IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 3 Issue 7, July 2016
ISSN (Online) 2348 – 7968 | Impact Factor (2015) - 4.332
www.ijiset.com

Video Streaming Based on Frame Skipping and Interpolation Techniques

Fadlallah Ali Fadlallah
Department of Computer Science
Sudan University of Science and Technology
Khartoum-SUDAN
fadali@sustech.edu

Othman O. Khalifa
Department of Electrical and Computer Engineering
International Islamic University-Malaysia
Kuala Lumpur, Selangor-Malaysia
khalifa@iium.edu.my

Aisha Hassan Abdalla
Department of Electrical and Computer Engineering
International Islamic University-Malaysia
Kuala Lumpur, Selangor-Malaysia
aisha@iium.edu.my

Abstract- in this paper we proposed a video frame skipping and interpolation (VFS-I) techniques to streaming video. This technique is divided into main section: transmitter and receiver. The transmitter is employed Frame Skipping Mechanism (FSM) to drop/skip similar frames from video streaming sequence to reduce the amount of frames that need to be transferred over the network. This method specified frames to-be-kipped according to motion estimation between every neighboring frame. After that, re-encode the remaining video frames, and sent it to the receiver. The receiver employed Reconstructing video-Frames Interpolation (RFI) to rebuild the frames that were dropped in the transmitter side. The motion compensated interpolation of the key-frame (remaining frame) is used to reconstruct video frames back. A unified design of frame-skipping and interpolation mechanism are made for incrementing and improving compression process efficiency that have not considered in the previous works. One of the main advantages of the VFS-I is characterized by flexibility and scalability of video streaming to the requirements of the network bandwidth and end user**.**

Index Terms – Video Transcoding, Video Streaming, Frame skipping, Interpolation frame, Decoder, Encoder.

1. Introduction

The basic difference that we are going to have between the still image coding and video coding, the still image does not have temporal information, it is got only spatial redundancy, in the case of video we having the temporal aspects of the signal, which is not only vary in space but also varying in time. The great deal of similarity of the signal, that varies in time. Because, the big done in that the video captured in particular frame rate such as: 30 frame/sec; the difference between this successive frames are highly limited. So, there are great deal of temporal redundancy that is presents an addition to the spatial redundancy[1]. The Macroblock-level temporal and spatial redundancy reduction, which are every often used in a video encoder and decoder compression. In addition, frame-level redundancy reduction has been exploited as an encoding tool for temporal redundancy reduction [2]. For frame-level redundancy reduction, some technologies such as global motion compensation (GMC), temporal scalable video coding (T-SVC) [3], frame interpolation and frame skipping have been proposed.

GMC effectively encodes a frame-level motion in a scene, which can explain in a good coding efficiency, but the production of computational complexity at the encoder side.

T-SVC is made to deal with this situation, where the network bandwidth may vary significantly over time by providing control over the frame level rate of the video encoding with multiple layers. Frame skipping is employed to skip some frames from video sequences for improving the encoder process. The research on frame skipping has been conducted mostly in the area of rate control and transcoding. To ensure quality of service in real-time video transfer and utilizes encoder to meet the bandwidth requirements of the network. Some researchers applied frame skipping as an encoding tool for temporal redundancy reduction in a rate control[4]. In such a case, frame skipping was not targeted to maintain visual quality. In transcoding, frame skipping is used for getting additional coding efficiency, which can lead image degradation caused by handling the previously encoded bitstreams, because the original video data does no longer exist.

The objective of frame interpolation is to reconstruct video data with the more frame rates and the better visual quality from the given video bitstream. Most of the previous research employed motion compensated interpolation (MCI) with refinement of motion vectors (MVs) for frame interpolation [4-5]. Improve of MV through MCI provides best visual quality but at the expense of increased computational complexity at the decoder.

