Multimedia

Chapter-11

11

ADVANCED COMPRESSION TECHNIQUES

Unit Structure

1. Objectives
2. Introduction
3. JPEG Coding
4. ZIP Coding
5. Overview of Image and Video Compression Techniques
6. Summary
7. Unit End Exercises

1. OBJECTIVES

After studying this Unit, you will be able to:

1. Understand advanced compression techniques.
2. Describe JPEG coding
3. Describe ZIP coding
4. Understand other coding techniques

2. INTRODUCTION

A high-definition uncompressed video data stream requires about 2 billion bits per second of data bandwidth. Owing to the large amount of data necessary to represent digital video, it is desirable that such video signals are easy to compress and decompress, to allow practical storage or transmission.

The term data compression refers to the reduction in the number of bits required to store or convey data—including numeric, text, audio, speech, image, and video—by exploiting statistical properties of the data. Fortunately, video data is highly compressible owing to its strong vertical, horizontal, and temporal correlation and its redundancy.

Transform and prediction techniques can effectively exploit the available correlation, and information coding techniques can take advantage of the statistical structures present in video data. These techniques can be lossless, so that the reverse operation (decompression) reproduces an exact replica of the input.

In addition, however, lossy techniques are commonly used in video data compression, exploiting the characteristics of the HVS, which is less sensitive to some color losses and some special types of noises.

Video compression and decompression are also known as video encoding and decoding, respectively, as information coding principles are used in the compression and decompression processes, and the compressed data is presented in a coded bit stream format.

3. JPEG CODING STANDARD

The JPEG (Joint Photographic Experts Group) became an international standard in 1992.

In the sequential mode JPEG compression process is composed of following steps:

- Preparation of data
- Source encoding steps involving forward DCT and quantization
- Entropy encoding steps involving RLE and Huffman encoding

DECOMPRESSION PROCESS

- Entropy decoding steps involving RLD and Huffman decoding
- Source decoding steps involving inverse DCT and de-quantization

1. BLOCK PREPARATION

An image is represented by 1 or more 2D array of pixel values these blocks are in preparation of next steps where DCT is applied to each block instead of the entire image.

2. DISCRETE COSINE TRANSFORM (DCT)

The objective of this is to transform each block from the spatial domain to the frequency domain. We know that the synthesis equation of the DFT is given by the relation.

$$x[i] = \sum_{k=0}^{N/2} \overline{\mathrm{Re}\, X}[k] \cos(2\text{®} ki/N) + \sum_{k=0}^{N/2} \mathrm{Re}\, X[k] \sin(2\text{®} ki/N)$$

where *N* is the total number of samples, *k* is the variable, *i* is the variable indicating the number of input sample considered in the time or(space) domain, *x*[*i*] is the actual time(or space) domain signal,

$\mathrm{Re}\,\overline{X}[k]$ is proportional to the k-th entry and the frequency domain. If we ignore the imaginary components and consider the entire time (or space) domain signal to be compiled of real numbers we get

$$x[i] = \sum_{k=0}^{N-1} \mathrm{Re}\,\overline{X[k]} \cos(2\pi k i / N)$$

The above expression is called the synthesis equation of a one dimensional DCT, also known as inverse DCT. The forward DCT is just the reverse of the above relation, where we express the frequency domain in terms of time (or space) domain signal. Mathematically it can be expressed as

$$\mathrm{Re}\,X[k] = \sum_{i=0}^{N-1} x[i]\cos[\pi i(k + 1/2)/N)$$

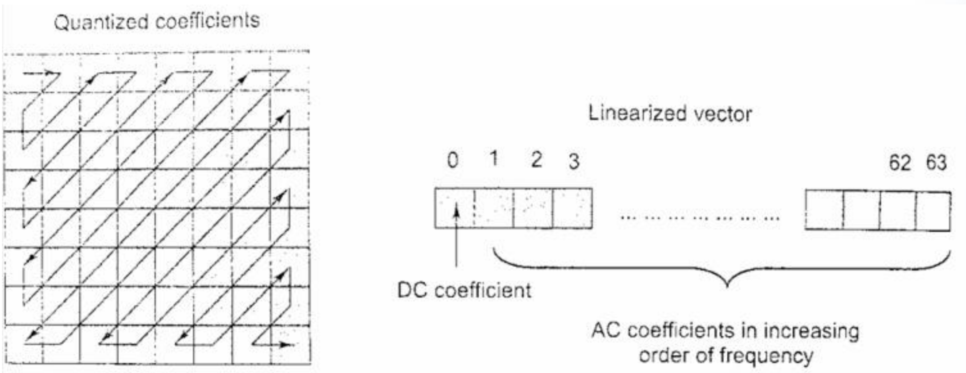
$$or,$$

$$\mathrm{Re}\,X[k] = \sum_{i=0}^{N-1} x[i]\cos[\pi i(2k + 1)/2N)$$

The expression for 2D forward DCT is written as

$$F[i, j] = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} P[x, y]\cdot\cos[\pi i(2x+1)/2N]\cdot\cos[\pi j(2y+1)/2N]$$

