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

# Streaming Video Based on Temporal Frame Transcoding.

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

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

***Abstract* - In a previous couple of years different types of interactive and multimedia services are appeared such as distance learning, video conferencing, video telephony and others, which are used as one of the best means of communication in everyday life, for work or practical purpose. However, the most challenging task is to deliver these services in a different quality for users who have different hardware capabilities (processing, display, decoder, memory, etc.) interconnected via a heterogeneous network. One of the most problematic network services is a video streaming. The video streaming consumes a lot of network bandwidth and any loss or delay of the data affecting the video quality. Hence, transcoding technology is required to adjust and reduce the constraints of the network and user terminal. This paper is proposed a new video streaming technique to reduce the number of video frames that you need to send across networks. By employing Smart Frame Skipping Module (SFSM) to skips some frames of video in the server and sent the adapted video stream while receiving video the client will re-construct the frames which skipped of the video in the server. Also, to test and evaluate the proposal flexibility and adaptability the operational parameters were used with different values applied to the samples videos.**

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

## I. INTRODUCTION

Todays, advanced types of interactive and multimedia services are proposed, which are used as one of the best means of telecommunication in everyday life, for work or practical purpose. The most challenge is to deliver these services to multiple users, over different networks. For that, multimedia content needs to be adapted dynamically according to the user environment. Hence, the *video transcoding technology* is needed to meet these tasks.

Video transcoding is a process that converts a compressed video stream to another with different features, without performing the full decoding and re-encoding process[1][2][3]. To allow different devices with various capabilities (processing, memories, decoding, etc.) interconnect with each other over heterogeneous networks with different characteristics hence, different kinds of transcoding are needed.

Video transcoding as illustrated in Fig.2 can play several roles and functions, like format conversion, bitrate conversion (quality transcoding), resolution scaling (spatial transcoding), bitrate conversion (quality transcoding), and frame rate conversion (temporal transcoding)[4][5].

1) Format conversion: can change the syntax of video coding standard from one to another.
2) Spatial transcoding: decreases the spatial resolution of the compressed video, to faces the problem of limited size in many access terminals.
3) Quality transcoding: works on the bit allocation for each frame, by enhancing the quantization parameters, according to the target bitrate.
4) Temporal transcoding: is a process, which skips/drops some frames in order to adjustment the frame rate of the video sequence, without reducing the video quality of not skipped frames.

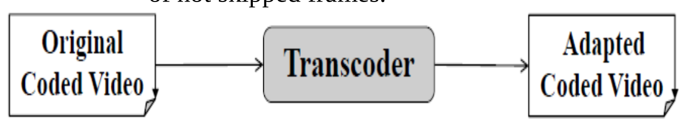


Fig.1: Video Transcoding scheme.

Generally, transcoding is divided into two main sections and additional functions as depicted in fig.2: homogeneous transcoding and heterogeneous transcoding. Homogeneous transcoding converts bitstreams into the same standard with frame-rate adjustments, an adaptive quantization, and a resolution conversion. A heterogeneous transcoding performs conversion between different video coding standards[4] [6].

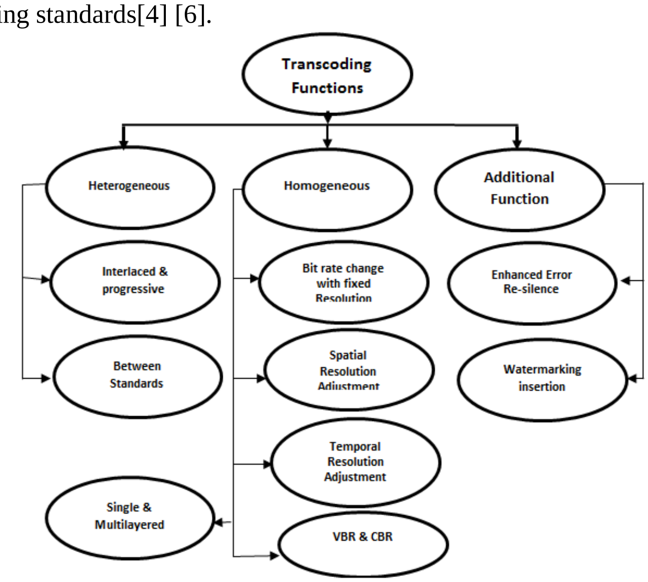


Fig.2: Classification of Transcoding functions.

