IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)
Volume 3, Issue 3 (Sep. – Oct. 2013), PP 69-73
e-ISSN: 2319 – 4200, p-ISSN No. : 2319 – 4197
www.iosrjournals.org

A Comparative study on Image and Video Compression Techniques

Gangadhar Tiwari [1], Debashis Nandi [2], Madhusudhan Mishra[3]
[1, 2]*(IT Department, NIT Durgapur India)* [3]*(ECE Department, NERIST, India)*

Abstract: *In the present era of Internet multimedia data especially Images and Videos are the most widely used digital format for data transmission. However due to their large data sizes and constraint of low bandwidth capacity of communication channel it is very difficult to transmit them at optimum speed maintaining the signal quality. Compression therefore, is a vital tool that not only reduces the data size thereby leading to faster data transmission but also protects it to some extent from transmission errors. A large variety of Image and Videos compression techniques are employed each having their own strengths and weaknesses. This paper is an effort to present an overview of image and video compression techniques, their working and comparison.*
Keywords: *Image Compression, Fidelity Criterion, Entropy, JPEG 2000, Video Compression*

I. Introduction

Compression is a tool to convert data to a different form by removing the existing redundancies. It makes the data human unreadable and reduce the size also. Compression of the signal enables faster transmission and less memory space is used. Moreover, the digital data protection against transmission errors ensures data reliability. In this presentation the Section-2 discusses the Image Compression Model and Fidelity Criterion. Section-3 discusses Information Theory. Section-4 is devoted on various Image compression techniques. Section-5 discusses MPEG Video Compression and Section-6 draws the conclusion and future scope in this field.

II. Image Compression Model

As discussed above, compression is performed by removing the redundant information. In images there are three sources of redundancy viz. **Coding Redundancy (**binary code representing grey values), **Inter pixel Redundancy (**correlation between adjacent pixels of image) and **Psycho visual Redundancy (**unequal sensitivity of the human eye to different visual information) [1]. Although image compression models differ in data compression, the Fig.1 below depicts a sample Image Compression Model. The original image is represented by *F(m,n)* while the compressed one by *F'(m,n)*. The source encoder is used to remove redundancy in the input image. The channel encoder is used as overhead to combat channel noise for e.g. introduction of a parity bit. The channel could be communication link or a storage/retrieval system. The job of the channel and source decoders is to basically undo the work of the source and channel encoders.

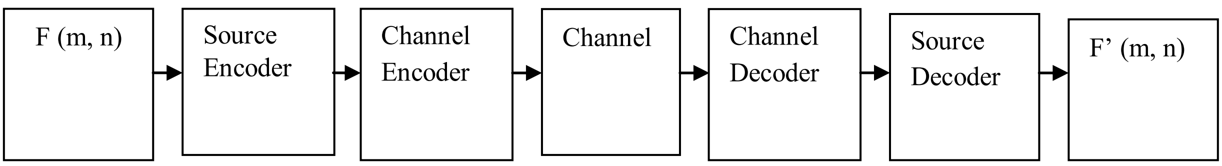


Figure-1 Image Compression Model

2.1 Fidelity Criterion for Image Quality Metrics

Lossless and lossy compression uses various methods to calculate compression quality. Standard criteria like compression ratio, execution time, etc are used to estimate the compression in both cases, however in lossy compression, it should calculate both the type and amount of degradation induced in the reconstructed image. The aim of image quality measurement is to precisely measure the difference between the original and reconstructed images. The objective quality measure like Peak Signal to Noise Ratio (PSNR), measures the difference between the individual image pixels of original and reconstructed images. The Structural Similarity Index (SSIM) is highly adapted for extracting structural information. The SSIM index is a reference metric which is computed on various image windows. The measure between two windows a and *b* of common size *N*N* is given in eq. (1):

$$SSIM(a,b) = \frac{(2\mu_a\mu_b + c_1)(2\sigma_{ab} + c_2)}{(\mu_a^2 + \mu_b^2 + c_1)(\sigma_a^2 + \sigma_b^2 + c_2)} \quad (1)$$

with μ_a the average of a; μ_b the average of b;σ_a^2 the variance of a; σ_b^2 the variance of b; σ_{ab} the covariance of a and b; $c_1=(k_1^2L^2)$ and $c_2=(k_2^2L^2)$ two variables to stabilize the division with weak denominator; L is the dynamic range of the pixel-values and $k_1=0.1$ and $k_2=0.03$ by default. The resultant SSIM index is a decimal value between -1 and 1, and value 1 is only available if two data sets are the same.

