CLARITY
clarity-centre.org

The Basics

AV Capture, Representation and Compression

Digital AV Capture

CLARITY
clarity-centre.org

An image is captured when a camera scans a scene

Colour => Red (R), Green (G) and Blue (B) array of samples

Density of samples (pixels) gives resolution

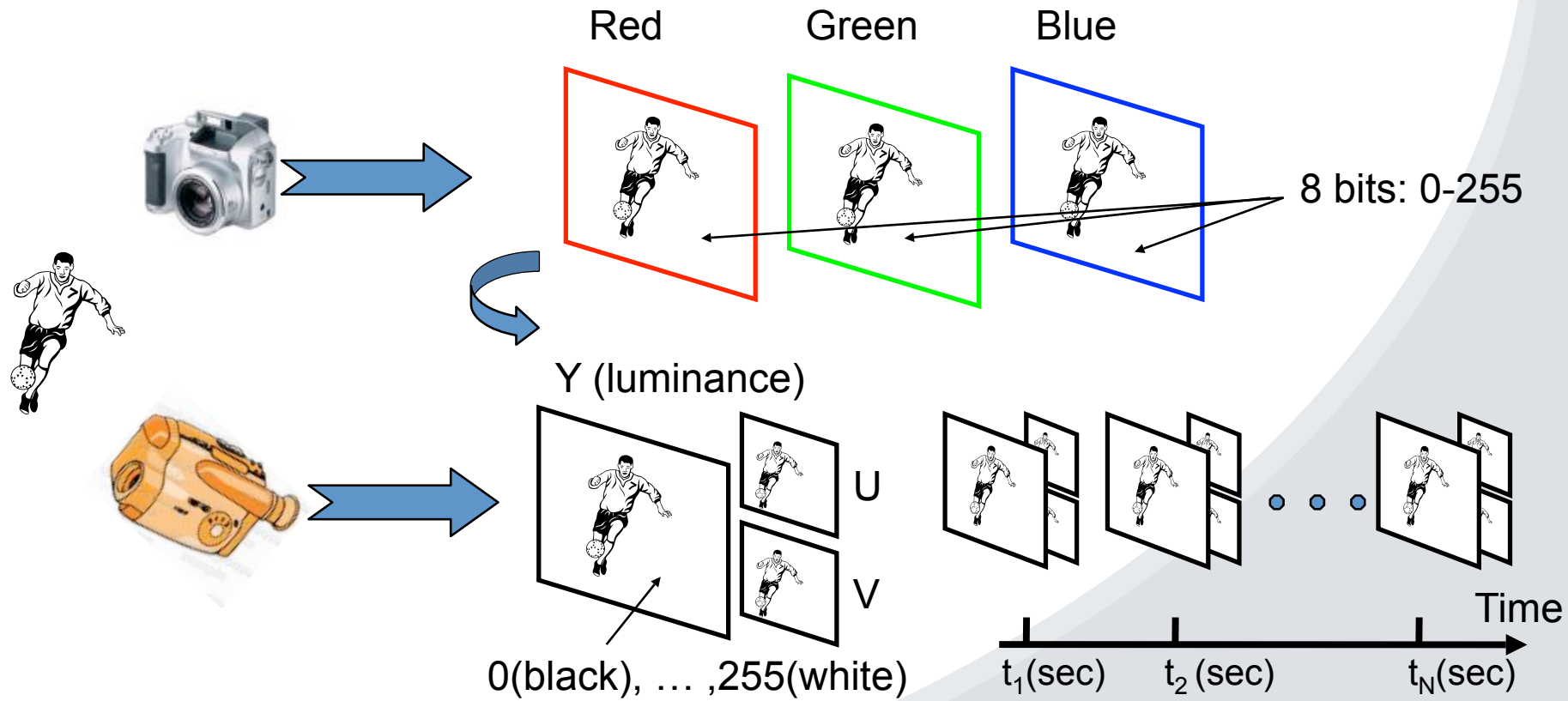
A video is captured when a camera scans a scene at multiple time instants

Each sample is called a frame giving rise to a frame rate (frames/sec) measured in Hz

