An overview Survey on Various Video compressions and its importance

Preeja Priji[1], Rashmi S Nair[2]

[1]M.Tech. Student, Department of CSE, Mohandas College of Engineering and Technology, Anad Thiruvananthapuram Kerala, India

[2]Asst. Professor Department of CSE, Mohandas College of Engineering and Technology, Anad Thiruvananthapuram Kerala, India

Abstract—*With the rise of digital computing and visual data processing, the need for storage and transmission of video data became prevalent. Storage and transmission of uncompressed raw visual data is not a good practice, because it requires a large storage space and great bandwidth. Video compression algorithms can compress this raw visual data or video into smaller files with a little sacrifice on the quality. This paper an overview and comparison of standard efforts on video compression algorithm of: MPEG-1, MPEG-2, MPEG-4, MPEG-7*

Keywords—Video compression, DCT, IDCT, MPEG, H.261, H263, MPEG 4 , MPEG 7.

I. INTRODUCTION

The evolution of internet technology paved way to the need for digital video storage and transmission. Video segments are part of all web pages. Video can be stored in DVD, Magnetic drives and other storage devices or it can be saves in server and transmitted to the viewer on demand. Video can also be real-time. Before moving to different video compression techniques, it is required to understand the structural characteristics of the digital video and their methods used in defining the format.

Video can be defined as a sequence of images which are displayed in order. We can call these images, a frame. Small changes in the frames such as slight difference of colour cannot be noticed to normal human eye, thus video compression standards ignore all these details while encoding the video. A little of ignorable details are lost during compression. This type of compression is termed as lossy compression. Very high compression ratios are obtainable with the use of lossy compression. Typical frame rate of video is 30 Frames Per Second. Consecutive frames repeat a lot of same information. If a building is displayed for 1 second, 30 images will contain that tree. This type of repetition can be avoided by compression and the related frames can be defined based on previous frames. The first frame contains original image. The succeeding frames are compressed using only the information in that frame also called intra-frame or using other frame's information also known as inter-frame. Intra-frame method of coding allows random access operations such as fast forward and makes fault tolerance possible. If a part of the next frame is damaged, the succeeding intra-frames can be displayed because it only depends on the intra-frames.

Colours can be represented as combination of green, red and blue. All Images can be represented using the colour space. However, RGB is not suitable for compression because it doesn't consider the human eye perception. A colour space of YUV where only Y contains the grayscale image information. Human eye sensitivity is more on change in Y and this technique is used in compression. YUV also uses the NTSC, PAL, SECAM composite colours of TV standards.

Compression ratio is ratio between sizes of original video to size of compressed video. Better compression ratios are obtained with prediction based on other pixels. In spatial prediction, pixels can be obtained the same image and in temporal prediction, prediction of pixel can be obtained from previously transmitted images. Hybrid coding is prediction in temporal dimension with a required decorrelation technique in spatial domain. Motion compensation establishes a correspondence between the elements of adjacent images in the sequence. The prime application of motion compensation is to provide an appropriate predicting method for a given image from an image of reference.

II. VARIOUS COMPRESSION TECHNIQUE

Some of the various video compression techniques and current trends in compression are explained below

1. H.261

H.261 or Clarke95 has been developed for transmitting ofvideo at a rate of 64Kbps and its multiples. Video calling and visual conferencing are some applications of H.261. The H.261 frames can be of two types

I-frames: They are coded without reference to any previous frames.

P-frames: They are coded using a previous frame as a reference for prediction.

H.261 standard is like JPEG still image compression standard. H.261 uses motion-compensated temporal prediction. H.261 coder has a layered structure with four layers. Four layers are named as picture, group of block or GOB, macro-block or MB and block layer. Each block is 8 x 8. The layers are multiplexed for series transmission. Each layer has header frames. H.216 frame format is known Common Intermediate Format or CIF.