## II. THE FRAME SKIPPING TECHNIQUES

To adjust video streaming over the network, we need to decrease the number of frames that can be sent over the networks. For this emerged the need to skips some frames in order to adjustment the video frame rate.

The choice of frames to be skipped is an important issue because it greatly influences the quality of video sequences. Depending on the research topic, in order to adjust the operations on video frame rate, the study relied on video transcoding. Video transcoding performing various encoder parameters to achieved different aims such as improving quality of videos or decreasing bandwidth consumption [7]. Temporal transcoding is a process employed to skips/drops some frames in order to adjustment frame rate of a video sequence, without reducing the video quality of not skipped/dropped frames.

In addition, there are another study works on the topic of video-frame dropping which are not classified under video transcoding. In the next paragraph we will display a number of research related to the subject of our research area.

Most frame skipping techniques are based on motion information to skip not necessary frames such as in figure.3. In [8] proposed a new variable skipping-frame approach. Their work is based on motion detection to perform skipping-frame on the live-encoding. The frames select to be encoded depend on the correlations of the successive frames in an adaptive length sliding window. The length of the sliding window depends on the target bitrate in order to decide the number of skipped frames. So, to estimate the amount of motion, temporal differences between corresponding components of two consecutive video frames are used for motion detection. From the results, this algorithm provides a very low computational complexity, and is appropriate for real-time services.

There are several formulas for measuring the motion information; specially, how much motion is in a frame. This motion information measure is also called *motion activity*. Is defined in [5] as follow:

$$MA_n = \sum_{i=1}^{N} |(u_n)|_i + |(v_n)|_i \qquad (1)$$

In (1) where $N$ is the macroblock numbers in frame $n$, $|(u_n)|_i$ and $|(v_n)|_i$ are the motion vector components of macroblock $i$ in frame $n$.

The motion activity of the previous frames, and the number of transcoded frames, are computed from the comparing between the motion activity with a dynamic threshold value. If the motion information of a frame is greater than the threshold, this frame cannot be skipped. According to the motion activity, a frame rate control scheme dynamically adjusts the number of skipped frames.

After frame skipped, re-encoding error is possible in motion compensated macroblocks, even if applied error compensation technique. The main goal is to reduce the impact of re-encoding error and also preserving the motion smoothness. The scheme used to preserve frame re-encoding error is called Cumulative Frame Skipping (FSC), is defined in [9][10][11]as follow:

$$FSC_n(MA_n, RE_{n-1}) = \frac{\sum_{i=1}^{N}(MA_n)_i}{\sum_{i=1}^{N}(RE_n)_i} \qquad (2)$$

In (2) where $N$ is the macroblock number in the current frame $n$, $(MA_n)_i$ is the motion activity of the frame $n$, and $RE_{n-1}$ is the accumulated re-encoding errors due to transcoding, for the motion compensated macroblock of the current frame. This metric is compared with a dynamically tuned threshold according to target and outgoing frame rate. A large value of the accumulated re-encoding errors reduces the value of $FSC_n(MA_n, RE_{n-1})$, causing the skipping of the frame.

The motion change happened at the current frame as a consequence to the change between the motion of transcoded frame and in the original video sequence[5][11].

As show in Fig .3 (a), in the original video sequence the motion change that happened at the frame(r + 1) is defined as:

$$mv_{org} = \frac{mv_2 - mv_1}{t} \qquad (3)$$

Where $t$ is an interval time, if the frame (r + 1) is transcoded, the motion change that happened in this frame is the same in the original sequence and in the transcoded sequence. So, no jerky effect is introduced by transcoding as shown in fig 3 (a).