Fidelity Criterion is used to calculate the amount of data lost during compression. It is categorized into Subjective and objective. **Objective fidelity criterion** involves a quantitative approach to error criterion for e.g. root mean square error, mean square signal to noise ratio. It is useful in analyzing error estimation involved in a compression scheme; however our eyes do not always see things as they are. **Subjective fidelity criteria** are quality evaluations based on a human observer. There are absolute comparison scales, which are based only on the decompressed image, and there are relative comparison scales that involve screening the original and decompressed images together. However subjective fidelity criteria vary from person to person.

III. Information Theory

Shannon (1940) pioneered Information theory that answers questions like what is the minimum amount of data needed to represent an image without loss of information. Or, what is the best compression model? The basic principle is that the generation of information may be viewed as a probabilistic process. The input generates one of N possible symbols from the source alphabet set A={a, b, c,…..z},{0,1},{0, 2, 4….280}, etc. in unit time. The source output can be denoted as a discrete random variable E, which is a symbol from the alphabet source along with a corresponding probability (z). When an algorithm scans the input for an occurrence of E, the product is a gain in information denoted by $I(E)$, and quantified as:

$$I(E) = \frac{1}{P(E)} \qquad (2)$$

This relation indicates that the amount of information attributed to an event is inversely proportional to the probability of that event. Another important concept is the Entropy of a source (H(z)), and defined as the average amount of information gained by observing a particular source symbol. Mathematically, it is expressed as: For a random variable with outcome, {a_i=1,……n}, the Shannon entropy, a measure of uncertainty is denoted by equation (3)

$$H(A) = -\sum_{i=1}^{n} p(a_i)\log_b p(a_i) \qquad (3)$$

where $p(a_i)$ is the probability mass function of outcome (a_i). Basically, this allows an algorithm to quantize the randomness of a source [2]. The amount of randomness is vital because the more random a source is the more information is needed to represent it.

IV. Image Compression Techniques

Although, there doesn't exists any hard and fast rules for what is the best algorithm for what situation; some basic parameters like Bits/pixel Interlace method and Transparency are to be kept in mind while selecting the compression techniques [3]. The review below discusses some of the major image compression algorithms:

4.1 Graphics Interchange Format Compression (GIF)

It uses a compression algorithm called "LZW," which is a lossless algorithm. However, GIF only supports a maximum of 256 colors. It has a running time of ***$O(m^2)$***, where ***m*** is number of colors between 2 and 256**.** The first step in GIF compression is to "index" the image's color palette. This decreases the number of colors in image to a maximum of 256. The smaller the number of colors in the palette, the greater the efficiency of the algorithm. Another factor that affects GIF file size is interlacing. If an image is interlaced, it will display itself all at once, incrementally bringing in the details.

4.2. Portable Network Graphic Compression (PNG)

It uses the LZ77 compression algorithm. PNG is an open format and has a better average compression than GIF and other features like alpha transparency. It also supports 24-bit images. PNG is a lossless algorithm and has a running time of ***$O(m^2 \log m)$,*** where ***m*** is number of colors in the image. LZ77 compression replaces repetitions with references to previous occurrences. PNG also offers filtering options to rearrange pixel data before compression.

