REAL-TIME VIDEO QUALITY CONTROL FOR MULTIMEDIA NETWORK

by

BIAO JIANG

A dissertation submitted to the Graduate Faculty in Electrical Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2013

This manuscript has been read and accepted for the
Graduate Faculty in Engineering in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy

Tarek N. Saadawi, Professor, The City College of New York

Date

Chair of Examining Committee

Ardie D. Walser, Professor and Acting Associate Dean, The City College of New York

Date

Executive Officer

Mehmet Ulema, Professor, Manhattan College

Micheal Conner, Professor, The City College of New York

Yi Sun, Assistant Professor, The City College of New York

Yingli Tian, Professor, The City College of New York

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

REAL-TIME VIDEO QUALITY CONTROL FOR MULTIMEDIA NETWORK

by

Biao Jiang

Adviser: Professor Tarek N. Saadawi

In this thesis, we propose a new approach for video quality control for multimedia networks. Our new approach is based on video quality measure that combines both the network quality of service (QoS) as well as the user quality of experience (QoE). The proposed approach improves the end-to-end traditional video quality control for multimedia network by including the human perception of video data, which is major concern for the video client, along with the network quality of service (QoS) measurements. In our approach we use packet loss rate as quality of service (QoS) parameter, and self-reference complex wavelet video structural similarity index (SRCW-VSSIM) as quality of experience (QoE) parameter. Compared with traditional QoS only video quality control technique, the proposed video quality control technique for multimedia networks is based on including both QoS and QoE parameters. We will show that the proposed QoS-QoE based video quality control algorithm can reflect both the condition of the network environment and the human perception of the received networked video data stream. According to both QoS and QoE parameters, rather than using only QoS parameter, video quality control

action will satisfy the user needs more than relying only on the network conditions. Since our proposed QoE parameter SRCW-VSSIM can be obtained with no reference (NF) video data, it satisfies the requirement of real-time video transmission.

In addition to our video quality control technique, we introduce machine learning approaches for combined QoS-QoE based video quality control techniques for real-time streaming service. The proposed schemes are based on statistical learning technique, Support Vector Regression (SVR), to predict combined QoS-QoE parameter in the near future. The character of machine learning technique makes this scheme proactive, and be able to trigger the rate control action to adjust the video streaming rate before network conditions start deteriorating. QoS-QoE based video quality control indicator (QQVQCI), defined as the combined QoS-QoE parameter for real-time video quality control, mixed with QoE index are used to generate training dataset to predict QQVQCI in the near future.

Theoretical analysis as well as simulation results are presented.

Acknowledgment

First and foremost, I would like to express my sincere gratitude to my advisor, Professor Tarek N. Saadawi, for the continuous support of my Ph.D study and research, for his guidance, motivation, enthusiasm, patience, immense knowledge and understanding. His consistent support and guidance helped me in all the time of research and writing of this thesis. I could not have imagined writing this thesis without the help from my advisor. One simply could not wish for a better or friendlier advisor.

Besides my advisor, I would like to thank the rest of my dissertation committee members, Professor Mehmet Ulema, Professor Yingli Tian and Professor Yi Sun for their valuable suggestions and insightful comments.

My special thanks goes to Professor George Kranc. He helps me understand how to play the role as an instructor as well as a teaching assistant. His spirit of teaching and optimism inspired me, and encouraged me to get over all obstacles on the way of my doctoral study.

My sincere thanks also goes to my colleagues and friends in the Center for Information Networking & Telecommunications CCNY, including Qihua Yang, Orhan Celebi, Ahmed Abdelal, Joseph Soryal, Nuraj Pradhan, Xijie Liu and many others for their friendship and supports. I would like to give my special thanks to Qihua Yang and Orhan Celebi for their collaborations in support vector regression and video quality assessment parts of this dissertation.

Last but not least, I would like to thank my family: my wife, my parents and my lovely son. Without their love and support, I cannot make it so far.

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

Nowadays, huge amount of video is streamed over IP-Based multimedia networks, such as the Internet. However, both service providers and end users still suffer from unreliability of packet transmission. Ensuring the quality of video streaming becomes a major concern of general public. In order to meet user satisfaction, there is a need to monitor and control video quality. Currently, applications, such as video conference and video streaming, require a guaranteed Quality of Service (QoS) to work properly. Therefore current real-time video quality control (VQC) algorithms attempt to adapt streaming rate to avoid severe frame delay, frame distortion and frame loss.

The main video quality control approaches can be classified into two types: formula-based approach and measurement-based approach. Formula-based approach attempts to describe traffic, analyze and predict network condition based on mathematical models. Measurement-based method gathers path resource information, such as available bandwidth, packet loss, delay, and applies these statistics to control the source sending rate in order to satisfy the QoS requirement.

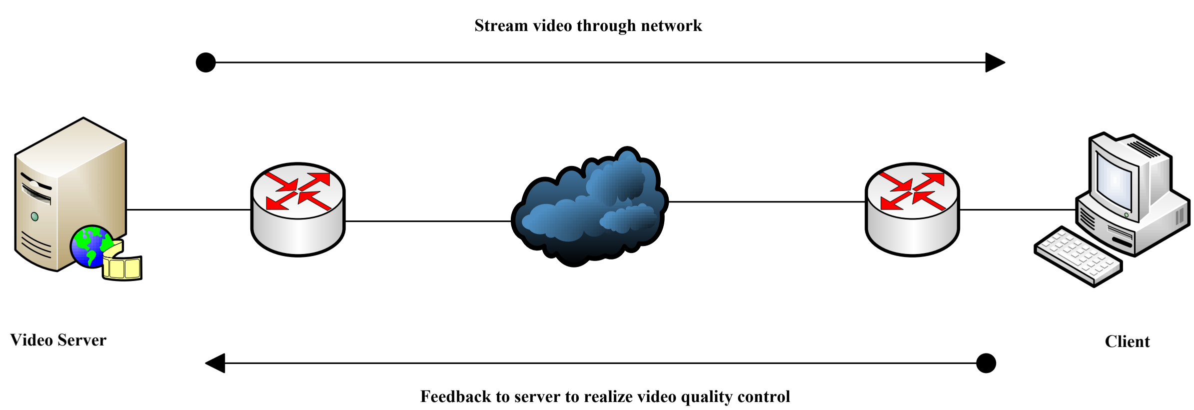

Figure 1. A networked video transmission system

Figure 1 shows the diagram of a networked video transmission system. However, both methods of video quality control attempt to adjust the video sending rate to adapt to the network condition only, and don't include the human perception. We concede that network condition has significant influence on video transmission, especially when severe congestion happens, but human perception, or referred to as quality of experience (QoE) [1], represents the major consideration for networked video data, and it should be properly included in video quality control algorithm to trigger proper actions to meet the needs of both QoS and QoE.

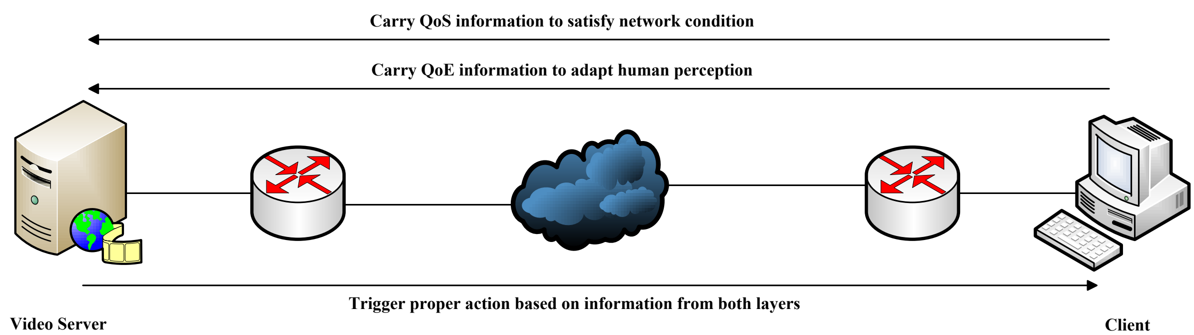

Figure 2. Cross-layer video quality control technique

As shown in Figure 2, Our key contribution to video quality control technique is that we bring our innovated real time QoE parameter to correlate with QoS parameter, and introduce a cross-layer based real time video quality control indicator. Compared with traditional QoS only video quality control technique, our cross layer design has the ability to trigger proper action to satisfy the needs of both human perception and network condition. To better design our cross-layer video quality control indicator, we should choose appropriate QoE parameter to represent human vision perception in real time, and proper QoS parameter to represent network condition.

Currently, quality of experience (QoE) is an intense research area, and different QoE assessment methods have been proposed to describe subjective video service experience. Generally, most of prior research on QoE assessment is divided into three categories: subjective assessment, objective assessment, and hybrid respectively.

Subjective assessment is considered the most accurate approach to assess perceived quality, since it is the indicator given directly by humans. Mean opinion score (MOS) [2] is the output of subjective assessment, and it rates the perceived quality using 5 grades: Excellent, Good, Fair, Poor and Bad. But its high cost of manpower limits the use of MOS.

Because of the limitation of subjective assessment, objective assessment is intensively researched. Instead of using humans to rate the perceived quality, objective assessment applies objective information, such as network parameters, to evaluate quality of experience automatically. A typical objective assessment is Peak Signal-to-Noise Ratio (PSNR) [3]. However, PSNR can hardly be applied to real-time video transmission based on the need for original video data. Meanwhile, PSNR cannot reflect the real human perception accurately, since error visibility is not in proper proportion to PSNR. Some other objective assessments try to predict quality of experience by using network parameters, such as packet loss, delay, jitter, but certain methods still can hardly correlate human perception with its index.

Hybrid assessment, for example, pseudo subjective quality assessment (PSQA) [4], mixes both subjective method and objective method, and try to balance the drawbacks of both methods, but introduce both drawbacks in certain extend. Due to the disadvantages of current QoE assessments, a real time, objective, and human perception focused QoE assessment is highly needed.

Current objective QoE assessment is classified into three types: Full-Reference (FR), Reduced-

Reference (RR), and No-Reference (NR). FR and RR methods, such as Mean Square Error (MSE) and Peak Signal Noise Ratio (PSNR), need additional network resources to access to full or portion of the original video signal [5], [6]. This can be considered as a serious drawback when it comes to real-time multimedia communication, because of the limited bandwidth and unavailability of reference data. At the same time, most of current FR or RR methods are error sensitive methods, and this can be considered as another main drawback, since error sensitive methods may not properly reflect the real human vision perception. We will discuss in details this part in the following chapter.

NR method becomes a very suitable option for real-time QoE assessment, since it does not require the original video [7]. However, with no reference video signal, NR method suffers from less accuracy when compared with FR and RR methods. NR methods can be divided into three types: No-Reference Pixel (NR-P), No-Reference Bit-stream (NR-B), and hybrid of NR-P and NR-B. Current NR methods mainly focus on video coding and transmission. However statistics from video coding and transmission system, such as coding rate and packet loss rate, hardly linear correlate with human vision perception. For example, with certain packet loss rates due to transmission errors, human vision perception may still be regarded as acceptable. In essence, human vision perception should be the core concern of QoE assessment.

As part of our QoS-QoE cross-layer design video quality control indicator, our QoE parameter, Self-Reference Complex Wavelet Video Structural Similarity Index (SRCW-VSSIM), can evaluate human perception in real-time. SRCW-VSSIM is a No-Reference QoE assessment method, since there is no need of original video data, it satisfies the requirement of real-time measurement. Also, instead of tracking error statistics, SRCW-VSSIM directly evaluates human vision perception.

Meanwhile, on QoS side, packet loss is an ideal parameter to apply to our proposed approach, since networked video is packetized, and transmitted through packet switched network. Packet loss can show the network condition clearly, so packet loss is chosen as the QoS parameter for our algorithm.

Based on the chosen QoE and QoS parameters, we propose a QoS-QoE cross-layer based video quality control approach. Such approach is an objective video quality control algorithm that works in real-time and reflects both human perception and network condition.

Although our innovated cross-layer based video quality control approach successfully take both network condition and human perception into consideration, it still suffers from its passive in nature, since actions triggered according to current video service quality are reactive, and may lead to severe situations. Obviously, reactive algorithm cannot completely satisfy the requirement of end users, since users hope that the video streaming maintains a constant quality output. The support vector regression (SVR) technique just meets the needs for proactive video quality control.

As a powerful machine learning technique, support vector regression (SVR) allows computer to evolve behaviors based on training data. Multiple linear or nonlinear inputs are applied as training examples, and then generate the output prediction. SVR has been proved to be well performed in various fields such as weather forecast, vehicular traffic, financial market. However, very limited research has been done on IP-based network video quality prediction.

Our innovative approach here proposes to apply SVR algorithm and develops a measurement and a machine learning mixed approach to predict the QoS-QoE cross-layer based video quality control indicator (QQVQCI) in the near future. According to the parameter predicted by our approach, proper actions can be triggered to prevent poor user experience proactively

The remainder of this thesis is organized as follows. Chapter 2 presents a survey of previous work in real-time video quality control area. We develop a QoS based video quality software application in chapter 3, and discuss the limitations of current video quality control technique. In chapter 4, we propose our research on similarity structural based real time video quality measurement. In chapter 5 we propose a novel video quality measurement technique, and apply our new QoE measurement to QoS-QoE based cross-layer video quality control approach. In this chapter, we also apply support vector regression to our video quality control technique to make it proactive method. Chapter 6 we evaluate the performance of our video quality measurement approach and real-time video quality control algorithm by simulation. In chapter 7, we conclude our work and discuss the future research.

2. PREVIOUS WORK

2.1 Previous Video Quality of Service (QoS) Control Work

With rapid development of communication networks, high volume of video streaming is possible to be transmitted by various consumer applications. Meanwhile, quality of networked video becomes the key concern of both video service providers and video service receivers. Since video transmission weight a very heavy portion in total network data flow, control the quality of the video is intensively researched. Figure 3 shows a block diagram of current video quality control techniques. Current video quality control technique only takes statistics from network layer, QoS, into account, and service provider can adjust video streaming rate based on the network condition.

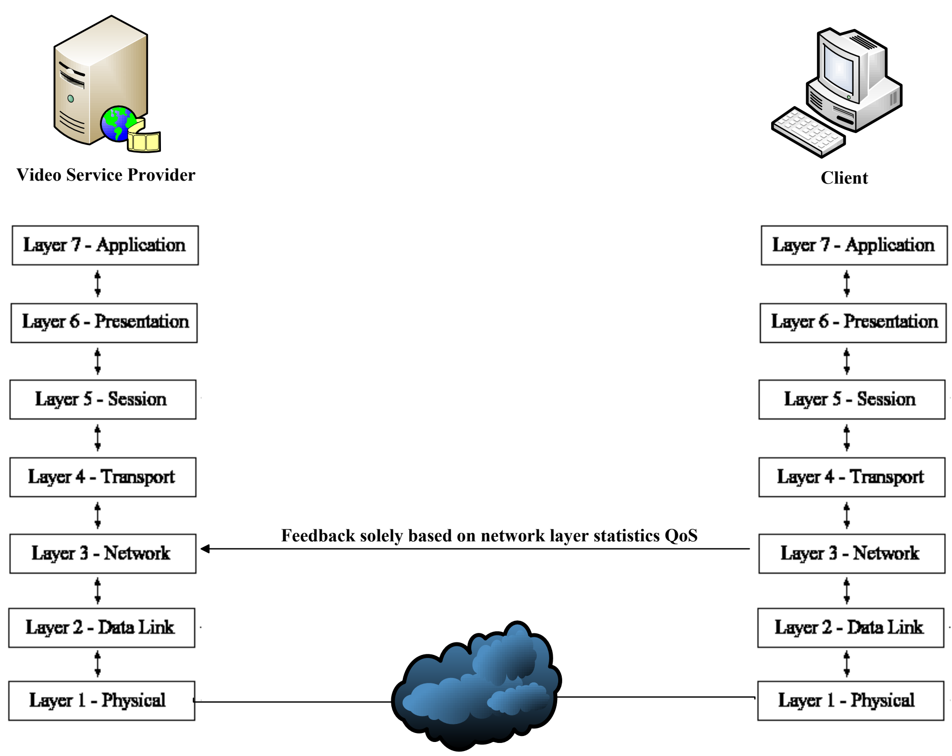

Figure 3. Current QoS only video quality control

Most of prior research on this problem is divided into two categories. One category is formula-based approach. Formula-based approach attempts to describe traffic, analyze and predict network condition based on mathematical models. Floyd et al. [8] propose an equation-based congestion control approach for unicast application. The approach lays on the TCP-Friendly rate control (TFRC) protocol. In [9], Suki et al. discuss the relationship of TFRC congestion control protocol to video rate control optimization. In [10], Huang et al. develop a feedback control system model for video streaming systems, which takes into account the interactions among video rate control, RED active queue management, and received video quality. The authors also derived a P controller that stabilizes both homogeneous video and heterogeneous video system.

The other category is measurement-based approach. Measurement-based method gathers path resource information, such as available bandwidth, packet loss [11], delay, and applies these statistics to control the source sending rate in order to satisfy the QoS requirement, which is believed to ultimately contribute to the user's QoE [12]. A multi-layer active queue management method is proposed by Kang et al [13], the authors allow the video to mark their own packets with different priority, and use the proposed queue management method to control the router to drop the less-important in order to stable the video quality when congestion happen. In [14], Kim et al. propose video quality control system which can control video service quality through the monitoring of end-to-end available bandwidth for video streaming service like IPTV in NGN convergence network. In [15], Jammeh et al. propose a delay-based congestion avoidance approach for video communication with fuzzy logic control, and this approach use delay and computational intelligence to replace packet loss and throughput modeling as input to proposed algorithm.

Although major measurement-based approaches focus on QoS parameter, such as bandwidth, delay and packet loss, a few researchers make their efforts to build mathematical model of QoE based on network measurement. In [16], Kim et al. propose a QoE assessment model for video streaming service using QoS parameters. Kitamura et al. [17] consider the relationship between the QoE of Video streaming and QoS, and propose mirco-second resolution to capture the precise behavior which effects the codec system's performance. In [18], Suzuki et al. estimate QoE from MAC-level QoS in audio-video transmission with IEEE 802.11e EDCA. Even though, all these methods suffer from uncertain relationship between QoE and QoS.

Most of current video quality control work is reactive in nature, since they are based on current measurement, no matter it is QoS based measurement or QoE based measurement. The introduction of support vector regression can solve the problem in a large extent.

In 1996, Vapnik et al. [19] proposed a new version of Support Vector Machine (SVM) for regression, and called it support vector regression (SVR). Different from tradition support vector machine, the prediction model generated by SVR depends only on a subset of the training data, since training data points lie beyond the margin are ignored by the cost function. Work [20] provides an extensive description of SVR.

