ISSN No. 0976-5697

Volume 2, No. 1, Jan-Feb 2011

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

An Analytical Study on Compression Ratio of Video Frames in Stereoscopic Images

Dr.M.Mohamed Sathik
Reader in Computer Science
Sadakkathella Appa College(Autonomous)
Tirunelveli-627 011,South India
mmdsadiq@gmail.com

T.Ramaprabha*
M Sc M Phil
Lecturer, Dept of Computer Science
Sarah Tucker College (Autonomous)
Tirunelveli– 627007 ,South India
ramaradha1971@gmail.com

Abstract : The increasing use of digital video has led to the use of video data compression to allow compact storage and fast video transfer rates. A wide variety of methods are used to compress video streams. Video data contains spatial and temporal redundancy, making uncompressed video streams extremely inefficient. Spatial redundancy is reduced differences between parts of a single frame; this task is known as intraframe compression and is closely related to image compression. Likewise, temporal redundancy can be reduced differences between frames; this task is known as interframe compression, including motion compensation and other techniques. The most common modern standards are MPEG-2, used for DVD and satellite television, and MPEG-4, used for home video. In this paper I analysis the ratio of different video frames in stereoscopic images.

Keywords*:* video, Compression. Bit rates, Video frames, Compression Ratio.

I. INTRODUCTION

A. *Bit rate*

Bit rate is a measure of the rate of information content in a video stream. It is quantified using the bit per second (bit/s or bps) unit or Megabits per second (Mbit/s). A higher bit rate allows better video quality. For example VideoCD, with a bit rate of about 1 Mbit/s, is lower quality than DVD, with a bit rate of about 5 Mbit/s. HD (High Definition Digital Video and TV) has a still higher quality, with a bit rate of about 20 Mbit/s. Variable bit rate (VBR) is a strategy to maximize the visual video quality and minimize the bit rate. On fast motion scenes, a variable bit rate uses more bits than it does on slow motion scenes of similar duration yet achieves a consistent visual quality. For real-time and non-buffered video streaming the bandwidth is fixed, e.g. in videoconferencing delivered on channels of fixed bandwidth, a constant bit rate (CBR) must be used.

II. VIDEO FRAMES

Video plays for hours at a time but is actually compressed in short sequences. The length depends on the video format and target platform. Fifteen frames for a group of pictures (GOP) is typical. There are three kinds of frame in a GOP: Intra-frames (I-frames) at the start; Predicted frames (P-frames) at the end; and Bidirectionally coded frames (B-frames) in-between. The I-frame is coded first, just like a still photograph. The image is divided into 16 × 16 pixel macroblock. Macroblock are grouped into horizontal slices that help with dropout reconstruction. Some bit rate saving results immediately from culling similar macroblock and only buffering unique blocks in Fig 1. Then P-frame content is analyzed. Only new blocks not present in the I-frame are retained. The collection of macroblock describes the frames at each end of the GOP. Now, the intervening B-frames can be coded more efficiently.

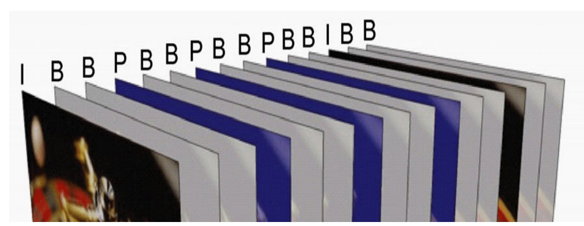

Figure 1. video frames

B-frame macroblocks are discarded if they duplicate any I- and P-frame blocks already collected. The buffer maintains these unique macroblocks that are referred to by different frames. The last frame of the GOP must be delivered earlier than it is presented for display so the B-frames can be reconstructed. Frame reordering immediately causes some coding latency because of the GOP length. If latency is a problem (perhaps for video conferencing), use shorter GOPs or omit P-frames altogether. Motion JPEG encodes I-frames only. It won't achieve compression ratios as high as MPEG, but it does produce editable content. I-frames could encode as small as 40KB. A GOP with just 15 I-frames would occupy 600KB. A single P-frame might save 35KB.. The whole GOP might encode in less than 60K. So, P- and B-frames yield a useful 10:1 compression if we can tolerate latency

III. VIDEO COMPRESSION TECHNIQUE

Video compression is about reducing and removing redundant video data so a digital video file can be effectively sent and stored. The process involves applying an algorithm to the source video to create a compressed file that is ready for transmission or storage. To play the compressed file, an inverse algorithm is applied to produce a video that shows virtually the same content as the original source video. The time it takes to compress, send, decompress and display a file is called latency. The more advanced the compression algorithm, the higher the latency, given the same processing power. A pair of algorithms that works together is called a video codec (encoder/decoder). Video codec's implement different standards are normally not compatible with each other; that is, video content that is compressed using one standard cannot be decompressed with a different standard. This is simply because one algorithm cannot correctly decode the output from another algorithm but it is possible to implement many different algorithms in the same software or hardware, which would then enable multiple formats to be compressed.