4.3. Tagged Interchange File Format Compression (TIFF)

The main advantages of TIFF over PNG and GIFF are extendibility, portability and revisability. TIFF enables for exchanging image information along with image editing applications. The compression algorithms include run length encoding, Huffman encoding and LZW. TIFF has the ability to decompose an image by tiles rather than scan lines. The limitations include; there are no provisions for storing vector graphics, text annotation, etc.

4.4 Joint Photographic Experts Group Algorithm (JPEG)

The JPEG algorithm was created to compress photographs with minimal data loss and high compression ratios. JPEG has four modes of operation:
1. **Sequential encoding-** Each image component is encoded in a single left-to-right, top-to-bottom scan.
2. **Progressive encoding-** Image is encoded in multiple scans for applications where transmission time is longer.
3. **Lossless encoding-** The image is encoded to guarantee exact recovery of every source image sample value.
4. **Hierarchical encoding-** The image is encoded at multiple resolutions so that lower-resolution versions may be accessed without decompressing the image at its full resolution.

JPEG compression and decompression consist of 4 distinct and independent phases [4].
1. The image is divided into 8 x 8 pixel blocks.
2. A discrete cosine transform is applied to each block to convert the information from the spatial domain to the frequency domain. The 2D discrete cosines transform equation is represented by equation (4)

$$F(u,v) = \frac{1}{4C(u)C(v)\sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y)(\frac{\cos(2x+1)u\pi}{16})(\frac{\cos(2y+1)v\pi}{16})} \qquad (4)$$

where ***f(x,y)*** is the 8-bit image value at coordinates ***(x, y),*** and ***F(u,v)*** is the new entry in the frequency matrix.
3. The frequency information is quantized to remove unnecessary information. The quantization equation (**5**) is used for each block in the image.

$$F_{Quantize}(u,v) = (\frac{F(u,v)}{Q(u,v)}) + 0.5 \qquad (5)$$

By adding 0.5 to each value, we essentially round it off.
4. Standard compression techniques compress the final bit stream.
5. Decompressing a JPEG image is basically the same as performing the compression steps in reverse in opposite order.

4.5 JPEG file interchange format (JFIF)

JFIF provides ease of use and has a simple format that only transports pixels. JFIF image orientation is top-down. This means that the encoding proceeds from left to right and top to bottom. Spatial relationship of components is defined with respect to the highest resolution component.

4.6 Joint Bi-level Image Experts Group Compression (JBIG)

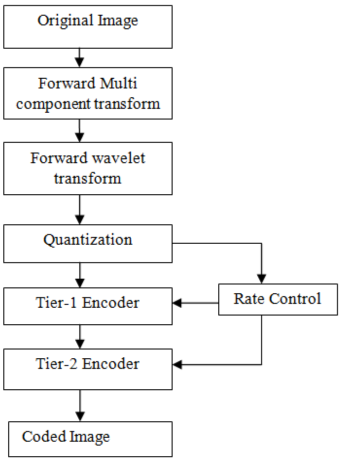
JBIG is a method for lossless compression of bi-level (two-color) image data [5]. JBIG also supports both sequential and progressive encoding methods. JBIG encodes redundant image data by comparing a pixel in a scan line with a set of pixels already scanned by the encoder. The advantage of JBIG is platform-independence. However, it is not free.

4.7 JPEG Tiled Image Pyramid (JTIP)

JTIP cuts an image into a group of tiled images of different resolutions. The highest level of the pyramid is called the **vignette** which is 1/16 the original size and is used for browsing. The next tile is called the **imagette** which is ¼ the original size and is used for image comparison. The next tile is the full screen image which is the only full representation of the image. Below this tile would be the high and very high definition images. The primary problem with JTIP is how to adapt the size of digital image to the screen definition or selected window.

