Coding and Data Compression

Lec.7

Video Compression

Video compression is important to understand the mechanism of working sites that broadcast videos such as YouTube. Without it, we could not broadcast HD videos.

For e*xample,* we will assume that you are watching video at 1920 x 1080 pixels. This is close to two million pixels in each frame. In this case, you have 30 frames per second. This is approximately 62 million pixels per second. If each pixel needs 24 bits or three bits of information, the video size will be 178 megabytes per second. If not compressed, it will be 51 GB if it is five minutes long.

Fortunately, the five-minute video size on YouTube is not equivalent to 51 GB, but equivalent to 27 MB, which is about 700 times smaller.

How did they do this? This is the secret of video compression.

With the rapid advances in computers in the 1980s and 1990s came multimedia applications, where images and sound are combined in the same file. Such files tend to be large, which is why compressing them became an important field of research. Digital video is the case where a (digital) camera generates a digital image, i.e., an image that consists of pixels.

Digital video, on the other hand, has the following important advantages:

1. It can be easily edited**:** This makes it possible to produce special effects. The images of an actor in a movie can be edited to make him look young at the beginning and old later.

Software for editing digital video is available for most computer platforms. Users can edit a video file and attach it to an email message.

2. ***It can be stored on any digital medium,*** such as disks, flash memories, CD-ROMs, or DVDs.

3. It can be compressed. This allows for more storage (when video is stored on a digital medium) and also for fast transmission. Sending compressed video between computers makes video telephony, video chats, and video conferencing possible.

Digital video is, in principle, a sequence of images, called frames, displayed at a certain frame rate (so many frames per second, or fps) to create the animation.

This rate, as well as the image size and pixel depth, depend heavily on the application. Surveillance cameras, for example, use the very low frame rate of five fps. The next table shows some typical video applications and their video parameters.

The table illustrates the need for compression. Even the most economic application, a surveillance camera, generates 5×640×480×12 = 18,432,000 bits per second! This is equivalent to more than 2.3 million bytes per second, and this information has to be saved for at least a few days before it can be deleted.

Application	Frame rate	Resolution	Pixel depth
Surveillance	5	640×480	12
Video telephony	10	320×240	12
Multimedia	15	320×240	16
Analog TV	25	640×480	16
HDTV (720p)	60	1280×720	24
HDTV (1080i)|	60	1920×1080	24
HDTV (1080p)	30	1920×1080	24

Video Parameters for Typical Applications.

The following is a short historical overview of this field:

DVI: the field of video compression had its origin in 1987, when Digital Video Interactive (DVI) made its appearance.

DVI was never adopted as a standard by any international body. In 1992 it change and became known as Indeo, a software product by Intel.

It was used for years to compress videos and other types of data on personal computers until it was surpassed by MPEG and other codecs.

MPEG-1: This international standard, the first in the MPEG family of video codecs was developed by the ISO/IEC from 1988 to 1993. Video compressed by MPEG-1 could be played back at a bitrate of about 1.2 Mbps . MPEG-1 is still used today for Video CDs (VCD).

MPEG-2: This popular standard evolved in 1996 out of the shortcomings of MPEG-1. The main weaknesses of MPEG-1 are (1) resolution that is too low for modern television, (2) inefficient audio compression, and (3) weak security. Today, a movie DVD is almost always compressed in MPEG-2.

MPEG-4: was first introduced in 1996. It builds on the experience gained with MPEG-2 and adds coding tools, error resilience, and much complexity to achieve better performance. This codec supports several profiles, the more advanced of which are capable of rendering surfaces. In addition to better compression, MPEG-4 offers facilities to protect private data.

H.264: (2006) is part of the huge MPEG-4 project. Specifically, it is the advanced video coding (AVC) part of MPEG-4. It is developers hope that it will replace MPEG-2 because it can compress video with quality comparable to that of MPEG-2 but at half the bitrate.

Video Compression

Video compression is based on two principles:

- The first is the ***spatial redundancy*** that exists in each frame because of pixel correlation.
- The second is the fact that most of the time a video frame is very similar to its immediate neighbors. This is called ***temporal redundancy***.

A typical technique for video compression should therefore start by encoding the first frame using a still image compression method. It should then encode several sequential frames by identifying the differences between a frame and its predecessor, and encoding these differences. If a frame is very different from its predecessor, it should be coded independently of any other frame.

