Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 2, Issue. 4, April 2013, pg.58 – 64

RESEARCH ARTICLE

Perceptual Video Quality Measurement Based on Generalized Priority Model

S. Bharath[1], S. Jaganath[2], J. Prakash[3]

[1]Assistant Professor, Department of ECE, PSNA College of Engineering and Technology, Dindigul, India
[2]Assistant Professor, Department of ECE, PSNA College of Engineering and Technology, Dindigul, India
[3]Assistant Professor, Department of ECE, PSNA College of Engineering and Technology, Dindigul, India

[1]bharath.psna@gmail.com; [2]s.jaganhsd@gmail.com; [3]jeyavelprakash@gmail.com

Abstract— We consider factors not only in a packet, but also in its locality, to account for possible temporal and spatial masking effects. We apply our visibility model to packet priority for a video stream, when the network gets jam-packed at an in-between router; the router is able to choose which packets to drop such that visual quality of the video is minimally crashed. To show the effectiveness of our visibility model and its corresponding packet priority method, experiments are done to compare our perceptual-quality-based packet priority approach with existing Drop tail & hint track, Mean square error priority methods. The result shows that our priority method produces videos of higher perceptual quality for different network conditions. Our model was developed using data from high encoding-rate videos, and designed for high-quality video sent over a mostly reliable network; however, the experiments show the model is valid to different encoding rates.

Key Terms: - Packet dropping policy; packet loss; perceptual video quality; video coding; visibility model

I. INTRODUCTION

The growing popularity of transmitting compressed video over the Internet increases the need for quality assessment methods that can accurately characterize how the network is affecting the video quality seen by the end-user. Transmitting video in digital form is the direct result of the benefits offered by digital compression. The potential impact of multimedia information is currently restricted by the bandwidth of the existing communication networks. Quality of Service (QoS) of Networks are of more importance, especially due to the need of Synchronization (frame-rate of video must be maintained) of video streams. The Burst nature of compressed video stream is the major hurdle for Network technologies to maintain QoS.

Video quality measurement in the network can be categorized into three different types based on the accessibility of information about the original (reference) video. Full-reference (FR) methods evaluate the video quality with access to the original video, providing the most precise measurements on the video quality difference. Reduced-reference (RR) metrics extract partial information about the original video at the sender and are sent reliably to the receiver to estimate the video quality. No-reference (NR) methods only use information available in the bit stream or the decoded pixels without reference video information [1].

Packet losses in the network (for example, due to congestion) can significantly damage video quality during transmission. Therefore, considerable research has been conducted to understand the relationship between packet losses and visual quality degradation. Although PSNR (Peak Signal to Noise Ratio) and MSE do not always reflect perceptual quality well, they are commonly used to measure video quality. The relation between PSNR and perceptual quality scores is considered in [2]. It finds that packet losses are visible when the PSNR drop is greater than a threshold, and the distance between dropped packets is crucial to perceptual quality.

For packet prioritization, one can assign low priority to packets that cause low loss visibility. When the buffer in a network node is congested, it can opt to discard the low-priority packets and, hence, minimize the degradation to perceived video quality for the end user.

Packet losses were introduced in MPEG-2 bit streams and concealed using zero-motion error concealment (ZMEC). Viewers were asked to observe the videos and respond to the visible glitches that they notice. Using the subjective test results and a set of factors that were extracted from the videos, the Classification and Regression Trees (CART) algorithm was applied to classify the losses as visible or invisible. This work was extended in [3] and [4] to model the probability of packet loss visibility using a generalized linear model (GLM) [5]. The visibility for H.264 packets is discussed in [6]. We derived visibility models specifically for individual and multiple packet losses based on RR factors.

The goals of this paper are twofold: to develop a visibility model for different GOP structures and encoding rates, and to demonstrate the effectiveness of a packet prioritization application based on our visibility model. Using the packet priorities, an intermediate router can intelligently drop low-priority packets.

In this paper the model is built based from subjective experiments and Drop-Tail policy using NS-2 (Network Simulator) [7].

This paper is organized as follows: Section II describes the experiment settings for the three different subjective tests Section III discuss possible applications of the proposed model. Section IV introduces the attributes of packet and the error between them, to predict packet loss visibility. In Section V, provide GLM and incorporate significant factors. Section VI presents the experiment results Section VII concludes the paper.

II. SUBJECTIVE DATASETS

The major purpose of this work is to develop a generalized and robust visibility model for packet loss impairments. The results of three prior subjective experiments [4], [6], [8] in which the video clips are generated by using various codecs and settings as summarized in Table I.