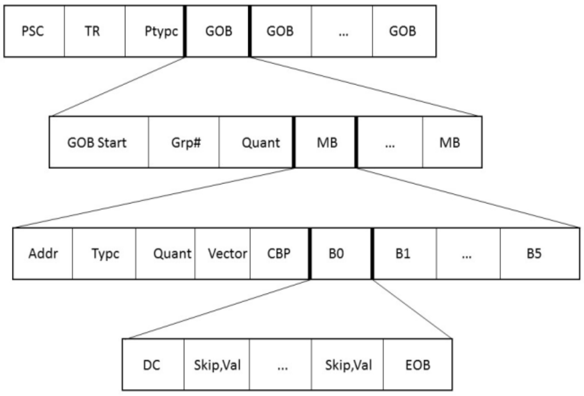

Fig.1: Bit stream structure of H.261

2. H.263

H.263 or Rao96 was designed for very low bit rate coding applications. Encoding of H.263 is done by using Block motion-compensated DCT structure. H.263 encoding has higher efficiency compared to H.261 encoding. Optimization in H.263 is done by using an encoding specification called test model or TMN. The basis of H.263 is on H.261 but its significantly optimized for low bitrates. Each picture is partitioned into macro-blocks. Macro block contains 16 x 16 luminance block and 8 x 8 chrominance blocks of Cb and Cr. All macro block can be coded as intra or inter blocks. Spatial redundancy is exploited by Discrete Cosine Transform coding and temporal redundancy by motion compensation. H.263 consists of motion compensation with half-pixel accuracy and bidirectional coded macro blocks. These features are not used in MPEG-1 and MPEG-2 because they are useful for low only bit rate applications.

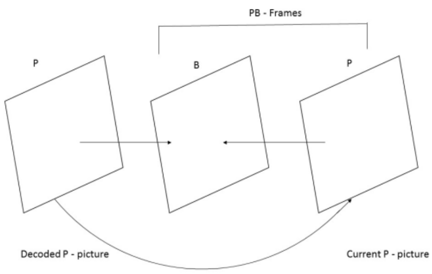

Fig.2: Prediction in PB frame

3. *MPEG*

MPEG or Moving Picture Experts Group is an ISO/IEC working group developing compression, decompression, and representation of moving pictures and audio in international standards.

3.1. MPEG-1

MPEG video compression standard is a layered, DCT-based video compression standards results in VHS quality compressed video stream which has a bit rate of nearly 1.5Mbps at a resolution of nearly 352 x 240. MPEG video sequences consist of different layers that provide the ability to randomly access a video sequence and also provide a barrier against corrupted information.

MPEG frames are encoded in any of three different ways: Intra-coded or I-frames, Predictive-coded or P-frames, or Bi-directionally-predictive-coded or B-frames. I-frames are encoded as discrete frames, without depending on the adjacent frames. Thus, within the video stream they providerandomly accessible points. Because of this, when compared to other frames I-frames have the worst compression ratio. With respect to a past I-frame or P-frame, P-frames are coded, which results in a smaller encoded frame size than the I-frames. The B-frames require a preceding as well as future frame, which may be either P-frames or I-frames, in order to be decoded.

3.2. MPEG-2

MPEG-2 is designed for diverse applications that require a bit rate of up to 100Mbps. Digital high-definition TV or HDTV, interactive storage media or ISM, cable TV or CATV is sample applications. For supporting these diverse applications multiple video formats can be used in MPEG-2 coding. MPEG-2 is possible to extract a lower bit stream to get lower resolution or frame rate. Decoding MPEG-2 is very expensive; bit stream scalability allows flexibility in the required processing power for decoding. MPEG-2 is forward, backward, upward, and downward compatible. The decoder can decode the pictures generated by a lower resolution is the encoder upward compatibility. A decoder can decode the pictures generated by a higher resolution encoder is called Downward compatibility. A new generation decoder can decode the pictures generated by an existing encoder is done in a forward compatible system, and existing decoders can decode the pictures generated by new encoders in a backward compatible system.

Since MPEG-2 is more oriented towards television applications Video sequence layers are similar to MPEG-1 the only improvements are frame/field motion compensation and DCT processing, scalability. 2 additional chrominance blocks are used in macro blocks in MPEG-2 when 4:2:2 input formats are used. 8 x 8 block size is retained in MPEG-2; in scaled format blocks, can be 1 x 1, 2 x 2, 4 x 4 for resolution enhancement. B and P frames have frame and field motion vectors.

In real time with current processors, MPEG-2 decoding cannot be done.

