VIDEO QUALITY AND TRAFFIC CHARACTERISTICS OF SKYPE VIDEO CALLS IN WAN

Kumara W. G. C. W.[1], Jayasekara J. M. N. D. B.[2], Gunarathne T.[3]

[1] MINE Lab, Department of Computer Science and Information Engineering, National Central University, Taiwan. Email: chinthakawk@gmail.com
[2] Faculty of Applied Sciences, Rajarata University of Sri Lanka, Sri Lanka. Email: nuwanjayasekara@gmail.com
[3] Department of Mechatronics, Faculty of Engineering, South Asian Institute of Technology and Medicine (SAITM), Sri Lanka. Email: tharanga.saitm@gmail.com

ABSTRACT

Long distance (WAN) voice and video calls have been used around the world for around many years now and are still being used intensively with the popularity of various hand held devices like smart phones, pads etc. and the expansion of high speed internet access technologies making it possible to call almost all around the world anywhere anytime. And many other applications are also increasingly use the voice and video over WAN like e-learning, video conferencing etc. But still there are many issues around the WAN video and voice calls due to changing traffic behaviors of the Internet. Here we present the results of an analysis of Skype video calls carried out over the Internet between Sri Lanka and Taiwan. Objective full reference and no reference analysis method were used during the analysis with the Foreman standard video sequence and it is found that the video call is not full satisfactory.

***Key words*:** video call, VoIP, Quality of user Experience (QoE)

1. INTRODUCTION

With the rapid growth of the Internet expansion in Sri Lanka in recent years, now use of Internet related applications is rapidly increased. As per the Statistical overview report – 2012 December from Telecommunications Regulatory Commission Sri Lanka (TRCSL) [9] we can clearly see that Internet subscribers both fixed and mobile has increased around 700% during last four years as shown in Figure 1. While Internet is being used for many applications mainly for browsing, email, entertainment etc., voice over IP (VoIP), video calls and video conferences are extensively used for business and personal communications, still Skype being one of famous applications for the voice/video calls.

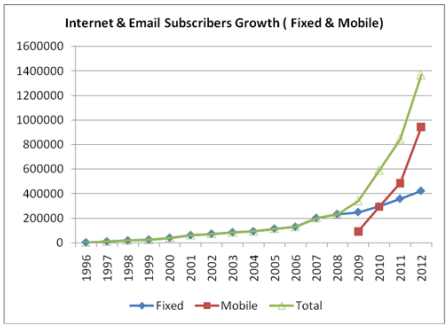


Figure 1: Internet and email subscribers' growth in Sri Lanka [9]

Many research efforts are carried out to analyze the Skype traffic characteristics and audio/video quality [1, 2]. Jing Zhu [1] has studied the on traffic characteristics and video quality of Skype video calls in a LAN simulating delay characteristics of a WAN.

Video quality assessment can be mainly carried out in two ways as subjective and objective, and this too has researched by researches for decades [3-8]. Subjective assessment uses human subjects and their experience about the quality of the video while objective assessments are basically use the mathematical based algorithms to evaluate the quality. Objective quality assessment then can be divided into three main categories as full reference (FR), reduced reference and no reference (NR). Full reference compares the received video with the original video, reduced reference compares only some characteristics and no reference use characteristics of the received video only [4].

Following are the seven parameters measured during the experiment in short [10].

- Peak-to-peak Signal-to-Noise Ratio (PSNR):

$$PSNR = 10 \cdot log_{10} \frac{MaxErr^2 \cdot w \cdot h}{\sum_{i=0,j=0}^{w,k} (x_{i,j} - y_{i,j})^2}$$

where, MaxErr = maximum possible

absolute value of color components difference, w = video width, h = video height

- SSIM (Structural similarity) Index is based on measuring of three components (luminance similarity, contrast similarity and structural similarity) and combining them into result value. We used the SSIM (fast) parameter only.
- MSU Blocking metric: This metric was created to measure subjective blocking effect in video sequence.
- MSU Blurring metric: This metric compares power of blurring of two images. If value of the metric for first picture is greater, than for second it means that second picture is more blurred, than first one.
- Delta: The mean difference of the color components in the correspondent points of image.

$$d(X,Y) = \frac{\sum_{i=1,j=1}^{m,n}(X_{i,j} - Y_{i,j})}{mn}$$

where *m* and *n* are the number of pixels in width and height of the frame.

- Mean Sum of Absolute Difference (MSAD): The mean absolute difference of the color components in the correspondent points of image.

$$d(X,Y) = \frac{\sum_{i=1,j=1}^{m,n}|X_{i,j} - Y_{i,j}|}{mn}$$

- Mean Squared Error (MSE):

$$d(X,Y) = \frac{\sum_{i=1,j=1}^{m,n}(X_{i,j} - Y_{i,j})^2}{mn}$$

Table 1 shows the reference method used during the experiment and the interpretation of the resultant parameter values [10].

