Video Quality Assessment and Management in Content Distribution Networks

Mu Mu, Andreas Mauthe
Computing Department, Lancaster University
m.mu@lancaster.ac.uk, andreas@comp.lancs.ac.uk

ABSTRACT

The presented research deals with (objective) video quality assessment and how it can be used within CDN. In order to guarantee a high perceived video quality an objective video content assessment and management service considering human visual system characteristics, audio-visual content and network impairments is being devised. Within this research aspects of automatic content assessment using human visual characteristics and human quality perception are brought together with core communication topics such as QoS, network topology, multi services networks, etc. The result of this work will be a service within a content infrastructure that assess where and how quality degradation occurs in the CDN. Using this service adaptation tools and services can, for instance, be instantiated in the CDN in order to provide the most optimal quality experience to the user. Alternatively, the communication path can be adjusted according to the experienced traffic characteristics and application requirements.

The poster introduces some initial work on objective video quality assessment and management. Further, it outlines the envisaged OVQA service architecture and then finishes with conclusion.

1. INTRODUCTION

Audio-visual (AV) content is still one of the most challenging content types in content distribution infrastructures. It is not only challenging in terms of bandwidth and timing requirements but also because of the user quality perception. In recent years is has become apparent that there is no linear correlation of resources and encoding schemes, traffic characteristics and the actually perceived quality. This is especially critical in heterogeneous environments with networks ranging from high-speed networks to wireless low capacity networks and devices ranging from HDTV sets to handheld devices. The goal is to provide the most optimal user experience to all participants in a CDN. Thus, content distribution has to support a large number of end users in heterogeneous content infrastructure while satisfying their various quality requirements. To support the optimal user experience in such an environment becomes challenge for content distribution networks.

The quality of video content can be assessed with two methods: subjective assessment and objective assessment. Subjective measurement is derived entirely from testers' opinion by controlled experiments. However, testers' feelings can be easily affected by many unrelated elements in the experiment. Furthermore, in order to keep the results generic, the number of samples must be enormous which makes the experiments time-consuming, costly and non-repeatable. Objective measurement methods intend to eliminate human participation by correlating the user's perception to a set of objective quality indicators (metrics) and algorithms. A number of the objective assessment methods have been developed and accepted by the academic community. However, they either come with high computational complexity or lack of general correlation to subjective assessment results.

In the following, we first outline the pros and cons of existing OVQA methods and then we take a brief view of how video characteristics and encoding techniques effect objective assessment. In chapter 4, the key impairments from network transmission will be discussed. After that, we introduce several factors that affect the visual content quality. In chapter 6, we discuss our ways and means to derive the HVS-based NR video assessment model and an outlook of how can we initiated it as a service over CDN.

2. OBJECTIVE QUALITY ASSESSMENT

Three types of objective assessment methods which are exclusively based on video signal parameters, network impairment and human visual system have been studied.

Signal parameters method like MSE/PSNR compares two video streams in pixel by pixel manner. PSNR is widely adopted in the literature but shows low correlation to viewer ratings of video quality in subjective testing performed by VQEG [1].

When video content is been delivered through content distribution network, the transport impairments usually degrade the video quality in various ways. The network impairment method is based on network monitoring. MDI is one of these network impairment based methods which considers Delay Factor and Media Loss Rate that have been introduced on transmission path [2]. This method helps researchers and service providers to find where and how in a content distribution network the content quality degrades but the overall result can not fully represent the perceived quality.

Human visual system based methods such as VQM [3] and MPQM [4] simulate the key features of human eyes and nerves system. For these methods, a full-reference assessment model is usually built and then the "near subjective" assessment tests can be repeatedly used with different parameters to derive relations (e.g. functions) between certain video parameters and perceived quality. These relations are then consolidated into algorithm and no-reference/reduced reference assessment model.

3. VIDEO CONTENT

The first thing that needs to be considering for video content is the taxonomy. Video streaming may be considered as continuing group of pictures. But characteristics of individual pictures and motion between pictures are distinctive for different video content. For example, the video content in a web conference will have more requirements on interactivity than the one in a broadcast TV program. We can categorize videos into groups based on characteristics as well as compression and transmission requirements.

Modern encoding techniques not only consider compression schemes like Intra Coding for Spatial Prediction and Inter Coding for Temporal and Spatial Prediction but also extended features for loss resilience such as redundant slices, experience enhancement such as de-block filtering [5].

When video content is encoded, there are different transport mechanisms over CDN: MPEG-TS/UDP, RTP/UDP or TCP. Each mechanism has its advantages and disadvantages facing different content type and network status.

4. NETWORK IMPAIRMENTS

Delay - Depending on the application classes large delays exceeding a certain level may be regarded as packet loss rather than a giant delay. Most of the current objective and subjective research efforts aim to find the upper bound of delay limitation. Though, the lower bound of delay (minimum delay noticed) should be taken into account and play a key role in video assessment.