In the video compression literature:

A frame that is coded independently is ***called intra frame (or just intra)*** that is apply for each frame independently. And we can use the same operation that used for image compression as JPEG.

While a frame that is coded from its predecessor is called ***inter frame (or just inter)***

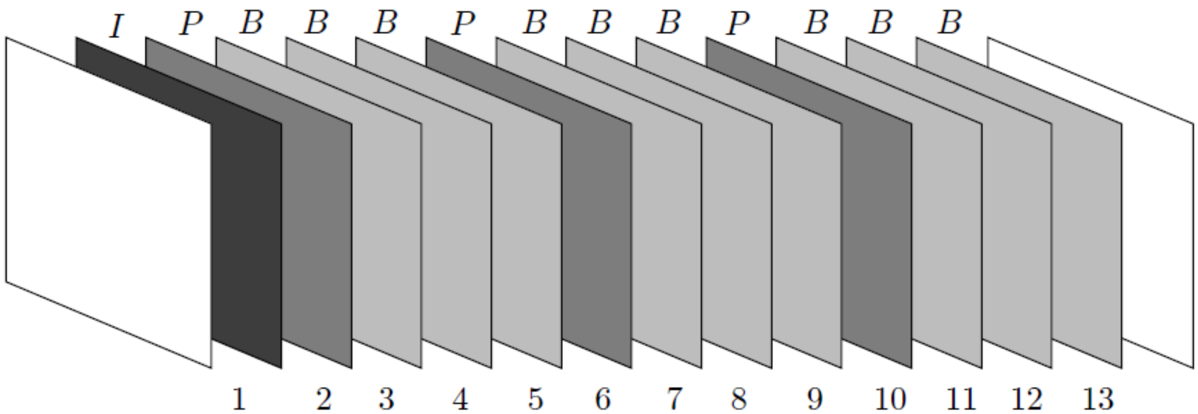
Video compression is normally lossy. Encoding a frame Fi in terms of its predecessor Fi−1 introduces some distortions. As a result, encoding the next frame Fi+1 in terms of (the already distorted) Fi increases the distortion.

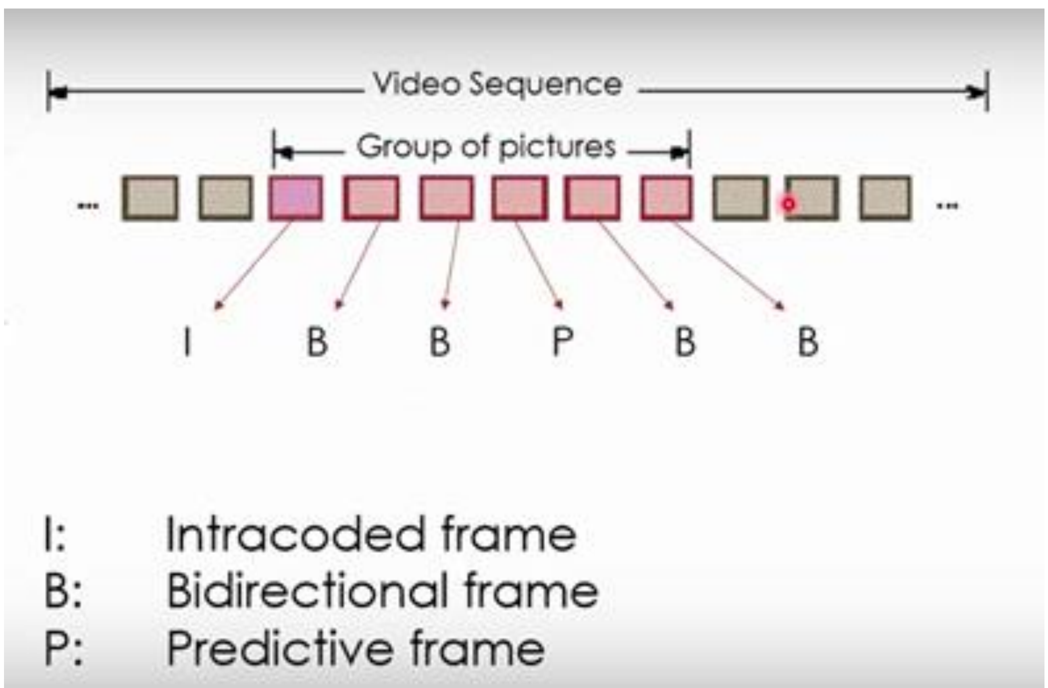
Even in lossless video compression, a frame may lose some bits. This may happen during transmission or after a long stay. If a frame Fi has lost some bits, then all the frames following it, up to the next intra frame, will be decoded incorrectly, possibly leading to accumulated errors. This is why intra frames should be used from time to time inside a sequence, not just at its beginning. An intra frame is labeled I, and an inter frame is labeled P (for predictive).