11.3.3 QUANTIZATION



The next step in the process is quantizing the coefficient numbers that were derived from the luminance and chrominance values by the DCT. Quantizing is basically the process of rounding off the numbers. This is where the file compression comes in. How much the file is compressed depends on the quantization matrix.

The quantization matrix defines how much the information is compressed by dividing the coefficients by a quantizing factor. The larger the number of the quantizing factor, the higher the quality (therefore, the less compression). This is basically what is going on in Photoshop when you save as JPEG and the program asks you to set the quality; you are simply defining the quantizing factor.

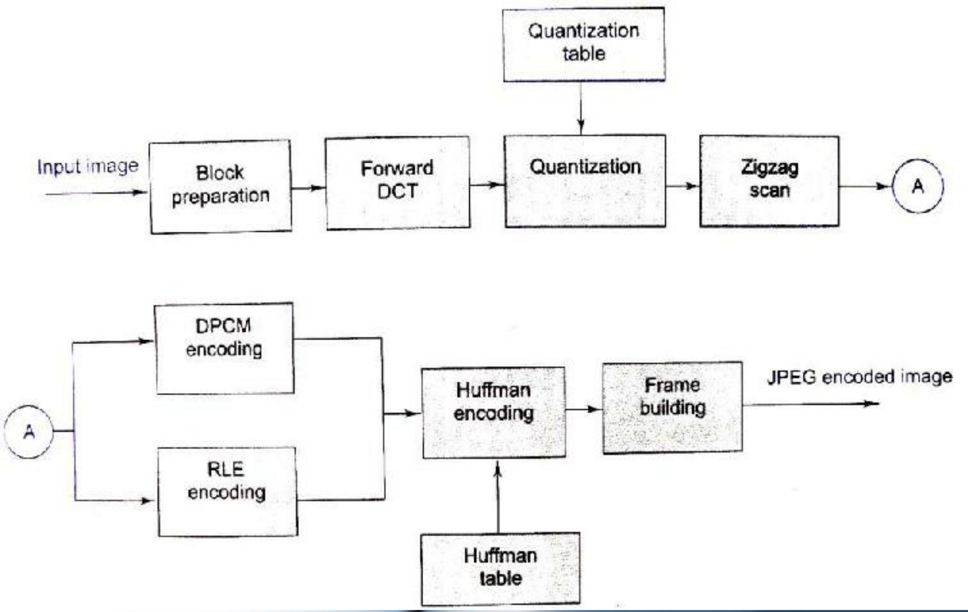
Once the numbers are quantized, they are run through a binary encoder that converts the numbers to the ones and zeros computers love so well. You now have a compressed file that is on average about one-fourth of the size of an uncompressed file.

4. ZIGZAG SCAN

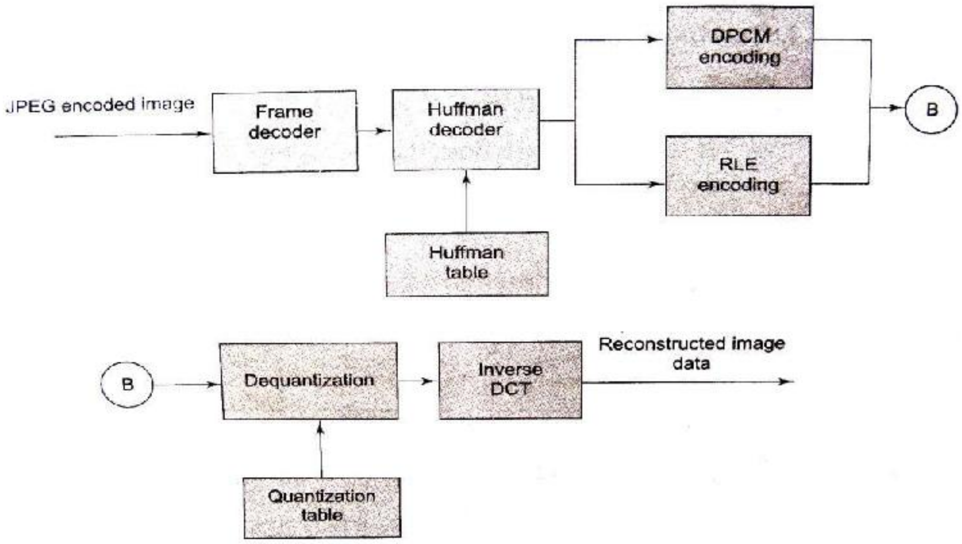
After the DCT stage, the remaining stages involve entropy encoding. The entropy coding algorithms operate on one dimensional string of values i.e. a vector. The output of the quantization stage is a 2D array, hence to apply an entropy scheme, the array is to be converted to a 1D vector. This operation is known as vectoring.

