Conference Proceeding Issue Published in International Journal of Trend in Research and Development (IJTRD), ISSN: 2394-9333, www.ijtrd.com

Image Transformation and Video Compression Techniques

[1]D. Viji and [2]S. Kothbul Zeenath Banu,
[1,2]Assistant Professor, Department of Computer Application, Krishnasamy College of Science, Arts and Management for Women, Cuddalore, India

Abstract: Image and Video Processing are fortune topic of the field of research and development. Image processing is any form of signal processing for which the input is an image, such as all types of image files, photographs or frames of video, the output of image processing can be either an image or a set of characteristics related to the image. Image transformation techniques used to transmit an image. Video processing is a particular case of signal processing, where the input and output signals are video files or video streams. Framing and Video Compressing techniques are used in Television, VCRs, DVDs, video codec, video players and other devices. In This paper, we present Image transformation and Video Compression techniques and processing.

Keywords: *Image transformation, Fourier Transform, video compression, MPEG, Macroblocks.*

I. INTRODUCTION

Images can occupy a large amount of memory both in RAM and in storage. Image compression reduces the storage space required by an Image and the bandwidth needed when streaming that image across a network. Before transforming the image it should be compressed, because it occupies low storage. To transmit some data we need to work out frequencies and amplitudes are sending instead of the actual raw data. The signals in dictionary are called basis functions and the corresponding amplitudes needed are called coefficients. So now to send a picture, that transmitting end has to work out which set of images could be added together to give the picture. Then the transmitter sends the indexes of those images to the receiver. The receiver then looks up the pictures and then adds them together to give the received image.

Digital video compression/decompression algorithms (codecs) are at the heart of many modern video products, from DVD players to multimedia jukeboxes to video-capable cell phones. Understanding the operation of video compression algorithms is essential for developers of the systems, processors, and tools that target video applications. In this article, we explain the operation and characteristics of video codecs and the demands codecs make on processors. Here codecs is differ from one another and it has own significance. Basically video clips are made up of sequences of "individual images" or "frames", video compression algorithms share many concepts and techniques with still-image compression algorithms.

Image & its Types:

An art or a picture or graphical representation of the external form of a person or thing.

JPG/JPEG: JPG (Joint Photographic Group) optimized for photographs and similar continuous tone images that contain many, numbers of colors . JPG used to analyzing images and discarding kinds of information that the eye is likely to notice. It stores 24 bit color information. Joint Photographic Expert Group (JPEG) is an excellent way to store 24-bit photographic images, such as those used for imaging and multimedia applications. JPEG 24- bit (16 million color) images are superior in appearance to 8-bit (256 color) images on a Video Graphics Array (VGA) display and are at their most spectacular.

TIFF: The TIFF (Tagged Image File Format) is a flexible format that can be lossless or lossy¬ compression [8]. It normally saves 8 bits or 16 bits per color (red, green, blue) for 24-bit and 48-bit totals, respectively. The details of the image storage algorithm are included as part of the file. TIFF is used as a lossless image storage format that uses no compression at all. TIFF files are not used in web images. They produce big files, and more importantly, most web browsers will not display TIFFs.

GIF: Graphics Interchange Format (GIF) is useful for images that have less than 256-(2^8) colors, grayscale images and black and white text. GIF file works only on images with 8-bits per pixel or less, which means 256 or fewer colors. To store 24 bits image in GIF format that must first convert the image from 24 bits to 8 bits. GIF is a lossless image file format.

PNG: Portable Network Graphics (PNG) is a file format for lossless image compression. An image in a PNG file can be 10% to 30% more compressed than in a GIF format. It allows making a trade-off between file size and image quality when the image is compressed. It produces smaller files and allows more colors. PNG also supports partial transparency. Partial transparency can be used for many useful purposes, such as fades and antialiasing for text

.

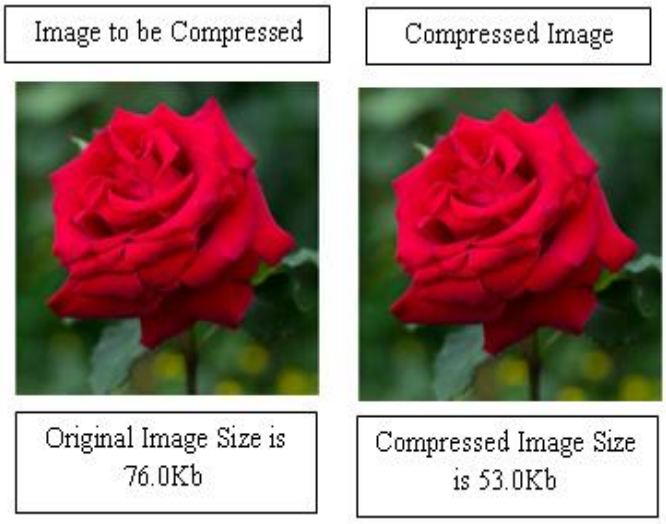
BMP: The Bitmap (BMP) file format handles graphics files within the Microsoft Windows OS. Typically, BMP files are uncompressed, hence they are large; advantage is that their simplicity, wide acceptance, and use in Windows program.