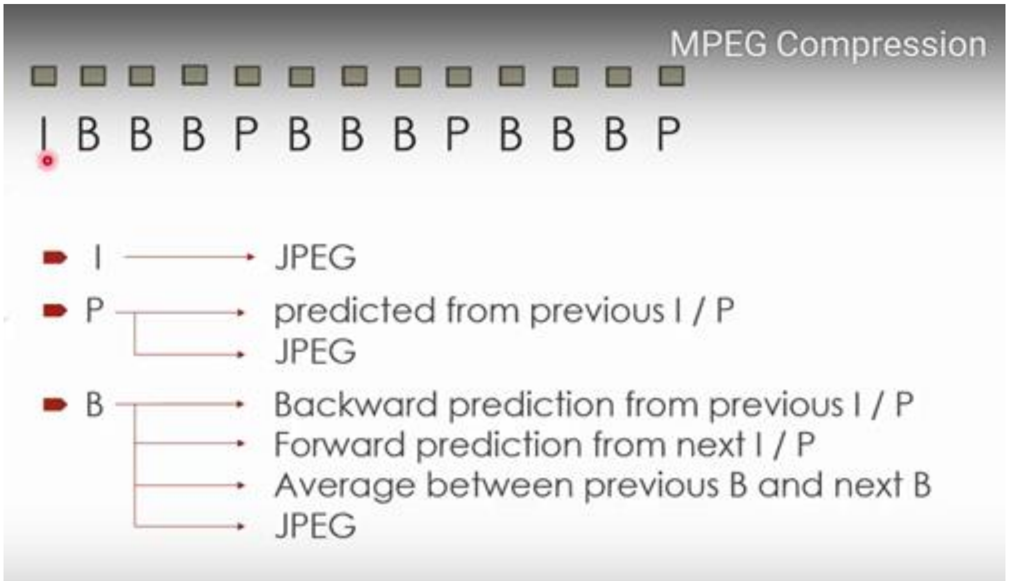
With this in mind it is easy to imagine a situation where the encoder encodes frame 2 based on both frames 1 and 3, and writes the frames on the compressed stream in the order 1, 3, 2. The decoder reads them in this order, decodes frames 1 and 3 in parallel, outputs frame 1, then decodes frame 2 based on frames 1 and 3.

A frame that is encoded based on both past and future frames is labeled B (for bidirectional).

Video compression requires many steps and computations, so researchers have been looking for optimizations and faster algorithms, especially for steps that involve many calculations.







MPEG

Starting in 1988, the MPEG project was developed by a group of hundreds of experts under the auspices of the ISO (International Standardization Organization) and the IEC (International Electrotechnical Committee). The name MPEG is an acronym for Moving Pictures Experts Group.

MPEG is a method for video compression, which involves the compression of digital images and sound. There currently are several MPEG standards.

MPEG-1 is intended for intermediate data rates, on the order of 1.5 Mbit/sec.

MPEG-2 is intended for high data rates of at least 10 Mbit/sec.

MPEG-3 was intended for HDTV compression but was found to be redundant and was merged with MPEG-2.

MPEG-4 is intended for very low data rates of less than 64 Kbit/sec.