4.8 JPEG 2000

JPEG 2000 algorithm relies on wavelets to convert the image from the spatial domain to the frequency domain**.** It compresses wavelet data in a bit plane-by-bit plane approach because of its intrinsic embedding and feasible coding performance [11]. Wavelets are much better at representing local features in a function, and thus create less loss in the image. Also, wavelet can examine the image at multiple resolutions, and determine exactly how to process the image. Moreover, JPEG 2000 considers an entire image at once; this makes the process very fast [6]. The structure and functioning of JPEG 2000 can be best understood by Fig.2 (converting original image to coded image) and Fig.3 (converting coded image to reconstructed image) as below**:**



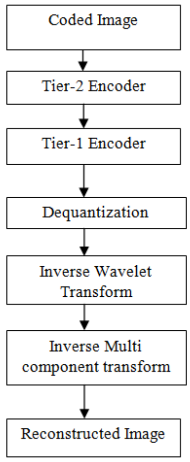


Fig.2 Original image to Coded image conversion **Fig.3 Coded image to reconstructed image conversion**

V. Moving Pictures Expert Group (MPEG) Compression

MPEG is used to compress video files. It is a DCT based scheme with Huffman encoding [7]. The primary technique used by MPEG for compression is transform coding with an 8x8 DCT spatial domain blocks. The basic idea is to transform a stream of discrete samples into a bit stream of tokens which takes less space. MPEG links the streams with layering so that the data types synchronized and multiplexed. The key MPEG schemes are

5.1 MPEG-1

It works on high bit rates in the 128 Mbps range. It handles progressive non-interlaced signals. MPEG1 has parameters of Source Input Format (**SIF**) pictures and a coded bit rate less than 1.86 Mbps. e.g. MP3 audio files.

5.2 MPEG-2

It was developed for lower bit rates in the 64 Mbps range that can handle interlaced broadcast video. It de-correlates multi channel discrete audio signals that have a higher redundancy factor then regular stereo sound. The two levels of MPE**G-** 2 are the SIF Low Level and Main Level [8].

5.3 MPEG-4

It was developed for low bit rates in the 32 Mbps range that would handle the new videophone standard (H.263). MPEG4 has the ability to pick the subjects of a video out of the scene and compress them separately. The MPEG syntax provides an efficient way to represent image sequences in more compact coded data. During compression, macro block predictions are formed out of arbitrary 16 x 16 pixel areas from previously reconstructed pictures. There are no boundaries that limit the location of a macro block prediction within the previous picture. Picture coding macro block types are *(I, P, B)*. All macro blocks within a ***I*** picture must be coded Intra. However, macro blocks within a ***P*** picture may either be coded as Intra or Non-intra. Finally, macro blocks within the ***B*** picture can be independently selected as either Intra, Forward predicted, backward predicted, or both forward and backward predicted. The macro block header contains macro block type, which can flip these modes on and off.

5.4 Multimedia Hypermedia Experts Group (MHEG)

MHEG is a data interchange format with three primary levels of representation to organize and compress multimedia files. It includes MHEG classes, objects and run-time objects. MHEG has data structures which are reused in multiple classes. The MHEG class hierarchy has a root class called H-object class. It defines two data structures common to all other classes and is inherited by all lower level classes. Class identifier identifies the type of each encoded class. MHEG will store its reusable objects in a database. There is a content class that describes objects to be presented to the user. The action class determines the behavior of the basic objects. The link class defines a logical relationship between the action object and a content object. A composite class allows composite objects to be part of other composite objects. The container class provides a set of objects that are transferred as a whole set. The descriptor class encodes information about objects in a presentation and uses its information to determine if there are available resources for the presentation. Finally, the script class communicates with external functions or programs.

5.5 MPEG-4 Part-X

Also called H.264 standard/MPEG-4AVC aims at coding video sequences at approximately half the bit rate compared to MPEG-2 at the same quality. It also aims at having significant improvements in coding efficiency using CABAC entropy coder, error robustness and network friendliness. Parameter set concept, arbitrary slice ordering, flexible macro block structure, redundant pictures, switched predictive and switched intra pictures have contributed to error resilience / robustness of this standard [9]. H.264 main profile which is the subset of high profile is designed with compression coding efficiency as its main target. Fidelity range extensions provide a major breakthrough with regard to compression efficiency.