RAW: RAW refers to a family of bitmap image formats (output) that are options available on some digital cameras. These formats usually use a lossless or nearly-lossless.

Image Transformation:

Images can occupy a large amount of memory both in RAM and in storage. Image compression reduces the storage space required by an Image and the bandwidth needed when streaming that image across a network. Before transforming the image it should be compressed, because it occupies low storage. To transmit some data we need to work out frequencies and amplitudes are sending instead of the actual raw data. The signals in dictionary are called basis functions and the corresponding amplitudes needed are called coefficients. So now to send a picture, that transmitting end has to work out which set of images could be added together to give the picture. Then the transmitter sends the indexes of

those images to the receiver. The receiver then looks up the pictures and then adds them together to give the received image.



Transform methods in image processing:

An image transform can be applied to an image to convert it from one domain to another. Viewing an image in domains such as frequency or Hough space enables the identification of features that may not be as easily detected in the spatial domain.

Common image transforms include:

- Hough Transform, used to find lines in an image.
- Radon Transform, used to reconstruct images from fan-beam and parallel-beam projection data.
- Discrete Cosine Transform, used in image and video compression.
- Wavelet Transform, used to perform discrete wavelet analysis, denoise, and fuse images.
- Discrete Fourier Transform, used in filtering and frequency analysis.

Hough transform: The **Hough transform** is a feature extraction technique used in image analysis, computer vision, and digital image processing. The purpose of the technique is to find imperfect instances of objects within a certain class of shapes by a voting procedure. This voting procedure is carried out in a parameter space, from which object candidates are obtained as local maxima in a so-called accumulator space that is explicitly constructed by the algorithm for computing the Hough transform.

Radon transform: The **radon transform** takes a projection of a 2D **image** along a direction to give the 1D profile. Kind of like if you stood the **image** on its edge and took an x-ray of it. In fact it's used in 3D medical imaging for CT reconstruction. Radon further included formulas for the transform in three dimensions, in which the integral is taken over planes (integrating over lines is known as the X-ray transform). It was later generalized to higher-dimensional Euclidean spaces, and more broadly in the context of integral geometry.

Discrete cosine transform: The **discrete cosine transform** (**DCT**) represents an **image** as a sum of sinusoids of varying magnitudes and frequencies. For this reason, the **DCT** is often used in **image compression** applications. For example, the **DCT** is at the heart of the international standard lossy **image compression** algorithm known as JPEG.

Wavelet Transform: The discrete **wavelet transform** returns a data vector of the same length as the input is. Usually, even in this vector many data are almost zero. The continuous **wavelet transform** in contrary returns an array one dimension larger than the input data. For a 1D data we obtain an **image** of the time-frequency plane.

Discrete Fourier Transform: Fourier **Transform** is an important **image processing** tool which is used to decompose an **image** into its sine and cosine components. The output of the **transformation** represents the **image** in the Fourier or frequency domain, while the input **image** is the spatial domain equivalent.

Video Compression:

Video: Video is an electronic medium for the recording, copying, playback, broadcasting, and display of moving visual media. Video was first developed for mechanical television systems, which were quickly replaced by cathode ray tube systems which were later replaced by flat panel displays of several types.

In Digital Video, video processing is a also signal processing like image processing, where the input and output signals are video files or video streams. Video processing techniques are used in television sets, VCRs, DVDs, video codecs, video players and other devices.

For example—commonly only design and video processing is different in TV sets of different manufactures.