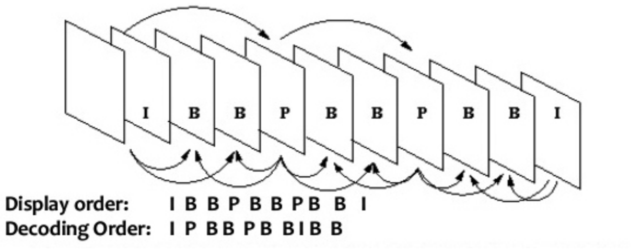

Fig. 3: MPEG Frame Order

3.3. MPEG-3

MPEG-3 was intended for HDTV and EDTV for higher bit rates and it later merged with MPEG-2.

3.4. MPEG-4

MPEG-4 was designed by MPEG group encouraged by the Success of digital television, interactive graphics applications and interactive multimedia. MPEG-4 allows the user to interact with the objects in the scene within the limits set by the author and brings multimedia to low bit rate networks.

MPEG-4 uses media objects to represent audio as well as visual content. Media objects are combined to form compound media objects. Before transmission, MPEG-4 synchronizes and multiplexes the media objects to provide QoS and at receiver's machine, it allows interaction with the constructed scene. MPEG-4 organizes the media objects in a hierarchical form, where lower level has primitive media objects which can be used to represent 2 or 3-dimensional media objects. It also defines a coded representation of objects for text, graphics, synthetic sound, talking synthetic heads.

MPEG-4 standard's visual part describes methods for compression of image, video, textures for texture mapping of 2-Dimesional and 3-Dimensional meshes, implicit 2-Dimensional meshes, time-varying geometry streams that animate meshes. MPEG-4 provide algorithms for random access to all types of visual objects, spatial, temporal and quality scalability, content based scalability of video, textures, and images. MPEG-4 has parametric descriptions for animation streams face and body, parametric descriptions of the human face and body for synthetic objects.

MPEG-4 supports coding of video objects with temporal and spatial scalability which allows decoding a part of a stream and construct images with reduced decoder complexity or quality, reduced temporal resolution, reduced spatial resolution., or with equal but reduced temporal and spatial resolution quality. Recovery tools are used to recover the lost data, after synchronization data. For concealing the lost data error concealment tools are used. Efficient resynchronization is one of the key to good data recovery and for the error concealment.

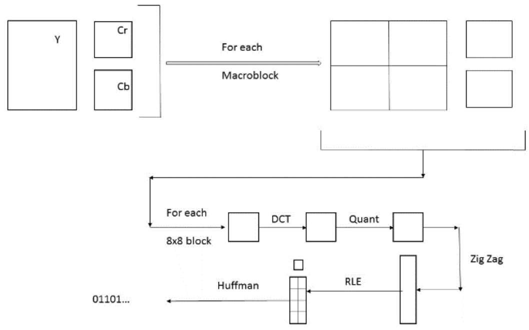

Fig .4: MPEG Coding Diagram

3.5. MPEG-7

MPEG-7 main purpose is to specify a set of descriptors to describe various forms of multimedia and also standardize ways to define other descriptors and structures for the descriptors and their relationship. This information will associate with the content to allow fast and efficient search. MPEG-7 standardizes a language to specify description schemes.

3.5.1 J.81

J.81 specifies the coding and transmission of digital TV signals at bit rates of 34-45Mbps in the format specified by recommendation ITU-R 601 and provides very high quality which is suitable for transparent compression necessary for contribution applications. Net video capacity ranges 26-31Mbps for Europe and depends on the number of optional channels used.

3.5.2 Fractal-Based coding

Fractal-Based coding is a newly introduced and promising technique. In an image, pixel's values that are close are correlated. Advantages of this observation was taken by Transform coding and Fractal compression. Advantage of these types of spatial structures cannot be taken by Transform coders by representing constant regions and straight edges efficiently using fractal coding and it tries to reconstruct the image.

3.5.3 Model-based Video Coding

Model based scheme defines 3-Dimensional space structural models of the scene where Coder and decoder use an object model. The coder and decoder can use same model to analyse and generate the image. Researches in model-based video coding or MBVC focuses on head tracking, local motion tracking, head modelling, and expression analysis, synthesis. Model-based video coding has been mainly used for video telephony and videoconferencing. MBVC is used in the applications such as virtual space teleconferences and speech driven image animation of talking heads.