Jitter - Variations of network transfer delays are known as jitter. Jitter is normally caused by changes of route or changes of conditions on router network nodes' buffer. The impacts of jitter are getting smaller on modern applications, which are always equipped by a de-jitter buffer to smooth transmission delay with introducing unobjectionable buffer delay.

Loss - Packet loss is defined as the ratio of number of lost packets during transmission to the total number of transmitted packets. Packet loss comes from congestion resulted from insufficient or non-optimal usage of network resource. For connection-oriented transport protocol, loss of packets will cause retransmission which will delay the arrival of packets, while for connection-less protocol like UDP, the loss may directly affect application performance.

Bandwidth - Bandwidth is the capacity of carrying data on a certain path of the network. If the traffic on the network exceeds this capacity, packets will get lost or delayed. For some applications the impact of bandwidth limit even happens before transmitting the packets. Network adaptive video compression is an example of this case. Besides encoding rate, frame rate and resolution are also elements that are impacted by bandwidth limitation.

5. HUMAN VISUAL SYSTEM

Human involved subjective assessment is time-consuming, costly and non repeatable. By modeling characteristics of human visual system, we can build a full reference assessment model that evaluates the video quality by comparing only parts of the video information that is critical for human eyes and nerve system. This full reference model can be used for
1, direct video assessment if the reference content is available and the algorithm has been optimize or
2, facilitate a no reference assessment model by providing correlations between perceived quality and each quality metrics.

There are five factors that affect the visibility of visual tasks [6]:

- Contrast - relationship between the luminance of an object and the luminance of the background.
- Size - The larger an object, the easier it is to see.
- Time - There is a time lag in the photochemical processes of the retina, therefore the time available for viewing is important.
- Luminance - proportion of incident light reflected into the eye.
- Color - related both to contrast and luminance factors.

VQM and MPQM are two well defined HVS-based assessment models. They both show correlation with subjective assessment but only under certain assumptions and conditions that may not generally apply to reality.

6. BUILDING THE ASSESSMENT MODEL

Combining the advantages of both HVS-based method and network impairment based method; we aim at an assessment method that evaluates video quality using metrics such as video content characteristics, encoding techniques, transmission damage and user preference. The relations between each metric and video quality will be studied separately and the dependencies between metrics will be discovered. Metrics of network impairments are about to be loose coupled with others so for CDN diagnosis and management. All the relations (e.g.functions) will be derived from a controlled experiment with FR HVS based test bed (Figure 1).

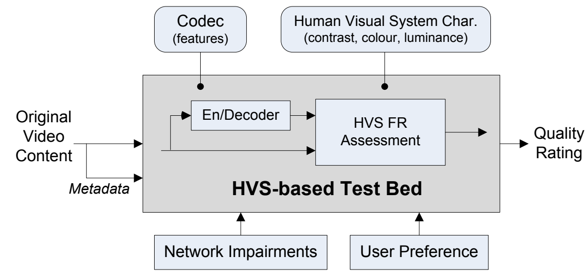

Figure 1. HVS-based Test Bed

When the assessment model has been built from the HVS-based test bed, we may instantiate it in the CDN in order to provide the most optimal quality experience to the user. Alternatively, the communication path can be adjusted according to the experienced traffic characteristics and application requirements. Issues include constraint based routing protocol, adaptive video content and content distribution and delivery.

7. CONCLUSION

Original the Internet was not designed for media streaming and thus impairments are not easy to predict and manage. Peer-to-peer transport and control layers enable a more effective content delivery network but meanwhile introduce many uncontrollable effects. Furthermore, the video compression and transport system has been extended with large numbers of tool sets which makes the designing of an efficient objective video assessment model even more complicated. The simulation model of the human visual system is believed to be a superior way for video assessment, management and CDN diagnosis. However, the human visual system is extremely complex and many properties have not yet been well understood.

8. REFERENCE

[1] VQEG, "Final report from the video quality experts group. on the validation of objective models of video quality assessment, phase ii (fr-tv2)," September 2003.

[2] J. Welch, J. Clark, "A Proposed Media Delivery Index (MDI)," RFC 4445, April 2006.

[3] M. Pinson, S. Wolf., "A New Standardized Method for Objectively Measuring Video Quality," in *IEEE Transactions on Broadcasting*. vol. 50, September 2004, pp. 312-322.

[4] C. J. Branden Lambrecht, O. Verscheure, "Perceptual Quality Measure using a Spatio-Temporal Model of the Human VIsual System," *SPIE,* vol. 2668, pp. 450-461, March 1996.

[5] ITU-T, "H.264 : Advanced video coding for generic audiovisual services."

[6] "Structure and Function of the Human Visual System," Cornell University.