- TV (full motion video) is 25Hz
- Mobile video telephony is 8-15 Hz … jerky

Audio is captured when a microphone temporally samples sound waves

Digital AV Capture



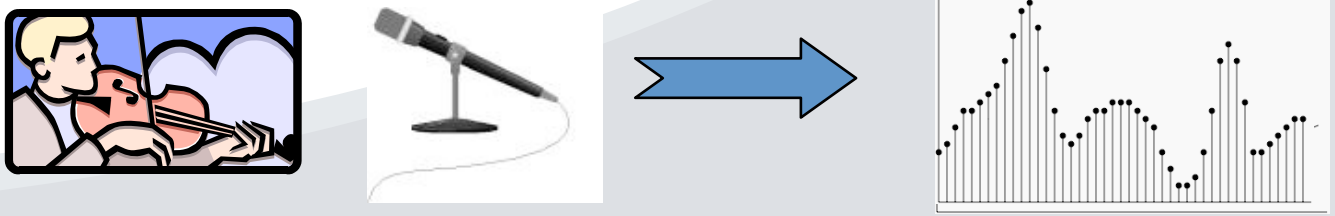


Image Data (RGB)

CLARITY
clarity-centre.org

Colour still image:

420 x 315 pixels, 8 bits/ pixel = 387KB

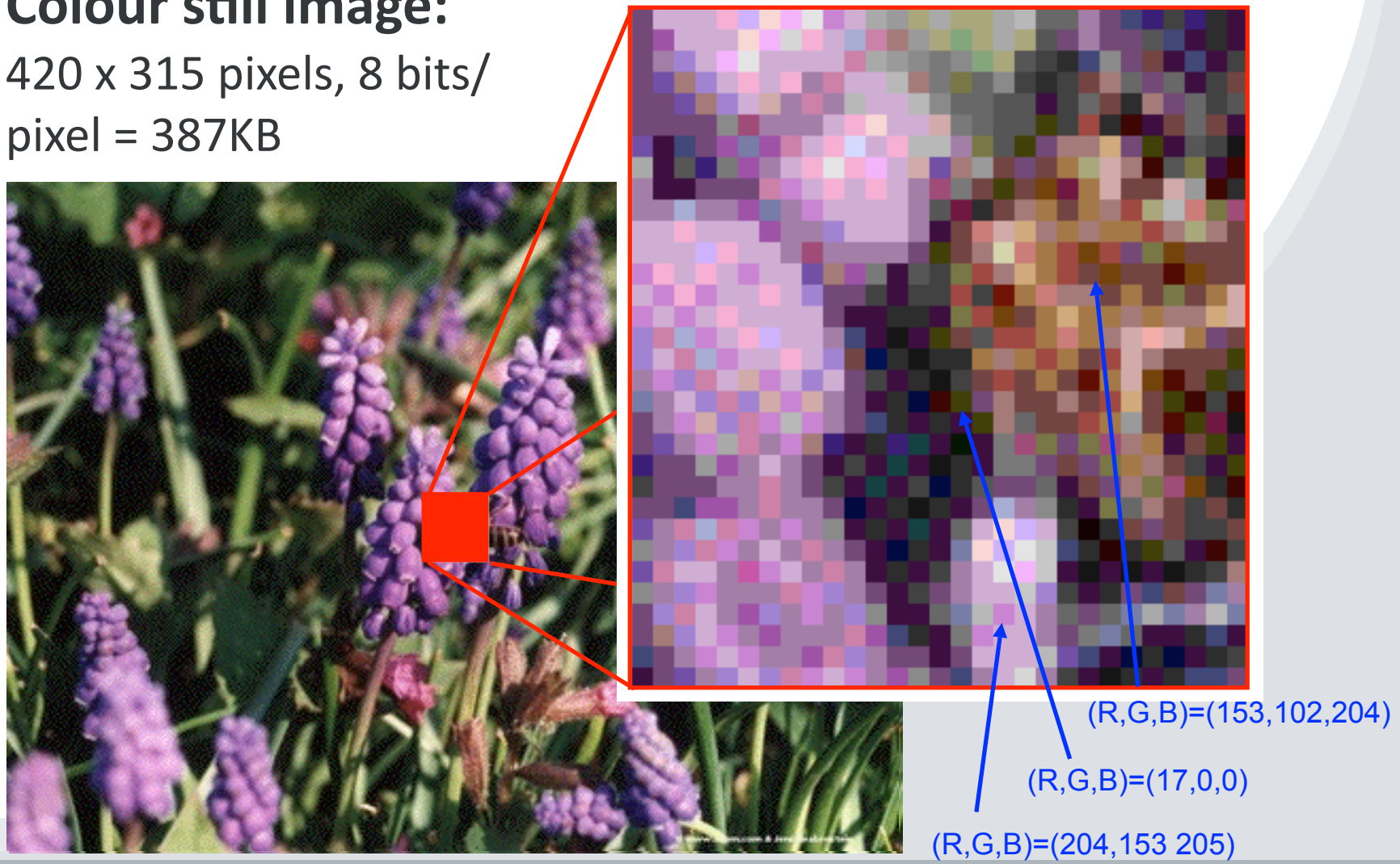
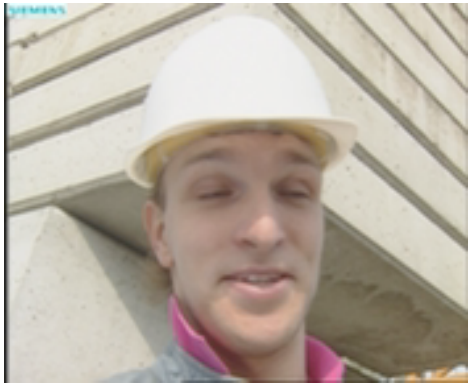