3.5.4 Scalable video coding

Nodes are presented in Multimedia communication systems with limited computation power to be used for decoding and heterogeneous networks. Such cases, we need to decode at a variety of bit rates. Scalable coders have this property. To provide scalability in video communication systems layered multicast has been proposed.

A promising approach to provide scalable video coding is Spatio-temporal resolution pyramids. Open and closed loop pyramid coders provide inclusion of multiscale motion compensation and efficient video coding. Spatial doe sampling and interpolation operations, fast and efficient codec's can be implemented by using simple filters. Morphological filters are also used to improve image quality.

Multistage quantization scheme is used in Pyramid coders. Closed-loop pyramid coders are more suited for practical applications than open-loop pyramids coders since they are less sensitive to the sub optimal bit allocations and simple heuristics can be used.

To utilize multistage motion compensation several ways are used. One way is efficiently computing motion vectors and using hierarchical group estimation encode them. Scalability of a video transmitted over heterogeneous networks is utilized by reducing the bit rate of video data in case of congestion. Without knowing the content of the packet or informing the sender, the network layer can reduce bit rate by using priorities.

3.5.5 Wavelet-based Coding

For low bit-rate coding Wavelet transform techniques has been used. Wavelet-based coding performs more advantage than traditional DCT-based coding. Based on the application of these techniques much lower bit-rate and reasonable performance are reported to still images. Better performance can be assured by combining wavelet transforms and vector quantization. The image is decomposed into a Multi-frequency channel representation decomposes from Wavelet transform.

Additional video coding research applying the wavelet transform on a very low bit-rate communication channel is performed. By overlapped motion compensation the efficiency of motion compensated prediction can be improved. Multiple frequency bands generated from the wavelet transform, multi-frequency motion estimation is available for the transformed frame and also provides a representation of the global motion structure. The lower-frequency bands motion vectors are predicted with the more details of higher-frequency bands. With the segmentation technique that utilizes edge boundaries of the zero-crossing points in the wavelet transform domain this hierarchical motion estimation can also be implemented. It is possible to get better visual quality at rates as low as 16kbps as per the results.

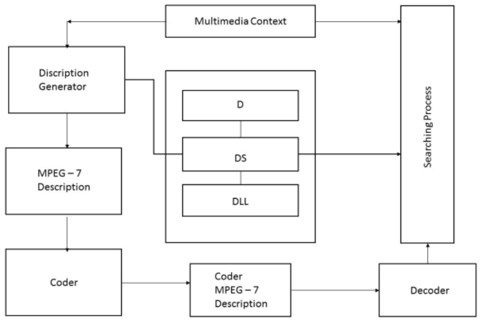

Fig.3.3: MPEG 7 Block diagram

III. ADVANTAGES AND DISADVANTAGES OF VARIOUS VIDEO COMPRESSION TECHNIQUES

Some of the advantages and disadvantages of the discussed techniques are given below

Table.4-1: Advantages and Disadvantages of Various Video Compression Techniques

Video coding standards	Advantage	Disadvantage
H.261	Very low bit rate. Can stream in any network	Quality is very low
Mpeg-1	Video does not consume much space	Resolution is not good enough for HQ videos
Mpeg-2	Video has more quality and can be used for digital TV	High definition not possible
H.263	Lower bit rate and higher quality.	Compression is not efficient
Mpeg-4	Can Stream HD, FHD UHD and 4K and small size facilitates internet streaming	Cannot be used for studio work due to high compression ratio
Mpeg-7	Metadata makes content search easier	Not widely used.
H.264/AVC	Small File size for longer recording time and faster transmission. Good for real-time surveillance application	Computationally expensive due to larger prediction units

Different techniques are used for different applications. These applications make use of the most required feature for its optimal performance.

IV. OVERVIEW OF THE FEATURES

An overview of features of various compression techniques are given below.

Table.4.2: Current and emerging video compression standards.

