A Design of Bandwidth Adaptive Multimedia Gateway for Scalable Video Coding

Yi-Mao Hsiao, Su-Wei Yeh, Jai-Shiarng Chen and Yuan-Sun Chu
SoC Laboratory, Department of Electrical Engineering,
National Chung Cheng University, Chia-Yi 621 Taiwan
93mowmow@vlsi.ee.ccu.edu.tw

Abstract—Delivering multimedia streaming on Internet remains several limitations, such as bandwidth fluctuations and network conjunction. We present a multimedia gateway design including dynamic bandwidth estimation, SVC extractor and buffer management to resolve the limitations. The dynamic bandwidth estimation not only selects the appropriate SVC sub-stream, but also measures the change of bandwidth and packet loss with multimedia transmission. When the network congestion happened, the buffer management picks the important one from all packets. Weighted Random Early Detection (WRED) is used to increase peak signal-to-noise ratio (PSNR) of video stream. Comparing four WRED schemes, we figure out that “Disjoint” setting can get higher average of PSNR. The result shows that PSNR is elevated in the system proposed by us.

Keywords—Gateway, SVC

I. INTRODUCTION

Internet nowadays under a rapid growth of necessity not only transmits texts, messages or files, but also video, voice and multimedia streaming. In order to support real-time services, The User Datagram Protocol (UDP) becomes the transport layer protocol. There is no retransmission mechanism in UDP so that data loss and delay will happen during delivery. If these limitations do not affect video display, little data loss or delay are also acceptable. Multimedia streaming data are passed by several routers from source to receiver side over Internet. Some packets are dropped because the buffer is filled in the router or the time of passing has exceeded. As figure 1 shows, when a packet is in a router, the packet is queued to wait for proceeding. If the time decreases, router will pass one by one more quickly. If the time is near zero, the packet is dropped. Every router has it’s own packet queue. When the congestion happened, the queue is full, and the ongoing packet is dropped. So, video display quality becomes worse and the peak signal-to-noise ratio is reduced. In order to handle the translation of different protocols, gateway plays an essential role in packet rate translation, error sequestration and signal transforming.

Gateway is used for translation of IP network packets and public switched telephone network (PSTN) signals. Yoo designs a media gateway that can handle network, telephone and mobile phone’s traffic at the same time [1]. Under the increasing necessity of multimedia information, Xu proposes MeGaDip protocol named media gateway discovery protocol to speed up the search time for multimedia resource [2]. Kang proposes multimedia gateway architecture for High definition TV. The system contains two parts: media gateway and service broker. The gateway functions as transcoding; the broker classifies service based on user’s network status and then passes relative information to media gateway. If the classes are increased, the load of media gateway also increases, and network bandwidth will affects the quality of video [3]. Wien constructs a real-time system for adaptive video streaming based on Scalable Video Coding (SVC). The system contains server, adaptation nodes and client. Server provides source media streaming when nodes request media. Adaptation nodes control the flow to provide appropriate media streaming for client. They present an unequal erasure protection scheme for protection of packet losses in an error prone environment [4]. Foh proposes cross-layer rate control for scalable video transmission over the IEEE 802.11e network. The system is composed of optimal bit allocation, SVC packetization with relative priority index, bandwidth estimation in Enhance Distributed Channel Access (EDCA), and adaptive Quality of Service (QoS) mapping. Optimal bit allocation is an essential component that provides different quality for different bandwidth situation. The idle gap concept is used for bandwidth estimation. A node has to own the right to use the channel; otherwise, it returns into idle status. The idle time is used as idle gap for bandwidth estimation. In order to support QoS in IEEE 802.11e, the parameter settings are mapped to different access categories (ACs) [5].

Fig. 1 Data transmission over server and client sides.

In this paper we propose a design of multimedia gateway for Scalable Video Coding that supports appropriate bit stream for bandwidth adaptive. Section II asserts related background. The proposed system architecture is offered in section III. Section IV contains implementation and analysis. Finally, section V is the conclusion.
II. RELATED BACKGROUND

A. Scalable Video Coding

In order to make H.264/AVC more flexible, JVT creates new video coding standard named Scalable Video Coding [6], which supports three scalabilities: Temporal, Spatial, and Quality Scalability. In the first scalability, video frames are composed of base layer with one or more enhanced layers. The depth of enhanced layers is based on requirement, and are all B-frame (so they are called enhance B-frame). In the second, there are usually two or more layer space resolutions from the minimal QCIF, CIF to 4CIF. The third is different from motion vector and take the remainder from quantization parameter. Then, the remainder is used for enhanced layer that the percentage of transmission decides the quality scalability.

