



A Comparative Analysis of Video Compression Standard and Algorithms: State of the Art

M. O. Momoh¹, J. N. Ndunagu², V. K. Abanihi¹, S. M. Yusuf³, M. I. Abdullah³

¹Department of Electrical & Computer Engineering, Edo University Iyahmo, Nigeria

²Department of Computer Science, National Open University of Nigeria, Abuja, Nigeria

³Department of Computer Engineering, Ahmadu Bello University Zaria, Nigeria

ABSTRACT

The advancement in the internet technology fosters the application of video streaming which is gaining more awareness amidst the populace. The application of video streaming includes internet broadcasting, video conferencing and video lecture among others. In video transmission, either through wired or wireless, a lot of images are needed to create an illusion of continuity. Therefore, there is need for effective utilization of storage space and allocated bandwidth. In lieu of this, digital video standards and methods have been proffered by researchers and manufacturers in order to ensure compatibility among video codec from different manufacturers and to simplify the development of newer applications. This paper presents a comparative study of the emerging and current video compression standards with a demonstration of some of the compression algorithms that are being employed in digital video compression.

ARTICLE INFO

Article History

Received: November, 2019

Received in revised form: Dec., 2019

Accepted: February, 2020

Published online: April, 2020

KEYWORDS

Modeling, Simulation, Solar absorption, cooling system.

INTRODUCTION

Advancement in technology has led to the progress made in digital video communication. Nowadays, digital video is the main stream in several applications such as DVD, HDTV, video telephony and teleconferencing (Apostolopoulos *et al.*, 2002; Zou *et al.*, 2015; Mahini *et al.*, 2017). These applications of digital videos were realized due to the technological advancement in the field of computing, communication and video compression algorithms (Wang *et al.*, 2013; Kumar & Kumar, 2016). The adoption and deployment of these technologies were feasible due to the economies of scale brought about by competition and standardization (Ponlatha & Sabeenian, 2013; Adedokun *et al.*, 2019; Momoh *et al.*,

2019). Majority of the standards employed in video compression are based on the concept of redundancy reduction in digital video (Mittal & Vetter, 2016; Momoh *et al.*, 2019). Digital video consists of multiple images called frames displayed overtime. Frames in a digital video are 2D representation of the 3D real world. These frames are 2D array of pixels where each pixel value denotes the color and intensity of a specific spatial location at a specific time. The pixel values are representation of RGB component. The RGB color space is often used in capture and display digital video. A lot of images are required to create an illusion of continuity during the display of digital video through a medium (Liu *et al.*, 2008; Taksande *et al.*, 2015; Kumar & Kumar, 2016), for instance,



considering a video sequence at a standard definition television with a resolution of 720 X 480 at a frame rate of 27 FPS and for RGB which consist of 3 bytes per pixel. The memory required to store the video for a second will be $720 \times 480 \times 27 \times 3 = 28\text{MB}$ while it requires 112GB for one hour. In order to deliver this over a network, (wired or wireless) the bandwidth requirement is $28 \times 8 = 224 \text{ BMPS}$. Besides the large memory and bandwidth requirement, using uncompressed video will lead to an increase in the cost of hardware and systems that process digital video (Liu *et al.*, 2008; Zhang *et al.*, 2015). Compressed files are easier to transfer because there is a sizable amount of reduction in the size of data to be transferred. This results in reduction in the time and bandwidth utilization needed for the file transfer and thus providing good video quality even over a slow network.

VIDEO FRAMES

Video frames or frames are individual pictures in a sequence of images. These frames are what make up a video. A second video may contain about thirty frames. In video processing techniques, videos are being compressed using different algorithms. These algorithms are referred to as frame type or picture type. Each algorithm has its advantages and disadvantages. These algorithms are Intra-coded frame (I-frame), predicted frame (P-frame) and bi-directional predicted frame (B-frame) (Kumar & Kumar, 2016).

i. **I-frames:** These types of video frames are used as reference for other types of frames. They are complete picture (MPEG, PNG etc). The I-frames are known as the key frames. It does not require any other frame to decode; however, they are least compressible when compared to other types of frames (Kumar & Kumar, 2016; Zhang *et al.*, 2015).