Video coding standards	Year developed	Primary intended application	Bit Rate
H.261	1990	Video telephony and teleconferencing over ISDN	p x 64 kb/s
Mpeg-1	1991	Video on digital storage media	1.5 Mb/s
Mpeg-2	1994	Digital television	2-100 Mb/s
H.263	1996	Video telephony over PSTN	33.6 kb/s and higher
Mpeg-4	1998	Object-based coding, synthetic content, interactivity, video streaming, HD, FHD, UHD,4K	64Kb/s-720000Kb/s
Mpeg-7	2001	Real-time and non-real-time applications, to tag the contents and events of applications vid eo streams for more intelligent processing in video management software or video analytic	64Kb/s-720000Kb/s
H.264/AVC	2003	Improved video compression.	10‟s to 100‟s of kb/s

V. CONCLUSIONS

Video compression has become a necessity since it reduces the need for large storage and high bandwidth. The standardization efforts for major compression algorithms including H.261, H,263, MPEG 1, MPEG 2, MPEG 4 and MPEG 7 were analyzed in this paper. Even though there exist many different criteria for video coding, scalable video coding is best for heterogeneous systems, wavelet coding is suitable for low bit-rate systems and there are some video coding algorithms yet to be researched further. Fractal-based, model-based and segmentation-based video coding are some of the unsearched techniques.

REFERENCES

[1] Chang Wen Chen, Ya-Qin Zhang, Visual Information Representation, communication, and image processing, 1999 Marcel Dekker, 551 pages.
[2] Roger J. Clarke, Digital Compression of Still image and video, 1995 Academic Press, 453 pages
[3] Joan Mitchell, William B. Penne baker, Chad, E. Fogg and Didier J. Legall, MPEG video compression standard, International Thompson publishing, 470 pages.
[4] K. R. Rao, J. J. Hwang, Techniques and standards for image, video and audio coding, 1996 Prentice Hall, 563 pages.
[5] Overview of MPEG-4 standard, 52pages http://drogo.cselt.stet.it/mpeg/#The_MPEG_standards.
[6] Eduardo Asbun, Paul Salama, Ke Shen, Edward J. Delp., Very low bit rate wavelet-based scalable video compression. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, No. 1, February 1999, pp. 09122.http://dynamo.ecn.purdue.edu/~ace/delppub.html.
[7] MPEG-7: Context and Objectives at http://drogo.cselt.stet.it/mpeg/standards/mpeg-7/mpeg-7.html.
[8] H. Kalva, A. Eleftheriadis, A. Puri, and R. Schmidt, Stored file formats for MPEG-4 Contribution ISO-IEC JTC1/SC29/WG11 MPEG97/2062, April 1997, Bristol, UK (39th MPEG meeting) http://www.ee.columbia.edu/~eleft/publications.html.
[9] ISO. "ISO/IEC 15938-1:2002 - Information technology -- Multimedia content description interface -- Part 1: Systems". Retrieved 2009-10-31.
[10] MPEG."About MPEG- Achievements". chiariglione.org. Archived from the original on July 8, 2008. Retrieved 2009-10-31
[11] MPEG. "Terms of Reference". chiariglione.org. Archived from the original on February 21, 2010. Retrieved 2009-10-31.

[12] MPEG. "MPEG standards - Full list of standards developed or under development". chiariglione.org. Archived from the original on April 20, 2010. Retrieved 2009-10-31.
[13] Network Dictionary. "Complete Protocol dictionary, glossary and reference - M". Retrieved 2009-12-26.
[14] ISO/IEC JTC 1/SC 29 (2009-10-30). "MPEG-7 (Multimedia content description interface)". Retrieved 2009-11-10.
[15] ISO/IEC JTC1/SC29/WG11 (October 2004). "MPEG-7 Overview (version 10)". Chiariglione.org. Retrieved 2009-11-01.
[16] Lux, Mathias. "Caliph & Emir: MPEG-7 photo annotation and retrieval." Proceedings of the 17th ACM international conference on Multimedia. ACM, 2009.
[17] WG11 (MPEG) (March 2002). "Overview of the MPEG-4 Standard". Retrieved 2010-04-11.
[18] WG11 (1997-11-21), Text for CD 14496-1 Systems (MS Word .doc), retrieved 2010-04-11
[19] "MPEG-4 Systems Elementary Stream Management (ESM)". July 2001. Retrieved 2010-04-11.
[20] "MPEG Systems (1-2-4-7) FAQ, Version 17.0". July 2001. Retrieved 2010-04-11.