Image Data (YUV)

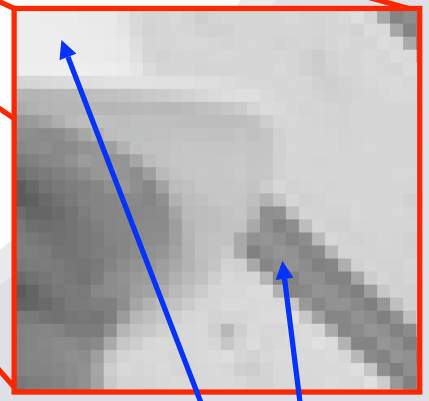


RGB

Y (luminance)

U (col. diff.)

V (col. diff.)



RGB & YUV are known as colour spaces
Many different colour spaces exist

E.g. HSV
Hue: the color type (such as red, blue, or yellow)
Saturation: the intensity of the color
Value: the brightness of the color

See: [A.K. Jain, Fundamentals of Digital Image Processing, Prentice Hall, 1988]

Video Data

CLARITY
clarity-centre.org

Desktop PC

CIF (352 x 288), 8 bpp, 30hz = 8.7 MB/sec
30 sec clip = 261 MB

Video to mobile device

QCIF (176 x 144), 8 bpp, 30 hz = 2.2 MB/sec
30 sec clip = 65 MB

High Definition TV (HDTV)

1280 x 720, 24 bpp, 50 hz = 0.4 GB/sec
2.5 hour movie = 3.4 TB

Video Data

CLARITY
clarity-centre.org

Desktop PC

CIF (352 x 288), 8 bpp, 30hz = 8.7 MB/sec
30 sec clip = 261 MB

Video to mobile device

QCIF (176 x 144), 8 bpp, 30 hz = 2.2 MB/sec
30 sec clip = 65 MB

High Definition TV (HDTV)

1280 x 720, 24 bpp, 50 hz = 0.4 GB/sec
2.5 hour movie = 3.4 TB

Compression

CLARITY
clarity-centre.org

When captured, audio/video data is referred to as as “raw” or “uncompressed”

Undergo software/hardware process to compact data:
Termed “compression” or “encoding”
Results in a bitstream that can be stored or transmitted
Requires a (less) complex process to uncompress/decode before it can be displayed/heard

Terminology:

Encoded data … termed compressed domain
Fast (real-time), simply parsing the bitstream

Raw data … termed uncompressed domain
Slower (requires decode) but less restrictive
i.e. greater range of more expressive features possible

Compression

CLARITY
clarity-centre.org

Two types:

Lossless: doesn't change data "simply" reorganizes

Used in medical applications (e.g. X-Rays) and document scanning (e.g. FAX)

Lossy: throws some data away during encoding

Used in most multimedia applications

Popular AV compression standards

JPEG (still images)
JPEG 2000 (enhanced functionality/quality)
MPEG-1 (video from CD-ROM)
MPEG-2 (digital TV, DVD)
MPEG-4 (mobile and content-based functionality)
H.264 (advanced video coding)

