Imperial College London

Digital Image Processing

Fundamentals of Image Compression

DR TANIA STATHAKI

READER (ASSOCIATE PROFFESOR) IN SIGNAL PROCESSING
IMPERIAL COLLEGE LONDON

Imperial College London

Compression

- New techniques have led to the development of robust methods to reduce the size of the image, video, or audio data.

- Such methods are extremely vital in many applications that manipulate and store digital data.

- Informally, we refer to the process of size reduction as a compression process. We will define this process in a more formal way later.

- On the architecture front, it is now feasible to put sophisticated compression processes on a relatively low-cost single chip; this has spurred a great deal of activity in developing multimedia systems for the large consumer market.

Compression

- Compression is a process intended to yield a compact digital representation of a signal.

- In the literature, the terms *source coding, data compression, bandwidth compression,* and *signal compression* are all used to refer to the process of compression.

- In the cases where the signal is defined as an image, a video stream, or an audio signal, the generic problem of compression is to minimise the bit rate of their digital representation.

- There are many applications that benefit when image, video, and audio signals are available in compressed form. **Without compression, most of these applications would not be feasible!**

Why are signals amenable to compression?

- **There is considerable statistical redundancy in the signal.**
  1. Within a single image or a single video frame, there exists significant correlation among neighbour samples. This correlation is referred to as *spatial correlation.*
  2. For data acquired from multiple sensors (such as satellite images), there exists significant correlation amongst samples from these sensors. This correlation is referred to as *spectral correlation.*
  3. For temporal data (such as video), there is significant correlation amongst samples in different segments of time. This is referred to as *temporal correlation.*

- **There is considerable information in the signal that is irrelevant from a perceptual point of view.**

- **Some data tends to have high-level features that are redundant across space and time; that is, the data is of a fractal nature.**

Imperial College London

Why do we need compression standards ?

- Multimedia information comprising image, video, and audio has become just another data type.

- This usually implies that multimedia information will be digitally encoded so that it can be manipulated, stored, and transmitted along with other digital data types.

- For such data usage to be pervasive, it is essential that the data encoding is standard across different platforms and applications.

- This will foster widespread development of applications and will also promote interoperability among systems from different vendors.

- Furthermore, standardisation can lead to the development of cost-effective implementations, which in turn will promote the widespread use of multimedia information.

- This is the primary motivation behind the emergence of image and video compression standards.