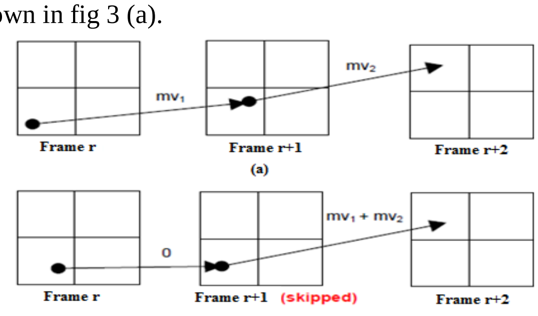


Fig 3: Motion-Based Skipping Frame.

However, if the frame(r + 1) is skipped, a direct copy of the previous coded frame r is used for displaying. In this case, the motion change happened at the frame(r + 1) as illustrated in Fig.3 (b) can be expressed as:

$$mv_{trs} = \frac{mv_1 + mv_2 - 0}{t} \qquad (4)$$

if the frame(r +1) is skipped, $mv_{org}$ and $mv_{trs}$ are different, and this difference causes an undesired jerky effect. In order to compute this difference, the Square Difference of Motion Change (SDMC) is defined as:

$$SDMC = (mv_{trs} - mv_{org})^2 \qquad (5)$$

The aim of (5) is to reduce the difference of motion change between the transcoded and the original video sequence, decreasing the jerky effect caused by frame skipping. If *SDMC metric* is higher than a threshold, it means that a large motion jerky will happen by skipping this frame, so the frame is transcoded. If it is smaller than the threshold, the frame is skipped, since skipping this frame will not introduce visual quality degradation.

- If SDMC metric > threshold → this frame doesn't skip.
- If SDMC metric < threshold → skip this frame.

Authors of [12] presented a new method to select video frames to-be-skipped. The method used both the perceived audio and video play out quality in consideration of the selection process. This algorithm is used three types of polices to perform dropping frames, RND – used to dropping frames randomly, D (2), and D (5) - both do not drop a consecutive frame. Furthermore, from their results, it presented that the method provided a better performance of perceiving audio-video play out quality is very similar if audio information is used in selecting video frames to be drop.

Ref [13]- [14]- [15] proposed a method for selective frame dropping for adaptive network bandwidth. This method based on a (GoP) group of pictures coding structure. This method offers a dynamic frame dropping creation of each GoP relying on the previous GoP results. From experimental results it displayed that the method provides a great real-time performance and lower computational difficulty. However, the experimental also show that the method is concentrated on decreasing the network bandwidth whereas the quality of video is of lesser significance. Finally, the frame reconstruction is not considered as a way to enhance the video quality transmitted.

The research on the subject of the video transcoding and dropping video frames, it is not the only one in this topic. There is the other research on the subject of dropping video frame (s) such as [5] - [10]. But these studies do not contribute to achieving the objective of this research, and for that excluded. However, one of the most the importance weakness point of the previous research is that they are deals with the video stream on structure of mathematical modeling not on the content of the video. the proposal objectives are to structure video stream into different parts with based on the residual frames movement, while dropping frame with less important according to the similarity between frames. This will decrease the amount of frames which we require to send over the network while after receiving the video the client-side will rebuilt the dropped frames in the server-side and return to the number of the original video frame.

## III. Results Analysis

Table.1 shows the four video clips that are selected as samples for this research are: Akiyo; News, Foreman, and Tennis videos. All four video samples are (176x144) QCIF frame-resolution and all of 300 frames, as depicted in the following.

Table.1: The video clips types

| Video | Format | Size | Frames # |
|---|---|---|---|
| Akiyo.avi | QCIF | 176x144 | 300 |
| News.avi | QCIF | 176x144 | 300 |
| Foreman.avi | QCIF | 176x144 | 300 |
| Tennis.avi | QCIF | 176x144 | 300 |

Table.2 shows that each video file 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 a last part of Foreman-video is a huge, we can see later.

Table.2: The Characteristics of Video files

| Test sequence | Characteristics |
|---|---|
| Akiyo.avi | Slow |
| News.avi | Slow |
| Foreman.avi | Moderate |
| Tennis.avi | Fast |

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.3 shows the proposal operational parameters that are used to deal with the video clip samples.