ii. **P-frames:** These frames are predicted from an earlier frame (I-Frame or P- Frame). A video frame cannot begin with a P-frame as its first frame, that is, a P- frame needs a referencing frame for its reconstruction. The P- frames are also known as the predicted or delta frame. The P frame only holds the part that changed from the previous frame. For example, in two successive frames, only the part that changes in the first frame need to be encoded. It needs not to encode the unchanging part, thereby, it is more compressible than the I-frames as much storage space is being saved (Kumar & Kumar, 2016; Zhang *et al.*, 2015).

iii. **B-frames:** These frames are commonly known as the bi-directional predicted frames. It uses the previous and forward frame in order to decode. It is highly compressible when compared with I-frame and P-frame as it saves more space (Kumar & Kumar, 2016; Zhang *et al.*, 2015).

COMPRESSION

Compression is an important aspect of computer and communication which is employed in transferring data from the source to destinations. Compression in computer terms simply means reducing the physical size of data so that it occupies less storage space. In the internet, transferring of data is time dependent, so the use of compression is necessary in order to reduce the amount of data that needs to be transferred (Ponlatha & Sabeenian, 2013). Compressed files are therefore, easier to transfer due to significant reduction in the size of data to be transferred. It also leads to reduction in time needed and decrease in the bandwidth utilization in transferring data, thus providing good video quality even over a slow network (Adedokun *et al.*, 2019; Momoh *et al.*, 2019).



VIDEO COMPRESSION

Video compression is the process of reducing the frame size so as to save cost of transmission, increase transfer speed and data compatibility with the internet. Video compression can be achieved by exploiting the similarities in intra frame or inter frame (Gangurde & Nikam, 2017; Mittal & Vetter, 2016). In a single frame, there is spatial redundancy (Kumar & Kumar, 2016) which exists due to high correlation that within a frame as the amplitude of nearby pixels are often correlated. More so, successive frames exhibit temporal redundancy (Kumar & Kumar, 2016) since they typically contain the same object. The ultimate goal of video compression is to reduce redundancy.

VIDEO COMPRESSION STANDARD

Video compression standards provide several benefits. It ensures interoperability and communication between encoders and decoders from different manufacturers. For these, it reduces the risk for both manufacturers and consumers which leads to quicker acceptance and widespread use. In addition, these standards are designed for a large variety of application and the resulting economics of scale lead to reduced cost and further widespread use.

There are two families of video compression standards. They are, the International Telecommunication Union-Telecommunication (ITU-T) and the international organization for standardization (ISO) (Taksande *et al.*, 2015).

International Telecommunication Union-Telecommunication (ITU-T)

The first video compression standard to gain popularity under the ITU-T was the ITU-H.261. It was designed for video conferencing over the integrated

service network (Taksande *et al.*, 2015). In 1990, it was adopted as standard. It was designed to operate at $p=1,2,3,\dots,30$ multiples of the baseline ISDN data rate. The ITU-T initiated a standardization effort in 1993, with the goal of video telephony over the public switch telephone network (PSTN) where the overall data rate was about 33.6kb/s. The H.263 was the video compression standard portion of the standard and it was firstly adopted in 1996. An enhanced H.263 was developed in 1997, which is known as H.263 version 2. Now, a new algorithm, formerly referred to as H.26L is currently being finalized as H.264/AVC (Taksande *et al.*, 2015).

International Organization for Standardization (ISO)

The ISO established the Moving Picture Expert Group (MPEG) in 1998. The standard was developed with the aim of compressing moving pictures (video) and audio on digital storage media (CD-ROM). In 1991, they finalized the MPEG-1 and it achieve approximately VHS quality video and audio at about 1.5Mb/s (Taksande *et al.*, 2015). The second phase known as MPEG-2 which is an extension of MPEG-1 was developed for digital television and for higher bit rates (Apostolopoulos *et al.*, 2002; Gangurde & Nikam, 2017). MPEG-3 was envisioned for higher bit rate application such as HDTV but was wrapped into MPEG-2 because it can also recognize such application that is why MPEG-3 never gains popularity. In order to attain more functionality such as higher compression efficiency, error resilience features, natural and synthetic content, content based interactivity etc. The joint Video Team from both ITU and ISO MPEG finalized on H.26L standard. It was adopted by the two teams and they called it H.264 and MPEG-4 part 10, Advanced Video Coding (AVC) (Taksande *et al.*, 2015). The emerging and current

video compression standard is given in Table 1

Table 1: Emerging and current video compression standard (Taksande et al., 2015)