Examples of data compression

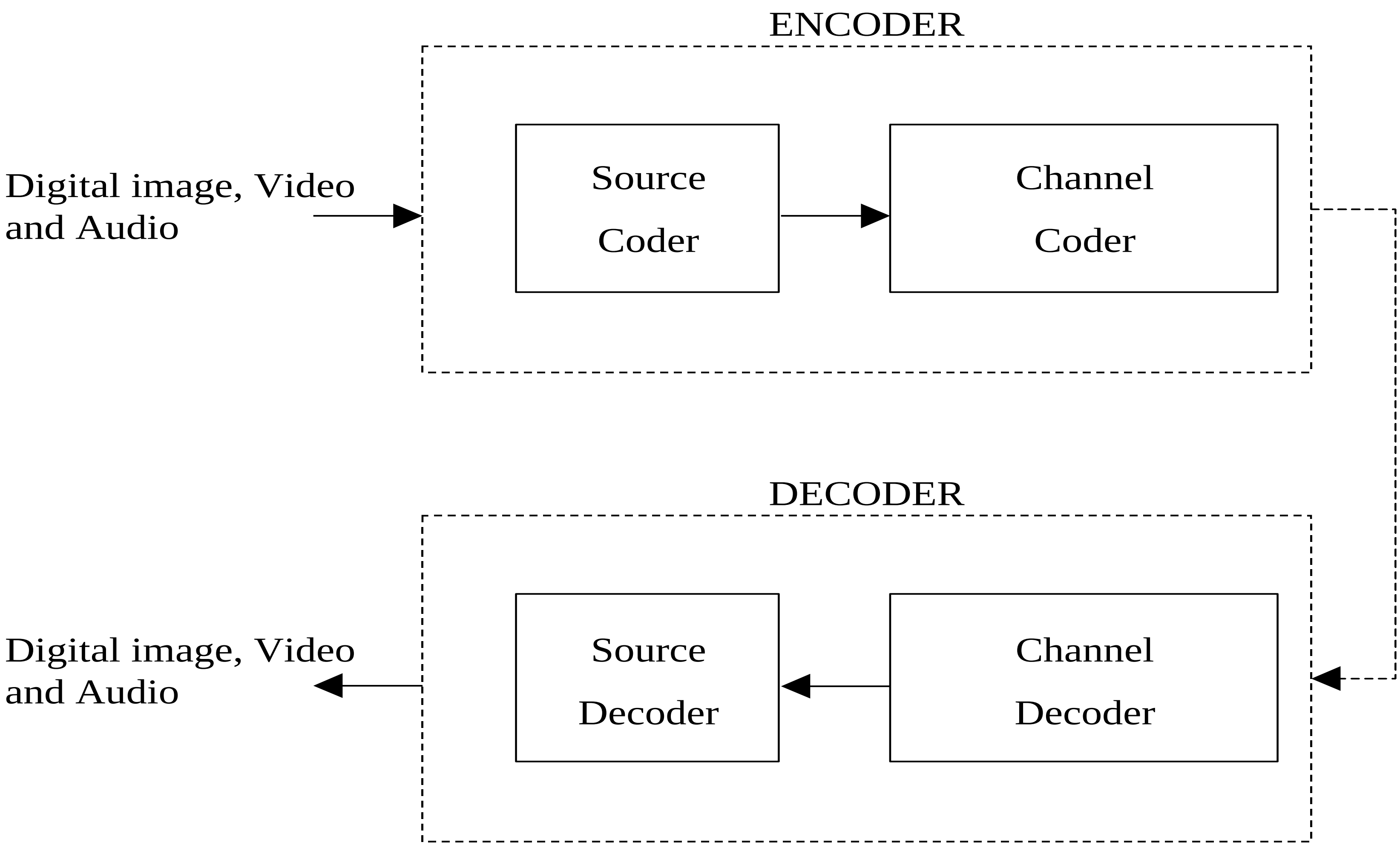
Example 1: Let us consider **facsimile image transmission**. In most facsimile machines, the document is scanned and digitised. Typically, an 8.5x11 inches page is scanned at 200 dpi **(the number of individual dots that can be placed in a line within the span of 1 inch)**; thus, resulting in 3.74 Mbits. Transmitting this data over a low-cost 14.4 kbits/s modem would require 5.62 minutes. With compression, the transmission time can be reduced to 17 seconds. This results in substantial savings in transmission costs.

Example 2: Let us consider a video-based CD-ROM application. **Full-motion video**, at 30 fps and a 720 x 480 resolution, generates data at 20.736 Mbytes/s. At this rate, only 31 seconds of video can be stored on a 650 MByte CD-ROM. Compression technology can increase the storage capacity to 74 minutes, for VHS-grade video quality.

Applications for image, video, and audio compression

Application	Data	Rate
	Uncompressed	Compressed
Voice 8 ksamples/s, 8 bits/sample	64 kbps	2-4 kbps
Slow motion video (10fps) framesize 176x120, 8bits/pixel	5.07 Mbps	8-16 kbps
Audio conference 8 ksamples/s, 8 bits/sample	64 kbps	16-64 kbps
Video conference (15fps) framesize 352x240, 8bits/pixel	30.41 Mbps	64-768 kbps
Digital audio 44.1 ksamples/s, 16 bits/sample	1.5 Mbps	1.28-1.5 Mbps
Video file transfer (15fps) framesize 352x240, 8bits/pixel	30.41 Mbps	384 kbps
Digital video on CD-ROM (30fps) framesize 352x240, 8bits/pixel	60.83 Mbps	1.5-4 Mbps
Broadcast video (30fps) framesize 720x480, 8bits/pixel	248.83 Mbps	3-8 Mbps
HDTV (59.94 fps) framesize 1280x720, 8bits/pixel	1.33 Gbps	20 Mbps

Generic compression system



Imperial College London

Source coder – Compression ratio

- The source coder performs the compression process by reducing the input data rate to a level that can be supported by the storage or transmission medium.