5.6. AVS CHINA Part7 Intra Frame

Audio-video coding standard (AVS) is the digital video codec standard developed by China. It is an advanced second generation source coding standard, compatible with MPEG-2 standard and has a coding efficiency similar to that of H.264 with lower computational complexity [10]. It is divided into nine parts as shown below

Part-1	System for Broadcasting
Part-2	High Definition Video
Part-3	Audio
Part-4	Conformance Test
Part-5	Reference Software
Part-6	Digital Rights Management
Part-7	Mobility Video
Part-8	System over IP
Part-9	File Format

5.7 DIFFERENCE BETWEEN THE CODECS

The main difference between the codecs is at the transformation stage. JPEG use discrete cosine transform (DCT) to de-correlate the image. JPEG2000 de-correlates image data via the global discrete wavelet transform (DWT) while H.264 and HD Photo choose the block-based coding framework very similar to the discrete cosine transform (DCT). The major difference between H.264's and HD Photo's transformation stage is the way the two coders handle inter-block de-correlation. While H.264 relies heavily on adaptive spatial prediction of the current block from its neighbors, HD Photo employs an overlap operator which performs preprocessing of pixels along the block boundaries before feeding them into the core DCT-like 4x4 block transform. Another difference is at the entropy coding stage. H.264/AVC employs intra prediction in spatial domain. AVS China follows the same technique. This avoids propagating the error due to the motion compensation in inter-coded macro-blocks. On the other hand, all the previous video coding standards like H.263 and MPEG-4 visual use intra prediction in transform domain coding.

VI. Conclusion

In this review, different compression techniques and their performance is analyzed for e.g. H.264, JPEG, JPEG 2000 etc. Both objective and subjective methodologies for the comparison of various still image coding techniques are used. While PSNR gives an objective quality metrics of image, SSIM gives subjective quality which takes human visual perception (HVS) into account. It is found that the performance of H.264 is better. AVS China performs better than JPEG, but less compared to JPEG2000. Although the compression techniques discussed above can help image and video compression to a great extent but there exists artifacts for each of them as discussed above. So from security perspective it is not tolerable because the true message and its integrity are lost. Thus multimedia compression is still in its developing stage and has a lot to be researched.

References

[1]. A.E Cawkell, Introduction through -A Quick Comparison of Image Compression Techniques-A Guide to Image Processing and Picture Management, Gower Publishing Limited, 1994.

[2]. C.E. Shannon, A Mathematical Theory of Communication, The Bell System Technical Journal,27, 1948,379–423.

[3]. Harris et. al, The JPEG Algorithm for Image Compression: A Software Implementation and some Test Results, Signals, Systems and Computers, Conference, IEEE,2, 2002, 870-875

[4]. Neelamani et al, JPEG compression history estimation for color images, IEEE Transactions on Image Processing, 15, 2006,1365–78

[5]. Jiang et al, JPEG image compression using quantization table optimization based on perceptual image quality assessment, Asilomar Conference on Signals, Systems and Computers, 2011, 225 – 229

[6]. Ebrahimi et al, Secure JPEG 2000-JPSEC, IEEE International Conf. on Acoustics, Speech, and Signal Processing,4, 2003, 716-19

[7]. Tim Borer, Open Technology Video Compression for Production and Post Production, British Broadcasting Corporations, 2007.

[8]. Rao et al, JPEG 2000: Video/Image Processing and Multimedia Communications, International Symposium on VIPromCom, IEEE, 2002, 1 - 6

[9]. Kalra et al, Video codec comparative analysis between H.264 and DIRAC PRO/VC-2 ,Canadian Conference, IEEE, 2011, 951 -955

[10]. Navakitkanok, Improved rate control for advanced video coding (AVC) standard under low delay constraint, International Conference on Information Technology: Coding and Computing, 2, 2004, 664 – 66.

[11]. Jimenez-Rodriguez et al, Visually Lossless JPEG 2000 Decoder, Data Compression Conference (DCC), 2013, 161- 170.