Table 1: Interpretation of the test parameters [10]

Metric	(FR)/ (NR)	Interpretation
PSNR	FR	100 for equal frames, higher values are better
SSIM (fast)	FR	Higher values are better, 1 for equal frames
MSU Blocking metric	NR	Lower value corresponds to lower blocking.
MSU Blurring metric	NR (but 2 files are necessary in case of measure visualization)	Lower value corresponds to higher blurring
Delta	FR	0 means equal frames, positive and negative values mean deviation, lower absolute values are better
MSAD	FR	0 means equal frames, lower values are better
MSE	FR	Lower values are better, 0 for equal frames

2. METHODOLOGY

We used the FR and NR objective quality assessment procedure in this paper. Foreman standard AVI video file with 320×240 spatial resolution, 25 fps, 4:2:0 YUV color space and 12 s long was used.

Steps of the test procedure are given in Figure 2. First using the SplitCam[1] software the reference video file was fed into Skype and video was played continuously in a loop. In our study the sender was in Taiwan and the receiver was located in Sri Lanka and both participants were behind NAT and were using the private IP addresses. We didn't use any special parameter adjustments in network during the test; it was only a general Skype video call in Internet. At the receiver Supertintin[2] software was used to record the received video. 30 seconds waited before the record the 10 seconds long videos at the receiver letting network traffic to settle down and five videos were recorded and analyzed. Wireshark[3] software was used to capture all the packets exchanged during the call. At last received video was compared with the original (reference) video using the MSU Video Quality Measurement (VQM) Tool[4]. A mathematical explanation of the full reference video assessment is presented in Jing Zhu's work [1].

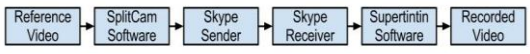


Figure 2: Steps of the test procedure

3. RESULTS

[1] http://www.splitcamera.com

[2] http://www.supertintin.com

[3] http://www.wireshark.org

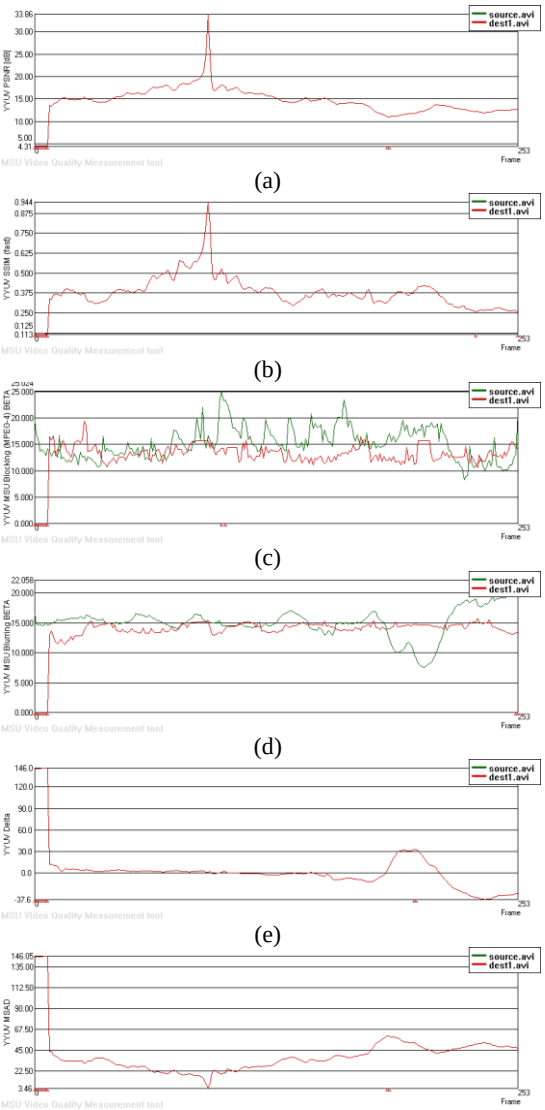
[4] http://compression.ru/video/quality_measure/video_measurement_tool_en.html

Table 2 provides the average parameter values provided by the MSU VQM tool during the experiment.

Table 2: Parameter values of the experiment

Parameter		Average value
PSNR		13.5247
SSIM (fast)		0.3826
MSU Blocking metric	Reference	15.0808
	Test	14.4812
MSU Blurring metric	Reference	15.0831
	Test	13.7794
Delta		0.9641
MSAD		38.2605
MSE		2956.5030

Figure 3 shows the example graphs provided by the MSU VQM tool for each parameter.