As a powerful tool, SVR has been applied in various areas. In environmental parameter prediction, Lu [21] estimated forest air quality parameters using SVR. In financial data estimation, Chen [22] studied the use of SVM and SVR for predicting financial distress and bankruptcy. However, limited research has been done on network related estimation problems using SVR prediction technique. Reference [23] however uses SVR for prediction of network flows, and [24] uses SVM to predict round trip latency to random network destinations.

SVR is also applied to predict the QoE of a video. A No-reference video quality measurement with SVR has been proposed in [25]. Reference [26] uses SVR to predict image quality automatically. In [27], Wang et al. extract the video features from both bitstream and pixel domains, and use SVR to predict the mean opinion score (MOS).

We include SVR in our video quality control technique to demonstrate that it is possible to implement our real time cross layer design video quality indicator to predict the proper actions before either QoE or QoS of video service degrading dramatically.

2.2 Previous QoE Assessment Work

In order to achieve the goal of our design, we need to embed proper QoE parameter and QoS parameter into our indicator. QoS parameter has been intensively researched for decades, and several QoS parameter can be chosen and embedded into our indicator. Here we choose packet loss as our QoS parameter, since real time video streaming is transmitted through packet switched network, and lost packets can be considered as the major impact to video streaming service. Compared with QoS parameter, current QoE parameter can hardly satisfy the goal of our design, since our design requires a real time and human perception based QoE parameter. In this section, we will review the previous research on the QoE of a video.

In the last ten years, video application has been developed rapidly. Several well developed video compression standards [28]-[30] as well as high performance transmission systems [31]-[35] bring video services, such as IPTV, P2P streaming, video conference, video on demand, and video surveillance, closer to general public than ever before. A recent forecast [36] shows that mobile devices will contribute 66% video transmission of the global mobile data traffic by 2014. This factor motivates the video service providers to match the expectations of end user of video service. Service providers believe that reliable video quality assessment is extremely important based the needs for meeting the promised quality of service (QoS) and improving the end user's quality of experience (QoE) [37].

We can see from Figure 4 as the detailed classification of video quality measurement. Subjective video quality measurement is widely considered as the most accurate tool to describe the human perception of video quality. A very popular way to subjectively measure video quality is mean opinion score (MOS). Audiences are involved to view the video flow and give response to the quality of the video based on a score system, which has a scale of 1 to 5 to map the Excellent, Good, Fair, Poor and Bad. However, because of the manpower, the method becomes very expensive in cost, so it is impossible to be used as a major tool to measure the video quality.

Other than the subjective method, objective method turns out to be the only choice of human being. Objective methods try to predict human vision perception by using objective statistics. ITU has classified objective video quality measurement into five main categories [38] depending on the type of input data that is being used for quality assessment: media-layer model, parametric packet-layer model, parametric planning model, bitstream-layer model and hybrid model. Here we focus on media-layer model. Media-layer objective video quality assessment methods can be further classified into three categories: full-reference (FR), reduced-reference (RR), and no-reference (NR) [39].

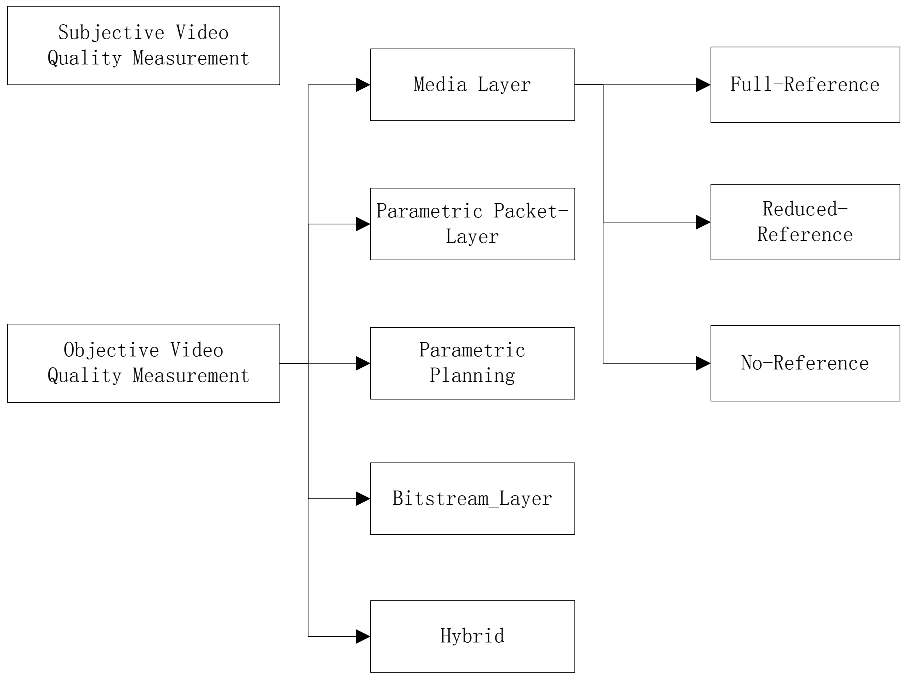

Figure 4. Classification of Video Quality Measurement

When a reference or partial information is used to assess the video quality, we call the video quality method full-reference or reduced-reference method. These two methods are normally applied to non-real-time application, since real-time scenario requires large bandwidth to ensure the service quality.

Traditional full-reference method includes mean square error (MSE) and peak-signal-to-noise ratio (PSNR). Given a reference frame f and a received frame g, and assume that both of them have the same dimension of $M \times N$, so the MSE is defined as:

$$MSE(f,g) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(f_{ij} - g_{ij})^2$$

and PSNR defined as:

$$PSNR(f,g) = 10log_{10}(255^2/MSE(f,g))$$

In statistics, the mean squared error (MSE) of an estimator is one of many ways to quantify the difference between values implied by an estimator and the true values of the quantity being estimated [40]. According to the definition, we can tell that MSE is an error sensitive tool when applied to video quality assessment based on its mathematical character. PSNR can be evaluated the same way since it is built on top of MSE. MSE and PSNR have their own advantage that they are very sensitive to errors, but this advantage can be consider as the weakness of them, because human vision system (HVS) is not very sensitive to error [41]. For example, a few errors can be found when comparing original video and test video, MSE or PSNR will indicate that video quality is poor, however, human may consider the video quality is good due to the visibility of these errors. Even MSE and PSNR suffer certain drawbacks, but they can still indicate severe distortion every accurately.

Other than MSE and PSNR, a few new video quality measurement tools are proposed, and these measurement use the model closer to human vision system, so they can adapt the human perception better than traditional methods. According to natural visual statistics, Wang et al. [42] proposed the Video Structural Similarity (VSSIM) index (this is discussed in chapter 4), and VSSIM uses the Structural Similarity (SSIM), which is used as an image quality assessment (IQA) tool, to compare original frame and test frame to indicate the perceptual structural information loss in HVS. The author also tried other versions of SSIM to applied to video quality assessment (VQA), for example MultiScale-SSIM (MS-SSIM) [43] and the Speed SSIM [44]. These newly invented methods highly improved the performance of full-reference video quality measurement from error visibility to human perception.

Different from error sensitivity-based methods, Wang et al. [45], [46] proposed Structural Similarity (SSIM) and Video Structural Similarity (VSSIM), which follow the new philosophy that structural distortion can be an estimate of perceived visual distortion. As FR methods, SSIM and VSSIM have been proved closer to human vision perception than error sensitive method. However, VSSIM does not satisfy the requirement of real-time QoE assessment because of its full reference method and complexity.

Wang et al. [47] propose Structural Similarity in complex wavelet domain (CW-SSIM) for image quality assessment (IQA) to avoid the drawbacks that SSIM in spatial domain is highly sensitive to translation, scaling and rotation. This core character of CW-SSIM will be applied in our no-reference QoE design. Meanwhile, its less complexity makes it friendly to real-time algorithm.

Reduced-reference method can be sorted into two categories: frequency domain and pixel domain. In frequency domain, DCT transform, wavelets transform, and other transforms are applied to detect the distortion in different frequency regions. Applying the spatio-temporal model of the human visual system, Lambrecht and Verscheure created the MPQM [48]. In pixel domain, Hekstra et al. [49] find that the HVS is sensitive to edges and local changes in luminance, and propose Perceptual Video Quality Metric (PVQM). The author linearly combined three distortion indicators: edginess, temporal decorrelation, and color error to measure the perceptual quality. Some other methods [50], [51] also use partial information of original video data to evaluate the QoE of a video. Reduced-reference method has the advantage over the full-reference method due to the less needs of original video data, but still hard to be applied to real-time application.

No-reference is becoming more and more important due to blooming development of real-time applications. Real-time scenario makes it impossible to access original video data, so well-developed FR and RR methods cannot contribute under certain situation. No-reference video quality assessment has three major branches: no-reference pixel (NR-P), no-reference bitstream (NR-B), and hybrid of them. Video coding and transmission are focuses of current no-reference video quality assessment research, while such statistics from video coding and transmission system, for example, coding rate and packet loss rate, can hardly tell the human vision perception linearly.

The first no-reference method [52]-[54] concentrates on the measurement of blocking artifacts. The limitation of this method is that it solely takes blocking artifacts into account. Gastaldo et al. [55]-[57] propose to extract features from the compressed bit-stream rather than to process the decoded video. This research inspired a new direction of no-reference video quality metrics. In [58], the authors extract a set of features from MPEG-2 bitstream and use tree classifier and Generalized Linear Model (GLM) to determine the packet loss visibility to human perception. In [59], the authors consider a NR-P method, where they addressed the problem of estimating which portions of a frame have been lost during transmission. Then they use this information to compute the MSE distortion for H.264/AVC video. Yang et al. [60] use information extracted from the compressed bit stream without resorting to complete video decoding to measure networked video. Kim et al. [61] propose a no-reference video quality assessment method with estimation of dynamic range distortion.

3. QoS BASED VIDEO QUALITY CONTROL

In this chapter, we develop our own application to better understand the QoS only video quality control technique. During testing, we identified some limitations of current QoS only video quality control technique, and we will prove our cross-layer real time video quality indicator can overcome these limitations.

In recent years, the design of video distribution system is an intensive research area. Well-designed streaming application has to face two main challenges:

1) How to adapt the needs from users with different heterogeneous capabilities such as buffers size, processing speed rate, reception rate.

2) How to adapt the dynamic network condition, such as transmission delay, packet loss rate.

Thus a successful video streaming application should keep tuning the streaming rate to prevent network congestion and avoid overwhelming the client buffer. Therefore, a proper choice to solve the problem is that client side detects its network condition and asks to tune streaming rate to achieve better human perception.

Our software implementation focuses on two scenarios:

1) End-to-End unicast streaming. As shown in Figure 5, receiver keeps monitoring and collecting network quality of service (QoS) parameter, such as packet loss rate. Server side will receive the update of network condition from receiver side, and will adjust the source sending rate once the pre-determined threshold is reached.

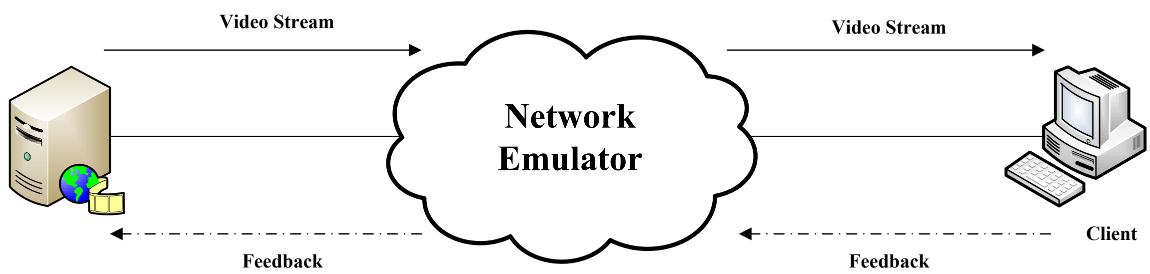

Figure 5. Unicast rate control technique

2) Multicast streaming. As shown in Figure 6, service provider creates multiple multicast streams with different rates, and allows users to switch between different multicast groups. Receivers dynamically joining and leaving the multicast groups. The same as the unicast scenario, receiver monitors and collects packet loss rate information, and choose the multicast group with proper stream rate to join.

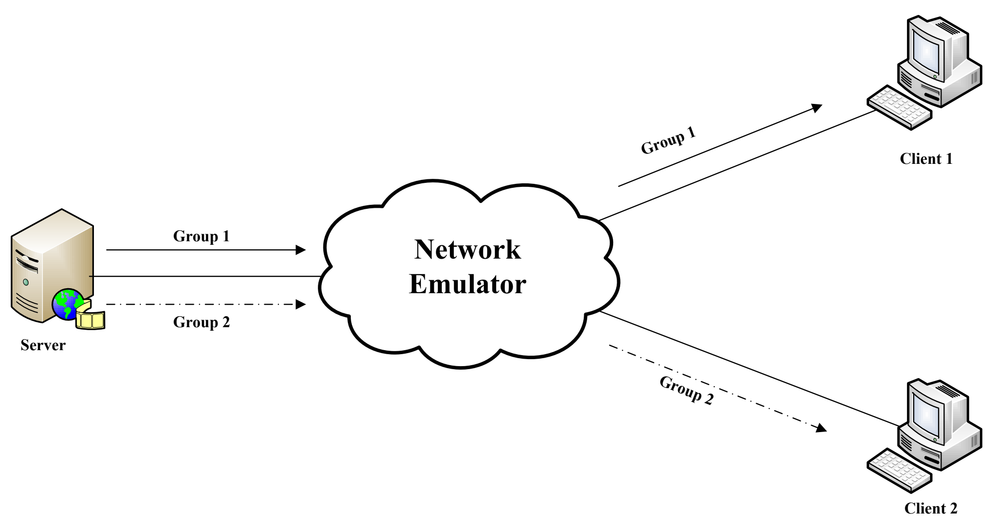

Figure 6. Multicast video quality control technique

Other than the regular QoS based video quality control, our demo can also use external software to determine the network situation, and tune the streaming rate or switch to proper multicast group.

As an example, we work with Telcordia Technologies, and enable its product, Network Resource and Performance Estimation (NRPE), to provide network measurement service. The NRPE software has been implemented in C and computes a congestion indicator through a technique known as Differential Performance Packet Probing (DP3).

3.1 Testing System Setup

We setup our testing environment using following equipment:

a) Three end system work stations: one acts as the video server (station 1 in the Figure 7), the other two are the clients (i.e., users, station 2 and 3 in the Figure 7)

b) One workstation equipped with OPNET (System in the Loop, SITL, should be included), and the emulated network is three routers shown in Figure 7.

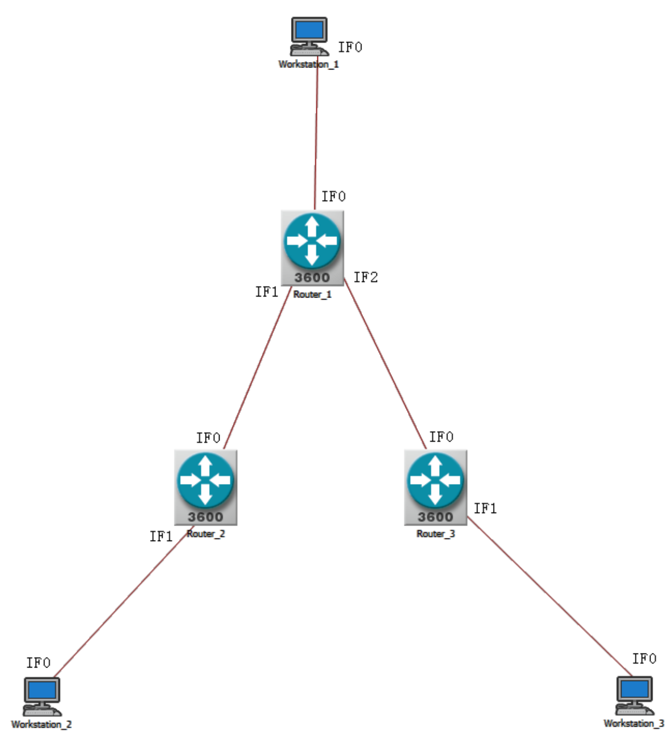

Figure 7. Implementation of the Test Environment

VMware is also needed on the three workstations, since external network measurement tool NRPE runs on Linux while our multi-streaming software runs on Windows. In OPENT, configure IP address, subnet mask and default gateway for Windows platform, Linux virtual platform and certain interface of the router, and make sure that priority queue is configured for all routers. For example, if the testbed is implemented as shown in Figure 7, the IP addresses for the three workstations and three routers should be configured as follows:

Workstation	Windows System	Linux virtual system
Workstation_1	IF0 (Interface 0) IP: 192.168.1.1 Subnet mask: 255.255.255.0 Default Gateway: 192.168.1.3	IF0 (Interface 0) IP: 192.168.1.2 Subnet mask: 255.255.255.0 Default Gateway: 192.168.1.3
Workstation_2	IF0 (Interface 0) IP: 192.168.2.1 Subnet mask: 255.255.255.0 Default Gateway: 192.168.2.3	IF0 (Interface 0) IP: 192.168.2.2 Subnet mask: 255.255.255.0 Default Gateway: 192.168.2.3
Workstation_3	IF0 (Interface 0) IP: 192.168.3.1	IF0 (Interface 0) IP: 192.168.3.2

	Subnet mask: 255.255.255.0 Default Gateway: 192.168.3.3	Subnet mask: 255.255.255.0 Default Gateway: 192.168.3.3

Table 1. IP addresses for the three workstations

Router	Interface	Routing parameters
Router_1	IF0	IP: 192.168.1.3 Subnet mask: 255.255.255.0
	IF1	IP: 192.168.4.1 Subnet mask: 255.255.255.0
	IF2	IP: 192.168.5.1 Subnet mask: 255.255.255.0
Router_2	IF0	IP: 192.168.4.3 Subnet mask: 255.255.255.0
	IF1	IP: 192.168.2.3 Subnet mask: 255.255.255.0
Router_3	IF0	IP: 192.168.5.3 Subnet mask: 255.255.255.0
	IF1	IP: 192.168.3.3

		Subnet mask: 255.255.255.0

Table 2. IP addresses for the three routers

3.2 Graphical User Interface for Unicast Scenario and Multicast Scenario

In this section, we will walk through the graphical user interface (GUI) of our application. Since our software demo is designed for video quality control under unicast and multicast scenarios, we will introduce GUI for each scenario.

Scenario 1: Play video stream using Unicast.

Configure Unicast Server (UC SERVER).

On the server side, for example workstation_1, choose "UC SERVER" mode. The interface below will show up. Click "Add" button to choose video files that user want to play for both receiver with feedback and receiver without feedback. Configure IP address and port number, and user can also adjust the initial sending rate. Click "play" button on the right to start the server.