Video coding format	Primary intended application	Bit rate
MPEG-1	Storage of digital video (CD-ROM)	1.5Mb/s
MPEG-2	DTV	2-20Mb/s
MPEG-4	Video streaming, synthetic content, interactivity, object based coding	Variable
H.261	Teleconferencing over ISDN, Video over IP	64kb/s
H.263	Video telephony over PSTN	33.6kb/s and up
H.264/MPEG	Improved video compression	10's to 100's of kb/s

The emerging and current video compression standards are shown in Table 1. Currently, MPEG-4 and H.263 V2 are now used for video streaming and video communication and the emerging H.264/MPEG-4 part 10 AVC will probably

gain more popularity. There is also an incoming format known as H.265 which is still at an infant stage that aims to achieve high compression efficiency and lower the computational overhead in the process of video compression.

COMPRESSION ALGORITHM

The aim of compression algorithm is to reduce the number of parameters that is used to represent a signal, so as to lower the transmission cost, achieve a high compression efficiency, compatibility with the internet while delivering reasonable

quality picture to the user (Liu *et al.*, 2008; Kumar & Kumar, 2016). Some of the compression algorithms that have been employed towards reducing the redundancies in digital videos are as follows:



(a) Smooth image



(b) coarse image

Fig 1: Video frame of a smooth and coarse image



Table 1: Pixels value of the smooth image

246	246	246	245	240	243	241	243	248	247
246	246	245	199	200	199	196	247	247	247
242	242	245	202	204	203	203	247	247	245
241	241	241	186	181	192	206	241	241	245
242	242	243	167	187	174	206	242	242	243
239	244	244	202	198	204	206	205	238	238
239	239	239	201	200	204	206	204	238	238
240	240	237	203	204	204	204	238	238	238

Table 2: Pixels value of the coarse image

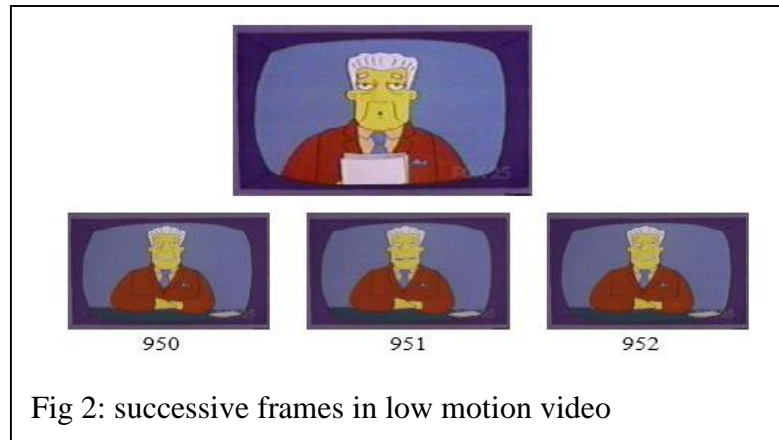
191	202	161	182	120	123	221	107	120	120
187	145	164	121	128	109	213	146	150	127
171	182	196	122	122	125	100	162	148	151
108	177	146	190	121	130	121	201	103	102
111	172	149	150	161	162	196	203	199	212
132	156	190	102	228	249	267	221	104	201
239	239	239	202	112	145	111	204	104	219
240	240	237	208	182	204	113	121	191	162

Spatial Redundancy: This is the type of redundancy that exists within a frame. Usually, in a frame, there exist similarities among the pixels. Exploiting the spatial redundancy will definitely lessen the amount of data that is to be encoded and transmitted as few bits can be used to represent the correlation during encoding. However, in a coarse image or busy scene

video where there are fewer or no correlation among adjacency pixels, spatial redundancy is ineffective (Ponlatha & Sabeenian, 2013; Kumar & Kumar, 2016). Fig 1 depicts a video frame of a smooth and a coarse image while its corresponding pixel value are given in Table 2 and 3 respectively. From Fig 1, it is observed that the smooth image has more correlation of pixel values than the coarse image as it is evident in

Table 1 and 2. The spatial redundancy techniques perform better when there are more correlations in adjacency pixels as

compared to when there are fewer or no correlations in adjacency pixels.