The increment of the complexity may be critical in a real-time environment. A disadvantage of frame skipping method is that it may expose image quality deterioration more severely than coding all the to-be-skipped frames with poor quality. Hence, the objective of the skipping frame is to determine skipping frames that will minimize the deterioration effect through frame interpolation. In other words, the advantage of frame skipping method is highly dependent upon that of frame interpolation (reconstruction). Therefore, the integration of skipping frame method and interpolation frame together, it is necessary to support not only a good coding efficiency, but to add visual quality in the frame level [1-5].

This paper proposes a new method of video frame skipping and interpolation (VFS-I). Video frame skipping is one of the effective methodologies to decrease the numbers of video frames. Frame skipping method skips only frames that have a relationship with adjacent frames to determine the key frames (remaining frames) [7].

2. VIDEO FRAMES SKIPPING:

The proposed of streaming video VFS-I technique is divided into two main sections: the transmitter and receiver, as illustrated in Figure 1. In a transmitter part, the frame skipping module specifies frames to-be-skipped through the similarity analysis between the adjacent frames of the video sequences. After filtered by frame skipping method, then, the remaining frames (or key-frames) and the corresponding timestamp are input to the video encoder as shown in figure 1. Frame skipping method can be defined as follows:

$$VFN(vs) \geq VFN(SFSM(vs)) \qquad (1)$$

Where VFN is video frame numbers, FSM is Frame Skipping Mechanism, and vs is the original video sequence, VFN($SFSM(vs)$) is the remaining video sequence (only the key frames), in the other words, the adaptive video in vs .

From figure 1, is shown that, the frame skipping by FSM can present a half compression process (encoder) depending on the number of frames skipped.

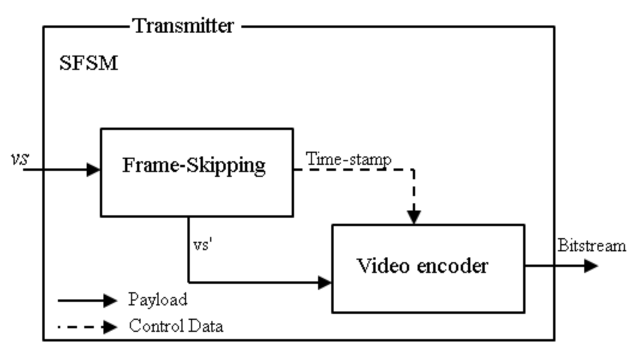

Figure 1: the video frame skipping model.

3. Frame Interpolation:

Figure 2 shows the receiver, which consists of a video decoder and interpolation frame. The video decoder decodes the input bitstream to rebuild key frames and produces some data controls such as motion, vectors (MV), frame rate, and information timestamp. Frame interpolation method, then, uses the output of the video decoder to reconstruct the interpolated frame. The resulting video sequence can be represented as follows:

$$VFN(vs'') = VFN(RFM(vs')) \qquad (2)$$

Where *RFM* is Reconstructed Frame Mechanism, vs' the adaptive video sequence with only the key frames, vs'' the final video sequence including the interpolated frames; generally, $VFN(vs'')$ should be equal to the $VFN(vs')$.

Also, figure 2 presents the frame interpolation module in VFS-I technique, which is based on a macroblock interpolation method. The reconstructed frame is as a result of macroblock-based interpolation in between two consecutive key frames. The VFS-I uses motion activity between two key frames is considered in the proposed interpolation based on MV values only from the second key frame, which is to capture the dominant motion of the second key frame for reconstructing interpolated frames.

An important feature of the VFS-I method is that we applied macroblock-based interpolation from the least active (in motion) macroblocks first. It is like a painting algorithm that fills the background first, foreground next. Also, we employed pixel-based padding using adjacent pixels after the macroblock-based interpolation since there may be cracks (i.e., some unfilled pixels) that are resulted by macroblock-based interpolation based on the motion activity of the second key frame.

