - Lower bandwidth
  - Has some built-in repair (header redundancy)
- MPEG-7
  - New standard
  - Allows content-description (ease of searching)
- MP3, for audio
  - MPEG Layer-3

MPEG Tools

- MPEG tools at:
  - http://www-plateau.cs.berkeley.edu/mpeg/index.html
- MPEG streaming at:
  - http://www.comp.lancs.ac.uk/
- FFMPEG
  - http://ffmpeg.sourceforge.net/index.org.html