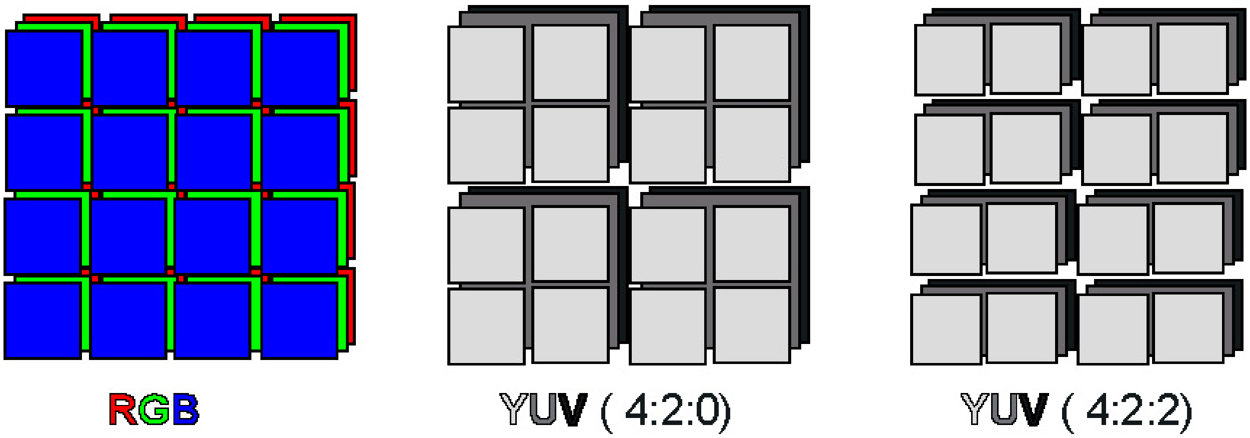
The MPEG Compression

The MPEG compression algorithm encodes the data in 5 steps

First a reduction of the resolution is done, which is followed by a motion compensation in order to reduce temporal redundancy. The next steps are the Discrete Cosine Transformation (DCT) and a quantization as it is used for the JPEG compression; this reduces the spatial redundancy (referring to human visual perception). The final step is an entropy coding using the Run Length Encoding and the Huffman coding algorithm.

Step 1: Reduction of the Resolution

The human eye has a lower sensibility to colour information than to dark-bright contrasts. A conversion from RGB-colour-space into YUV colour components help to use this effect for compression. The chrominance components U and V can be reduced (subsampling) to half of the pixels in horizontal direction (4:2:2), or a half of the pixels in both the horizontal and vertical (4:2:0).



Step 2: Motion Estimation

An MPEG video can be understood as a sequence of frames. Because two successive frames of a video sequence often have small differences (except in scene changes), the MPEG-standard offers a way of reducing this temporal redundancy. It uses three types of frames:

I-frames (intra), P-frames (predicted) and B-frames (bidirectional).

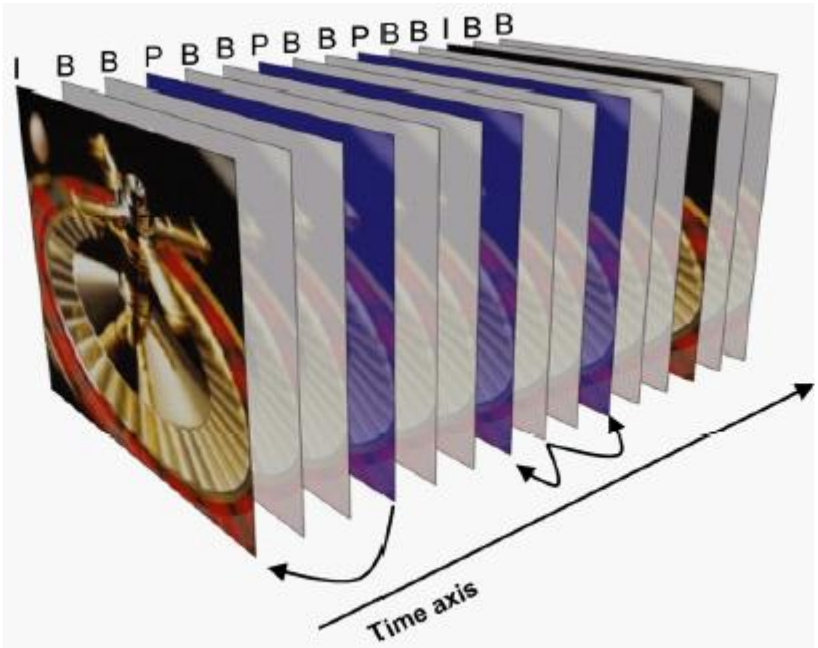


Figure: An MPEG frame sequence with two possible references: a P-frame referring to a I-frame and a B-frame referring to two P-frames.

Step 3: Discrete Cosine Transform (DCT)

Step 4: Quantization

Step 5: Entropy Coding

The entropy coding takes two steps: *Run Length Encoding (RLE)* and *Huffman coding.* These are well known lossless compression methods, which can compress data, depending on its redundancy.

All five Steps together