INTRODUCTION

A video consists of multiple images called frames taken quickly over a period of time (Apostolopoulos et al., 2002; Mahini et al., 2017; Adedokun et al., 2019). In the past, video was captured and transmitted in analog form. Digitization of video became possible due to the advent of integrated circuit and as such, digital video enable a revolution in the compression and communication video (Apostolopoulos et al., 2002). In the early 2000s, the internet experience a booming escalation of network bandwidth in combination with an algorithm for compression and more powerful computer system, video streaming became possible (Zou et al., 2015). The commercial video streaming applications such as QuickTime, ActiveMovie, and RealPlayer became reality almost 35 years ago.

**Keywords:** Spatial redundancy, temporal redundancy, spatio-temporal redundancy, statistical redundancy, perceptual redundancy

---

TOWARDS ACHIEVING AN OPTIMAL BANDWIDTH UTILIZATION IN LIVE VIDEO STREAMING

M. O. MOMOH¹, B. YAHAYA¹, A. A. ABDULRAZAQ², H. M KATAGUM¹, & U. ABUBAKAR¹

Department of Computer Engineering, Ahmadu Bello University Zaria, Nigeria ¹Department of Computer Engineering, University of Maiduguri, Borno, Nigeria.

**Abstract**

Video streaming has received a lot of attention from industry and academia due to an explosive growth of the internet and increasing demand for multimedia information on the web. However, the current best effort service does not guarantee effective utilization of bandwidth. One of the challenges which still remains a bottleneck to researchers is redundancy. Reducing or eliminating...
Redundancies in live video streaming can significantly improve the Quality of Service (QoS). This paper presents the various types of redundancies that exist in videos and the ways in which they can be exploited so as to achieve optimal utilization of video streaming can be reduced to the barest minimum even at low bandwidth.

After the internet was born (Zou et al., 2015). Large computing and bandwidth requirements associated with video encoding and transmission are factors which hinder the development of video streaming (Wang et al., 2013; Kumar & Kumar, 2016). Streaming refers to transport of live or stored media such as audio, video or any associated data over a network. A lot of images are required to create an illusion of continuity when transferring videos over the internet. The bandwidth requirement was just too high for the early days of the internet even for low resolution videos (Taksande et al., 2015). In reducing the number of bytes required per frames, compression is being employed. However, the computational complexity of compression grows with number of bytes that can be saved. Thus, and therefore, powerful computers are needed.

Live video streaming refers to the real time transport of live content over the internet i.e. the media is captured, compressed and transmitted on the fly (Kumar & Kumar, 2016). A lot of content providers now make a live broadcast of their media on the internet. The applications of video streaming are numerous which include internet broadcasting, video conferencing, video lectures amongst others. Such applications provide end users convenient, abundant and interactive multimedia service (Zhang et al., 2015). In 2014, more than 64% of the internet traffic resulted from data streaming (Azhar et al., 2016). Time constraint and variable bit rate property make video streaming application difficult, this is because the buffering time of video streaming depends on the bandwidth offered by the network (Apostolopoulos et al., 2002).

Network congestion is the main cause of video degradation which occurs due to heavy traffic, low bandwidth and delay in delivery, which in turn leads to poor video quality (Kumar & Kumar, 2016). Over the years, researchers are concerned with the higher demands of streaming due to the increasing needs of videos on the fly. Several algorithms have been employed in the compression techniques so as to reduce the amount of data to be transmitted over the network which in turn improved the bandwidth utilization and also saved transmission cost (Azhar et al., 2016).
The indexing based streaming algorithm attempts to also improve the compression rate by sending a reference frame, the difference buffer and indexing information that is required for the successful reconstruction of a new frame (Kumar & Kumar, 2016). This in turn improved the bandwidth utilization and save transmission cost. However, despite the use of the index based streaming algorithm, the difference buffer is still observed to have redundancy of pixels (spatial redundancy) in the generated difference buffer. (Canel et al., 2018) proposed a deep neural network based algorithm for detecting interesting frames in video by identifying the relevant ones. The algorithm has the ability to accurately select the desired number of relevant frames, thus, creating a uniform output frame rate from a non-uniform streaming of interesting frames. This is in contrast to a more dynamic algorithm that adjust some selected frames based on the content of the video. Although, the algorithm succeeds in its requirement of achieving exactly the desired reduction factor. However, the characteristic of the result depends on the size of the buffer. Using a very large buffer results in delay and memory overhead due to the polynomial complexity of the longest path algorithm. (Adedokun et al., 2019) implemented a spatio-temporal frame indexing algorithm to address the limitation of (Kumar & Kumar, 2016) the research yielded a promising result when both are compared. However, other forms of redundancies were not considered. The remainder of the paper is organized as follows: Section II presents the significance of compression, the various forms of redundancies are presented in Section III while Section IV concludes the paper.