Tests 1 and 2 use videos compressed by MPEG-2 at spatial resolution 720 x 480 with an adaptive GOP (group-of-picture) structure in which an I-frame is inserted at each scene cut. In these videos, there are usually 2 B frames between each reference frame, and the typical GOP length is 13 frames.

Test 3 uses videos encoded by H.264/AVC extended profile (JM 9.1) at spatial resolution 352 x 240 with a fixed IBPBPB-type GOP structure of 20 frames. The encoder in this case uses each I-frame of the current GOP as a long-term reference frame. For P frames, a long-term reference frame and a short-term reference frame (previously-coded P frame) are used for motion compensation. B frames use the future P frame and either the long-term or short-term reference frame for bidirectional prediction.

Test 1 uses default error concealment typical of a software decoder that is designed for speed rather than error resilience. Test 2 uses zero-motion error concealment (ZMECTest 3 uses Motion-Compensated Error Concealment (MCEC) [5], which incurs a lower initial error compared to ZMEC.

During the subjective test in all three tests, each packet loss was evaluated by 12 viewers. Viewers were told that they will watch videos which are affected by packet losses.

TABLE 1: SUMMARY OF SUBJECTIVE DATASET

Parameters	**Test 1[42]**		**Test 2[25]**	**Test 3[27]**
Spatial resolution	720x480		720x480	352x240
Frame rate	30	24	30	30
Duration of video in test(minutes)	7.3	8.9	72	36
Compression standard	MPEG-2		MPEG-2	H-264
GOP structure	I-B-B-P		I-B-B-P	I-B-P

I Frame Insertion	Scene adaptive		Scene adaptive	Fixed
GOP Length	<13	<15	<13	20
Concealment	default		ZMEC	MCEC
Losses	108	107	1080	2160
Losses in B frames	14 %		14 %	50 %
Full frame losses	20 %		30 %	0 %
Null Pred.error	0.14599		0.12236	0.041571
Initial Mean Sq.error	5.245		3.919	1.708

Whenever they see a visible artifact or a glitch, they should respond by pressing the space bar. Based on comments from viewers after the tests, the full-color full motion video was sufficiently compelling that they were immersed in the viewing process rather than searching for every artifact.

III. APPLICATIONS OF A PACKET-LOSS VISIBILITY MODEL

The first scenario is that, our visibility model is used for in-network quality monitoring of transmitted video, as described in [4]. In this application, the visibility model is computed for the specific loss pattern that is observed in the network.

The second scenario is our visibility model is used to prioritize packets for transmission by a video server. In this case, the goal is to label each packet with a priority, assigned using our visibility model, that describes the impact of losing this specific packet during transmission. Factors needed by the visibility model can either be extracted from the complete loss-free bit stream on the fly at the server when needed for transmission, or precomputed and stored with the specific packet in the server.

IV. FACTORS FOR PACKET-LOSS IMPAIRMENTS

To create a versatile model for packet loss visibility, it is crucial to understand the types of impairments induced by a packet loss, and whether these impairments depend on (a) the codec and its parameters, (b) the packetization strategy, (c) the decoder error concealment, and (d) the video content. In this section, the issues for describing attributes that affect the visibility of packet loss impairments, and describe the associated measurements, or factors. The signal is defined as (i).The original signal (ii) The compressed signal (iii) The Decompressed signal (iv) The error signal

A. Encoded Signal at Location of Loss

For the encoded signal, (t) the tendency of human observers to track moving objects with their eyes may enhance visibility of packet loss in smoothly moving regions, yet local signal variance and motion variability may hide the packet loss. For each macro block, we measure its motion vector (x,y) by forward motion estimation from the previous frame.

B. Encoded Signal Surrounding Location of Loss

The attributes of the encoded signal *surrounding* the location of the packet loss can also affect visibility. For a packet loss *after* a scene cut, the impairments can be masked by the change of the scenes. This is forward masking and it decreases visibility of packet loss. Backward masking also decreases the visibility of a packet loss *before* a scene cut.

C. Decoded Signal

The decoded signal (t), at the location of a packet loss has several attributes that affect packet-loss visibility. Due to imperfections in the error concealment of the lost packet, there can be spatial (vertical or horizontal) or temporal discontinuity with the neighboring MBs or frames; these are called *edge artifacts*.

D. Error Signal

The error caused by the impairment e(t), is completely characterized by its *support* and its *amplitude*. The error support is characterized by spatial support (size, spatial pattern and location) and temporal support (duration).