- The bit rate output of the encoder is measured in bits per sample or bits per second.

- For image or video data, a pixel is the basic element; thus, bits per sample is also referred to as bits per pixel.

- In the literature, the term *compression ratio* c_r, is also used instead of *bit rate* to characterise the capability of the compression system. An intuitive definition is

$$c_r = \frac{\text{source coder input size}}{\text{source coder output size}}$$

Imperial College London

Compression ratio

- The definition of compression ratio is somewhat ambiguous and depends on the data type and the specific compression method that is employed.

- For a still-image, size could refer to the bits needed to represent the entire image.

- For video, size could refer to the bits needed to represent one frame of video.

- Many compression methods for video do not process each frame of video, hence, a more commonly used notion for size is the bits needed to represent one second of video.

Compression requirements

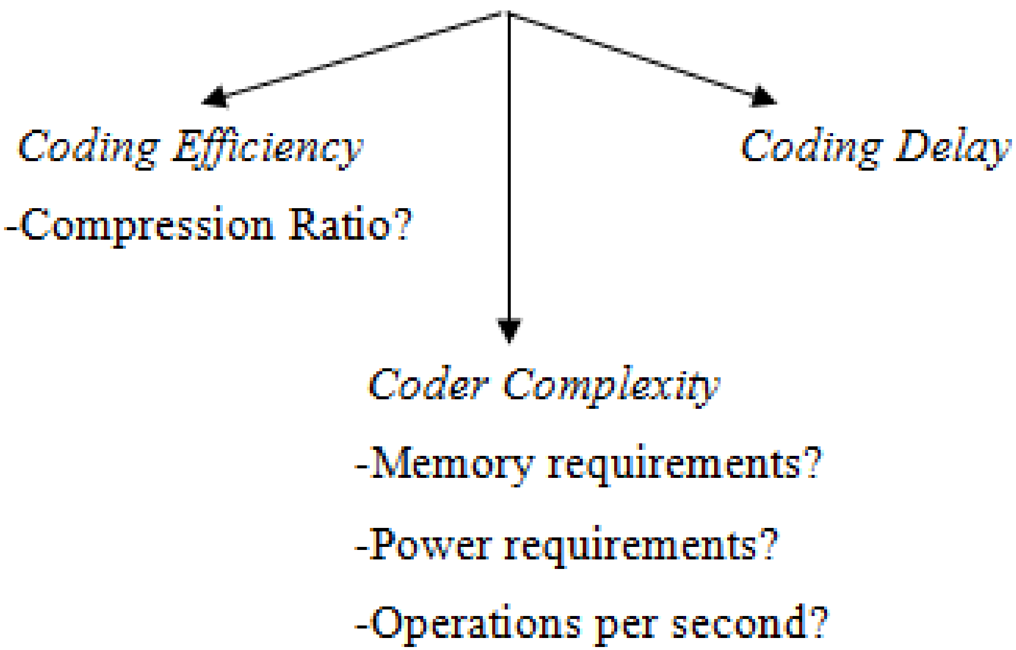
- **Specified level of signal quality**. This constraint is usually applied at the decoder.

- **Implementation complexity**. This constraint is often applied at the decoder, and in some instances at both the encoder and the decoder.

- **Communication delay**. This constraint refers to the end to end delay, and is measured from the start of encoding a sample to the complete decoding of that sample.

Note that, these constraints have different importance in different applications. For example, in a two-way teleconferencing system, the communication delay might be the major constraint, whereas, in a television broadcasting system, signal quality and decoder complexity might be the main constraints.