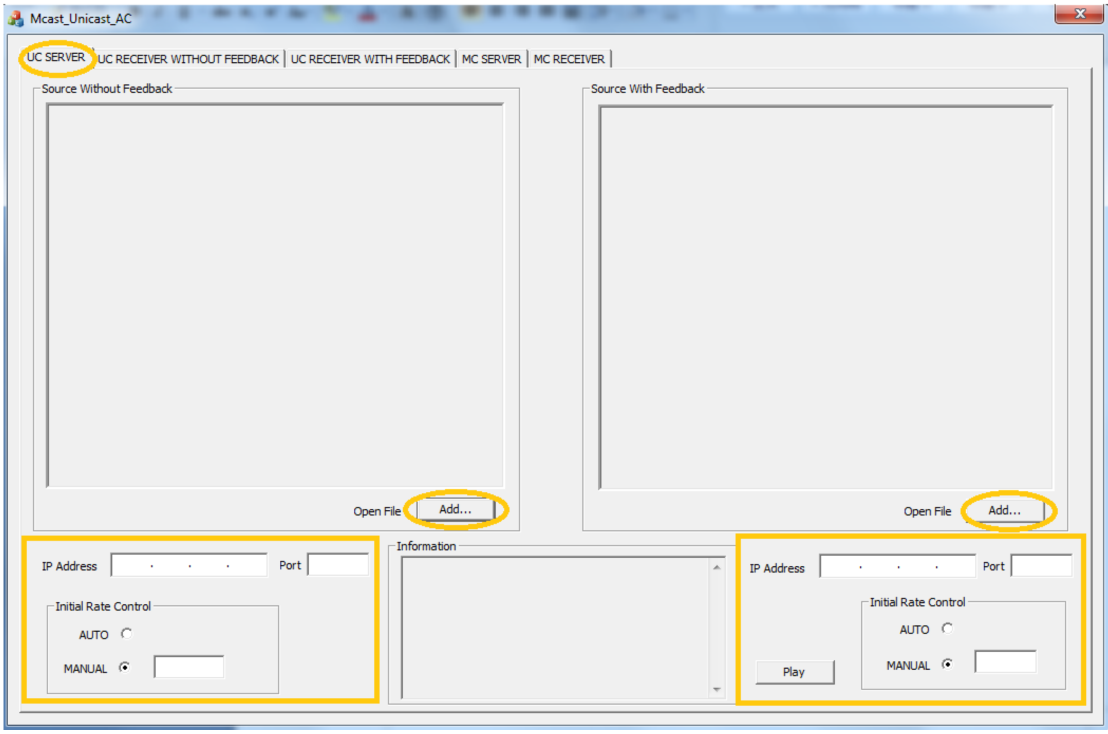

Figure 8. Interface of Unicast Server

Configure Unicast Receiver (UC RECEIVER WITHOUT FEEDBACK, UC RECEIVER WITH FEEDBACK)

On the receiver side, for example Workstation_2, choose “UC RECEIVER WITHOUT FEEDBACK” or “UC RECEIVER WITH FEEDBACK”, and configure IP address and port number of the server, click “Connect” to play the video streamed by the server.

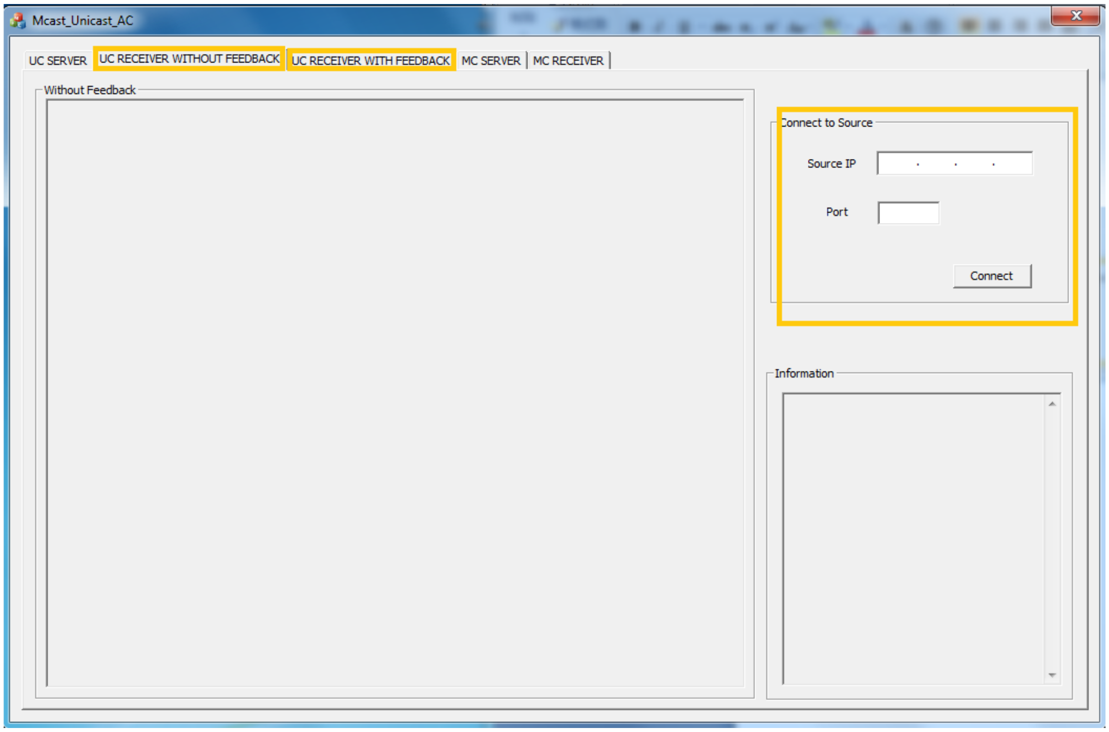

Figure 9. Interface of Unicast Receivers

Scenario 2: Play video stream using Multicast.

Configure Multicast Server (MC SERVER).

On the server side, for example Workstation_1, choose "MC SERVER" mode. The interface below will show up. Click "Add" button to choose video files that user want to play for both Group 1 and Group 2. Configure multicast IP address for group 1, for example 224.1.1.1, and port number, and user can also adjust the initial sending rate. Click "Create Group" button to start the server. Multicast IP address for Group 2 will automatically generated by the application, which is the next IP address to the Group 1's, for example 224.1.1.2.

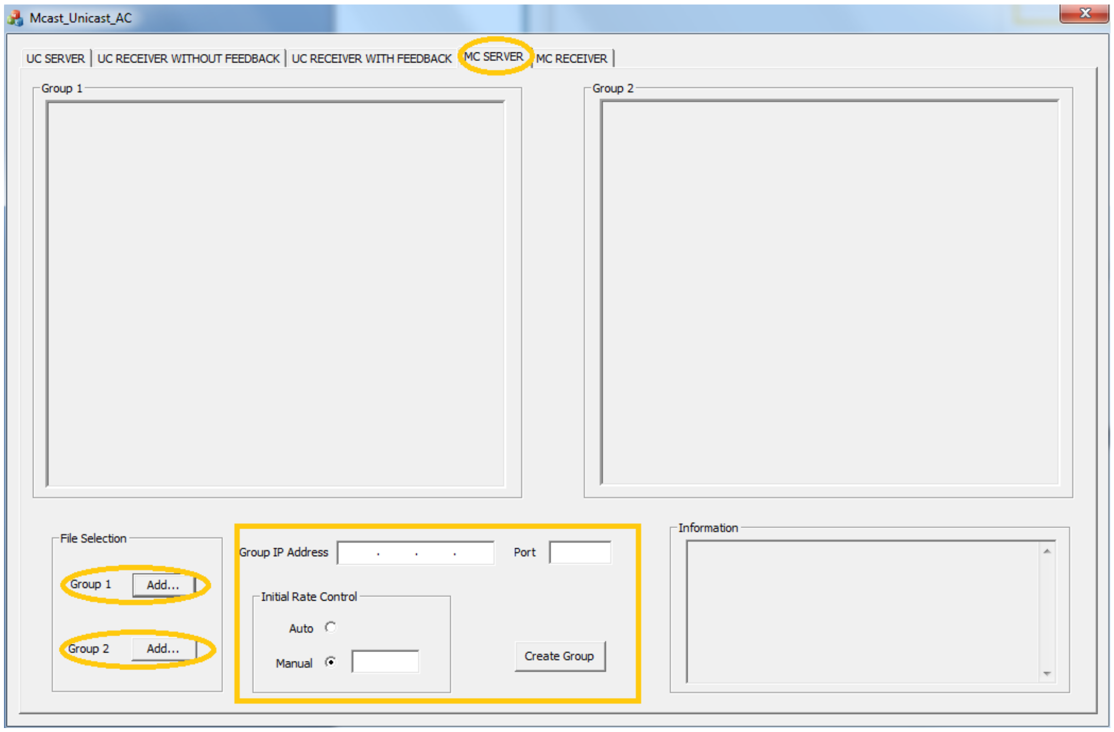

Figure 10. Interface of Multicast Server

Configure Multicast Receiver (MC RECEIVER).

On the receiver side, for example Workstation_2, choose “MC RECEIVER” mode, and configure multicast IP address and port number of Group 1 or Group 2. Click “Join Group” button to join certain multicast group. And click “switch” button to switch around these two groups.

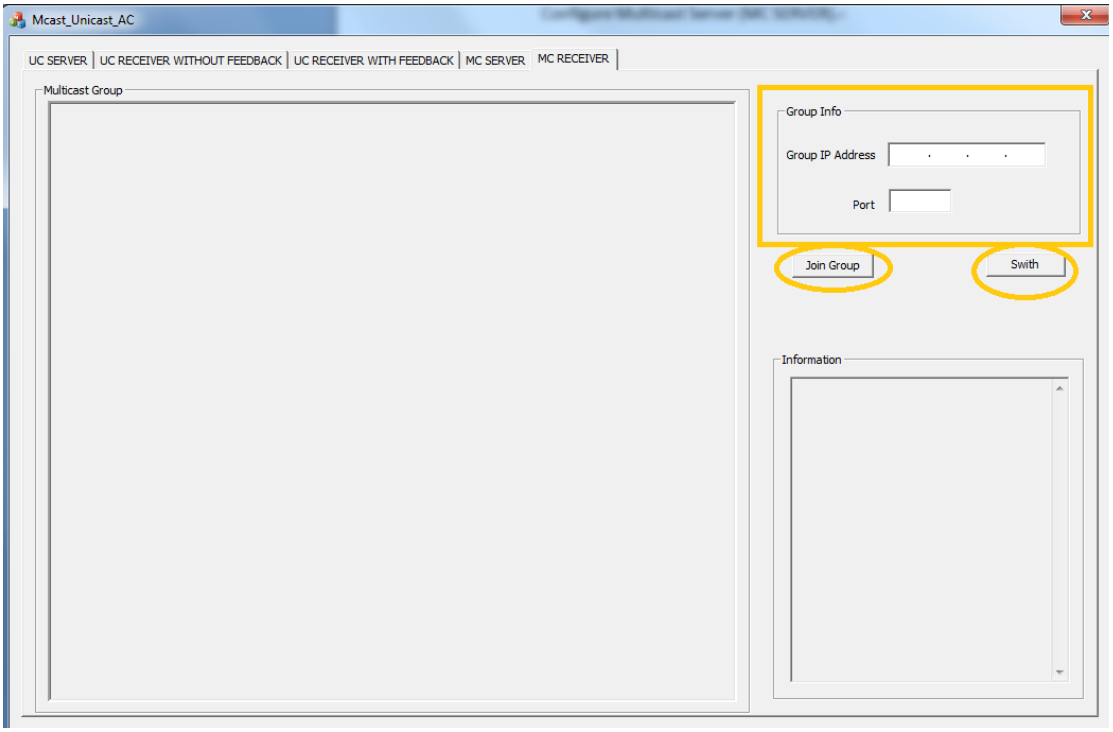

Figure 11. Interface of Multicast Receiver

3.3 Software Testing

In this section, we test our software under different network environment using build-in QoS measurement and external measurement tool. In the next section, we will raise some limitations found during test and discuss the solution to overcome these limitations.

3.3.1 Software Test Using Bulid-In QoS Measurement

We setup a complex wireless scenario using 18 mobile wireless routers, and these routers can move randomly. Another 4 fixed wireless routers connected with cross-traffic generator, which is composed by 4 virtual workstations. Two SITL ports are connected to two fixed wireless routers to allow actual laptops to run our software. Figure 13 and Figure 14 indicate the cross-traffic created by congestion generator.

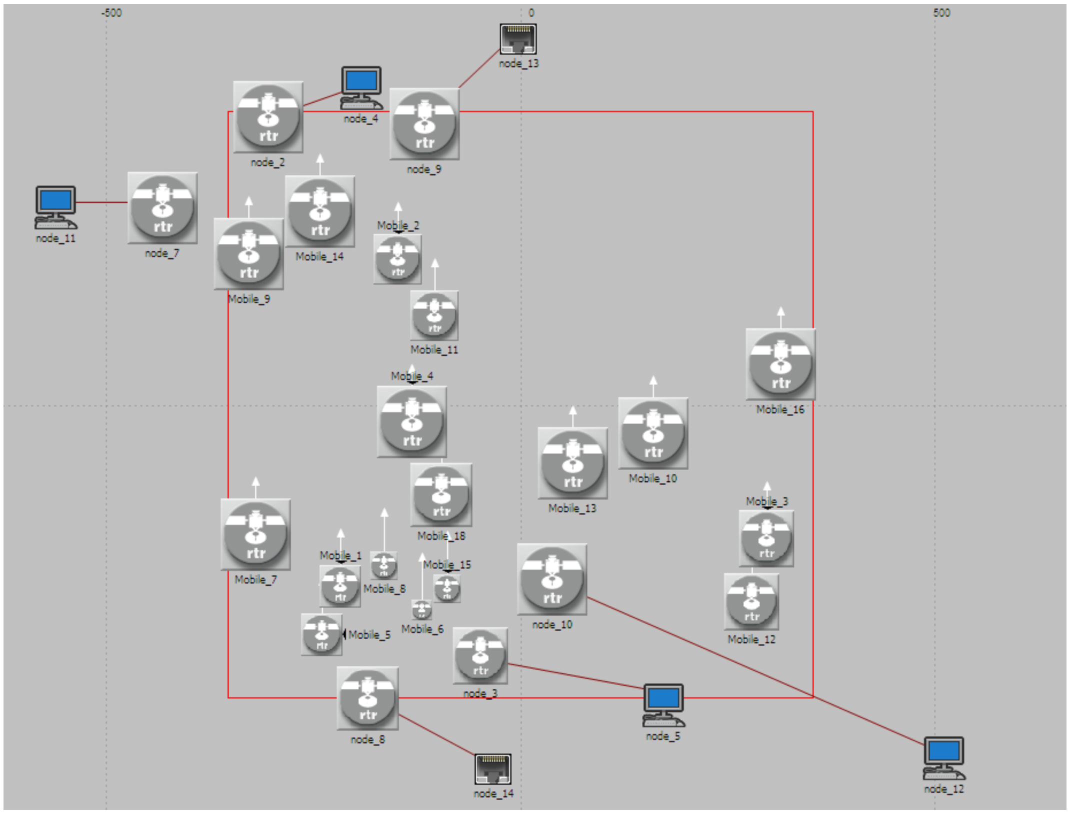

Figure 12. Topology to test software using build-in QoS measurement

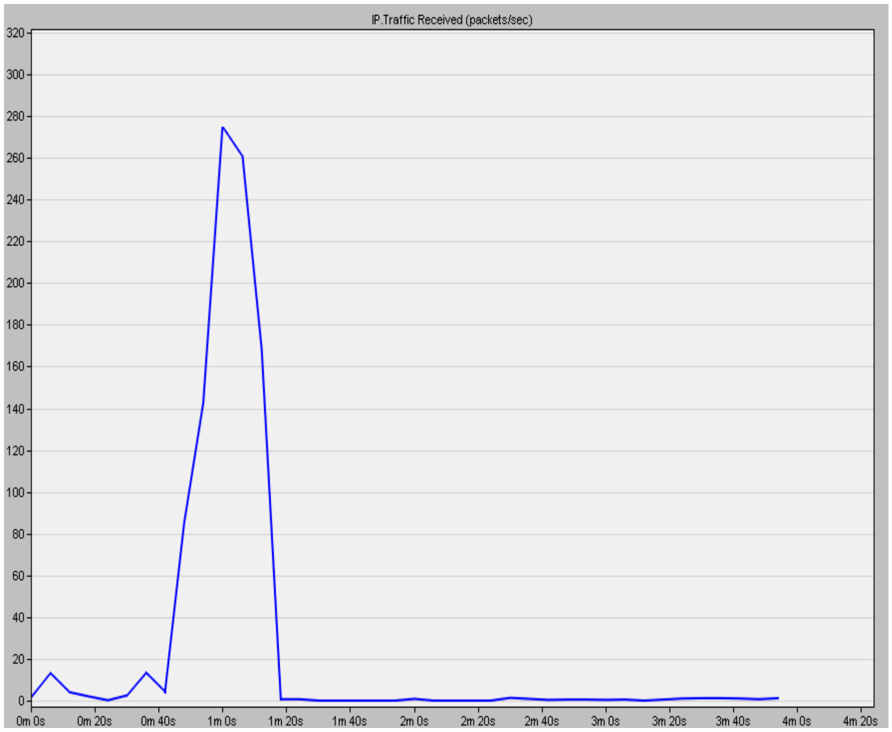

Figure 13. Cross-traffic between workstation 11 and workstation 12

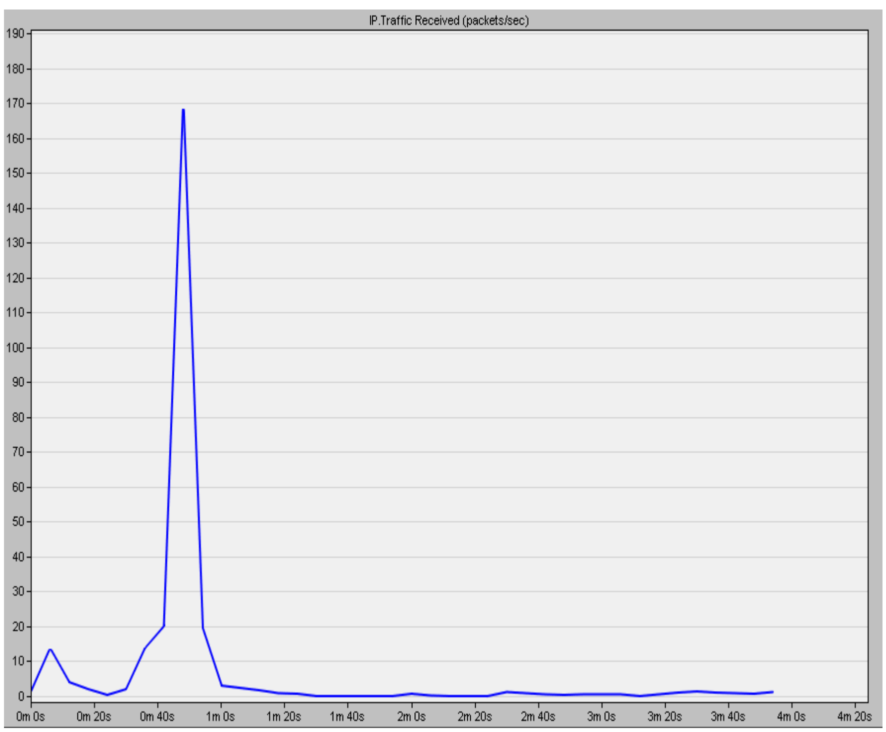

Figure 14. Cross-traffic between workstation 4 and workstation 5

In Figure 15, client side detected the network congestion, and notify video server to tune the stream sending rate. We can see from Figure 16, when network condition is good, video distribution system increased the stream rate and try to achieve a better video quality. When network congestion happen, our system can downgrade the video sending rate and release the network congestion.