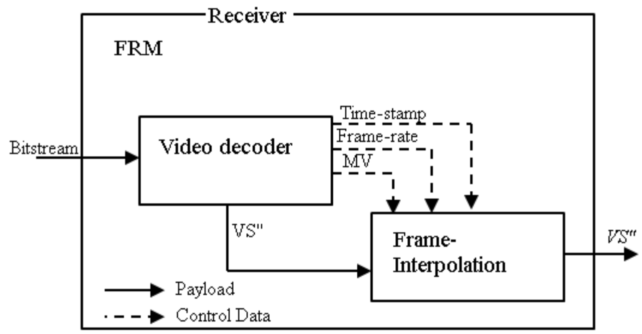

Figure 2: Reconstructed Frame interpolation

The timestamps of two adjacent key frames at ($t1$, $t2$) that could be employed to specify how many frames should be interpolated. The value of the timestamp is, for example, between 0 and 29 when the frame rate is 30. In frame rate decision, the total number of frames to be interpolated (nF) can be calculated as follows:

$$nF = \begin{cases} (t_2 - t_1 - 1) & when\ (t_2 - t_{12}) > 0 \\ (t_2 - t_1 - 1 + fr) & others \end{cases} \qquad (3)$$

Where fr is the frame rate from encoding condition; for example, when t_1 and t_2 are 3 and 8 respectively, there are three frames to be interpolated (i.e., nF is set to 4), whose timestamps are 4, 5, 6 and 7. To interpolate nF frames, actual reconstructing frame interpolation modules (RFI-M) including Block-based frame interpolation and Pixel-based frame padding are performed nF times.

The classification of macroblocks is performed in MV classification module, according to the corresponding MV values of the second key frame. In the case of intra and skip macroblocks, a zero MV can be assumed. A set of MVs in a frame can be represented in a point scan order as follows:

$$MV = \{(x1, y1)1^k, (x2, y2)2^k, \dots, (xi, yi)m^k \qquad (4)$$

$$k = 0,1,2, \dots$$

Where $(xi, yi)m^k$ is a MV of the *i-th* macroblock, *m* the total number of the

ISSN (Online) 2348 – 7968 | Impact Factor (2015) - 4.332
www.ijiset.com

macroblocks; when *k* is zero, it indicates a *MV* of 8x8; otherwise, *k* can be set to 1, 2, 3, 4 for 'k' MV.

4. VIDEO QUALITY EVALUATION

Currently, multimedia data are everywhere; and multimedia evaluation plays an essential role in the development and validity of image and video applications, like video compression and enhancing. There are two main types of quality assessments, subjective and objective quality assessment [8]. The subjective quality assessment is evaluated subjectively by a human viewer. However, it's time-consuming and expensive, and cannot be employed in real-time applications. To account for these drawbacks, a number of objective test methods have been developed, in which the quality rating is automatically calculated. Therefore, the expert must choose the objective metric very carefully in order to get appropriate results. On the other hand, the objective one can measure between un-distortion differences and distorted image quality signal. There are several aspects which are effect on multimedia quality such as blur/sharpness, noise, color fidelity, resolution, geometric distortions, frame dropping and freezing, and so on. There are three categories of Objective quality metrics according to their input data. The first representative of objective metrics is mostly used (PSNR) Peak Signal to Noise Ratio. This metric requires both original and distorted image in order to be computed [9]. Therefore, it is a full-reference method (FR). Other FR metrics are based on Structural Similarity or Image Evaluation based on Segmentation. Then the second category is reduced-reference (RR) metrics, it used when we have just little information about the original image. There are many examples of reduced reference metrics can be found e.g. in [4] or [5]. The final one is no-reference metrics. This it used when we do not need any information about the original image or video. Thus, they can be used in cases, where the FR and RR metrics cannot be. In this research, we were relying on the PSNR as means of video quality assessment tool.

5. RESULT ANALYSIS

In this research we used four video samples as subjects of this study are: Akiyo, News, Foreman, and Tennis videos. All four video samples are (176x144) QCIF frame-resolution and all of 300 frames; each video file sample has a different characteristic, such as slow motion, moderate motion, and fast motion. Such as, Akiyo-video mostly shows the face movement (motion); Foreman-video shows face movement and background movement change. Tennis-video shows the body movement and background movement. Moreover, the body movement in Tennis and in the last part of Foreman-video is a huge movement.