5. DPCM encoding

There is one DC co-efficient per block because of the small physical area covered by each block, the DC co-efficient varies from one block to the next. To exploit this similarity, the sequence of DC co-efficient is encoded in DPCM mode. This means the difference between the DC co-efficient of each block and the adjacent block is computed and stored.



JPEG Encoder



JPEG Decoder

11.4 ZIP CODING

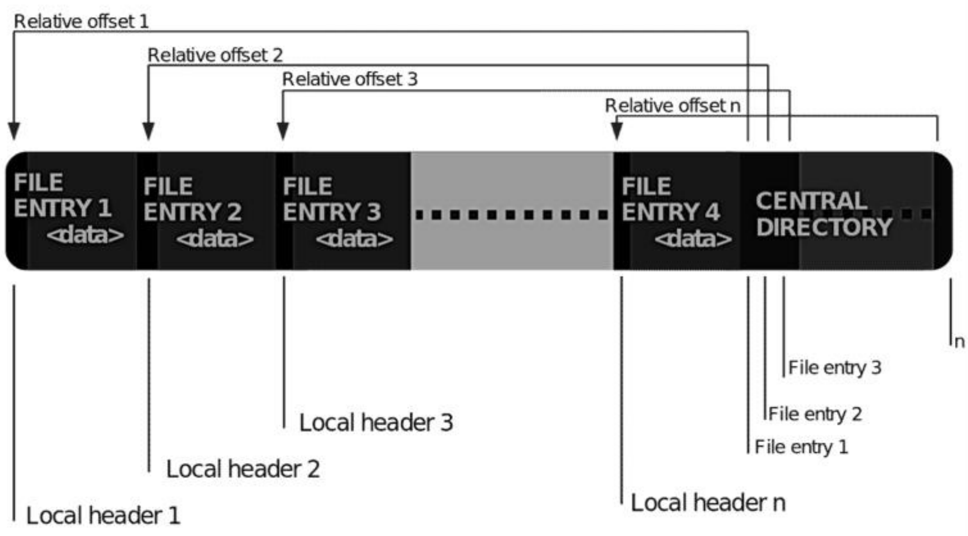
.ZIP files are archives that store multiple files. .ZIP allows contained files to be compressed using many different methods, as well as simply storing a file without compressing it. Each file is stored separately, allowing different files in the same archive to be compressed using different methods. Because the files in a .ZIP archive are compressed individually it is possible to extract them, or add new ones, without applying compression or decompression to the entire archive.

This contrasts with the format of compressed tar files, for which such random-access processing is not easily possible.

A directory is placed at the end of a .ZIP file. This identifies what files are in the .ZIP and identifies where in the .ZIP that file is located. This allows .ZIP readers to load the list of files without reading the entire .ZIP archive. .ZIP archives can also include extra data that is not related to the .ZIP archive.

This allows for a .ZIP archive to be made into a self-extracting archive (application that decompresses its contained data), by prepending the program code to a .ZIP archive and marking the file as executable. Storing the catalog at the end also makes possible hiding a zipped file by appending it to an innocuous file, such as a GIF image file.

The .ZIP format uses a 32-bit CRC algorithm and includes two copies of the directory structure of the archive to provide greater protection against data loss.