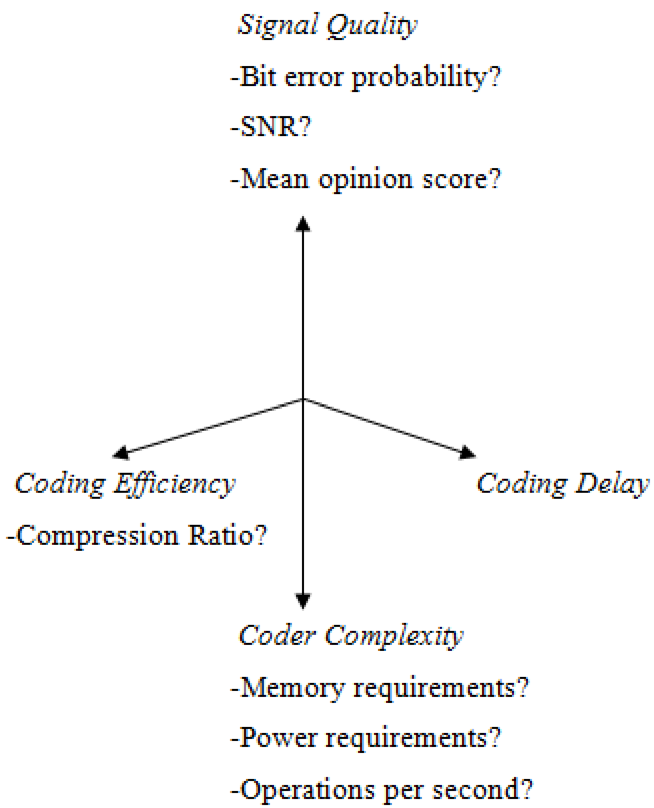
Lossless compression - Trade offs



Lossy compression

- The majority of the applications in image or video data processing do not require that the reconstructed data and the original data are identical in value.
- Thus, some amount of loss is permitted in the reconstructed data. A compression process that results in an imperfect reconstruction is referred to as a lossy compression process.
- This compression process is irreversible. In practice, most irreversible compression processes degrade rapidly the signal quality when they are repeatedly applied on previously decompressed data.
- The choice of a specific lossy compression method involves trade-offs along the four dimensions shown in figure below.
- Due to the additional degree of freedom, namely, in the signal quality, a lossy compression process can yield higher compression ratios than a lossless compression scheme.

Lossy compression - Trade offs



Lossy compression – Signal quality - *SNR*

- This term is often used to characterise the signal at the output of the decoder. There is no universally accepted measure for signal quality.
- One measure that is often cited is the signal to noise ratio , which can be expressed as

$$SNR = 10\log_{10}\frac{\text{encoder input signal energy}}{\text{noise signal energy}}$$

- The noise signal energy is defined as the energy measured for a hypothetical signal that is the difference between the encoder input signal and the decoder output signal.
- High or values do not always correspond to signals with perceptually high quality.

Lossy compression – Signal quality – Mean opinion score

- Another measure of signal quality is the mean opinion score, where the performance of a compression process is characterised by the subjective quality of the decoded signal.

- For instance, a five point scale such as *very annoying, annoying, slightly annoying, perceptible but not annoying,* and *imperceptible* might be used to characterise the impairments (levels of distortion) in the decoder output.

- In either lossless or lossy compression schemes, the quality of the input data affects the compression ratio. For instance, acquisition noise, data sampling timing errors, and even the analogue-to-digital conversion process affects the signal quality and reduces the spatial and temporal correlation. Some compression schemes are quite sensitive to the loss in correlation and may yield significantly worse compression in the presence of noise.

Imperial College London

Issues in compression method selection

- Lossless or lossy
- Coding efficiency
- Variability in coding efficiency
- Resilience to transmission errors
- Complexity trade-offs
- Nature of degradations in decoder output
- Data representation
- Multiple usage of the encoding-decoding tandem
- Interplay with other data modalities, such as audio and video
- Interworking with other systems