B. Peak Signal-to-Noise Ratio

In order to define the difference between before and after of video encoding/decoding processes, Peak Signal Noise Ratio measures this difference of two video frames, that is also shown in (1). MAXI is maximum possible pixel value of the image. And, samples are linear pulse-code modulation with B bits per sample shown in (2). If PSNR is higher that means the different of two video frames is fewer and the video quality is better shown in (3).

\[
PSNR(dB) = 10 \log_{10} \left( \frac{MAX_i^2}{MSE} \right) \quad \text{.........(1)}
\]

\[
MAX_i = 2^B - 1 \quad \text{.........................(2)}
\]

\[
MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|S(i, j) - D(i, j)\|^2 \quad \text{.........(3)}
\]

C. Random Early Detection and Weighted Random Early Detection

Random early detection (RED) queue mechanism is different from DropTail that packets are dropped before the queue is sufficient [7], and it defines an average queue length variable that is recalculated when every packet arrives shown in (4). The goal of RED is for early detection and preventing network congestion that it uses active queue. Figure 2 shows the relation of queue and dropping probability.

\[
\text{avg} = (1 - Wq) \times \text{old}_\text{avg} + Wq \times q \quad \text{.........(4)}
\]

avg: average queue length
q: current queue length
Wq: real queue length weight (0 < Wq < 1)

When a packet arrives with the average length is smaller than minth, the packet can be enqueued. If the average length is between minth and manth, the dropping probability is Pa. The highest dropping probability is maxp and it is linear growing. As (5) shows, if it exceeds manth, all incoming packets are dropped. But, the dropping probability is not Pb, it is Pb shown in(6). Pa is calculated by last enqueuing packet so that Pb becomes more average. WRED is proposed for different kind of packet with different dropping probability. Besides the same advantages with RED, WRED gives different dropping probability to different priority packets. When the queue length is longer enough, it drops low priority packet randomly. The longer queue length is, the higher dropping probability will see. The high priority packet is dropped even the congestion situation can not under control.

\[
Pb = \text{maxp} \times (\text{avg} - \text{minth}) / (\text{manth} - \text{minth}) \quad \text{..........(5)}
\]

\[
Pa = Pb / (1 - \text{count} \times Pb) \quad \text{..........................(6)}
\]

III. PROPOSED SYSTEM ARCHITECTURE

A media server restrictedly gets information from receiver sides, so we need a gateway in the middle of transmission system. The gateway collects network status from server and bandwidth condition from every client in figure 3. If the network quality or computing ability is not good enough in receiver side, it is hard to get the corresponding video quality. So, we propose a multimedia gateway design to resolve the communication limitation between server and receiver side, so that it can achieve bandwidth adaptive.

![Fig. 2. The relation of queue and dropping probability in RED.](image)

![Fig. 3. System architecture of media delivery system.](image)
We propose a media gateway design for SVC that contains dynamic bandwidth estimation, extractor and buffer management shown in figure 4. The video stream is encoded from server. The media gateway estimates bandwidth condition for receiver side. Then the corresponding sub streams are extracted by extractor. Comparing with bandwidth parameters, the sub stream is transmitted to the client by the media gateway.

A. Dynamic bandwidth estimation

The main purpose of dynamic bandwidth estimations is to measure bandwidth between media gateway and clients. As figure 5 shows, we assume that the bandwidth between server and media gateway is enough that there is no data loss when media stream is encoded and then transferred from server to media gateway. In the initial status system, bandwidth estimation is static so that it can avoid losing video packet. The time of static bandwidth estimation is five seconds. The dynamic bandwidth estimation is active during the transmission status. Because there is no retransmission for UDP protocol even the packet loss may happen, the dynamic bandwidth is still essential. Figure 5 shows the operation of bandwidth estimation. The server sends the media stream to media gateway. There is a short period for initial measurement to realize the bandwidth status and then the SVC sub stream is sent to clients with dynamic bandwidth estimation. So, the system can save extra bandwidth and time for its estimation.

B. Extractor

The extractor efficiently and speedily extracts SVC sub stream. The source stream for SVC encoder contains several formats with different frame rates—30fps, and 15 fps—and different resolutions—CIF(352x288), and QCIF (176x144). In our system, the extractor and dynamic bandwidth estimation are adaptive to video stream. The extractor not only extracts sub stream but also generates packet trace file for SVC stream so that we can get the information about length and types of packets.

C. Buffer management

The video streams are transmitted as UDP protocol so that some packets still loss during delivery process. The packets are classified according to different important priority of the dropping decision when network congestion happened. WRED is buffer management in proposed media gateway. Every queue has its own parameters, such as minth, maxth and maxp, to decide how to drop packets during network conjunction. In WRED we can choose dropping percentages for different priority packets. The priority is based on video frame types. The incoming packets are classified into I-frame, P-frame and B-frame, and then WRED calculates current avg and is based on minth and maxth to decide which packet to drop.