Technical representation: ZIP Compression

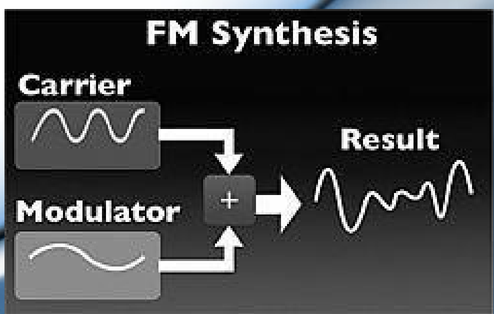
11.5 OVERVIEW OF AUDIO VIDEO COMPRESSIONS

FM Synthesis

In frequency modulation synthesis (or FM synthesis), the timbre of simple waveform (usually a sine wave) called a carrier, is changed by modulating it with another waveform in the audio range (higher than 20Hz). The result is a complex waveform with a different tone.

Video-game developers used FM synthesis in the past because the sounds created by FM synthesis took up less space than those created through PCM.

Nowadays FM synthesis is commonly found in modern DAWs as a plug-in. The sound is also considered 'retro' and is still used in certain specialized 'tracker' applications to reproduce the sound of a well-loved era of gaming.



A simplification of FM Synthesis, originally developed by John Chowning at Stanford, and patented by Yamaha.

MIDI

The Musical Instrument Digital Interface, or MIDI, is one the most innovative and far reaching standards developed in the last 30 years and is still heavily in use today.

It is an industry protocol enabling different manufacturers to control and trigger events (musical or otherwise) by sending messages back and forth on a serial connection. Keep in mind MIDI isn't audio but it can certainly control audio.

A MIDI interface and MIDI cables. MIDI data can also be used internally and transmitted wirelessly.

Games use MIDI like a player-piano roll to trigger sounds and music passages. When used in conjunction with FM synthesis, for example, MIDI can trigger a significant amount of music and sound for very little file size.
Recently, MIDI achieved industry recognition by winning a technical Grammy.

Dave Smith, the polyphonic synthesizer pioneer, and Ikutaro Kakehashi, a Yamaha engineer, were both honored for the creation of this amazing language spoken by synthesizers, samplers, drum machines, computers and sequencers and much more, that enabled digital music to advance to never-before-seen levels of sophistication.

VIDEO CODING STANDARDS

International video coding standards are defined by committees of experts from organizations like the International Standards Organization (ISO) and the International Telecommunications Union (ITU). The goal of this standardization is to have common video formats across the industry, and to achieve interoperability among different vendors and video codec related hardware and software manufacturers.

The standardization of algorithms started with image compression schemes such as JBIG (ITU-T Rec. T82 and ISO/IEC 11544, March 1993) for binary images used in fax and other applications, and the more general JPEG (ITU-T Rec. T81 and ISO/IEC 10918-1), which includes color images as well. The JPEG standardization activities started in 1986, but the standard was ratified in 1992 by ITU-T and in 1994 by ISO.

The main standardization activities for video compression algorithms started in the 1980s, with the ITU-T H.261, ratified in 1988, which was the first milestone standard for visual telecommunication. Following that effort, standardization activities increased with the rapid advancement in the television, film, computer, communication, and signal processing fields, and with the advent of new usages requiring contributions from all these diverse industries. These efforts subsequently produced MPEG-1, H.263, MPEG-2, MPEG-4 Part 2, AVC/H.264, and HEVC/H.265 algorithms. In the following sections, we briefly describe the major international standards related to image and video coding.

MPEG-4

MPEG-4, formally the standard ISO/IEC 14496, was ratified by ISO/IEC in March 1999 as the standard for multimedia data representation and coding. In addition to video and audio coding and multiplexing, MPEG-4 addresses coding of various two- or three-dimensional synthetic media and flexible representation of audio-visual scene and composition.

As the usage of multimedia developed and diversified, the scope of MPEG-4 was extended from its initial focus on very low bit-rate coding of limited audio-visual materials to encompass new multimedia functionalities.

Unlike pixel-based treatment of video in MPEG-1 or MPEG-2, MPEG-4 supports content-based communication, access, and manipulation of digital audio-visual objects, for real-time or non-real-time interactive or non-interactive applications. MPEG-4 offers extended functionalities and improves upon the coding efficiency provided by previous standards. For instance, it supports variable pixel depth, object-based transmission, and a variety of networks including wireless networks and the Internet.

Multimedia authoring and editing capabilities are particularly attractive features of MPEG-4, with the promise of replacing existing word processors. In a sense, H.263 and MPEG-2 are embedded in

MPEG-4, ensuring support for applications such as digital TV and videophone, while it is also used for web-based media streaming.