Different video compression standards utilize different methods of reducing data, and hence, results differ in bit rate, quality and latency. The general and important steps in video compression are:

[a] Motion Estimation

[b] Motion Compensation and Image Subtraction

[c] Discrete Cosine Transform

[d] Quantization

[e] Run Length Encoding

[f] Entropy Coding – Huffman Coding

A. *Motion Estimation*

Motion estimation is the process of calculating motion vectors by finding matching blocks in the future frame corresponding to blocks in the current frame. Motion estimation helps in detecting the temporal redundancy. Various search algorithms have been devised for estimating motion. The basic assumption underlying these algorithms is that only translational motion can be compensated for Rotational motion and zooming cannot be estimated by using block based search algorithms. It is known to be the most crucial and computationally intensive process in the video compression algorithm.

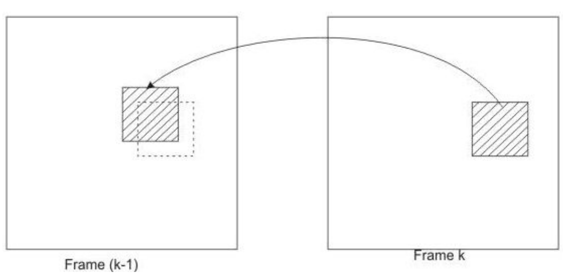

Figure 2. backward motion estimation

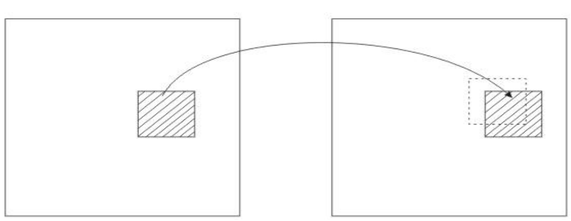

Figure 3 forward motion estimation

The motion estimation that we have discussed in Section-20.3 and Section 20.4 is essentially backward motion estimation, since the current frame is considered as the candidate frame and the reference frame on which the motion vectors are searched is a past frame, that is, the search is backward. Backward motion estimation leads to forward motion prediction.

Since most of the video streams have a frame rate ranging from 15 to 30 frames per second, there is never a very large motion of any object between two successive frames. Therefore most search algorithms search for matching block in the neighborhood of the position of the current block in the next frame. The region where matching block is searched for is called the search region.

The task to be performed by the search algorithm is to find the best match for a block in the current frame in the next frame. A typical block size is 8x8 or 16x16 pixels. The quality of match found will depend on the value of Mean Absolute Error. This is the average absolute pixel-wise difference between two blocks, reference block in the current frame and probable match found in the next frame. The matching block is figured out on the basis of the magnitude of the value of its mean error. Smaller the magnitude better is the match. The displacement of the block with the minimum MAE is taken as the motion vector.

Smaller the magnitude better is the match. The displacement of the block with the minimum MAE is taken as the motion vector.

B. *Motion Compensation and Image Subtraction*

The process of Motion Estimation and Motion Compensation is similar to DPCM. The idea is to reduce the bandwidth required for the video by sending only the difference frames instead of the actual frames. The motion vectors produced during Motion Estimation are utilized in the Motion Compensation process in order to produce the predicted image in the encoder just like it would be produced in the Decoder. The two images (current frame and the motion compensated frame) are now subtracted and the difference is sent to the receiver along with the motion vectors. The block diagram of the Encoder is given below in Figure 2. in order to illustrate the idea.

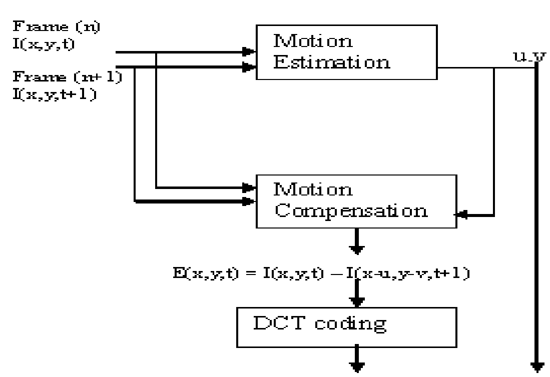

Figure.4 Block Diagram of Video Encoder

Thus the decoder can produce the exact copy of the future frame by first motion compensating the current frame using the motion vectors and then adding the difference image. The two measures are taken , they are SAD(Sum of absolute difference) and SSD(Sum of squared difference)

$$SAD = \sum \left| C_{i,j} - R_{I+U,J+V} \right|$$
$$I,J \in Currentblock$$

$$SSD = \sum \left(C_{i,j} - R_{I+U,J+V} \right)^2$$
$$I,J \in Currentblock$$

are adopted to evaluate how closely a candidate macro-block matches the current one. where Cij and R(i+u, j+v) are the pixels being compared in the current macro-block and the macro-block on the reference frame, respectively. Usually only the luminance plane is computed for SAD or SSD.