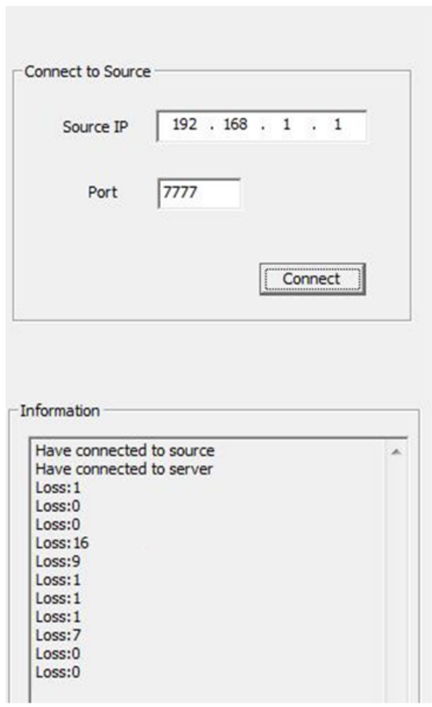

Figure 15. Client side network condition detection

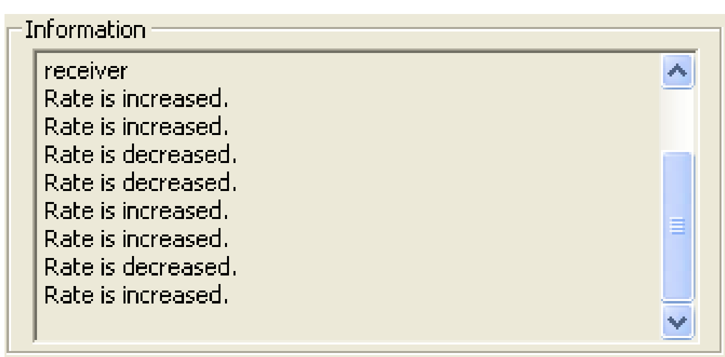

Figure 16. Server side tune stream sending rate

3.3.2 Software Test Using External Measurement Tool

We setup a simple wired scenario using three fixed routers. Our cross-traffic generator is composed by 2 virtual workstations. Two SITL ports are connected to two fixed routers to allow actual laptops to run our software. Since we will test our application in multicast mode, we create two streams with different video streaming rates.

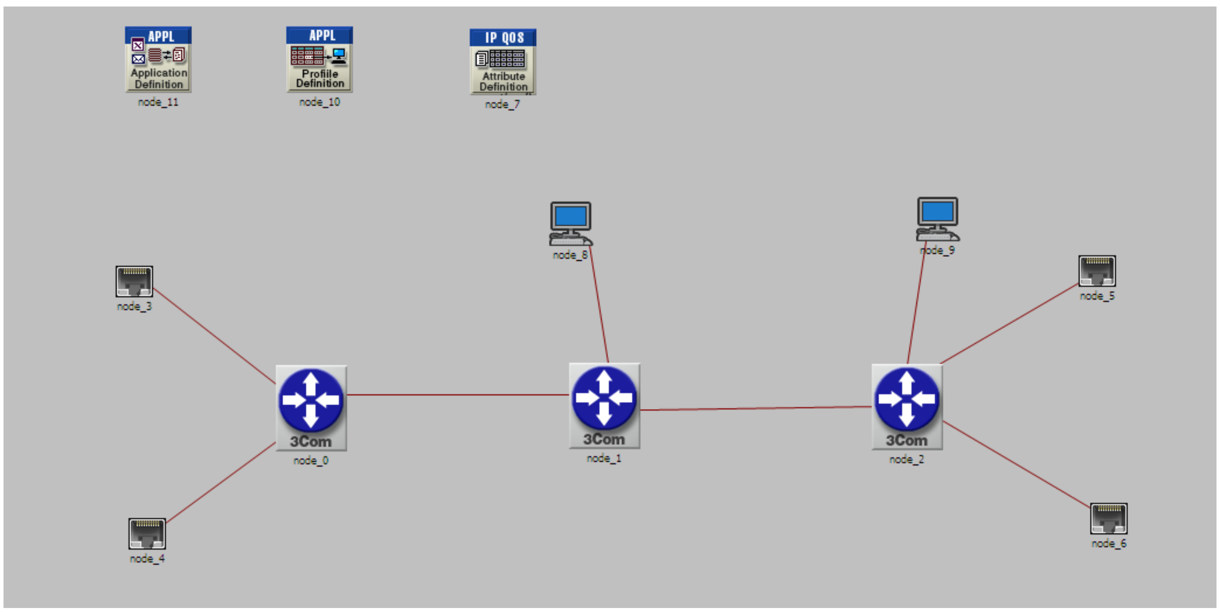

Figure 17. Topology to test software using external measurement tool

Figure 18 to Figure 20 show the external measurement tool, NRPE, detected network congestion level. External measurement tool at the user side periodically update the network congestion level with video software server side. Figure 21 shows the action made to switch multicast group in order to adapt the network condition.

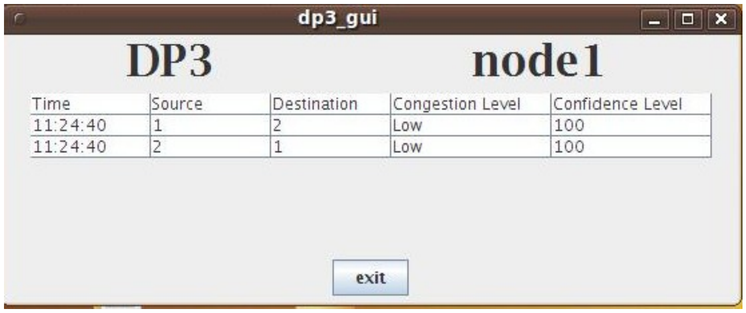

Figure 18. NRPE initial status

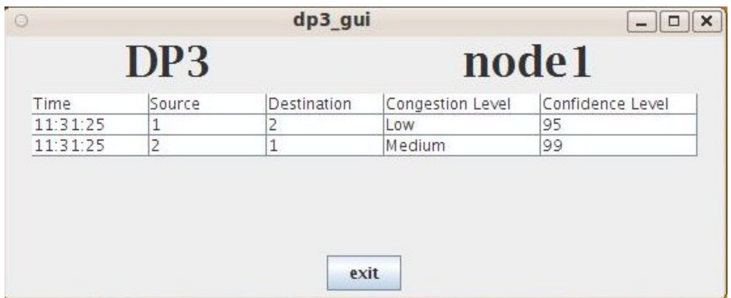

Figure 19. NRPE detected medium congestion

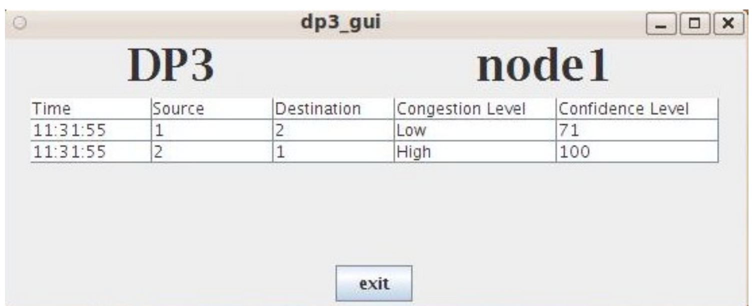

Figure 20. NRPE detected medium congestion

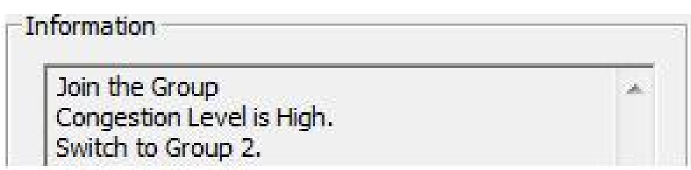

Figure 21. Multicast group switch to adapt network condition

3.4 Limitations of QoS Based Video Quality Control

Our software implementation well illustrates the basic idea of QoS only video quality control technique. During the period of video transmission, software measurement tool is gathering the network statistics using different tools, and then feedback to video service provider to trigger proper actions. Actions such as sending rate increasing, sending rate decreasing and group switching, can help video stream go through the network in real time without suffering from network congestion.

Although QoS based technique is widely considered as the ideal method to avoid sever quality loss of video transmission, we identify some scenarios that triggered action can hardly help relief network congestion or even make the network congestion getting worse. For instance, current video transmission service suffer from medium packet loss rate, however, client might consider the video quality as acceptable. This is because lost packets may belong to less important frames, or codec technique may recover the lost packets using its own algorithm. Traditional QoS based technique will trigger the action to lower the sending rate while it is not necessary at all. Another drawback is that QoS based technique cannot properly reach the needs of human perception when network condition allows to do so.

The nature of QoS only based video quality control technique makes it only sensitive to network condition not to human perception. However, video service provider do cares about user perception rather than network condition, so the introduction of QoE to video quality control technique becoming more and more desirable.

The introduction of QoE, along with QoS parameter, will enable the video quality indicator to trigger proper action, which takes both QoE and QoS into account. Once we choose proper QoE parameter, and correlate with QoS parameter, and design a new real time video quality indicator, video server can follow the indicator and make the right action.

The following chapters will introduce a novel QoE parameter, and then use this QoE parameter to correlate with chosen QoS parameter, so that the new indicator can overcome the drawback of QoS based video quality control technique.

4. Research on QoS-QoE Based Video Quality Control Technique

In this chapter, we start working on correlating QoE with real time video quality control technique. We design a new QoE parameter based on structural similarity index, and correlate with QoS parameter, packet loss rate, to indicate the networked video quality. However, due to the drawback of structural similarity index, this method can be only considered as an approximation. Even through, this research provides a direction that we can work on, and prove that introduction of QoE parameter can improve the performance of video quality control technique.

We will first provide a brief review of structural similarity index, and describe how we apply it to real time vide quality measurement. Simulation and emulation results at the end of the chapter show the performance of our approach. We also discuss limitations of this approach and find new direction to our research, which will be introduced in the following chapter.

4.1 Structural Similarity

Nowadays, IQA and VQA become focus on human vision perception rather than error visibility, because people realize that the efficiency of access to the information carried by image or video is definitely much more important than the sensitivity of error, which might be caused during transmission or decoding. In [45], the assumption that human visual perception is highly adapted for extracting structural information from a scene helps find a new direction to evaluate the IQA, referred to as Structural Similarity Index (SSIM). SSIM has been proved much closer to human vision perception and simpler than traditional error sensitive methods. However, SSIM in spatial domain has its own drawback, being over sensitive to translation, scaling, and rotation. A modified version of SSIM is proposed in [47], Complex Wavelet Structural Similarity (CW-

SSIM), and Complex Wavelet Transform (CWT) successfully overcome the drawback of the original method.

In [45], two images in spatial domain can be represented as $x=\{x_i \mid i=1,...M\}$ and $y=\{y_i \mid i=1,...M\}$, and SSIM between image x and image y is defined as

$$S(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)}, \qquad (1)$$

Where C_1 and C_2 are two small positive constants.

$$\mu_x=\frac{1}{N}\sum_{i=1}^{N}x_i\ , \qquad (2)$$

$$\mu_y=\frac{1}{N}\sum_{i=1}^{N}y_i\ , \qquad (3)$$

are used to construct the luminance comparison function.

$$\sigma_x=\left(\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i-\mu_x\right)^2\right)^{\frac{1}{2}}, \qquad (4)$$

$$\sigma_y=\left(\frac{1}{N-1}\sum_{i=1}^{N}\left(y_i-\mu_y\right)^2\right)^{\frac{1}{2}}, \qquad (5)$$

are used to construct the contrast comparison function.

$$\sigma_{xy}=\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i-\mu_x\right)\left(y_i-\mu_y\right), \qquad (6)$$

is used to construct the structure comparison.

4.2 Framework

It is clear that network layer packet loss can be correlated with the video quality damage. We concede that packet loss plays an important role in networked video quality, however, using only quality of service (QoS) parameter, such as packet loss, to define the current transmitted video quality is definitely not the best approach in measuring the multimedia quality. One of the major concerns of this issue is how compression and decompression works. Many modern decompression techniques have their own way to deal with limited number of lost packets, and these schemes have been proved effective. That is to say, sometimes, even real time stream transmission experience certain percentage of packet loss, codec can still retrieve the frame sent from server side using its complicated and successful recovery algorithm. This gives an exception that lost packets in the network layer is not proportional to the multimedia quality.

Therefore, our approach is based on using SSIM and number of packet loss information together to determine the picture quality, which then can be used in the application layer to trigger the proper action.

As we all know, SSIM is designed for static image and needs full reference, and even its application to video quality needs original video as a reference to compare with. While original video is impossible to obtain due to the limited bandwidth, our designed algorithm should avoid using original data. Fortunately, the fact that back to back frames from a video content are still very similar to each other provides a key to this problem. The similarity between frames varies because of the quickness of motion, so using a constant threshold is improper, and we use a running weighted average of SSIM calculation for back to back frames. Our simulation shows that SSIM values and change in SSIM values both show correlation with the number of packets lost (thus, number of frames lost) in the network layer. So, using these two parameters allows us to automatically detect quality loss in the video content without feedback from a human subject.

Our approach combines structural similarity, which is information gathered from the application layer, and number of packets lost in the network layer. To detect loss of quality we first calculate structural similarity index for every 2 frames (back-to-back). This calculation is limited to a pre-determined period of time over which the standard deviation of similarity indexes is observed. The standard deviation for a good quality video shows very little scattering and almost forms a line parallel to x axis. When the quality is bad due to traffic, the standard deviation changes rapidly and is highly scattered. Thus, using the standard deviation of similarity index values is a solid way of measuring perceived video quality in multimedia networks.

4.3 Simulation

4.3.1 SSI Measurement under light and heavy traffic scenarios

The testbed is based on SITL (System In The Loop) in OPNET Modeler, which provides an interface for connecting live network hardware or software applications to an OPNET discrete event simulation. Figure 22 shows the configuration of the testbed. Two SITL ports are connected with two real workstations, and inside the OPNET module, three virtual workstations work as traffic generators.

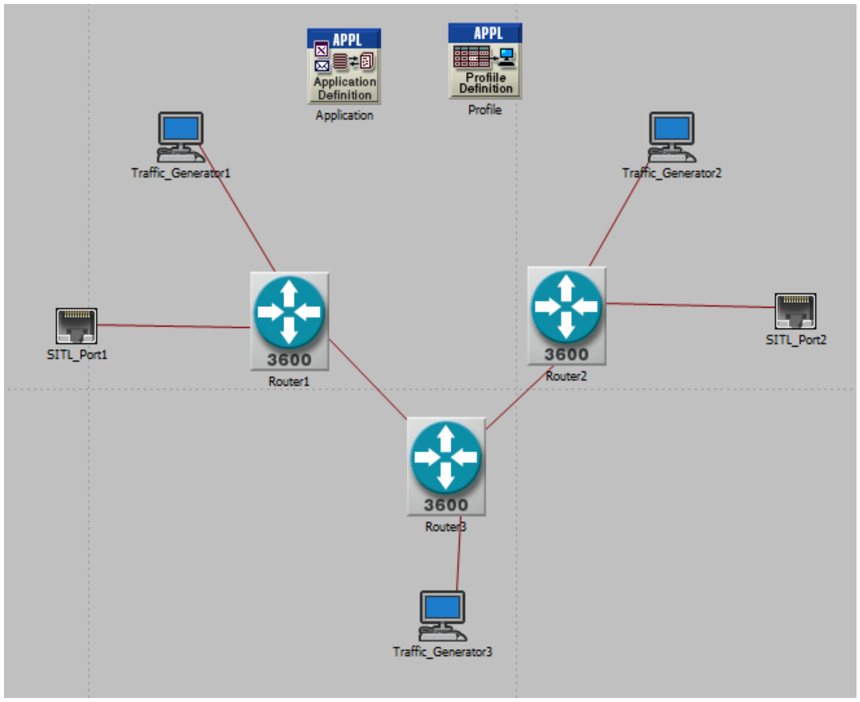

Figure 22. Testbed Configuration

Figure 23 shows the standard deviation of the similarity index (SSI) for video streamed through light background traffic (marked as solid line) and heavy background traffic (marked as dashed line). A smooth video stream has a standard deviation of the similarity index between 0.1566 and 0.1785, and its average is 0.1682. Video streamed through heavy background traffic has a standard deviation of the similarity index between 0.0975 and 0.2229, and its average is 0.1656.

The similarity index of the smooth video spread out almost uniformly from each other, while the same video streamed through heavy background traffic has the similarity index spread out unevenly from each other, even though both cases have the average of the standard deviation around 0.1669.

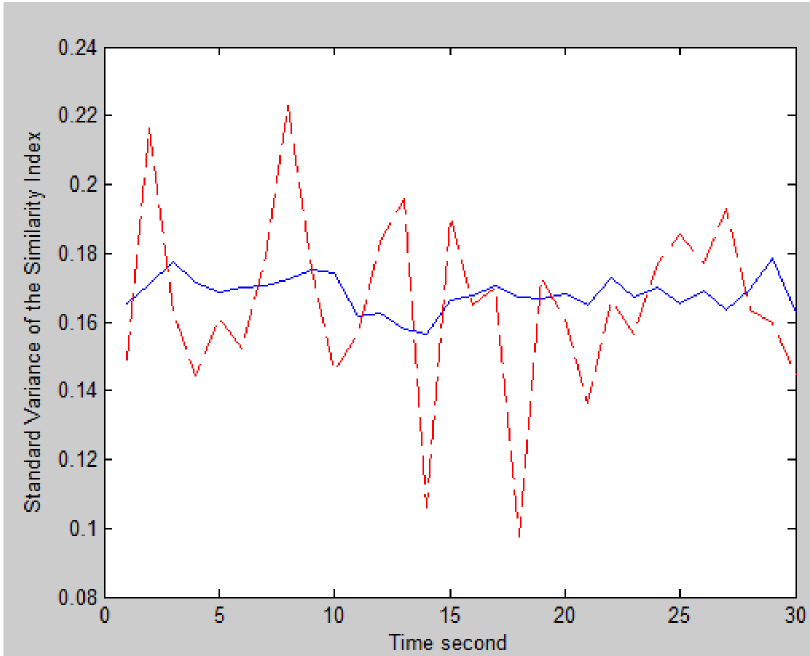

Figure 23. Standard Deviation of Similarity Index

For the scenario of the heavy traffic, the human perception at certain time with standard deviation greater than 0.18 is discontinuity, while human perception at certain time with standard deviation less than 0.15 results in a freeze frame, besides, $\Delta\sigma$, which is $\sigma - \overline{\sigma}$, describes the degree of the influence act on the quality of the video. Human perception is also recognized as QoE, which is quality of user experience. Based on the standard deviation of the similarity index, a quantitative measurement of QoE is proposed.