MPEG-4 distinguishes itself from earlier video coding standards in that it introduces object-based representation and coding methodology of real or virtual audio-visual (AV) objects. Each AV object has its local 3D+T coordinate system serving as a handle for the manipulation of time and space. Either the encoder or the end-user can place an AV object in a scene by specifying a co-ordinate transformation from the object's local co-ordinate system into a common, global 3D+T co-ordinate system, known as the scene co-ordinate system.

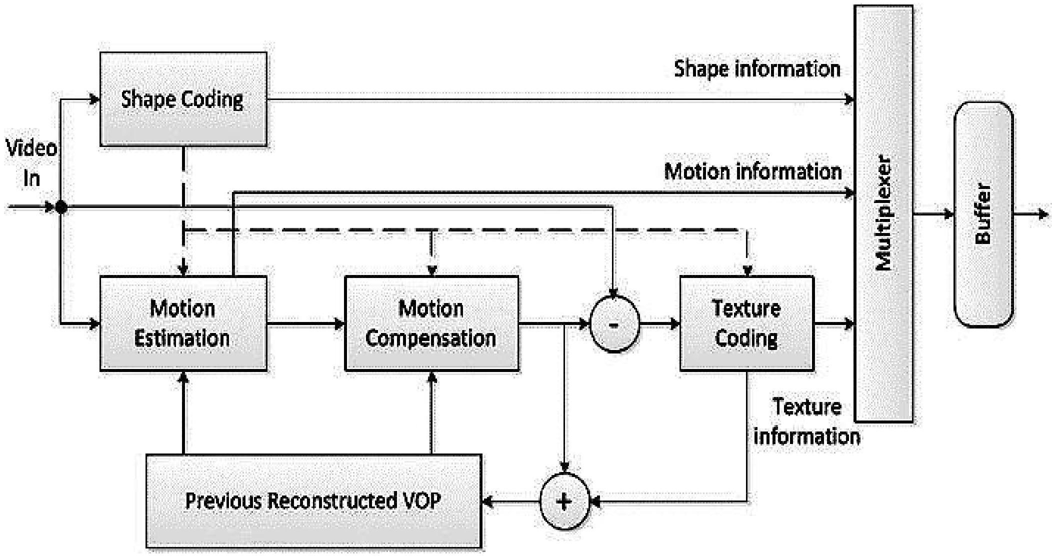
The composition feature of MPEG-4 makes it possible to perform bit stream editing and authoring in compressed domain.

One or more AV objects, including their spatio-temporal relationships, are transmitted from an encoder to a decoder. At the encoder, the AV objects are compressed, error-protected, multiplexed, and transmitted downstream.

At the decoder, these objects are demultiplexed, error corrected, decompressed, composited, and presented to an end user. The end user is given an opportunity to interact with the presentation. Interaction information can be used locally or can be transmitted upstream to the encoder.

The transmitted stream can either be a control stream containing connection setup, the profile (subset of encoding tools), and class definition information, or be a data stream containing all other information.

Control information is critical, and therefore it must be transmitted over reliable channels; but the data streams can be transmitted over various channels with different quality of service.



Video Object encoder structure: MPEG-4

11.6 SUMMARY

- The JPEG (Joint Photographic Experts Group) became an international standard in 1992.
- The larger the number of the quantizing factor, the higher the quality.
- .ZIP allows contained files to be compressed using many different methods, as well as simply storing a file without compressing it.
- In frequency modulation synthesis (or FM synthesis), the timbre of simple waveform (usually a sine wave) called a carrier, is changed by modulating it with another waveform in the audio range.
- MPEG-4, formally the standard ISO/IEC 14496, was ratified by ISO/IEC in March 1999 as the standard for multimedia data representation and coding.
- Unlike pixel-based treatment of video in MPEG-1 or MPEG-2, MPEG-4 supports content-based communication, access, and manipulation of digital audio-visual objects, for real-time or non-real-time interactive or non-interactive applications.

11.7 UNIT END EXERCISES

1. Write a note on ZIP Coding.
2. Write a note on JPEG Compression Technique.
3. What is Discrete Cosine Transformation?
4. Diagrammatically, explain JPEG Encoder – Decoder.

Reference:

I. Principles of Multimedia, Eighth reprint edition 2009, Ranjan Parekh, Tata McGraw-Hill Companies.

II. Wikipedia

❖❖❖❖

Thank You

eduClash.com
Just another Way To Learn