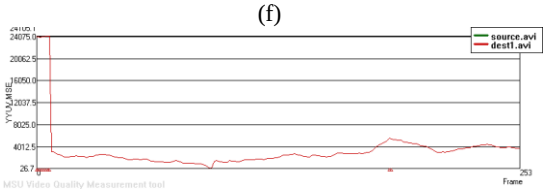
(a)

(b)

(c)

(d)

(e)



(f)

(g)

Figure 3: Performance graphs (a) PSNR (b) SSIM (fast) (c) MSU Blocking metric (d) MSU Blurring metric (e) Delta (f) MSAD (g) MSE

Figure 3 shows frames extracted from reference, test and the difference between the frames for a selected best match and a worst match for PSNR test. Relevant PSNR values for each frame are 33.8615 and 10.7984 respectively.

(a) (b) (c)

Figure 3: Selected frames for PSNR test (top: frame 91, bottom: frame 185) (a) reference frame (b) test frame (c) difference

Figure 4 shows the network traffic graph captured from the Windows Task Manger of the sender during the test. High traffic blocks shows the Skype call durations and changes slightly during Wireshark runs.

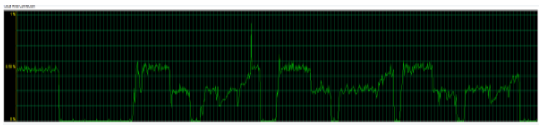


Figure 4: Network traffic of the sender

Average round trip times from the hosts in Taiwan and Sri Lanka to the visible public IPs of the peer is 212 ms and 293 ms respectively. Traceroute results showed that both visible public IP of the peer in other end is 18 hops away from each other in Taiwan and Sri Lanka respectively.

In analysis of the results, FR and NR objective video quality measurement provided very important results regarding the Skype QoE. If the reference frame and test frame is very similar to each other parameter values provide better results. Due to the various delays in the network and processing delays of the encoders/

decorders, frame distribution of the test video is different compared to the reference video as clearly shown in Figure 3. For the selected best and worst match it provides PSNR values of 33.8615 and 10.7984 respectively, confirming the mismatch of the frames.

Jing Zhu's [1] results were obtained in a laboratory setup while this this paper presents a real world experiment. Parameter values obtained in this experiment clearly shows the quality of the video call is not satisfactory comparing to the benchmarks provided by the MSU VQM tool as in Table 1. But still it is not fair to draw solid conclusions out of these results without an in deep analysis of the operations of the tool.

4. CONCLUSION

In this paper Skype video quality using the objective video quality assessment based tools was analyzed in the Internet between Taiwan and Sri Lanka. Video quality is low in the real environment than the values obtained during semi-simulated research results. Further researches in the video quality assessment of the WAN calls using Internet is encouraged for different video call applications like Skype and Google Hangout. Specially quality assessment of audio and video together with different parameters like spatial resolution, complex background information, more than two party calls etc. will contribute more to the development of the more user friendly advanced applications. Authors plan to study this further, specially analyzing the algorithms used in measuring the parameter values.

5. REFERENCES

[1] Jing Zhu, "On traffic characteristics and user experience of Skype video call," *Quality of Service (IWQoS), 2011 IEEE 19th International Workshop on*, vol., no., pp.1-3, 6-7 June 2011

[2] Dario Rossi, Marco Mellia and Michela Mio, "A detailed measurement of Skype network traffic", *The 7th international workshop on Peer-to-Peer systems (IPTPS 2008)*, Feb 2008

[3] José Luis Martínez, Pedro Cuenca, Francisco Delicado and Francisco Quiles, "Objective video quality metrics: A performance analysis", *14th European Signal Processing Conference (EUSIPCO 2006)*, Sep 2006

[4] Iain E. Richardson, *The H.264 Advanced Video Compression Standard*, 2nd Ed., Wiley, 2010, ISBN 0470989289, 9780470989289

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

[6] Chapter 41 in *The Handbook of Video Databases: Design and Applications*, B. Furht and O. Marqure, ed., CRC Press, pp. 1041-1078, September 2003

[7] Tutorial on Objective perceptual assessment of video quality: Full reference television, ITU-T, 2004, http://www.itu.int/pub/T-TUT-OPAVQ-2004-FRT

[8] Digital Video Quality: Vision Models and Metrics, Stefan Winkler, Illustrated Ed., Wiley, 2005, ISBN 0470024046, 9780470024041

[9] Statistical overview report – Dec 2012, Telecommunication Regulatory Commission Sri Lanka (TRCSL), http://www.trc.gov.lk/statistics/statistical-overview.html, visited on 10 Jan 2013

[10] MSU Video Quality Measurement Tool, http://compression.ru/video/quality_measure/video_measurement_tool_en.html, visited on 10 Jan 2013