4.3.2 QoE Measurements

The proposed measurement of QoE below will include elements from both network layer and human perception.

Packet loss ratio, p, is ration of number of lost packets to total number of sent packets within a time window, and is chosen as the measurement element from the network layer.

Figure 24 is the packet loss over times when video stream experience light background traffic. In this scenario, light background traffic only caused very few packet losses, and the video quality is fine.

Figure 25 is the packet loss over times when video stream experience heavy background traffic. In this scenario, quality of video stream is greatly influenced by the background traffic. Packet loss ratio is increased.

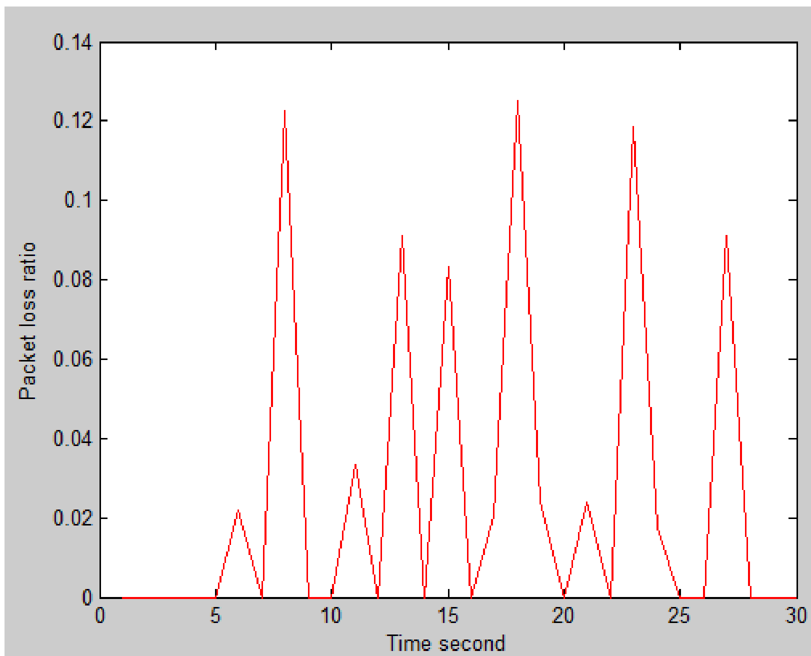

Figure 24. Packet loss ratio when video streamed through light background traffic

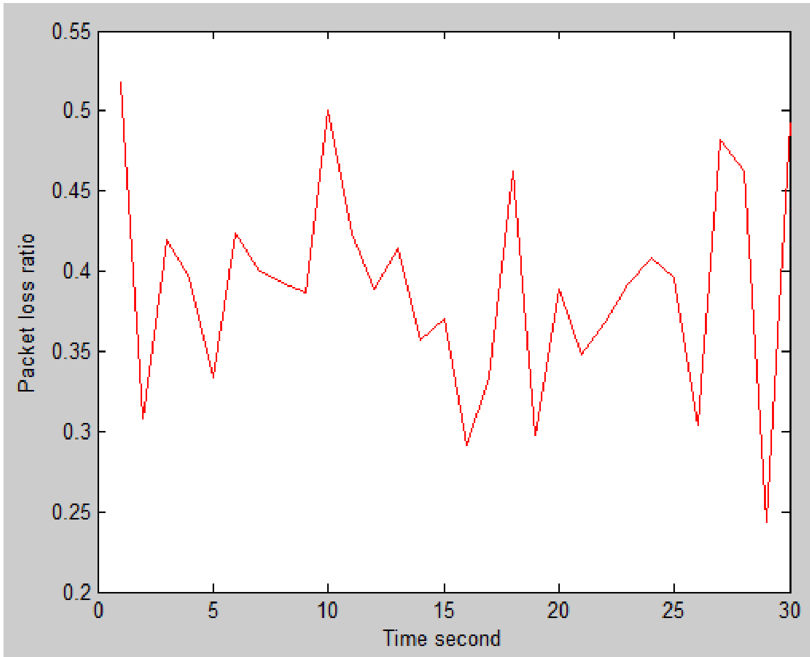

Figure 25. Packet loss ratio when video streamed through heavy background traffic

$\Delta\sigma$, difference between standard deviation of SSI and the average of the standard deviation of SSI, are chosen as the quantitative measurement of human perception.

$\Delta\sigma \times p$ is designed as the parameter for the QoE quantitive measurement, which is shown in Figure 26. For the light background traffic scenario, although there are some packets losses, the quality of the video is still good, and $\Delta\sigma$ is small. So $\Delta\sigma \times p$ is very small and near to 0 (marked as solid line). For heavy background traffic scenario, video quality is not as good as the first scenario, some frames are lost and freeze frame appears sporadically, and packet loss ratio becomes large and $\Delta\sigma$ also becomes large. Absolute value of $\Delta\sigma \times p$ varied from 0 to 0.03 (marked as dashed line).

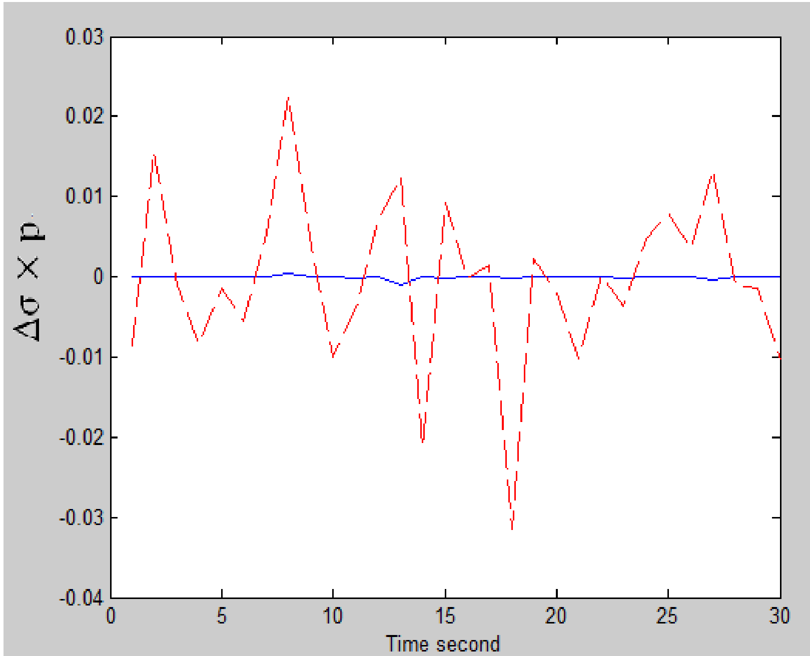

Figure 26. Δσ×p

4.3.3 Interpretation of new quantitative measurement.

When Δσ × p is very close to 0, either Δσ or p is close to 0, and that means the video quality is good. If Δσ is close to 0, quality of the video is considered fine, and packet losses are ignored. If p is close to 0, quality of the network is considered fine, and video quality can not be improved at network layer, or the original quality of the video is poor.

When the absolute value Δσ × p is large, both Δσ and p are away from 0, and that means bad video quality is caused by the network. Now, certain reaction on the network can improve the quality of user experience.

According to the human perception, when $0 < \Delta\sigma \times p < 0.06$, the video quality is under the tolerance, QoE can be considered good.

4.4 Drawbacks and Research Direction

The application of structural similarity index to back-to-back frames brings some drawbacks. Since structural similarity index is over sensitive to translation, scaling and rotation, when applied to back-to-back frames with acceptable quality, the continuity is not always stable. Our designed parameter sometimes cannot accurately reflect the quality change. Also, the designed indicator is still not clear enough to distinguish different combinations of QoS and QoE of video service.

Due to these limitations, we design another algorithm for QoE of a video, and redesign our real time video quality indicator based on the new algorithm of QoE. Chapter 5 will discuss our design in detail.

5. PROPOSED REAL-TIME QoS-QoE BASED VIDEO QUALITY CONTROL TECHNIQUE

In this chapter, we overcome the inaccuracy of method mentioned in previous chapter by inventing a new QoE parameter based on CW-SSIM. The redesigned indicator, which based on our new QoE parameter, is proved to be able to distinguish complex combination of network condition and human perception.

5.1 Complex Wavelet Structural Similarity (CW-SSIM)

Compared with traditional IQA and VQA method, spatial domain SSIM has the advantage of analyzing the structural information, making it more sensitive to human vision system (HVS) than the error itself, while we cannot overlook the drawback of SSIM. SSIM is highly sensitive to translation, scaling, and rotate which normally occur during coding, decoding and transmission. However in certain scenario these changes to image will not influence the structural information from the image. For example, comparing to the reference image, current image is shifted to the right by two pixels, SSIM will rate the image as an image with poor quality, but human vision system still recognize this image as a acceptable quality one, since human vision perception ranks it as acceptable. Similar argument can be discussed with regard to scaling and rotation.

In [47], a modified version of SSIM, which is SSIM in complex wavelet transform domain, is mentioned. If two images can be represented as two sets of coefficients extracted at the same spatial location in complex wavelet transform domain, $c_x = \{c_{x,i} \mid i = 1,...N\}$ and $c_y = \{c_{y,i} \mid i = 1,...N\}$, SSIM now can be written in this domain as:

$$\bar{S}(c_x, c_y) = \frac{2\left|\sum_{i=1}^{N} c_{x,i} c^*_{y,i}\right| + K}{\sum_{i=1}^{N} \left|c_{x,i}\right|^2 + \sum_{i=1}^{N} \left|c_{y,i}\right|^2 + K}, \quad (7)$$

Where $\bar{S}(c_x, c_y) \in (0,1)$, and greater value means image y has closer human vision perception to reference image x.

As discussed in [47], translation, scaling and rotation factor in 2-D spatial domain can be defined as:

$$\begin{pmatrix} \Delta t_1 \\ \Delta t_2 \end{pmatrix}, \quad (8)$$

$$\begin{pmatrix} \cos\Delta\theta & -\sin\Delta\theta \\ \sin\Delta\theta & \cos\Delta\theta \end{pmatrix}, \quad (9)$$

$$\begin{pmatrix} 1+\Delta s_1 & 0 \\ 0 & 1+\Delta s_2 \end{pmatrix}, \quad (10)$$

Where $\Delta\theta$ is small, so $\cos\Delta\theta \approx 1$ and $\sin\Delta\theta \approx \Delta\theta$, and therefore

$$\begin{aligned} c_y\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} &= c_x(\begin{pmatrix} \cos\Delta\theta & -\sin\Delta\theta \\ \sin\Delta\theta & \cos\Delta\theta \end{pmatrix}\begin{pmatrix} 1+\Delta s_1 & 0 \\ 0 & 1+\Delta s_2 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \\ &+\begin{pmatrix} \Delta t_1 \\ \Delta t_2 \end{pmatrix}) \approx c_x\begin{pmatrix} u_1 + (u_1\Delta s_1 - u_2\Delta\theta + \Delta t_1 - u_2\Delta s_2\Delta\theta) \\ u_2 + (u_2\Delta s_2 + u_1\Delta\theta + \Delta t_2 + u_1\Delta s_1\Delta\theta) \end{pmatrix} \\ &= c_x\begin{pmatrix} u_1 + \Delta u_1 \\ u_2 + \Delta u_2 \end{pmatrix} \end{aligned}, \quad (11)$$

We can consider this is only the linear phase shift in Fourier domain, and $\bar{S}(c_x, c_y) \approx 1$.

So translation, scaling, and rotation are proved to be not sensitive to CW-SSIM, meanwhile, CW-SSIM can still perform well as FR-IQA, and it still can provide structural information of image rather than the error itself. This is also the reason that we choose CW-SSIM, not SSIM, to evaluate the quality of video streaming. Only change of structural information, not regular motion of video stream, can sharply change CW-SSIM of back-to-back frames, while both changes of structural information and regular motion can influence SSIM of back-to-back frames.

5.2 Self-Reference Complex Wavelet Video Structural Similarity

Many factors can affect and/or impair the video quality. Due to current motion compensation blocked-based coding technique, networked video always suffers from blockiness, blurriness, color bleeding, ringing, false edges, jagged motion [62]. To detect these video quality issues, many techniques are introduced, but most of these techniques are FR or RR methods, which need full or portion of the original video stream. Obtaining the original transmitted video stream at the receiving end for comparison purposes with the actual received stream is difficult due to the highly needed bandwidth. At the same time, many real-time video communication applications, such as video conference, battle field real-time video communications, can never provide the original video data. In other words, it's essential to develop a technique that is based on no reference frame approach.

Since there are no reference video data techniques available, current NR methods are mainly based on coding rate and QoS parameter, however both of them can not reflect the video quality directly to human vision perception.

5.2.1 Proposed Approach to Determine Video Quality

At the receiving end, video streaming can be recognized as a set of frames. Continuous and clear video requires no random change between back-to-back frames. The differences between back-to-back frames are translation, scaling, and rotation. Scene shifting can be considered as translation of previous frame, and zoom-in or zoom-out can be treated as the scaling of previous frame, while scene rotation can be recognized as the rotation of the previous frame. For the special need of real-time video, especially the efficiency of information transmission, we only need to distinguish severe distortion, which leads to bad human perception and will influence the client to access the information carried by video. While slight distortion, which can be detected and probably corrected by error sensitive FR method, is not that crucial, since human perception in many cases may still ranks it as acceptable, and most of the information will be transmitted efficiently. We should not waste limited computational and network resource to detect and attempt to correct these distortions.

Back-to-back frames from video with slow and regular motion have high CW-SSIM, while those frames from video with fast motion tend to have relative low CW-SSIM. This can be explained since slow motion means slight changes between frames, and fast motion means large changes. However, for any continuous video (no sudden scene switch), the set of CW-SSIM for all back-to-back frames should be continuous, since the motion is continuous, whether it is slow motion, fast motion or mixture of these two. If the set of CW-SSIM is not continuous, or in other words, if the discrete degree of the set of CW-SSIM is large, that means some substantial unpredicted changes being introduced to the video, such as blockiness, blurriness, or false edges, In this case, we can conclude that a severe distortion has occurred. If the set of CW-SSIM is continuous, or with very small discrete degree, we can consider the video has good human perception, even there might be slight distortion.

5.2.2 New Algorithm

We assume real-time video can be represented as a set of frames, $v = \{v_i \mid i = 1,...M\}$, v_n and v_{n+1} are back-to-back frames. CW-SSIM between v_n and v_{n+1} by using equation (7) is $\overline{S}(v_n, v_{n+1})$. We set up a slide window, and the width of the window is M frames, then we describe the discrete degree by standard deviation:

$$SRCW-VSSIM = \sqrt{\frac{1}{M-1}\sum_{n=1}^{M-1}\left(\overline{S}(v_n, v_{n+1}) - \frac{1}{M-1}\sum_{n=1}^{M-1}\overline{S}(v_n, v_{n+1})\right)^2} \quad , (12)$$

So far, we can make the expectation of SRCW-VSSIM:

when real-time video quality is good, and we experience a fluent and clear video, SRCW-VSSIM is relatively small and close to 0, otherwise, video with poor human vision perception will lead SRCW-VSSIM relatively large. The simulation results are presented in chapter 6.

5.3 QoS-QoE Based Video Quality Control Approach

Most of the video quality control approaches, both formula-based and measurement-based, try to adjust the source sending rate according to quality of service (QoS) rather than the quality of experience (QoE). While QoS can provide a lot of useful information of the network, and sometimes can also tell the quality of the real-time video service, but QoS can hardly tell the exact QoE. For example, although we experience packet loss, less important frame lost won't affect the overall quality of the video, so it is unnecessary to tune the source sending rate down dramatically. Besides, even when network condition is good, proper action can hardly be decided without the information from QoE side.

However, using QoE alone to tune the video service is also unacceptable. Since we try to control the video quality by adapting the source sending rate, we have to pay attention to the network condition. If the original video quality is poor from the source side, QoE parameter will show that tuning is needed to be done, while the truth is that tuning is useless.

See from the two examples above, parameter generate from single layer cannot determine when and how to tune the network to meet the needs of user's satisfaction. So here we propose a new indicator including both QoS parameter, which is packet loss, and QoE parameter, which is SRCW-VSSIM, to trigger the proper action.

We plan to design an end to end cross-layer video quality control system as shown in Figure 27 and Figure 28. Video service client collects statistics of packet loss rate and SRCW-VSSIM as input to our QoS-QoE based video quality control indicator (QQVQCI). We define the packet loss rate as the ratio between the number of the lost packets and the number of transported packets during each interval.

$$pktlossrate = \frac{P_{Loss}}{P_{Loss} + P_{received}} \qquad (13)$$

P_{Loss} is the number of packet loss, and $P_{received}$ is the number of received packets. Indicator is sent to video service source, and video service source determines whether tuning is needed.