## VI. The Operational Parameters

Table.3 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 network are: B-size, R-Frame, and M-change.

**B-size**- a 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.

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

Table.3: Research Parameters.

| Parameters | Range | Impact on Quality |
|---|---|---|
| B-size | 8, 64 | High |
| R-Frame | 5,10 | High |
| M-change | 30-60 | High |

As depicted in Fig.4 explains that how our proposal deals with video file with all parameter combinations. Also, presents when the M-change reaches to 40, therefore the maximum number of dropping frames hits or near to the maximum value. As depicted in follows:

In case of R-Frame = 5 Approximately 145; (maximum skipping).

In case of R-Frame = 10 Approximately 190; (maximum skipping).

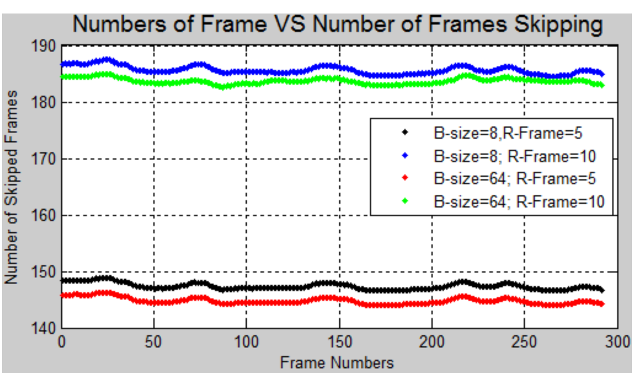


Fig.4: Snapshot of the Operational parameters and frames skipped of Akiyo-video clip.

As shown in fig.5, reach to the permitted number of dropping frames is quite easy because, the video focuses mainly on the motion of the face and can also easily reconstruction (compared to other two videos).

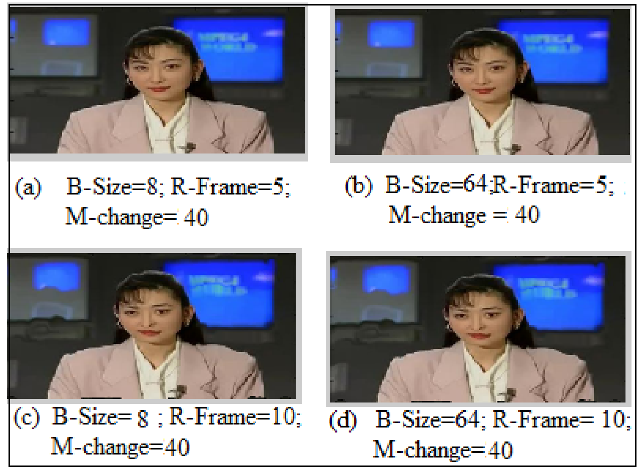


Fig.5 snapshot of frame 2 from Akiyo-video after reconstructed by using different parameters.

As shown in Fig.5 presents frame-2 of Akiyo-video that when the value of the R-Frame parameter increased also increased the number of skipped frames. This leads to increasingly worse image quality. In fact, the R-Frame parameter has a direct effect on frame quality rebuilt. The figure also showed that the value of R-Fame 5 has less influence on the quality of the video than the R-Frame of the value of 10 distorts the frame, but it can be understood. The B-size parameter did not have a significant impact in these samples because it is more important in dealing with the video that contains more movement.

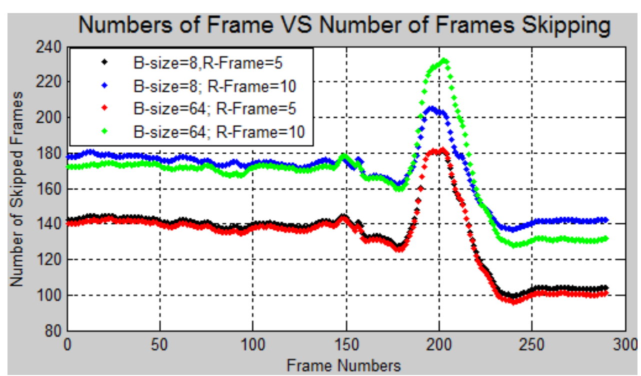


Fig.6: Snapshot of the Operational parameters and frames skipping of Foremen-video clip.