In terms of video codecs video filters are divided into three parts:

• **Prefilters:** used before encoding

- ❖ Video denoising,
- ❖ Size conversion (commonly downsampling)
- ❖ Contrast enhancement
- ❖ Deinterlacing
- ❖ Deflicking, etc

• **Intrafilters:** inside of codec

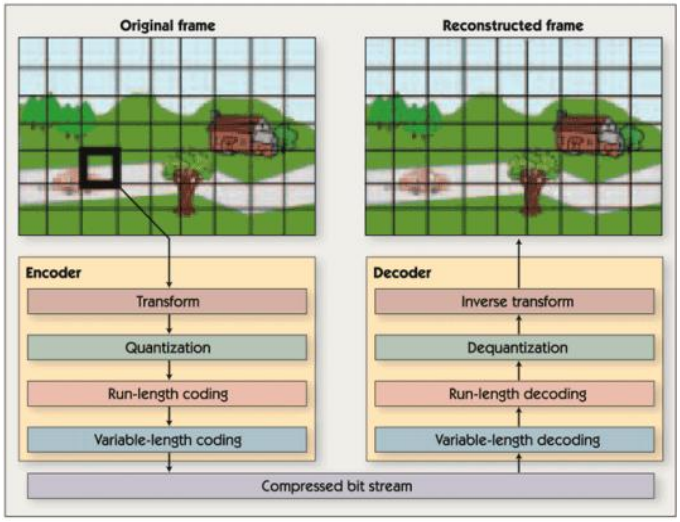
- ❖ Deblocking

• **Postfilters:** used after decoding Common

- ❖ Deinterlacing (to convert interlaced video to progressively scanned)
- ❖ Deblocking
- ❖ Deringing

Works of Video Compression:

Digital video compression/decompression algorithms (codecs) are at the heart of many modern video products, from DVD players to multimedia jukeboxes to video-capable cell phones. Understanding the operation of video compression algorithms is essential for developers of the systems, processors, and tools that target video applications. In this paper, the operation and characteristics of video codecs and the demands codecs make on processors. We also explain how codecs differ from one another and the significance of these differences. Starting with stills because video clips are made up of sequences of individual images, or "frames," video compression algorithms share many concepts and techniques with still-image compression algorithms. Therefore, we begin our exploration of video compression by discussing the inner workings of transform-based still image compression algorithms such as JPEG, which are illustrated in below figure

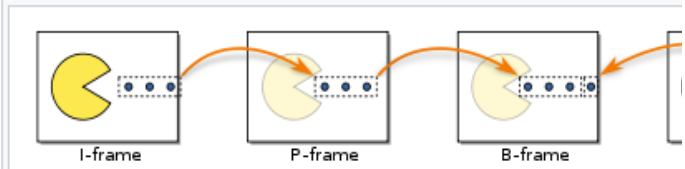


The image compression techniques used in JPEG and in most video compression algorithms are "lossy." That is, the original uncompressed image can't be perfectly reconstructed from the compressed data, so some information from the original image is lost. The goal of using lossy compression is to minimize the number of bits that are consumed by the image while making sure that the differences between the original (uncompressed) image and the reconstructed image are not perceptible—or at least not objectionable—to the human eye.

Different types of techniques are used for video compression among that here MPEG compression technique is used. MPEG is often considered synonymous with image compression. MPEG stands for Moving Pictures Experts Group, a committee that publishes international standards on video compression. MPEG is a portfolio of video compression standards. Image compression theory and implementation focuses on taking advantage of the spatial redundancy present in the image. Video is composed of a series of images, usually referred to as frames, and so can be compressed.

These different algorithms for video frames are called **picture types** or **frame types**. The three major picture types used in the different video algorithms are **I**, **P** and **B**. They are different in the following characteristics:

- **I**-frames are the least compressible but don't require other video frames to decode.
- **P**-frames can use data from previous frames to decompress and are more compressible than I-frames.
- **B**-frames can use both previous and forward frames for data reference to get the highest amount of data compression.



A sequence of video frames, consisting of two keyframes (I), one forward-frame (P) and one bi-directionally predicted frame (B).

Three types of pictures (or frames) are used in video compression: I, P, and B frames.

- An **I-frame** (Intra-coded picture) is a complete image, like a JPG or BMP image file.
- A **P-frame** (Predicted picture) holds only the changes in the image from the previous frame. For example, in a scene where a car moves across a stationary background, only the car's movements need to be encoded. The encoder does not need to store the unchanging background pixels in the P-frame, thus saving space. P-frames are also known as delta-frames.
- **B-frame** (Bidirectional predicted picture) saves even more space by using differences between the current frame and both the preceding and following frames to specify its content.