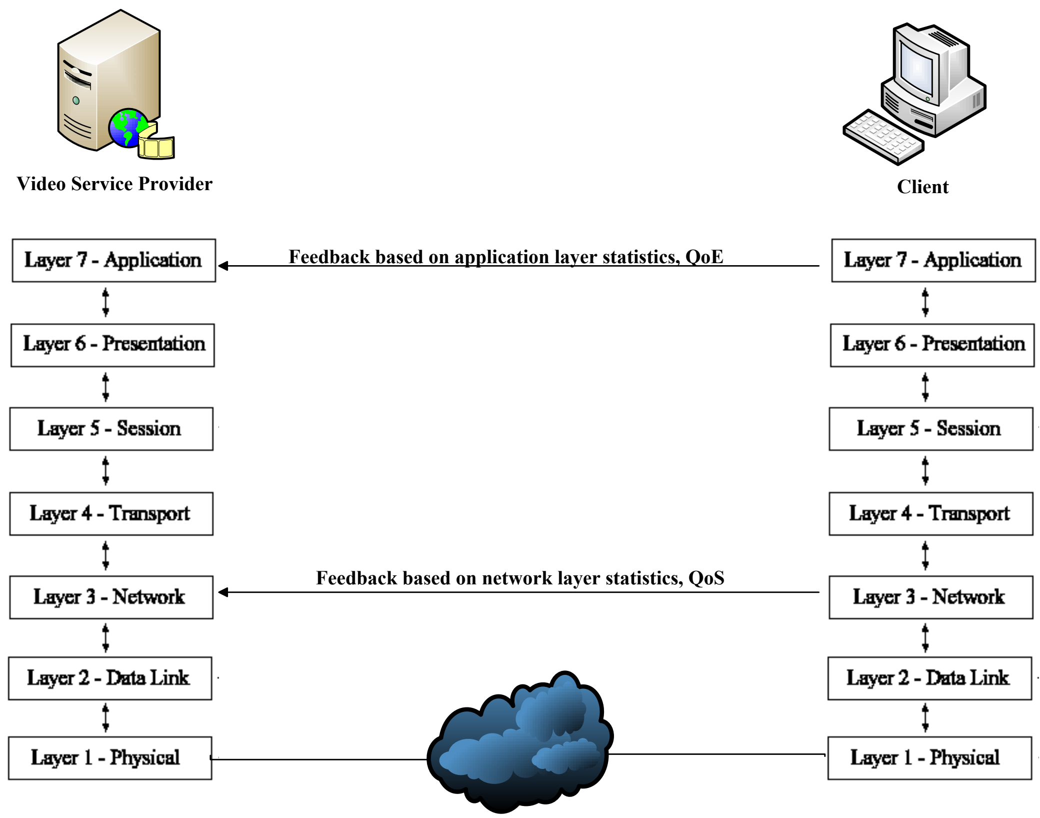

Figure 27. Proposed QoS-QoE based video quality control technique

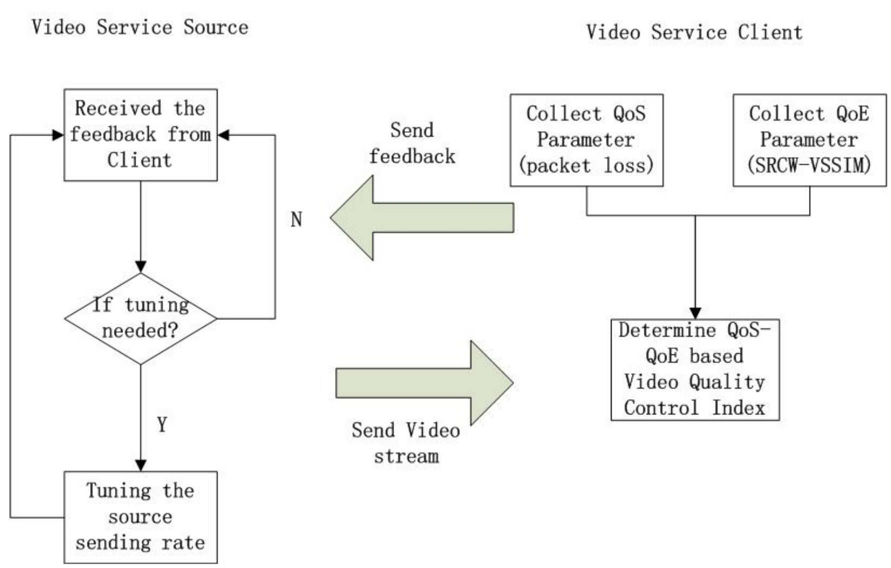

Figure 28. End to end video quality control system

Our QoS-QoE based video quality control indicator (QQVQCI) is defined as:

$$\mathrm{QQVQCI} = (pktlossrate + \delta) \times (SRCW\_VSSIM - \alpha) \quad (14)$$

where δ is a very small positive constant, and α is the threshold for SRCW-VSSIM.

	Good QoE	**Bad QoE**
Good QoS	Scenario 1 No Action is needed	Scenario 3 Action (sender increases the sending rate ect.)
Bad QoS	Scenario 2 No Action is needed	Scenario 4 Action (Network tuning, e.g. relief congestion, reduce source.)

Table 3. Scenarios for different QoS and QoE conditions

To demonstrate our cross-layer design video quality control system, we discuss the following scenarios as stated in Table 3:

1) When both network condition and human perception are good, packet loss rate is small and SRCW-VSSIM is less than the threshold α, so we expect our indicator be negative, and its absolute value is small . No further action should be done.

2) Another possible scenario is as follows: network condition may experience slight congestion, so some packets may be lost. Traditional QoS based video quality control approach considers this as the reason to lower the sending rate. However, this action may n't be necessary, since limited number of lost packets may not degrade the human perception of the video, for example, key frames may n't experience packet losses, and in this case human perception may still be acceptable. Instead of enhancing the quality of experience, the action hurts the human perception of the video data. QQVQCI successfully avoids this exception due to the introduction

of QoE as a factor of our indicator. Our indicator is a negative value since QoE index is less than, and video sender should not change the sending rate.

3) Sometimes, low sending rate leads to poor video quality, however, the network condition is good enough to allow a higher sending rate. Under this situation, our indicator is a small positive number. The reason is simple: one factor of our indicator, SRCW-VSSIM, is much greater than threshold α, but another factor, packet loss rate, stays low. Once the video service provider receives such indicator, video service provider should increase the sending rate to enhance the user's quality of experience.

4) If both network condition and human perception are bad, packet loss rate and SRCW-VSSIM are large at the same time. QQVQCI turns out to be a large value. This indicates the video source or network should take action to relieve congestion and achieve a better perceived video quality.

Our contribution is that we can accurately distinguish Scenario 2 and Scenario 3 from the traditional Scenario 1 and Scenario 4, and make proper actions based on statistics from both layers to tune network or source server to reach the user's needs correctly.

Generally speaking, the introduction of QoE parameter along with QoS parameter helps video sending rate tuning adapt not only network condition but also human perception. The proper actions according to our new indicator, QQVQCI, are as followed:

- When QQVQCI is a large positive number, source sending rate should be reduced to adapt the network condition.

- When QQVQCI is negative number, no further action should be done since human perception is within the tolerated range.

- When QQVQCI is a small positive number, source sending rate should be increased to achieve a better human perception, because current network condition allows to do so.

Proper thresholds of QQVQCI to trigger reasonable action are decided in the experiment chapter.

The simulation results for this chapter are presented in chapter 6.

5.4 Support Vector Regression Method

As a special implementation of the support vector machine for prediction, Support Vector Regression (SVR) method trains a regression function $f(x)$ that maps input features to output $\hat{y}_i$ to limit error from the obtained targets y_i within ε and as flat as possible. Regression function $f(x)$ is then applied to predict the future y_{i+1}. Compared with artificial neural network, a very popular forecasting method in recent years, SVR shows two advantages. First of all, SVR predicts based on the structural risk minimization principle, and the risk is measured according to certain loss function. Secondly, SVR maps low-dimensional input space to linear functions in high-dimensional feature space, and estimates the regression accordingly.

Consider the training data set $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_l, y_l)\}$ where $x_i \in R^n$, $y_i \in R$, n is the dimension of the input space, and l is the total number of training dataset. Linear function below is used to make regression in the feature space.

$$f(x) = \langle w, \varphi(x) \rangle + b \qquad (15)$$

Where $w \in R^n$ is the weight vector, $b \in R$ is the bias. $\varphi(x)$ is the function that maps the input feature from nonlinear low dimensional feature space to the high dimensional feature space.

Equation (15) can be used to estimate the optimal values of weight w and bias b only when the regression risk function is minimized:

$$min \left\{\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l} L(y_i, f(x_i))\right\} \qquad (16)$$

Where

$$L(y_i, f(x_i)) = \begin{cases} |y_i - f(x_i)| - \varepsilon, & if \ \ |y_i - f(x_i)| \geq \varepsilon \\ 0, & otherwise \end{cases}$$

Subject to

$$\begin{cases} y_i - \langle w, \varphi(x_i)\rangle - b \leq \varepsilon + \xi_i \\ \langle w, \varphi(x_i)\rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

Then, use the Lagrange multipliers to obtain the dual Lagrange form:

$$max \left\{ -\frac{1}{2}\sum_{i,j=1}^{l} (\alpha_i + \alpha_i^*)\left(\alpha_j - \alpha_j^*\right)\langle\varphi(x_i), \varphi(x_j)\rangle - \varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*)\right\} \qquad (17)$$

Subject to

$$\begin{cases} \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \end{cases}$$

Where α_i, α_i^* are the Lagrange multipliers.

The support vector regression function can be written as follows:

$$f(x) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) \langle \varphi(x_i), \varphi(x) \rangle + b \qquad (18)$$

Define the kernel function

$$k(x_i, x_j) = e^{(-\frac{(x_i - x_j)^2}{2\sigma^2})}$$

Finally, support vector regression function can be transformed:

$$f(x) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) k(x_i, x_j) + b \qquad (19)$$

To make SVR predict accurately, proper parameters should be chosen. Since our work focus on demonstrating the idea of application of SVR for real time video quality control, our parameters are optimized by libsvm [68].

6. SIMULATION RESULTS

In this chapter, section 6.1 to section 6.4 prove that our new QoE index, SRCW-VSSIM, is a sensitive real-time video quality measurement. Section 6.5 demonstrates that QQVQCI can be used to trigger proper actions in order to satisfy the needs for both network and human perception. Section 6.6 shows how SVR works with our QQVQCI parameter to predict real time video quality.

6.1 Implementation of Experiment

In order to accelerate the computational speed, we take every other frame as back-to-back frames, and use the same implementation of CW-SSIM in [63] to calculate the CW-SSIM between back-to-back frames. We choose window size as 10 frames, so the first ten frames are used to initialize our index. We will show the CW-SSIM, SRCW-VSSIM, and PSNR for each video sample. PSNR is calculated by MSU Video Quality Measurement Tool, [64].

For video with slight distortion, we use LIVE Video Quality Database, [65] [66] to show our results. Meanwhile, we use Sirannon, [67] to simulate heavy packet loss to video stream during transmission, and the sample video is provided by [67].

6.2 Experiment: Video with slight distortion

As shown in Figure 29 and Figure 30, original video without distortion has continuous CW-SSIM for back-to-back frames, and our SRCW-VSSIM stays at a very low level.

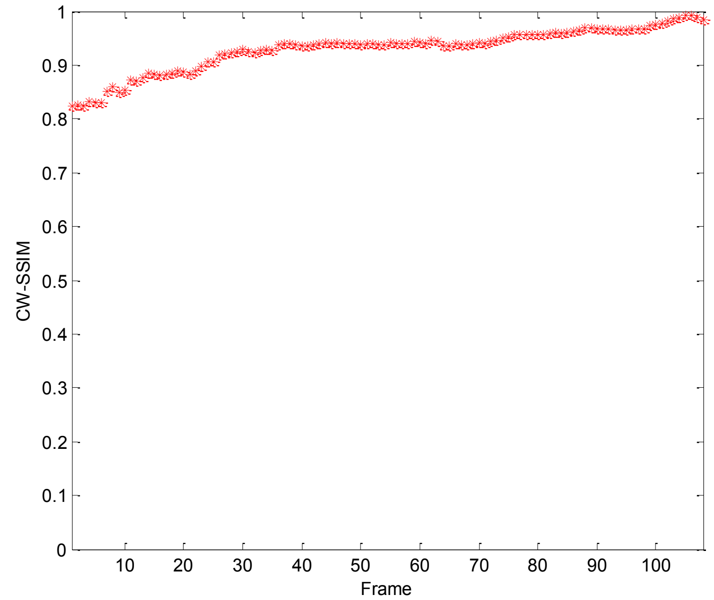

Figure 29. CW-SSIM of back-to-back frames for video sample 1 without distortion

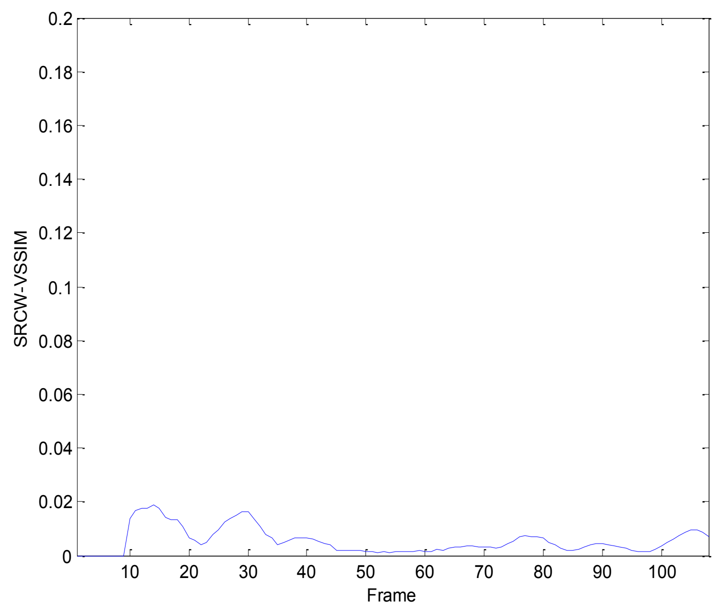

Figure 30. SRCW-VSSIM of video sample 1 without distortion

Figure 31 and Figure 32 show the CW-SSIM and SRCW-VSSIM for video with slight distortion. We can see some discontinuity points which represent worse quality video segment. Figure 33 shows the original frame 48 of video sample 1 while Figure 34 shows the one with slight distortion, and this matches the discrete point of Figure 31. SRCW-VSSIM is relatively higher when compared with the previous data. However, SRCW-VSSIM still stays at a very low level, less than 0.05. Human perception is still good even with the existence of some slight distortion. PSNR in Figure 32 can also tell that overall quality of the distorted video is accepted, and PSNR is generally above 25 dB.

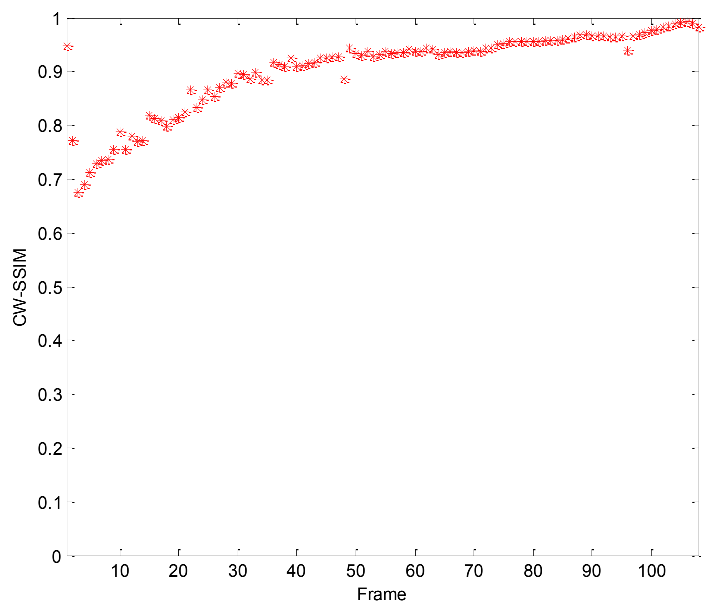

Figure 31. CW-SSIM of back-to-back frames for video sample 1 with slight distortion

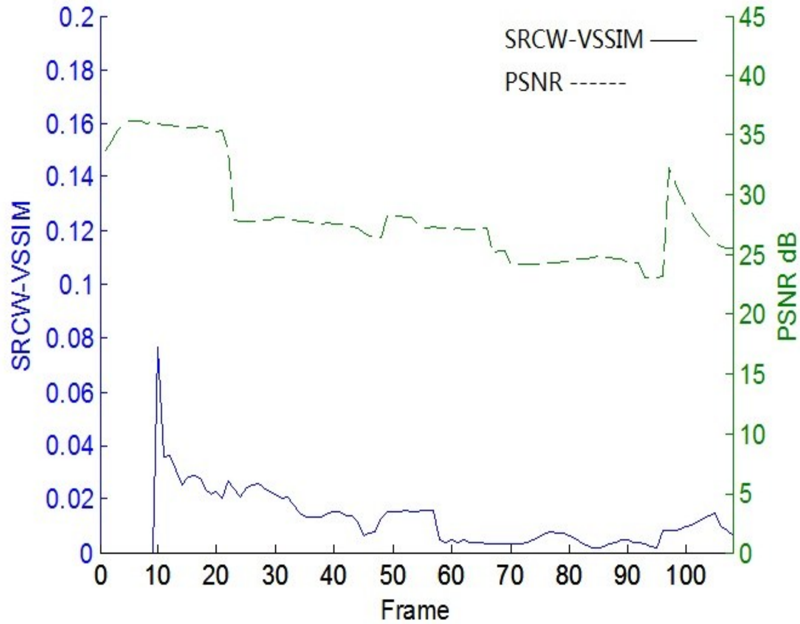

Figure 32. SRCW-VSSIM and PSNR of video sample 1 with slight distortion

Figure 33. Original frame 48 of video sample 1

Figure 34. Distored frame 48 of video sample 1

We repeat the procedures for the video sample 2, see Figures 35 to Figure 40, and experimental results can still support our proposed algorithm, and prove that our experiment is repeatable.