As illustrated in Fig.7 that the operational parameters of our proposal are applied on Foreman video clip It

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

illustrated that with Akiyo-video , reach to the maximum number of frames that can be skipped with all parameter combinations it easy this sample. But Foreman-video sample has a many of small movements and with the small size of the pixel block the proposal has a smallest of error when choosing frames to skip. And also, with the largest size of pixels block, small movements do not considerable changes to keep the frame. With Foreman video needs a larger value of M-change. Also this video is moderated (mixed) because it combined between simple movements and complex movements, as shown in the chart.

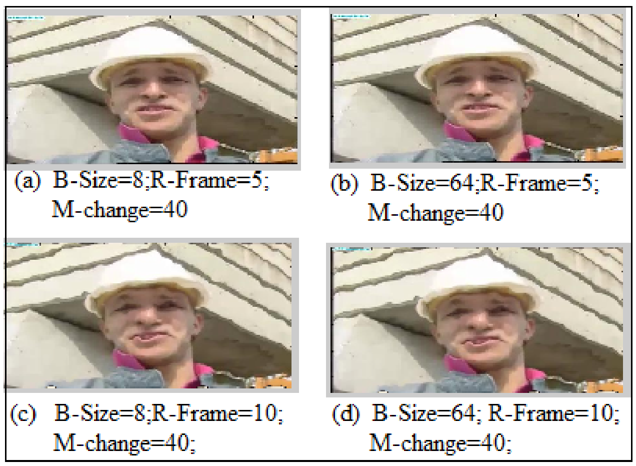


Fig.7 snapshot of frame 56 from Foreman-video after

The Fig.7 shows the frame-56 of Foreman-video that was dropped and re-built by operating different parameter values. It turned out in Akiyo video the relationship between the operational parameters and the quality of the video is similar. And also it turns out that with the low value of the parameter R-Frame is achieving the best results. If we look at the frame snapshots (a, b) in Fig.8; we can find more distorted frame whenever value of the parameter R-Frame rose up from 5 then to 10 the result was more distortion.

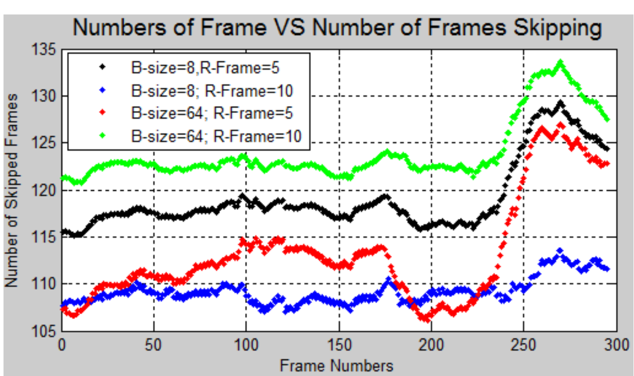


Fig.8: Snapshot of the Operational parameters and frames skipping of Tennis-video clip.

The Fig.8 shows that how our operational parameter handles the Tennis video sample. It is different from all the previous samples (Akiyo and Foreman) Videos, because, the amount and number of frames that are dropped totally different as shown in the fig.9. Besides, the permitted number of frames that can be skipped does not achieved with any combinations of operational parameters. This is because the video contains more intense movement of all previous test videos. Constitute a movement in some cases, more than half of the video-frame resolution. Our proposal performs best with large size of pixel block. Of course, if we look careful in Fig.9 whatever value of the B-size increased that means SFSM skips more frames. It can be concluded that with largest blocks of pixels to deal with large movement better, and R-Frame parameter is less important in intensive video movement.