Note: MP3 = MPEG-1 layer 3 audio encoding

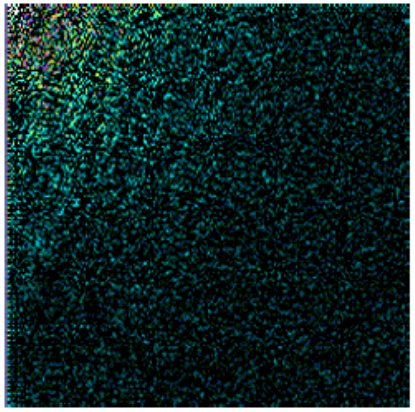
Image Compression

CLARITY
clarity-centre.org

Use frequency domain analysis

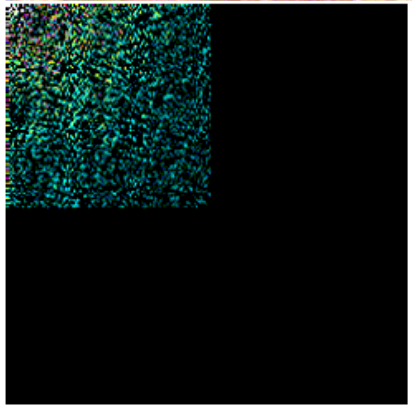
The discrete cosine transform (DCT)

- Spatial Domain
- Highly Textured
- Fine Detail



- DCT Domain
- Energy Compaction
- Less information to store or transmit over network

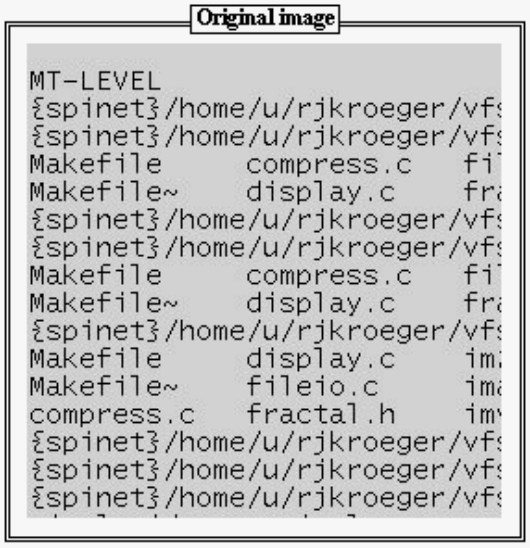
- Low Pass Filter
- Ignore high frequency info
- Even better compression