The testing was conducted on a laptop PC with Intel core™ i3 2.4 GHz CPU and 4.00 GB RAM running Microsoft Windows 7 home edition.
To calculate the results of this study accurately, we assume that the network works efficiently and the video stream will reach to the client without any frame loss.

Table.1 presents the operational parameters that are used to influence in a number of frames and a quality of video samples which required to be sent across a network.

Table 1: the operational Parameters

Parameter	Definition
B-size	The size pixel of block, the parameter used to determine the minimum value of the resolution of the original source of video.
R-Frame	The reference frame must be adding every-*n* frame. The parameter specify what is a minimum number of a frame must be kept in the original video.

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 3 Issue 7, July 2016
ISSN (Online) 2348 – 7968 | Impact Factor (2015) - 4.332
www.ijiset.com

M-Change	Specify the maximum number of changed motion vectors that allowed. This parameter used by FSM to decide what frame should skip from a video sequence, according to the motion vectors between two frames.

The relationship between operational parameter values and video frame quality is configured as: the B-Size value is 8, R-Frame value is 5; and the M-change value is 40. Figure (3, 4 and 5), show the PSNR values for all the video frames in all three test videos (Akiyo, Foreman, and Tennis) with given parameter values. It is noticed that the PSNR number never drops below 28.5.

Besides, most of the frames score above 30 and only the Tennis video has PSNR scores in the interval [28, 32] rather frequently.

This happens because Tennis is more motion intensive than other videos. In this example the frames of video Akiyo score. Figure 3(a) presents that all frames of video Akiyo score above 32 all the time. Also, Figure 3(b) shows Frame 158 with the lowest PSNR value and figure 3(c) shows Frame 98 has highest PSNR value.

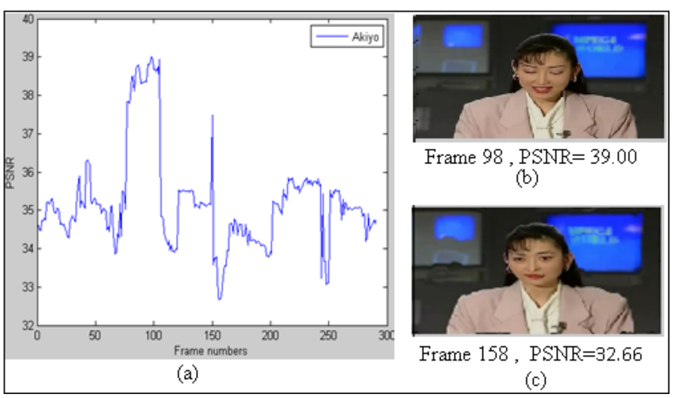

Figure.3: (a) PSNR vs. Frame numbers; (b) and (c) show frames with different quality values of Akiyo video clip.

It was shown that the parameter R-Frame has the largest effect on the quality of the reconstructed videos. Figure 4 (chart) shows how the average quality drops when R-Frame value is 5. It is noted that for this set of samples a block of 8 pixels was used since in this case it shows slightly better results. In this example it is noticeable that the average quality of all frames above 35 all the time (see figure 4). Also, Figure 4(b) shows Frame 221 with the lowest PSNR value 35.74; and figure 4(c) shows Frame 291 has highest PSNR value 42.34.

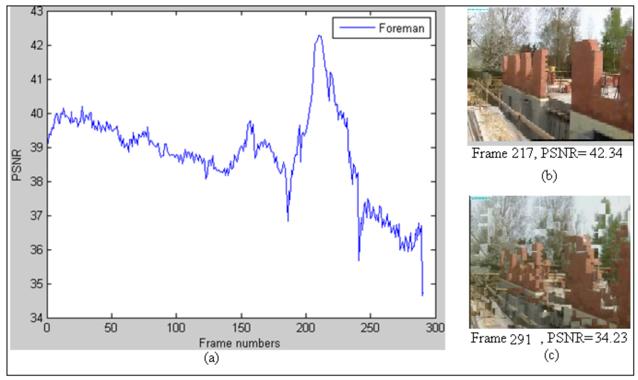

Figure.4: (a) PSNR vs. Frame numbers; (b) and (c) show frames with different quality values of Foreman video clip.