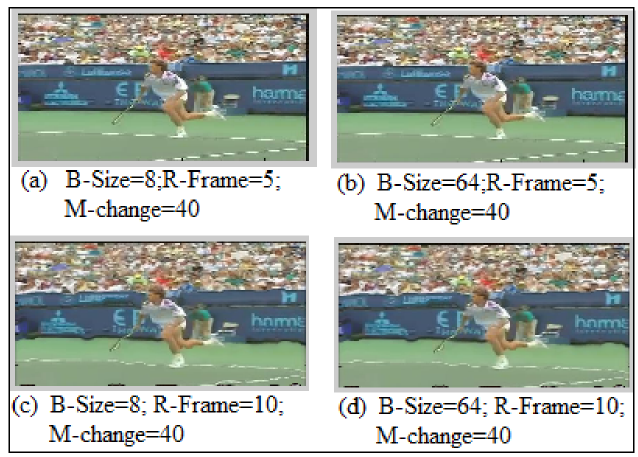


Fig.9 snapshot of frame 235 from Tennis-video after reconstructed by using different parameters.

## V. CONCLUSION

The video streaming is one of the most important issues, particularly, in the topic of global communication system and data processing. Our work divides a video stream into two parts, server and client side interconnected via network. The study focused on the provision of video data from the server to the client over the network as quickly with the least amount of possible loss of frames. The goal of this study is to decrease the number of video frames in the server side by using skipping techniques, while receiving video the client will rebuilt the frame(s) which are dropped in the server side by using interpolation techniques.

The proposal operational parameters were used to test and evaluate the flexibility and adaptability of the outgoing and receiving video streaming through identifying the range values of the operational parameters and then were applied on the video samples. Before the video sent, the server-side in our proposal reducing the number of input video frame (skip), based on the values of the operational parameters with the client will re-built the frames were dropped in server-side.

## IV. References

[1] L. Yuan, F. Wu, Q. Chen, S. Li, and W. Gao, "the Fast Close-Loop Video Transcoder With Limited Drifting Error," pp. 2–5.

[2] S. Moiron, "Video Transcoding for Media Adaptation," 2011.

[3] N. Negi, J. Shukla, I. Ntroduction, I. M. April, and W. S. Email, "Video Transcoding : An Overview of Various Architectures & Design Issues ARCHITECTURE DECODER AND ENCODER," vol. 2, no. 2, pp. 353–357, 2013.

[4] E. Peixoto, "Advanced Heterogeneous Video Transcoding," *Ph.D*, 2012.

[5] Francesca, "Temporal Video Transcoding in Mobile Systems," *Ph.D*, 2007.

[6] I. Ahmad, X. Wei, Y. Sun, and Y. Zhang, "Video Transcoding : An Overview of Various Techniques and Research Issues," vol. 7, no. 5, pp. 793–804, 2005.

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

[8] Z. Zhang, H. Shi, S. Wan, and A. F. R. Decision, "Dynamic frame-skipping scheme for Live Video Encoders," *IEEE Trans. Multimed.*, 2010.

[9] K. Fung, Y. Chan, and W. Siu, "Dynamic frame skipping for high-performance transcoding," *2001 IEEE*, pp. 22–25, 2001.

[10] M. A. Bonuccelli, F. Martelli, and F. Lonetti, "VIDEO TRANSCODING TECHNIQUES IN MOBILE SYSTEMS COORDINATOR : PARTICIPANTS :," no. December, 2005.

[11] 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.

[12] M. Furini and V. Ghini, "A Video Frame Dropping Mechanism based on Audio Perception," pp. 211–216, 2004.

[13] L. Huo, Q. Fu, Y. Zou, and W. Gao, "Network Adapted Selective Frame-Dropping Algorithm for Streaming Media," *IEEE Trans. Consum. Electron.*, vol. 53, no. 2, pp. 417–423, 2007.

[14] B. Zheng and M. Atiquzzaman, "TSFD : Two Stage Frame Dropping for Scalable Video Transmission over Data Networks ," vol. 00, no. C, pp. 43–47, 2001.

[15] H. Liu, W. Zhang, S. Yu, and X. Yang, "Channel-Aware Frame Dropping for Cellular Video Streaming," *2006 IEEE Int. Conf. Acoust. Speed Signal Process. Proc.*, vol. 5, pp. V–409–V–412, 2006.