V. PROPOSED APPROACHES - GLM

Our goal in this paper is to develop a model that predicts the probability of a lost packet being visible to viewers based on the factors discussed above. In our experiment and data analysis, we assume each viewer's response is an independent observation of the average viewer (for whom we are developing the model). Therefore, each viewer response can be considered independent and identically-distributed with probability p for seeing a particular packet loss. This leads us to the binomial allotment for modeling the packet loss visibility.

A *Generalized Linear Model* (GLM) is suitable for our purpose since it can be used to predict the probability parameter of a binomial distribution.

A. Introduction of Generalized Linear Models

GLMs are an extension of classical linear models .The probability of visibility is modeled using logistic regression, a type of GLM which is a natural model to predict the parameter of a binomial distribution A generalized linear model can be represented as

$$g(p_i) = \gamma + \sum_{j=1}^{P} x_{ij}\beta_j$$

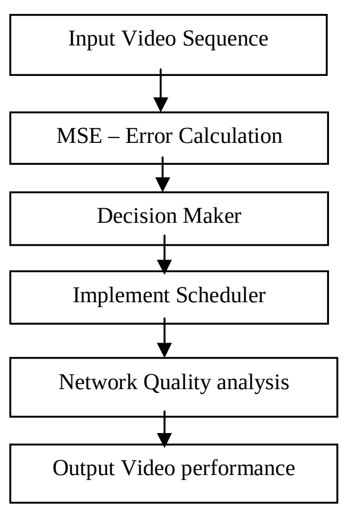

Fig 1: Block diagram

Where g(.) is called the link function, which is typically nonlinear, and are the coefficients of the factors. For logistic regression, the link function is the logic function, which is the canonical link function for the binomial distribution. The logic function is defined as

$$g(p) = \log\left(\frac{p}{1-p}\right)$$

Given N observations, one can fit models using up to N parameters. The simplest model (Null model) has only one parameter: the stable. At the other end, it is possible to have a model (full model) with as many factors as there are observations.

B. GLM Model Building Approach on Multiple Data Sets

A model is trained on a fraction of the data (*training set*) and then tested using the remaining data points (*testing set*). A partition like this is known as a *fold*, and we repeat for different folds with different training and testing partitions of the data. We select our training and testing sets based on the fact that we should achieve equal representation from all datasets including Dataset 1, which has the fewest samples (215). Specifically for each fold, we randomly choose 159 samples from each dataset to fit a model using 159 x 3 training data. Also, we have a testing set containing the remaining 56 samples from Dataset 1, the remaining 921 samples from Dataset 2, and the remaining 2001 samples from Dataset 3.We apply the method discussed in Section V-A to estimate the model coefficients from the *training set* for given factors, and then evaluate the performance error of the fitted model in the fold using the *testing set* as follows:

$$q_j = \frac{1}{3}\sum_{k=1}^{3}\left[\frac{1}{N_k}\left(\sum_{\substack{i\text{th packet loss}\\ \text{in testing set } k}} (p_i - p_i^{-})2\right)\right]$$

Where p is the predicted fraction of viewers who saw the packet loss and N is the number of samples in the testing set of dataset.

VI. RESULTS AND DISCUSSION

The aim is to develop an efficient packet dropping policy for the router. We propose the *perceptual-quality based packet prioritization* policy, denoted PQ, designed by our visibility model that prioritizes packets. At the server, we set a packet to be low priority when its visibility is less than 0.25, and high priority otherwise. The 1-bit high/low priorities can be signaled in the packet itself. The router can be, therefore, designed to drop packets of low priority to reduce traffic during network congestion.

The intermediate router with this capability is realizable in a DiffServ (Differentiated Services) network [10]. A one-bit prioritization scheme, called cMSE (cumulative MSE prioritization method), is designed. The cumulative MSE for a particular packet is measured by summing the MSE in all frames in a video affected by the packet drop, and a packet is assigned high priority if the cumulative MSE due to its loss is larger than a threshold and low priority otherwise.

TABLE 2: EXPERIMENTAL RESULTS

Codec	**Bit rate (Mbps)**	**Loss period (# of IP packets)**	**Acceptable average PER (Packet Loss w/zero retries)**
MPEG-2	15.0	24	<= 1.17 E-06
	17	27	<= 1.16 E-06
	18.1	29	<= 1.17 E-06

	8	14	<= 1.28 E-06
MPEG-4	10	17	<= 1.24 E-06
	12	20	<= 1.22 E-06

Max duration of an error event <= 16 ms; 1 error event per 4 hours
Max video/audio delay < 200/50 ms; max jitter < 50 ms

Furthermore the loss of an IP packet can mean the loss of a Packetized Elementary Stream header or a loss of a timestamp at the TS or Packetized Elementary Stream layer. The worst case for losing an internet protocol packet causes loss of 0.5 seconds worth of video.