Temporal Redundancy: This type of redundancy exists in corresponding pixels within successive frames. Higher compression is being achieved by exploiting these similarities. However, in a high motion videos or many scene changes videos, temporal redundancy may be ineffective because of the lesser pixels correlation in the successive frames (Ponlatha & Sabeenian, 2013; Kumar & Kumar, 2016). Temporal redundancy algorithm works by setting the first frame as a reference for the next, thereby instead of transmitting the entire frame during streaming, only the pixel values that

changes with respect to the previous frame are sent along with its indexing information for the reconstruction of the next frame. In temporal redundancy techniques, the corresponding changes in pixel values are usually stored in a buffer (difference buffer) while the indexing information are stored in the indexing buffer. However, this algorithm (temporal redundancy) does not consider any further redundancy in the difference buffer, though the redundancy in difference buffer are often negligible; also, the algorithm is not effective in a high motion videos. Fig 2 depicts the successive frames in a low motion video.

Spatio-temporal Redundancy: Spatio-temporal is a hybrid of both the spatial and temporal redundancy techniques which was proposed by (Adedokun *et al.*, 2019; Momoh 2019). This algorithm exploit both the spatial and temporal redundancy that occur in videos (low-motion videos) accounting for optimal bandwidth utilization which in turns reduces the amount of data/information that is to be encoded and transmitted.

This algorithm ensures no redundancy even at the difference buffer. However, there is an increase in time to build up frames due to increase in the indexing information that is required for the successful reconstruction of the frames.

Perceptual Redundancy: There are details in the picture that the human eyes cannot perceive. Such details can be discarded without affecting the quality of



the picture. The spatial and temporal details in a video sequence that are perceived by human depend on the visual system (Zhang *et al.*, 2015; Momoh *et al.*, 2019). The human visual perception does not truly represent the RGB color space. Thus, the YCbCr representation allows exploiting the characteristics of the visual perception better (Lee & kalva 2011).

Statistical Redundancy: In video compression, the binary codes make uses of the motion vector, transform coefficient and other data in compression (Lee & kalva 2011; Parodkar & Bade, 2015; Momoh *et al.*, 2019). The fixed length code (e.g 16 bits words) is widely used in coding due to its simplicity (Lee & kalva). The drawback with the fixed length coding is that it is always wasteful because of lack of uniform distribution among the values. Average code length can be reduced by assigning shorter code words to values with higher probability. Variable length coding is used to exploit these statistical redundancies and increase further compression efficiency (Duanmu *et al.*, 2017).

CONCLUSION

Digital video is gaining more recognition in the society which is attributed to the adoption of modern compression mechanisms for reducing storage and network bandwidth requirements. Though, none of the compression mechanisms or algorithms was able to achieve absolute compression because of the tradeoff between efficiency and complexity of the algorithm. Thus, digital video standards and applications have been proffered by researchers and manufacturers in order to ensure compatibility among video codec from different manufacturers and to simplify the development of newer applications. This paper presents a comparative study of the

The YCbCr color space (YUV) matches the human visual perception where Y gives the average brightness of a picture and Cb and Cr give the chrominance components (blue and red chrominance).

emerging and current video compression standards and demonstrate some of the compression algorithms that were being employed in digital video compression