C. *Discrete Cosine Transform*

Discrete Cosine Transformation (DCT) is a process of representing original data as a linear sum of basic cosine functions with different frequencies. Each image delivers one brightness and two color signals per pixel. The DCT converts these signals into frequency coefficients containing the color and brightness information. The signals can then be compressed more easily.

DCT:

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

where

$$C(u) = \frac{1}{\sqrt{2}} \text{ for } u = 0$$
$$C(u) = 1 \text{ for } u > 0$$
$$C(v) = \frac{1}{\sqrt{2}} \text{ for } v = 0$$
$$C(v) = 1 \text{ for } v > 0$$

IDCT :

$$f(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u,v) \cdot \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cdot \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

DCT based image coding is the basis for almost all the image and video compression standards. Discrete Cosine Transform is a derivative of the Discrete Fourier Transform (DFT), which is encountered very commonly in Digital Signal Processing. The fundamental operation performed by DCT is to transform the space domain representation of an image to a spatial frequency domain (known as DCT domain

D. *Quantization*

Quantization is the use of complex mathematical operations to ensure that image parts that are important to the human eye are represented precisely and irrelevant information is represented with less precision. Huffman encoding reduces the amount of transferred data using statistical distribution of quantized DCT coefficients. This method evaluates how often and with what probability certain values occur. Values that seldom occur receive a long code, while values occurring often receive a short code. The MPEG file consists of compressed video data, called the video stream, and compressed audio data, called the audio stream. It can also contain only one of the streams mentioned above. The human eye is not sensitive to the high frequency content in an image. Therefore removal of these spatial frequencies does not lead to any perceptible loss in image quality. This is the basic principle behind quantization. The spatial frequency content of the image is obtained by using the DCT operation, which is followed by a removal of the high frequency content that is the quantization process. The JPEG standard recommends standard values of quantization tables which are used to deemphasize higher frequencies in the DCT image. Quantization is a lossy process and some data is lost during quantization. This loss of information is irreversible.

E. *Run Length Encoding (RLE)*

Run-length encoding is the next stage of the compression process. It encodes the runs of zeroes. If pixel values are correlated to their neighbors, then there will be sequences of the same value. Instead of coding all the repeat values, just encode the first value and then give the run length of the sequence. Intuitively, one can understand how RLE can help in achieving compression. Suppose the data is 00000…0(ten times). Now instead of writing ten zeroes one can send only 0-10, which could be taken to mean that a zero occurs 10 times. This is how compression is achieved in Run-Length Encoding. Runs of zeroes are encoded in a 16 bit or 8 bit format. A higher compression can be achieved in Run-Length encoding if we somehow obtain longer strings of zeroes. This is achieved by performing RLE in a zigzag manner on a block. In the DCT image the higher frequency content is always found towards the lower right hand corner of the DCT image while the lowest frequencies are in the upper left hand corner of the image. During quantization the higher frequencies are reduced to zero and therefore the values in the lower right hand side are mostly zero. Therefore by performing RLE in a zig-zag manner, we try to obtain runs of zeroes out of the lower right hand side of the DCT domain representation.

F. *Huffman Encoding*

Huffman encoding is a form of entropy encoding and it is based on Shannon's Information theory. The fundamental idea behind Huffman encoding is that symbols, which occur more frequently, should be represented by fewer bits, while those occurring less frequently should be represented by more number of bits. This scheme is similar to the one utilized in Morse code. Shannon has proved that the entropy of the total message gives the most efficient code, with minimum average code length, for sending a message. Given n symbols S1 to Sn-1 with probabilities of occurrence P1 to Pn-1 in a certain message, the entropy of the message will be given by

$$Entropy = Pi \log 2(1/Pi)$$

Huffman encoding attempts to minimize the average number of bits per symbol and try to get a value close to entropy.