(1 frame affected)

(14 frames affected)

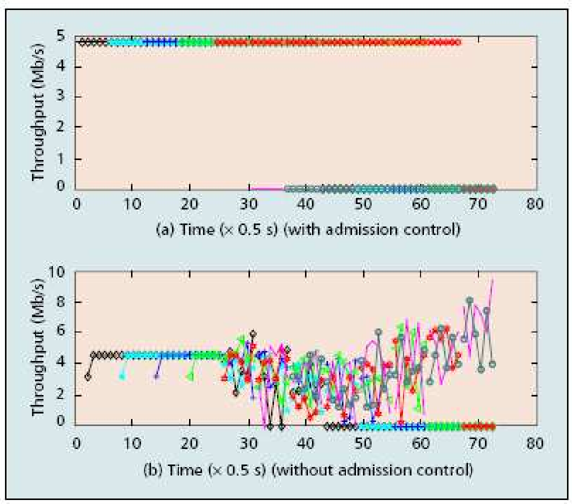

Fig.2 Throughput Variation

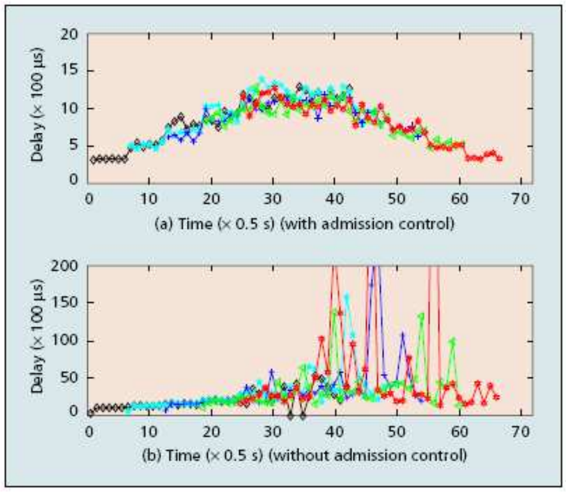

Fig.3 Delay Variation

VII. CONCLUSION

In this paper, we propose a generalized linear model to minimize the degradation in visual quality. The contributions of this paper are the following: (a) unlike earlier models, this visibility model is developed on datasets from multiple subjective experiments using different codecs, different encoder settings, and different decoder error concealment strategies. So the model has broad applicability. (b) We use our visibility model to prioritize video packets and design a policy for perceptual-quality based packet discarding. Although the model is designed for high-quality video transported over a mostly reliable network, the experiments show that the model performs well for videos with various encoding rates. (c) The analysis on packet loss rate across three different dropping policies shows that our policy achieves a better visual quality by dropping more, but perceptually unimportant, packets with smaller sizes. This emphasizes that evaluating video quality based solely on packet loss rate is inaccurate.

REFERENCES

[1] Yuan Zhi,David Poole, Pamela C. Cosman A "Versatile Model for Packet Loss Visibility and its Application to Packet Prioritization" IEEE Trans. Image Process, vo.1.19, no.3, Mar.2010
[2] T. Liu, Y.Wang, J. M. Boyce, Z.Wu, and H. Yang, "Subjective quality evaluation of decoded video in the presence of packet losses," Proc.IEEE ICASSP, pp. 1125–1128, Apr. 2007.
[3] S. Kanumuri, P. C. Cosman, and A. R. Reibman, "A generalized linear model for MPEG-2 packet-loss visibility," presented at the International Packet Video Workshop, Dec. 2004.
[4] S. Kanumuri, P. C. Cosman, A. R. Reibman and V. Vaishampayan, "Modeling packet-loss visibility in MPEG-2 video," IEEE Trans. Multimedia, vol. 8, no. 4, pp. 341–355, Apr. 2006.
[5] P. McCullagh and J. A. Nelder, Generalized Linear Models, 2nded.New York: Chapman & Hall, 1989.
[6] S. Kanumuri, S. G. Subramanian, P. C. Cosman, and A. R. Reibman, "Packet-loss visibility in H.264 videos using a reduced reference method," presented at the IEEE Int. Conf. Image Processing, Oct.2006.
[7] NS Project [Online]. Available: http://www.isi.edu/nsnam/ns/
[8] Y. Sermadevi and A. R. Reibman, Unpublished Subjective Test Results, 2002.
[9] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York: Springer-Verlag, 2001.
[10] J. F. Kurose and K. W. Ross, Computer Networking: A Top-Down Approach Featuring the Internet, 3rd ed. Reading, MA: Addison Wesley, 2004.