REFERENCES

- Adedokun, A. E., Abdulrazak, M. B., Omuya, M. M., Bello-Salau, H., & Sadiq, B. O. (2019). A Spatio-Temporal based Frame Indexing Algorithm for QoS Improvement in Live Low-Motion Video Streaming. *ATBU Journal of Science, Technology and Education*, 7(3), 305-315
- Apostolopoulos, J. G., Tan, W.T., & Susie, W. J. (2002). *Video Streaming: Concepts, Algorithms, and Systems*. Palo Alto, CA, USA: Hewlett-Packard Laboratories.
- Azhar, A. Z., Pramono, S., & Supriyanto, E. (2016). An Analysis of Quality of Service (QoS) In Live Video Streaming Using Evolved HSPA Network Media. *JAICT*, 1(1), 1-6.
- Duanmu, Z., Zeng, K., Ma, K., Rehman, A., & Wang, Z. (2017). A quality-of-experience index for streaming video. *IEEE Journal of Selected Topics in Signal Processing*, 11(1), 154-166. doi: 10.1109/JSTSP.2016.2608329
- Gangurde, T. P., & Nikam, S. (2017). A Survey on Live Video Streaming Over Peer to Peer Network. *International Journal of Computer Trends and Technology (IJCTT)*, 43(1), 5-7.
- Lee, J. B., & Kalva, H. (2011). *Video Coding Techniques and Standards-Introduction, The Need For*



- Compression, Perceptual Redundancies, The Human Visual System, Sensitivity To Temporal Frequencies. Retrieved, 6, 15.
- Liu, Y., Guo, Y., & Liang, C. (2008). A survey on peer-to-peer video streaming systems. *Peer-to-Peer Networking and Applications*, 1(1), 18-28. doi: 10.1007/s12083-007-0006-y
- Kumar, R. J. N., & Kumar, A. P. (2016). Improving QOS in Live Video Streaming for Low-motion Videos Through Frame Indexing Method. *International Journal of Communication Technology for Social Networking Services*, 4(1), 9-22. doi: <http://dx.doi.org/10.21742/ijctsns.2016.4.1.02>
- Mahini, H., Dehghan, M., Navidi, H., & Rahmani, A. M. (2017). Peer-assisted video streaming based on network coding and Beer-Quiche game. *AEU-International Journal of Electronics and Communications*, 73, 34-45.
- Mittal, S., & Vetter, J. S. (2016). A survey of architectural approaches for data compression in cache and main memory systems. *IEEE Transactions on Parallel and Distributed Systems*, 27(5), 1524-1536. doi: DOI 10.1109/TPDS.2015.2435788
- Momoh M. O., B. Yahaya, A. A. Abdulrazaq, H.M Katagum & U. Abubakar (2019) "Towards Achieving an Optimal Bandwidth Utilization in Live Video Streaming" Proceedings of the Academic Conference of Cambridge Research and Publications International on 21st Century Challenges and Opportunities: Multidisciplinary Approaches (Vol. 19 No.3) P 55-61, 19th – 20th December, 2019 University of Maiduguri, Maiduguri, Borno State, Nigeria
- Taksande, S., Joshi, K., Chikaraddi, V., & Raksha, S. (2015). Video Streaming Techniques and Issues. *International Journal of Advanced Research in Computer Science & Technology*, 3(1), 141-144.
- Parodkar, M. S., & Bade, D. (2015). Improve the performance of p2p live streaming. *International Journal of Computer Science and Mobile Computing*, 4(2), 28-32.
- Ponlatha, S., & Sabeenian, R. S. (2013). Comparison of video compression standards. *International Journal of Computer and Electrical Engineering*, 5(6), 549-554.
- Wang, H., Hu, R., Tu, W., & Zhang, C. (2013). The perceptual and statistics characteristic of spatial cues and its application. *International Journal of Computer Science Issues (IJCSI)*, 10(1), 621.
- Zhang, G., Hu, C., Xing, C., Wang, N., & Wei, X. (2015). A novel scheme for improving quality of service of live streaming. *Ninth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*. 39(4), 457-462. doi: 10.1109/3PGCIC.2014.93
- Zou, X. K., Erman, J., Gopalakrishnan, V., Halepovic, E., Jana, R., Jin, X., Sinha, R. K. (2015). *Can accurate predictions improve video streaming in cellular networks?* Paper presented at the Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications. 57-62.