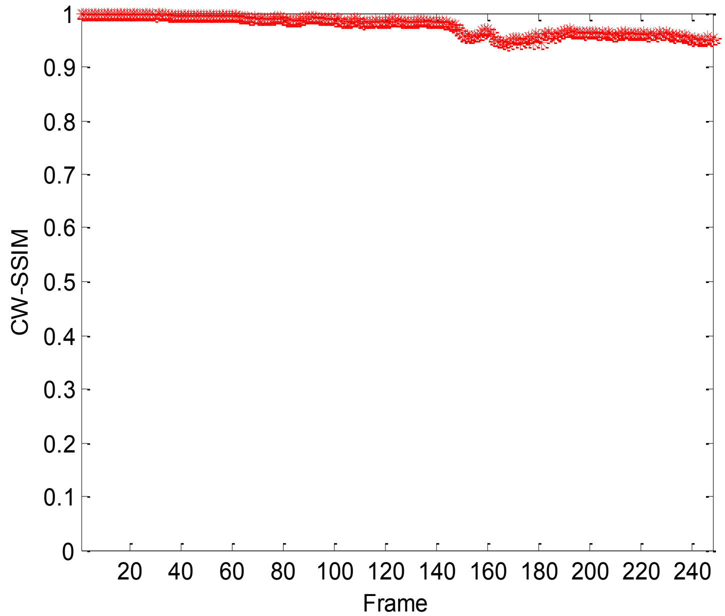

Figure 35. CW-SSIM of back-to-back frames for video sample 2 without distortion

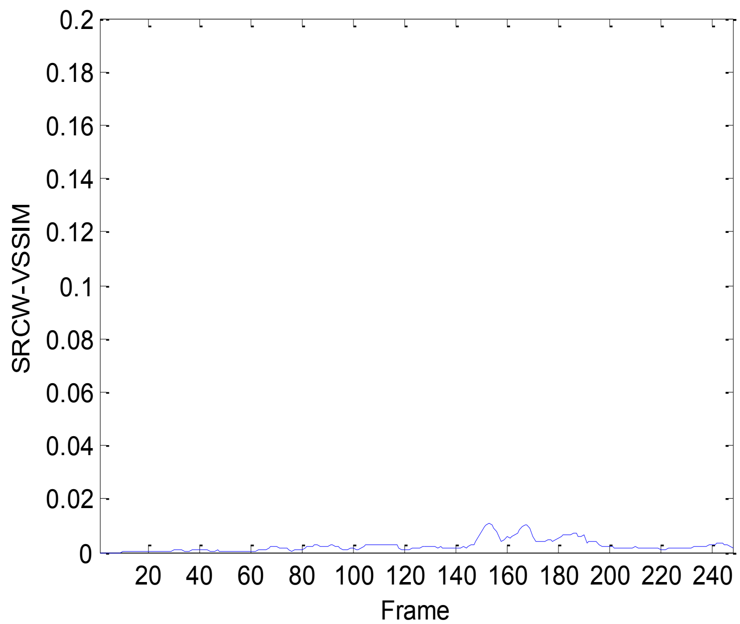

Figure 36. SRCW-VSSIM of video sample 2 without distortion

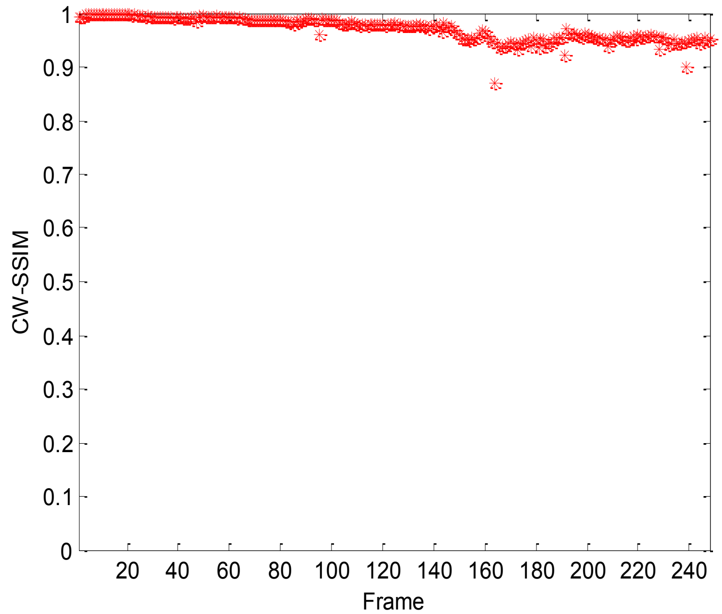

Figure 37. CW-SSIM of back-to-back frames for video sample 2 with slight distortion

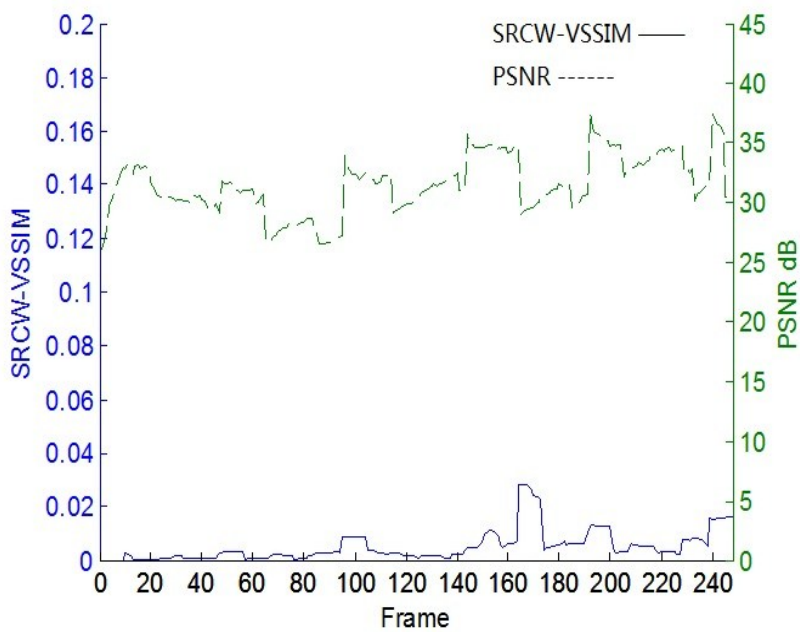

Figure 38. SRCW-VSSIM and PSNR of video sample 2 with slight distortion

Figure 39. Original frame 165 of video sample 2

Figure 40. Distorted frame 165 of video sample 2

6.3 Experiment: Video with heavy distortion

We can see from Figure 41 and Figure 42, for the original video, CW-SSIM is continuous and SRCW-VSSIM stays at a low level, less than or around 0.05. When heavy distortion is introduced, as shown in Figure 43 and Figure 44, CW-SSIM for video sample 3 becomes discontinuous. Figure 46 is a sample frame of heavy distortion while Figure 45 shows the original one. Meanwhile, SRCW-VSSIM increases to a very high level, greater than 0.05, when heavy distortion happens. At certain time, the human vision system (HVS) can easily detect video quality change, and information carried by video stream could hardly be accepted. When comparing with PSNR shown in Figure 44, all peak value of SRCW-VSSIM can match the PSNR less than 25 dB, which means unacceptable video quality.

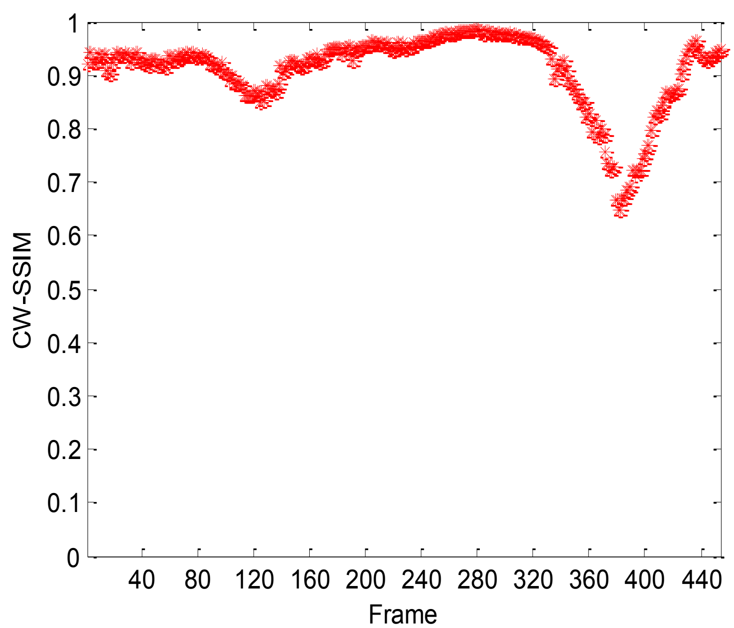

Figure 41. CW-SSIM of back-to-back frames for video sample 3 without distortion

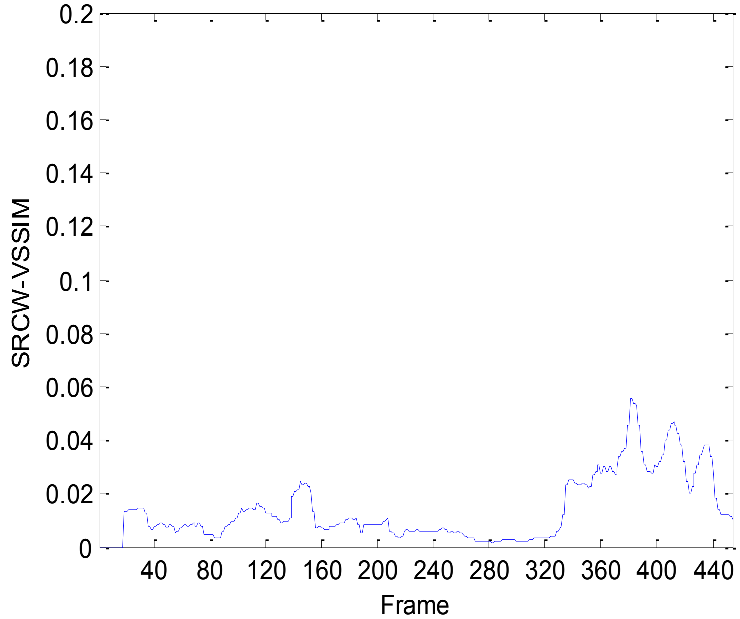

Figure 42. SRCW-VSSIM of video sample 3 without distortion

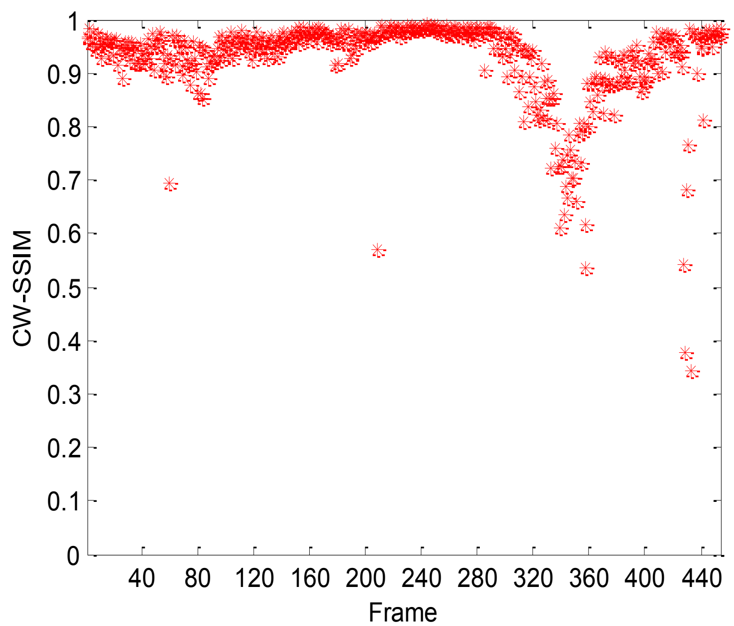

Figure 43. CW-SSIM of back-to-back frames for video sample 3 with heavy distortion

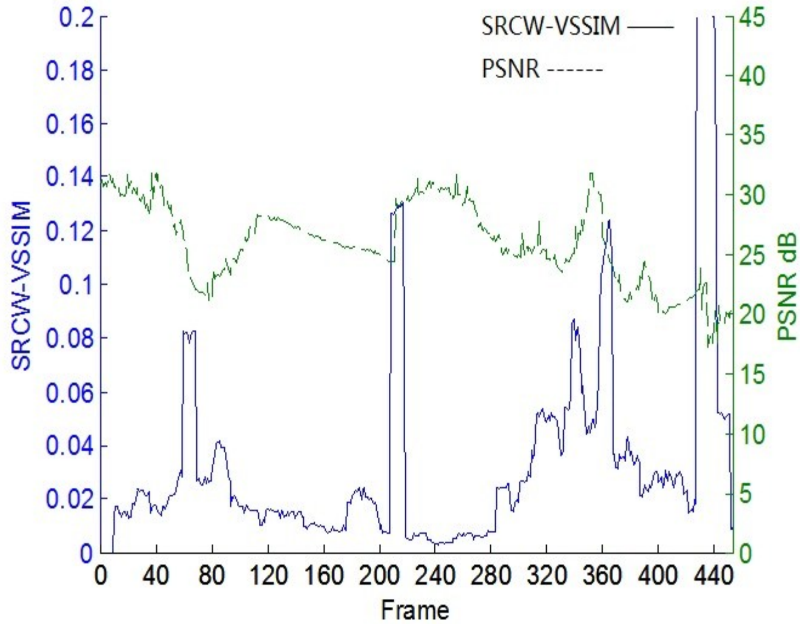

Figure 44. SRCW-VSSIM and PSNR of video sample 3 with heavy distortion

Figure 45. Original fame 209 of video sample 3

Figure 46. Heavy distorted fame 209 of video sample 3

6.4 Experiment Summary I

Conclusion can be reached based on the above experiment. Compared with PSNR, when SRCW-VSSIM is less than or around 0.05, human vision perception is acceptable, even the video is suffering from slight distortion. When SRCW-VSSIM changes sharply from a relatively low value (less than or around 0.05) to a relatively high value (around 0.1), human vision perception becomes worse, and video quality is heavily decreased. Our algorithm is proved to be a sensitive no-reference video quality measurement technique.

6.5 QoS-QoE based video quality control indicator (QQVQCI)

The parameter δ and α in (10) are set to be 0.01 and 0.05 (according to the simulation result of section 6.2 to section 6.3). We can see from Figure 47 and Figure 48, our indicator can detect network condition and human perception of quality of video clearly, and trigger the proper action to enhance the overall video transmission service. Indicator with value greater than 0.005 indicates poor network condition and human perception (Scenario 4 in Table 3), and reducing sending rate is needed. For instance, QQVQCI is greater than 0.005 near frame 365 (Arrow 4), hence server need to lower the sending rate to realize the optimization of perceived video quality. When the value of the indicator is around 0.005, for example near frame 340 (Arrow 3), network condition is good enough to increase the sending rate to enhance the video transmission service (Scenario 3 in Table 3). All other negative values of QQVQCI demonstrate that even when the network experience packet loss, since human perception of video quality still satisfies the user's requirement (Scenario 1 and Scenario 2 in Table 3), no further action is needed to tune the video sending rate (Arrow 1 and Arrow 2).

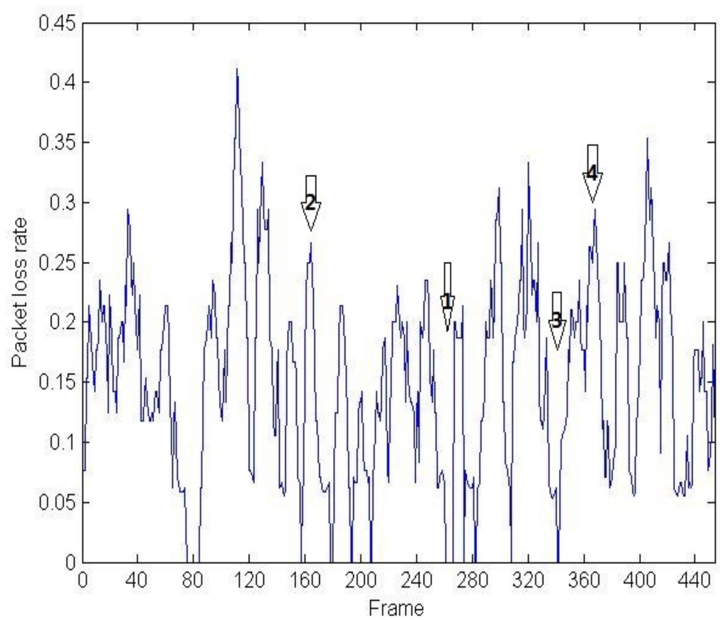

Figure 47. Packet loss rate of video sample 3 with heavy distortion

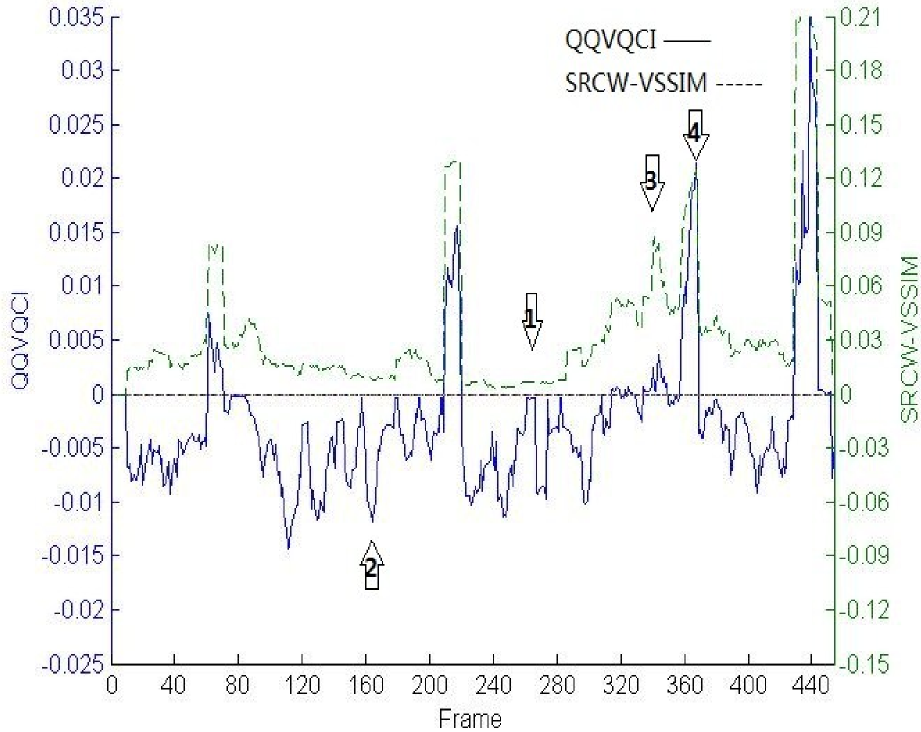

Figure 48. QoS-QoE based Video Quality Control Indicator and SRCW-VSSIM of Video sample 3 with heavy distortion

6.6 Support Vector Regression for Video Quality Prediction

We still use the same implementation to calculate our QQVQCI parameter, and Support vector regression is realized and optimized by libsvm [68]. We train the dataset by using the QQVQCI and SRCW-VSSIM of the first 290 frames. The dataset is later used to predict the QQVQCI of the following 128 frames. Figure 49 compares the actual QQVQCI and its prediction. We can see from Figure 49, stationary changes can be predicted relatively accurate. Although sharp changes can be predicted, it is not as accurate as stationary changes. Using Figure 50, we compare SVR predictor with traditional historical mean prediction tool. It is clearly that SVR prediction has its own advantage especially when sharp changes occur, and historical mean prediction tool cannot predict as accurate as SVR prediction, since simple average of previous dataset cannot provide the proper prediction.