Figure 5(a) shows that all frames of video Tennis score above 28 all the time. Also, Figure 5(b) shows Frame number 117 with the lowest PSNR value 28.89; and figure 5(c) shows Frame number 276 has highest PSNR value 35.34.

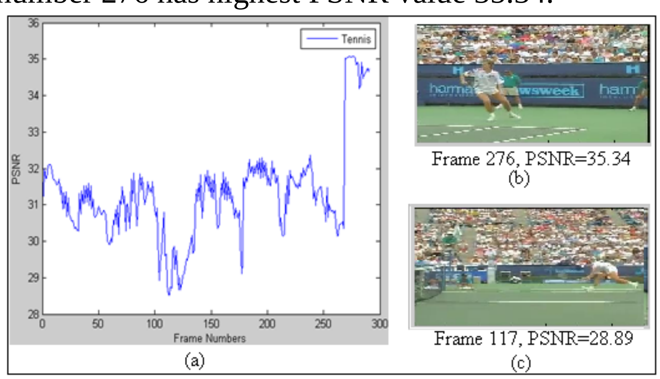

Figure 5: (a) PSNR vs. Frame numbers; (b) and (c) show frames with different quality values of Tennis video clip.

6. Conclusion

The video streaming is one of the most important contemporary issues, especially on the subject of the global communications system and data processing. Our work divides a video stream into two parts, transmitter and receiver side interconnected via a network. The study focused on the provision of video data from the transmitter to the receiver across the network as quickly with the least amount of possible loss of frames. The main objective of this research is to reduce the number of video frames to a lesser extent possible in the source side using skipping frames techniques while receiving video the destination will rebuild the frame(s) which are dropped in the server side by using interpolation techniques.

Operational parameters used to test and evaluate the quality of the outgoing and receiving the video stream through identifying a set of operational parameter values and then applied to the video samples. Before the video sent, the transmitter-side in our proposal reducing the number of input video frame, depending on the values of the operational parameters with the receiver-end will re-reconstruct the frames were skipped in transmitter-side.

7. References

[1] A. . Fallis, "Digital voice and picture communication," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013.

[2] S. Lee, M. Lee, K. Choi, and E. S. Jang, "Unified framework of frame skipping and interpolation for efficient video compression," *Proc. 2009 IEEE Int. Conf. Netw. Infrastruct. Digit. Content, IEEE IC-NIDC2009*, pp. 670–674, 2009.

[3] A. Puri, L. Yan, B. G. Haskell, and T. B. Labs, "TEMPORAL RESOLUTION SCALABLE VIDEO CODING Crawfords Comer Road," *Structure*, pp. 947–951, 1994.

[4] X. Wang, Y. Zhang, H. Li, and W. Zhu, "Adaptive rate control for dynamic bandwidth in video transcoding," *2008 11th IEEE Singapore Int. Conf. Commun. Syst.*, no. 1, pp. 773 – 7, 2008.

[5] J. Banelis and A. Proscevicius, "Streaming Video Based on an Intelligent Frame Skipping Technique," no. June, 2011.

[6] K.-T. Fung, Y.-L. Cham, and W.-C. Siu, "New Architecture for Dynamic Frame-Skipping Transcoder," *IEEE Trans. Image Process.*, vol. 11, no. 8, pp. 886–900, 2002.

[7] K. Frame and W. Video, "Adaptive Key Frame Selection Wyner-Ziv Video Coding," 2005.

[8] B. Ben Youssef and J. Bizzocchi, "Enhanced Pixel-Based Video Frame Interpolation Algorithms," *IEEE*, pp. 23–28, 2007.

[9] T. Na and M. Kim, "A Novel No-Reference PSNR Estimation Method With Regard to Deblocking Filtering Effect in H.264/AVC Bitstreams," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 2, pp. 320–330, 2014.