Macroblocks:

The pictures (frames) are segmented into macroblocks, and individual prediction types can be selected on a macroblock basis rather than being the same for the entire picture, as follows:

- I-frames can contain only intra macroblocks
- P-frames can contain either intra macroblocks or predicted macroblocks
- B-frames can contain intra, predicted, or bi-predicted macroblocks

Furthermore, in the H.264 video coding standard, the frame can be segmented into sequences of macroblocks called *slices*, and instead of using I, B and P-frame type selections, the encoder can choose the prediction style distinctly on each individual slice. Also in H.264 are found several additional types of frames/slices:

- SI-frames/slices (Switching I): Facilitates switching between coded streams; contains SI-macroblocks (a special type of intra coded macroblock).
- SP-frames/slices (Switching P): Facilitates switching between coded streams; contains P and/or I-macroblocks
- Multi-frame motion estimation (up to 16 reference frames or 32 reference fields)

Multi-frame motion estimation increases the quality of the video, while allowing the same compression ratio. SI and SP frames (defined for the Extended Profile) improve error correction. When such frames are used along with a smart decoder, it is possible to recover the broadcast streams of damaged DVDs.

Intra-coded (I) frames/slices (key frames):

- I-frames contain an entire image. They are coded without reference to any other frame except (parts of) themselves.
- May be generated by an encoder to create a random access point (to allow a decoder to start decoding properly from scratch at that picture location).
- May also be generated when differentiating image details prohibit generation of effective P or B-frames.
- Typically require more bits to encode than other frame types.

Often, I-frames are used for random access and are used as references for the decoding of other pictures. Intra refresh periods of a half-second are common on such applications as digital television broadcast and DVD storage.
For example, in videoconferencing systems it is common to send I-frames very infrequently.

Predicted (P) frames/slices:

- Require the prior decoding of some other picture(s) in order to be decoded.
- May contain both image data and motion vector displacements and combinations of the two.
- Can reference previous pictures in decoding order.
- Older standard designs (such as MPEG-2) use only one previously decoded picture as a reference during

decoding, and require that picture to also precede the P picture in display order.

- In H.264, can use multiple previously decoded pictures as references during decoding, and can have any arbitrary display-order relationship relative to the picture(s) used for its prediction.
- Typically require fewer bits for encoding than I pictures do.

Bi-directional predicted (B) frames/slices (macroblocks):

- Require the prior decoding of subsequent frame(s) to be decoded.
- May contain image data and/or motion vector displacements. Older standards allow only a single global motion compensation vector for the entire frame or a single motion compensation vector per macroblock.
- Include some prediction modes that form a prediction of a motion region (e.g., a macroblock or a smaller area) by averaging the predictions obtained using two different previously decoded reference regions. Some standards allow two motion compensation vectors per macroblock (biprediction).
- In older standards (such as MPEG-2), B-frames are never used as references for the prediction of other pictures. As a result, a lower quality encoding (requiring less space) can be used for such B-frames because the loss of detail will not harm the prediction quality for subsequent pictures.
- In H.264, B-frames may or may not be used as references for the decoding of other frames (at the discretion of the encoder).

CONCLUSIONS

Image Processing is the act of examining images for the purpose of identifying objects and judging their significance" Image analyst study the remotely sensed data and attempt through logical process in detecting, identifying, classifying, measuring and evaluating the significance of physical and cultural objects, their patterns and spatial relationship. Video processing and compression is a particular case of signal processing, where the input and output signals are video files or video streams. The primary objectives of on-going research on scalable video coding are to achieve high compression efficiency high flexibility (bandwidth scalability) and/or low complexity. Many algorithms for video compression which are designed with a different target in mind have been proposed. From that, this paper contained MPEG technique for video compression. And also discuss about the elements of Digital Video Processing.

References

[1] https://en.wikipedia.org>wiki>Video Compression _picture_types

[2] Wikipedia – Video processing http://en.wikipedia.org/wiki/Video_processing

[3] Bill Silver "An Introduction to Digital Image Processing" http://www.machinevisiononline.org/public/articles/cognex1.PDF

[4] Wikipedia – Computer vision http://en.wikipedia.org/wiki/Computer_vision

[5] Jae S. Lim. Two-Dimensional Signal and Image Processing. Prentice-Hall, 1990.

[6] A. Murat Tekalp. Digital Video Processing. Prentice Hall, 1995.

[7] S. Hecht. The visual discrimination of intensity and the weber-fechner law. General Phisiology, 1924.

[8] A. N. Netravali and B. Prasada. Adaptive quantisation of picture signals using spatial masking. In Processings of the IEEE, pages 536–548, April 1977.

[9] A.K. Jain. Fundamentals of Digital Image Processing. Prentice Hall, 1989.

[10] "Video Compression Fundamentals" https://www.oreilly.com/library/view/digital-video-processing/9780124157606/xhtml/CHP014.html