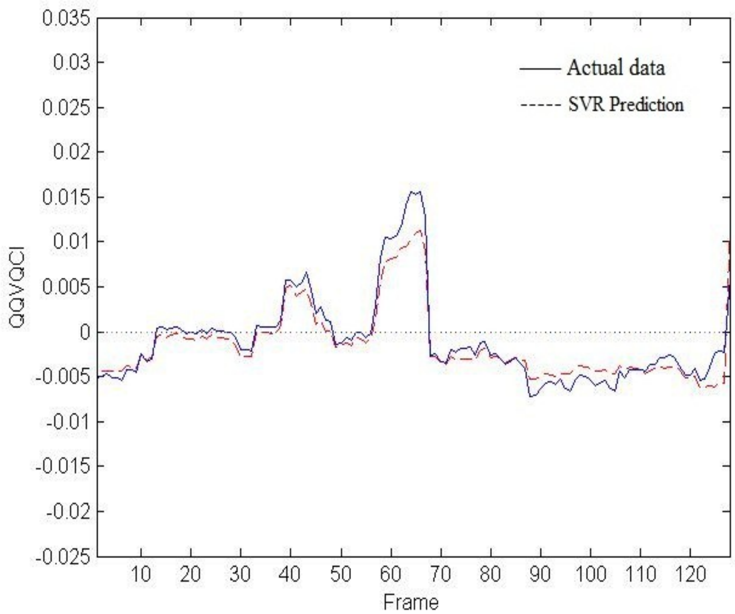

Figure 49. Actual QQVQCI and SVR Prediction of QQVQCI

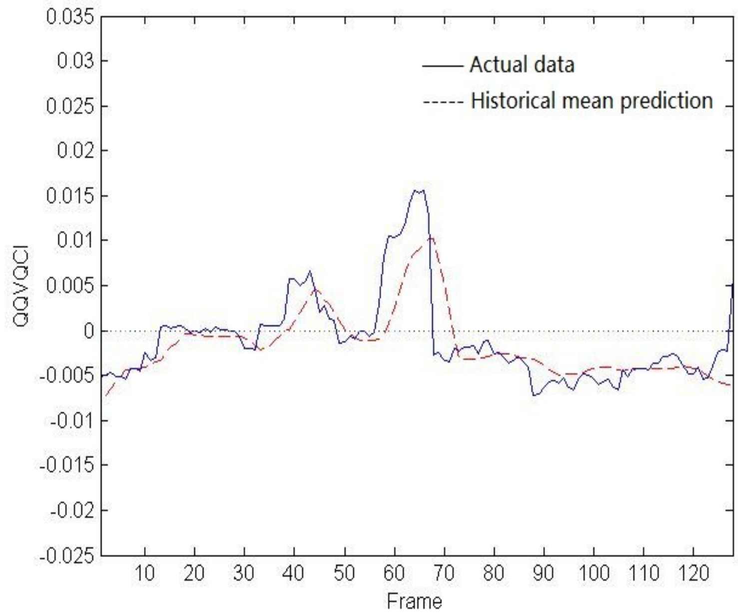

Figure 50. Actual QQVQCI and Historical Mean Prediction of QQVQCI

7. CONCLUSION AND FUTURE RESEARCH

We introduce a novel video quality control approach, which based on QoS-QoE cross-layer design. Our video quality control indicator, QQVQCI, gives more accurate information of the video transmission service, and successfully avoid the drawbacks of QoS-based video quality control algorithm. Compared with traditional single layer based video quality control algorithm, QQVQCI can look into both network condition and human perception, and trigger proper actions to balance the satisfaction of both layer's requirement. Especially, our newly introduced QoE index, SRCW-VSSIM, is reference free and closer to human perception, and this makes our video quality control indicator work perfectly under real-time video transmission environment. We introduce support vector regression technique to our novel combined QoS-QoE based video quality control algorithm. The application of support vector regression successfully converts reactive video quality control technique to proactive one, so that we are allowed to predict and pre-adjust multimedia sending rate.

Our future research will focus on the improvement of computing efficiency. Although we proved that SVR is an ideal prediction tool to switch our real time video quality control technique from reactive method to proactive method, current computing efficiency cannot fully reach the needs of real time communication. We will later integrate our indicator with SVR into our software application, and test in real time to evaluate the performance of our real time video quality control system.

REFERENCE

[1] “Definition of Quality of Experience (QoE),” Reference: TD 109rev2 (PLEN/12), ITU - International Telecommunication Union.

[2]“ITU-T Recommendation P.910. Subjective video quality assessment methods for multimedia applications, 2000.”

[3]A. N. Netravali and B. G. Haskell, “Digital Pictures: Representation, Compression, and Standards (2nd Ed),” 1995.

[4]G. Rubino, “The PSQA project,” INRIA Rennes-Bretagne Atlantique, http://www.irisa.fr/armor/lesmembres/Rubino/myPages/psqa.html.

[5] Objective Perceptual Video Quality Measurement Techniques for Digital Cable Television in the Presence of a Full Reference. ITU-T Rec. J.144, Geneva, Switzerland, Mar. 2004.

[6] M. H. Pinson and S. Wolf, “A new standardized method for objectively measuring video quality,” IEEE Trans. Broadcast., vol. 50, no. 3, pp.312-322, Sep. 2004

[7] J. Caviedes and F. Oberti, “No-reference quality metric for degraded and enhanced video,” in Proc. SPIE, vol. 5150. Jul. 2003, pp. 621–632.

[8] S. Floyd, M. Handley, J. Padhye, and J. Widmer, “Equation-based congestion control for unicast applications,” in Proc. ACM SIGCOMM '00, Stockholm, Sweden, Sept. 2000.

[9] A. Suki M. Arif, Suhaidi Hassan, Osman Ghazali, Shahrudin Awang Nor, “The Relationship of TFRC Congestion Control to Video Rate Control Optimization,” in Proc. 2nd Int. Conf. Network App. Protocols and Service (NETAPPS), 2010, pp. 31-36.

[10] Y. Huang, S. Mao, and S. Midkiff, “A Control-Theoretic Approach to Rate Control for Streaming Videos,” IEEE Transactions on Multimedia, Vol. 11, No. 6, October 2009.

[11] G. Van der Auwera and M. Reisslein, “Implications of smoothing on statistical multiplexing of H.264/AVC and SVC video streams,” IEEE Trans. Broadcast., vol. 55, no. 3, pp. 541–558, Sep. 2009.

[12] N. Staelens, S. Moens, W. Van den Broeck, I. Marien, B. Vermeulen, P. Lambert, R. Van de Walle, and P. Demeester, “Assessing quality of experience of IPTV and video on demand services in real-life environments,” IEEE Trans. Broadcast., vol. 56, no. 4, pp. 458–466, Sep. 2010.

[13] S. Kang, Y. Zhang, M. Dai, D. Loguinov, “Multi-layer Active Queue Management and Congestion Control for Scalable Video Streaming,” in Proc. 24th Intl. Conf. Distributed Computing Systems, 2004, pp. 768-777.

[14] Hyun long Kim, Seong Gon Choi, “Design of video streaming service quality control system using available bandwidth information,” in Proc. 13th Intl. Conf. Advanced Communication Technology (ICACT), 2011, pp. 1540-1544.

[15] E. Jammeh, M. Fleury, and M. Ghanbari, “Delay-based congestion avoidance for video communication with fuzzy logic control,” in Packet Video 2007, pp. 8-17.

[16] Hyun Jong Kim, Dong Geun Yun, Hwa-Suk Kim, Kee Seong Cho, Seong Gon Choi, “QoE Assessment Model for Video Streaming Service using QoS Parameters in Wired-Wireless Network,” in Proc. 14th Intl. Conf. Advanced Communication Technology (ICACT), 2012, pp. 459-464.

[17] Masahiko Kitamura, Takahiro Murooka, Mitsuhiro Teshima, Hisaya Hadama, Atsushi Takahara, “A Study on the Correlation between QoE of 4K Super High Definition Video Streamings and QoS of Network,” in Proc. 18th Intl. Conf. Comp. Commu. and Network, 2009, pp. 1-6.

[18] Takahiro Suzukit, Tasuku Kutsunatt, and Shuji Tasakatt, “QoE Estimation from MAC-Level 008 in Audio-Video Transmission with IEEE SOZ.lle EDCA,” in Proc IEEE 19th Intl. Symposium PIMRC, 2008, pp. 1-6.

[19] Drucker, Harris; Burges, Christopher J. C.; Kaufman, Linda; Smola, Alexander J.; and Vapnik, Vladimir N., “Support Vector Regression Machines”, Advances in Neural Information Processing Systems 9, NIPS 1996, 155–161, MIT Press.

[20] A.J. Smola and B. Schölkopf, “A tutorial on support vector regression,” NeuroCOLT2 Technical Report NC-TR-98-030, University of London, UK, 1998.

[21] W. Lu, W. Wang, A. Y. T. Leung, S. M. Lo, “Air Pollutant Parameter Foresting using Support Vector Machines”, in Proc. 2002 Int. Joint Conf. on Neural Network, May 12-17 2002.

[22] Mu-Yen Chen， Chia-Chen Chen， and Ya-Fen Chang, “Comparison of support vector machine and support vector regression: An application to predict financial distress and bankruptcy,” in Service Systems and Service Management (ICSSSM), Tokyo, Japan, Jun.2010.

[23] Li Yuan, Liu Hua and Liu Zhi-guo, “Network flow prediction based on grey-support vector regression technology,” In Education Technology and Computer (ICETC), Shanghai, China, Jun.2010.

[24] R. Beverly, K. Sollins, and A. Berger. “SVM Learning of IP Address Structure for Latency Prediction,” In Proceedings of the SIGCOMM Workshop on Mining Network Data, Pisa, Italy, Sept 2006

[25] Lian H. “No-reference video quality measurement with support vector regression”, International Journal on Neural System, Dec. 2009, 19(6): 457-64.

[26] Narwaria, M. Weisi Lin, “Objective Image Assessment Based on Support Vector Regression”, IEEE Transactions on Neural Networks, March 2010, Vol. 21, Issue 3, pp. 515-519.

[27] Beibei Wang, Dekun Zou and Ran Ding, “Support Vector Regression Based Video Quality Prediction,” in Proc IEEE Intl. Symposium ISM, 2011, pp. 476-481.

[28] M. Wien, H. Schwarz, and T. Oelbaum, “Performance analysis of SVC,” IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 9, pp. 1194–1203, Sep. 2007.

[29] B. Ciubotaru and G.-M. Muntean, “SASHA—A quality-oriented handover algorithm for multimedia content delivery to mobile users,” IEEE Trans. Broadcast., vol. 55, no. 2, pp. 437–450, Jun. 2009.

[30] G. Van der Auwera, P. David, and M. Reisslein, “Traffic characteristics of H.264/AVC variable bit rate video,” IEEE Commun. Mag., vol. 46, no. 11, pp. 164–174, Nov. 2008.

[31] M. Isnardi, “Historical overview of video compression in consumer electronic devices,” in Proc. Int. Conf. Consum. Electron.(ICCE), 2007, pp. 1–2.

[32] T. Wiegand, L. Noblet, and F. Rovati, “Scalable video coding for IPTV services,” IEEE Trans. Broadcast., vol. 55, no. 2, pp. 527–538, Jun. 2009.

[33] F. Speranza, A. Vincent, and R. Renaud, “Bit-rate efficiency of H.264 encoders measured with subjective assessment techniques,” IEEE Trans. Broadcast., vol. 55, no. 4, pp. 776–780, Dec. 2009..

[34] M. Pinson, S. Wolf, and G. Cermak, “HDTV subjective quality of H.264 vs. MPEG-2, with and without packet loss,” IEEE Trans. Broadcast., vol. 56, no. 1, pp. 86–91, Mar. 2010

[35] J. Monteiro, C. Calafate, and M. Nunes, “Evaluation of the H.264 scalable video coding in error prone IP networks,” IEEE Trans. Broadcast.,vol. 54, no. 3, pp. 652–659, Sep. 2008.

[36] K. Brunnstrom, D. Hands, F. Speranza, and A. Webster, “VQEG validation and ITU standardization of objective perceptual video quality metrics,” IEEE Signal Process. Mag., vol. 26, no. 3, pp. 96–101, May 2009.

[37] L. Karam, T. Ebrahimi, S. Hemami, T. Pappas, R. Safranek, Z. Wang, and A. Watson, “Introduction to the special issue on visual media quality assessment,” IEEE J. Sel. Topics Signal Process., vol. 3, no. 2, pp. 189–192, Mar./Apr. 2009.

[38] A. Takahashi, D. Hands, and V. Barriac, “Standardization activities in the ITU for a QoE assessment of IPTV,” IEEE Commun. Mag., vol. 46, no. 2, pp. 78–84, Feb. 2008.

[39] B. Ciubotaru, G.-M. Muntean, and G. Ghinea, “Objective assessment of region of interest-aware adaptive multimedia streaming quality,” IEEE Trans. Broadcast., vol. 55, no. 2, pp. 202–212, Jun. 2009.

[40] Lehmann, E. L.; Casella, George (1998). Theory of Point Estimation (2nd ed.). New York: Springer. ISBN 0-387-98502-6. MR 1639875.

[41] L. Guo and Y. Meng, “What is wrong and right with MSE?” in Proc. 8th Int. Conf. Signal Image Process., 2006, pp. 212–215.

[42] Z. Wang, L. Lu, and A. Bovik, “Video quality assessment based on structural distortion measurement,” Signal Process. Image Commun., vol. 19, no. 2, pp. 121–132, Feb. 2004.

[43] Z. Wang, E. Simoncelli, and A. Bovik, “Multiscale structural similarity for image quality assessment,” in Conf. Rec. 37th Asilomar Conf. Signals, Syst. Comput., 2003, vol. 2, pp. 1398–1402.

[44] Z. Wang and Q. Li, “Video quality assessment using a statistical model of human visual speed perception,” J. Opti. Soc. America A (Optics, Image Sci., Vision), vol. 24, no. 12, pp. B61–B69, Dec. 2007.

[45] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Trans. Image Processing, vol. 13, pp. 600–612, Apr. 2004.

[46] Z. Wang, L. Lu, and A. Bovik, “Video quality assessment based on structural distortion measurement,” Signal Process. Image Commun., vol. 19, no. 2, pp. 121–132, Feb. 2004.

[47] Zhou Wang and Eero P. Simoncelli, “Translation insensitive image similarity in complex wavelet domain,” Proc. IEEE Inter. Conf. Acoustic, Speech & Signal Processing Volume II, Pages 573-576, March 2005.

[48] C. van den Branden Lambrecht and O. Verscheure, “Perceptual quality measure using a spatio-temporal model of the human visual system,” in Proc. Int. Soc. Opt. Eng. (SPIE), 1996, vol. 2668, pp. 450–461.

[49] A. Hekstra, J. Beerends, D. Ledermann, F. de Caluwe, S. Kohler, R. Koenen, S. Rihs, M. Ehrsam, and D. Schlauss, “PVQM—A perceptual video quality measure,” Signal Process. Image Commun., vol. 17, no. 10, pp. 781–798, Nov. 2002.

[50] S Wolf, MH Pinson, “Low bandwidth reduced reference video quality monitoring system,” Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, Arizona, 2005, pp. 23–25

[51] I Gunawan, M Ghanbari, “Reduced-reference video quality assessment using discriminative local harmonic strength with motion consideration,” IEEE Trans Circ. Syst. Video Technol. 18(1), 71–83 (2008)

[52] T. Vlachos, “Detection of blocking artifacts in compressed video,” IEEE Electron. Lett., vol. 36, no. 6, pp. 1106–1108, Jun. 2000.

[53] K. Tan and M. Ghanbari, “Frequency domain measurement of blockiness in MPEG-2 coded video,” inProc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2000, vol. 3, pp. 977–980.

[54] Rohde & Schwarz. DVQ Digital Video Quality Analyzer [Online]. Available: http://www2.rohde-schwarz.com/product/DVQ.html

[55] P. Gastaldo, S. Rovetta, and R. Zunino, “Objective assessment of MPEG-video quality: A neural-network approach,” in Proc. IEEE Int. Conf. Neural Netw., Jul. 2001, vol. 2, pp. 1432–1437.

[56] P. Gastaldo, S. Rovetta, and R. Zunino, "Objective quality assessment of MPEG-2 video streams by using CBP neural networks," IEEE Trans. Neural Netw., vol. 13, no. 4, pp. 939–947, Jul. 2002.

[57] P. Gastaldo, G. Parodi, and R. Zunino, "DSP-based neural systems for the perceptual assessment of visual quality," in Proc. IEEE Int. Conf. Neural Netw., Jul. 2005, vol. 1, pp. 663–668.

[58] S. Kanumuri, P. C. Cosman, A. R. Reibman, and V. A. Vaishampayan, "Modeling packet-loss visibility in MPEG-2 video," IEEE Trans. on Multimedia, vol. 8, no. 2, pp. 341–355, Apr. 2006.

[59] G. Valenzise, S. Magni, M. Tagliasacchi, and S. Tubaro, "Estimating channel-induced distortion in H.264/AVC video without bitstream information," in Proc. 2nd Int. Workshop Quality Multimedia Experience, Jun. 2010, pp. 100–105.

[60] Fuzheng Yang, Shuai Wan, Qingpeng Xie, Hong Ren Wu, "No-Reference Quality Assessment for Networked Video via Primary Analysis of Bit Stream," IEEE Transactions on Circuits and Systems for Video Technology, Nov. 2010, Vol. 20, Issue 11, pp. 1544-1554.

[61] Yo-Han Kim, Junghyun Han, Hyuntai Kim, Jitae Shin, "Novel no-reference video quality assessment metric with estimation of dynamic range distortion," in Proc 12th Intl. Conf. Advanced Communication Technology (ICACT), 2010, pp. 1689-1692.

[62] Shyamprasad Chikkerur, Vijay Sundaram, Martin Reisslein, and Lina J. Karam, "Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison," IEEE Transactions on Broadcasting, Vol. 57, No. 2, June 2011

[63] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, M. K. Markey, "Complex Wavelet Structural Similarity: A New Image Similarity Index," IEEE Transactions on Image Processing, 18(11), 2385-401, 2009.

[64] MSU Graphics & Media Lab, "MSU Video Quality Measurement Tool," [Online]. Available: http:// compression.ru/video/quality_measure

[65] K. Seshadrinathan, R. Soundararajan, A. C. Bovik and L. K. Cormack, "Study of Subjective and Objective Quality Assessment of Video," IEEE Transactions on Image Processing, vol.19, no.6, pp.1427-1441, June 2010.

[66] K. Seshadrinathan, R. Soundararajan, A. C. Bovik and L. K. Cormack, "A Subjective Study to Evaluate Video Quality Assessment Algorithms," SPIE Proceedings Human Vision and Electronic Imaging, Jan. 2010.

[67] UGent-IBCN, "Sirannon 1.0.0: Modular Multimedia Streaming," [Online]. Available: http://xstreamer.atlantis.ugent.be

[68] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

List of Publication

Biao Jiang, Tarek Saadawi, "Real-Time Self-Reference Video Quality Measurement Technique," in WEBIST 2013, Aachen, Germany, May 2013

Biao Jiang, Tarek Saadawi, "Combined End-to-End QoS-QoE Parameters for Video Quality Control," in Journal of Visual Communication and Image Representation, Submitted

Biao Jiang, Tarek Saadawi, "Support Vector Regression Technique for Real Time Video Quality Control," in Computer Communication, Ready to submit

Qihua Yang, Biao Jiang, Tarek Saadawi, Ahmed Abdelal, Mitesh Patel, "Support Vector Regression Technique for Multimedia Quality Control in Multicast Networks," in NoF 2013, Pohang, South Korea, Oct 2013 Accepted

Orhan Celebi, Biao Jiang, Tarek Saadawi, Ahmed Abdelal, "Real Time Video Quality Measurement in Multimedia Networks," in ISM 2013, San Jose, USA, Dec 2013 Ready to submit

Qihua Yang, Tarek Saadawi, Ahmed Abdelal, Mitesh Patel, Biao Jiang, "Multicast Multi-Streaming and Real-Time Stream Rate Control Demo," in Globecom 2010, Miami, USA, Dec 2010