IV. SUMMARY OF VIDEO COMPRESSION STANDARDS

MPEG is the acronym for the Moving Pictures Experts Group. There are different MPEG standards implemented for use in specific applications. MPEG-1: It is a compression standard for audio and moving pictures, with support for bit rates up to 1.5 Mbit/sec. This is a popular standard for streaming videos as .mpg files over the internet. The ubiquitous MP3 format, standing for MPEG 1 Audio Layer 3, is a famed standard for audio compression. MPEG-4: It is a popular object-based compression standard for multimedia compression. Objects in a frame are tracked independently and compressed together to form an MPEG4 stream. The result is an efficient compression standard flexible over a vast range of bit rates. DV: DV is a high-resolution digital video format which employs lossy compression where certain redundant information in a file is permanently deleted, so that even when the file is uncompressed, only a part of the original information is still there. H.261 and H.263**:** H.261 is standard for duplex communication over ISDN lines and supports data rates in multiples of 64Kbit/s. Flash Video**:** The most popular compression format for videos on the internet .

V. CONCLUSION

The progression of the MPEG format: MPEG-1 delivers basic capabilities. MPEG-2 adds interlacing support for broadcast TV and DVD. MPEG-4 part 2 adds more sophisticated coding tools and alpha channels. MPEG-4 adds more efficient DCT computation and better result . The compression ratios possible in the video content can collect 25 percent of the macroblocks in an I-frame. I-frame 75 percent ,B- and P-frames 10 percent , Sub-sampling 50 percent DCT/Entropy 50 percent

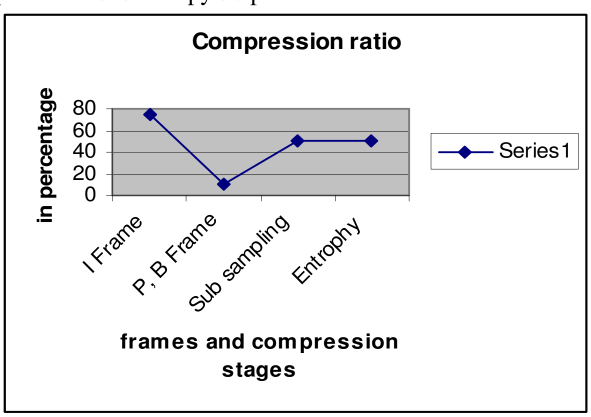

Figure: 5

That is about a 50:1 compression factor for a well-tuned compression system. Reducing picture size and frame rate for Internet streaming will improve the performance. In this paper I analyzed the compression ratio of different kind of frames along with the stages of video compression.

VI. REFERENCES

[1] Trucco, E; Verri, A; Introductory Techniques for 3-D Computer Vision (1998) Prentice Hall

[2] Hannah M J; Computer Matching of Areas on Stereo Images, Ph.D Thesis, Stanford University (1974)

[3] Rousseeuw P J; Least Median of Squares Regression, J. of Amer. Stat. Assoc., 79, pp. 871-880, 1984

[4] L. Irsara and A. Fusiello; Quasi-euclidean uncalibrated epipolar rectification Rapporto di Ricerca RR 43/2006, Dipartimento di Informatica - Universit? di Verona, 2006;

[5] Tordoff, B; Murray, DW. "Guided sampling and consensus for motion estimation." European Conference n Computer Vision, 2002.

[6] K. M¨uhlmann, D. Maier, J. Hesser, and R. Manner. Calculating dense disparity maps from color stereo images, an efficient implementation. IJCV, 47:79–88, 2002.

[7]. J. R. Ohm and E. Izquierdo M., "An Object-Based System for Stereoscopic Videoconferencing with Viewpoint Adaptation," Proc. On European Symposium on Advanced Imaging.

[8] J. R. Jain and A. K. Jain, "Displacement measurement and its application in interframe image coding," IEEE Trans. Comm., vol. 29, pp. 1799-1808, Dec. 1981.

[9] 2. W. Woo and A. Ortega, "Overlapped block disparity compensation with adaptive windows for stereo image coding," IEEE Trans. Circuits and Systems for Video Technology, vol. 10, no. 2, pp. 194~200, Mar. 2000.

[10] 3. D. R. Clewer, L. J. Lewer, C. N. Canagarajah, D. R. Bull, and M. H. Barton, "Efficient multiview image compression using quadtree disparity estimation," ISCAS 2001, vol. 5, pp. 295~298, May 2001.

[11] M. Orchard and G. Sullivan, "Overlapped block motion compensation: an estimation-theoretic approach," IEEE Trans. Image Processing, vol. 5, pp. 693~699, Mar. 1994.

[12] R. Rajagopalan, E. Feig, and M. Orchard, "Motion optimization of ordered blocks for overlapped block motion compensation," IEEE Trans. Circuit and Systems for Video Technology, vol. 8, no. 2, Apr. 1998.

[13] Final Committee Draft of Joint Video Specification, ITU-T Recommendation H.264 and ISO/IEC 14496-10 Advanced Video Coding (AVC), Oct. 2002

[14] International Telecommunication Union, "Video codec for low bitrate communication," ITU-T Recommendation