**Significance of video compression**

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 X 480 X 27 X 3= 28MB while it requires 112GB for one hour. In order to deliver this over a network, (wired or wireless) the bandwidth requirement is 28 X 8 =224 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. 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. One of the methods to achieve compression is by reducing or eliminating redundancies in video as it is evident from literatures that digital video has significant redundancies and reducing or totally eliminating this redundancies will results in compression.
Redundancies in video
Redundancies exist in different forms which are; 1) Spatial 2) Temporal 3) Spatio-temporal redundancies 4) Perceptual 5) Statistical

a) Spatial redundancy
Spatial redundancy (Intra frame redundancy): This is the type of redundancy that exists within a frame. Usually, in a frame, there exist similarities among the pixels especially in smooth images (Lee & Kalva 2011). The spatial redundancy simply takes advantages of similarities among neighboring pixels. Exploiting the spatial redundancy will definitely lessen the amount of data that is to be encoded and transmitted. However, in a coarse image, spatial redundancy is ineffective (Ponlatha & Sabeenian, 2013).

b) Temporal redundancy
Temporal redundancy (inter frame redundancy) refers to the correlation which exists between two successive frames (Kumar & Kumar, 2016; Adedokun et al., 2019). In a video, there exist a high correlation in successive frames especially in a low motion video. Exploiting the temporal redundancies will however result to the majority of the compression gains in a low motion video encoding (Lee & Kalva, 2011; Adedokun et al., 2019). Since two successive frames are similar, taking the difference between the two frames results in a smaller amount of data to be encoded. Predictive coding approach has been used in video coding techniques that make uses of the previous frame to predict the next frame in video (Kumar & Kumar, 2016; Canel et al., 2018). Efficient video compression is achieved by the correctness of the computation. Frame difference coding is the simplest form of predictive coding in which the previous frame is used as the prediction. The frame difference prediction begins to fail as the object motion in a video sequence increases resulting in a loss of correlation between collocated pixels in two successive frames (Lee & Kalva 2011). The index based algorithm in (Kumar & Kumar, 2016) was based on temporal exploitation, though due to the redundancy in the difference buffer optimal bandwidth utilization was never achieved. Temporal redundancy may also be ineffective in a high motion video or many scene changes video.

c) Spatio-temporal redundancy
Spatio-temporal redundancy was proposed to address the challenges of both spatial and temporal redundancies. The spatio-temporal frame indexing was proposed by (Adedokun et al., 2019) which exploit both the spatial and temporal redundancy that occur in video streaming (low-motion videos) accounting for optimal bandwidth utilization which in turn reduce the amount of data/information that is to be encoded and transmitted. However, the work only focuses on a low motion video, high
motion video was never considered. More so, there is an increase in time to build up frame due to the increase in the indexing information that is required for the successful reconstruction of the frames.

d) 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. The human visual perception does not truly represent the RGB color space. 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). Thus, the YCbCr representation allows exploiting the characteristics of the visual perception better (Lee & kalva 2011).

e) 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). 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
Real-time video streaming of events over a network continues to gain more popularity among the populace. One of the challenging issues which still remains a bottleneck is redundancy in videos. However, there is need to ensure the judicious utilization of allocated bandwidth without compromising the Quality of Service (QoS). This paper presents a study on various redundancies that exist in videos. Future work will exploits all these redundancies in videos in other to achieve high compression ratio. Thus, optimal utilization of bandwidth in video streaming can be achieved.

References