IV. IMPLEMENTATION AND ANALYSIS

We use UDP transport layer protocol to transmit video stream, as in figure 7 shows the system protocol stack. SVC encoder is in application layer. When media gateway receives the source, SVC stream, bandwidth estimation and extractor extracts sub stream and sends it to buffer management in network layer, then sends the sub stream to clients. SVC decoder decodes sub stream after receiving packets. We implement socket programming to do bandwidth estimation between media gateway and clients.
The formula shown in (7) is bandwidth estimation in UDP.

\[
\text{Bandwidth} = \text{datagram size} \times \text{number of datagram} / \text{period} \ldots (7)
\]

The extractor extracts sub stream from the source of SVC stream, using frame rate (temporal_level), resolution (dependency_id) and quality (quality_level). These parameters decide sub stream from extraction so that we can get the appropriate sub stream for bandwidth situation from the measurement result of the dynamic bandwidth estimation. In our system, media gateway receives SVC stream and then extract sub streams. Choosing appropriate sub stream for bandwidth condition, media gateway use UDP socket to send the sub stream to client.

![Fig. 8. Simulation network topology](image)

Four parameter settings are manifested in buffer management simulation: Disjoined, Staggered, Partially Overlapped and Fully Overlapped [8]. Disjoined WRED (D-WRED) uses three groups of \( \text{maxth} \) and \( \text{minth} \) with non-overlapping. Staggered WRED (S-WRED) is constantly in processes. Partially Overlapped WRED (PO-WRED) and Full Overlapped (FO-WRED) use partial and full overlapping to setup. We adopt NS-2 simulator to do simulation [9]. Figure 8 shows network topology. The bandwidth is 512 kbps. We transfer foreman H.264 video with 300 frames and 30 fps. The queue size is 100 packets and every packet size is 1000 bytes. As the parameter setting of four WRED setup shown in table I, we simulate the PSNR of video stream for these four setups. The packet loss and PSNR simulation in DropTail and WRED is shown in table II. There are packets loss comparison among I-frame, P-frame and B-frame. With different WRED setting, the packet dropping situations are different obviously. For D-WRED, S-WRED and PO-WRED, the minimal dropping average queue of B-frame is lower than the other two, so that the most frequent happening of packets dropping are B-frame. However, in FO-WRED there are full Overlapped, so that the packet loss situation is similar to DropTail even PSNR is lower.

### Table I

<table>
<thead>
<tr>
<th>Packet Type</th>
<th>B-min</th>
<th>B-max</th>
<th>P-min</th>
<th>P-max</th>
<th>I-min</th>
<th>I-max</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-WRED</td>
<td>20</td>
<td>40</td>
<td>45</td>
<td>65</td>
<td>70</td>
<td>90</td>
</tr>
<tr>
<td>S-WRED</td>
<td>30</td>
<td>50</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>90</td>
</tr>
<tr>
<td>PO-WRED</td>
<td>40</td>
<td>60</td>
<td>55</td>
<td>75</td>
<td>70</td>
<td>90</td>
</tr>
<tr>
<td>FO-WRED</td>
<td>70</td>
<td>90</td>
<td>70</td>
<td>90</td>
<td>70</td>
<td>90</td>
</tr>
</tbody>
</table>

### Table III

<table>
<thead>
<tr>
<th>Number of packets</th>
<th>Total</th>
<th>I</th>
<th>P</th>
<th>B</th>
<th>Avg. PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>DropTail</td>
<td>270</td>
<td>135</td>
<td>74</td>
<td>61</td>
<td>24.03</td>
</tr>
<tr>
<td>D-WRED</td>
<td>186</td>
<td>26</td>
<td>4</td>
<td>156</td>
<td>28.24</td>
</tr>
<tr>
<td>S-WRED</td>
<td>190</td>
<td>51</td>
<td>15</td>
<td>124</td>
<td>28.19</td>
</tr>
<tr>
<td>PO-WRED</td>
<td>232</td>
<td>80</td>
<td>52</td>
<td>100</td>
<td>26.04</td>
</tr>
<tr>
<td>FO-WRED</td>
<td>268</td>
<td>135</td>
<td>79</td>
<td>54</td>
<td>23.99</td>
</tr>
</tbody>
</table>

V. CONCLUSION

We propose a multimedia gateway design for SVC that is composed of dynamic bandwidth estimation, extractor and buffer management. The gateway can choose appropriate SVC sub stream for the clients of bandwidth condition. So, the media gateway achieves video quality of bandwidth adaptive. The simulation result shows that “Disjoint” setting of WRED saves more PSNR than others when network congestion happened. After simulations, PSNR has been improved in our proposed system.

### REFERENCES