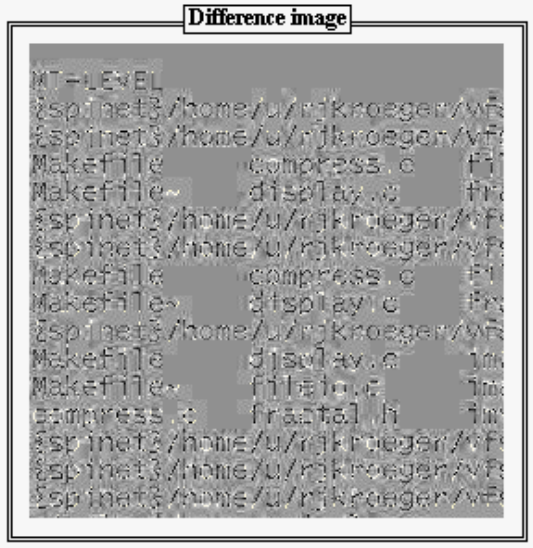


- Reconstruction
- Use Inverse Transform
- Low perceptual losses

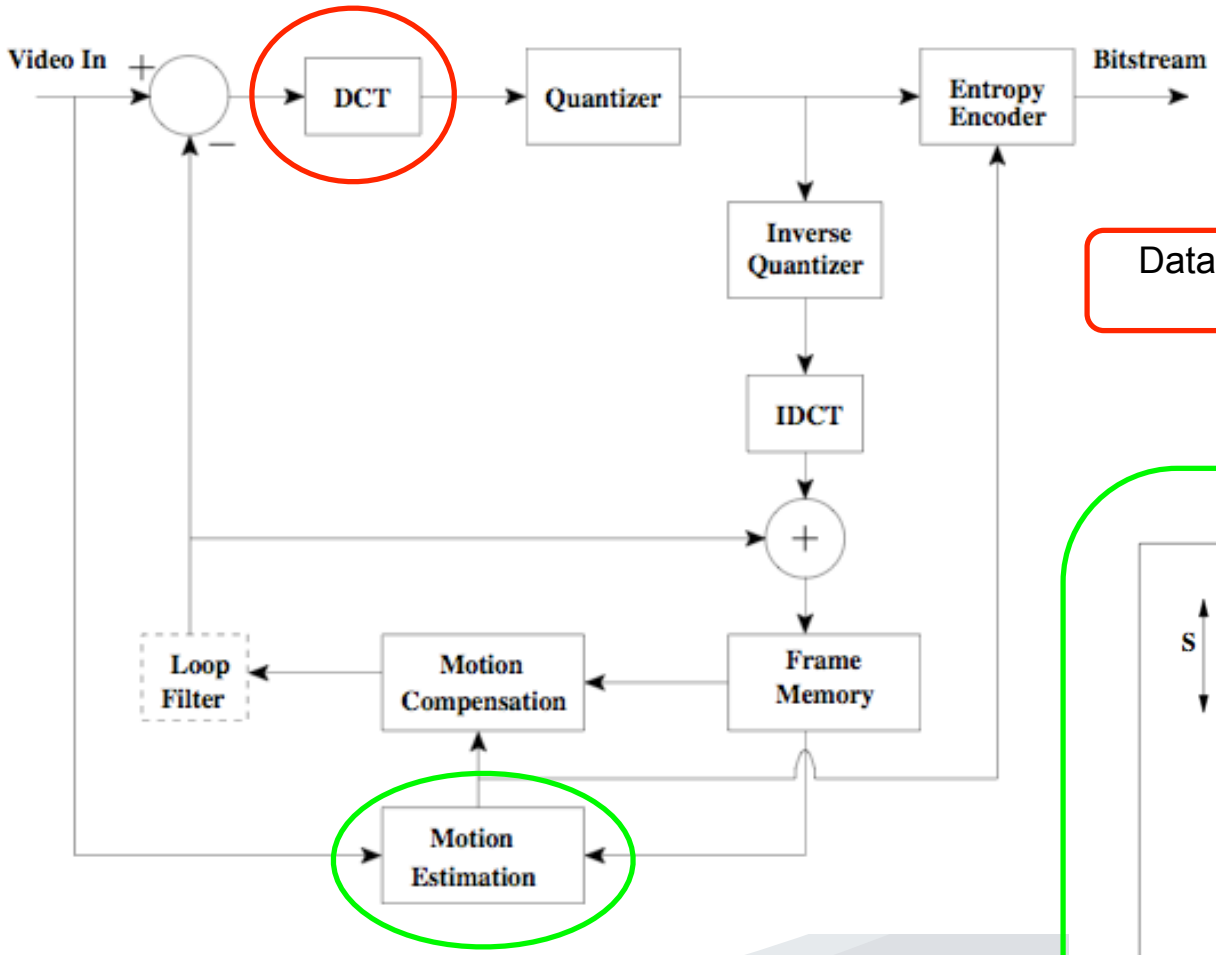
Image Compression

CLARITY
clarity-centre.org

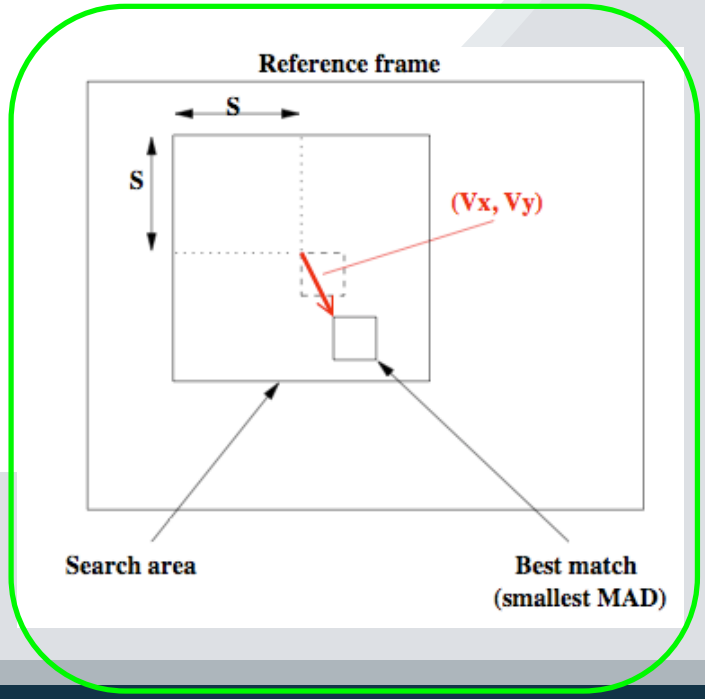




Video Compression

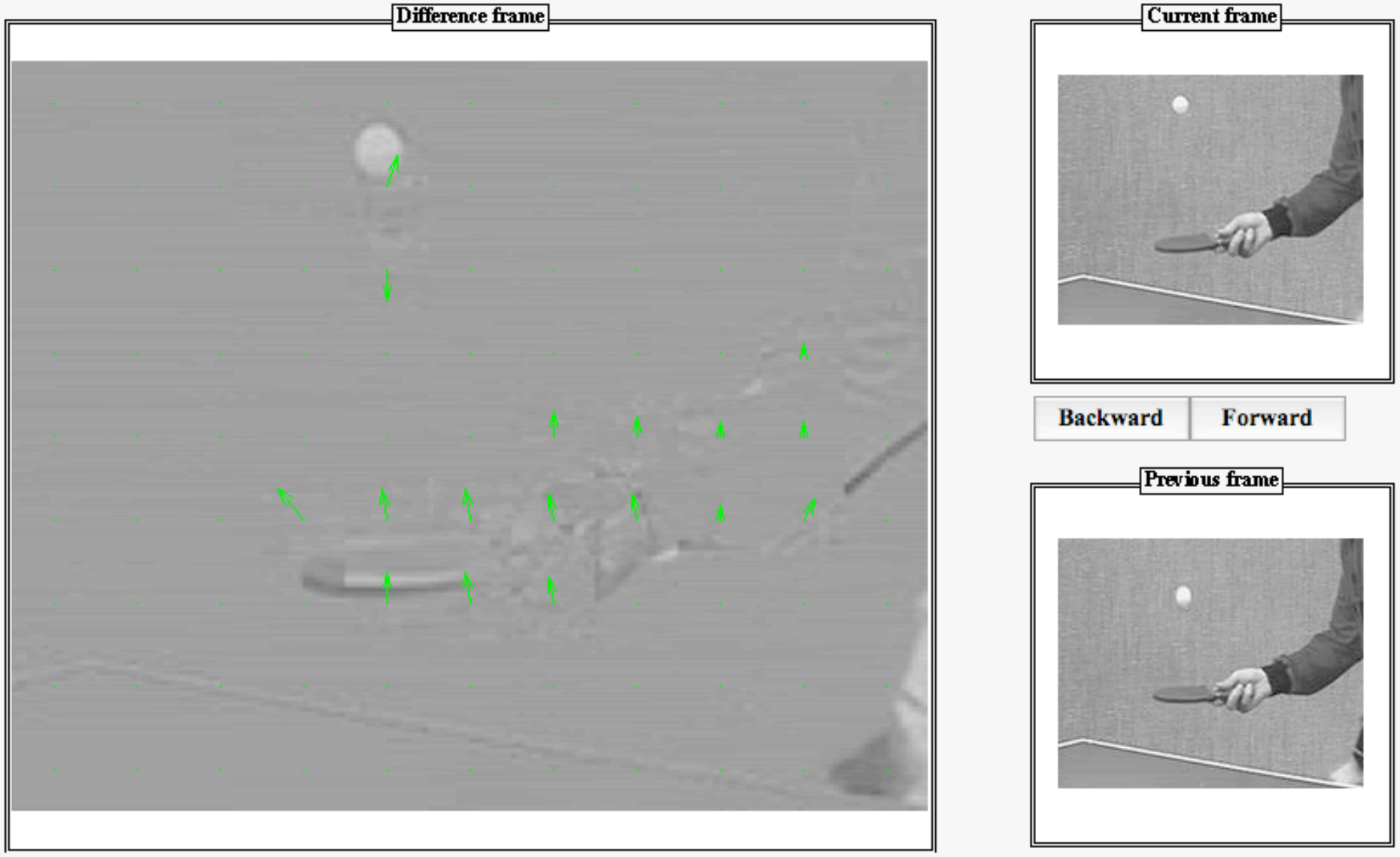


Data for DCT is motion compensated difference between frames



Video Compression

CLARITY
clarity-centre.org



Video Compression

CLARITY
clarity-centre.org