International Journal of Tomography and Simulation [ISSN 2319-3336];
Year: 2013, Volume: 22, Issue Number: 1;
[Formerly known as "International Journal of Tomography & Statistics"
(ISSN 0972-9976; 0973-7294)];

Aligned-PSNR (APSNR) to Improve Accuracy of Conventional PSNR for Video Stream over Wireless and Mobile Network

Herman.[1], Azizah Abdul Rahman.[2], Yoanda Alim Syahbana.[3]

Universiti Teknologi Malaysia
Faculty of Computer Science and Information System
Johor, Malaysia
[1]herman8@live.utm.my, [2]azizahar@utm.my, [3]asyoanda2@live.utm.my

ABSTRACT

Peak Signal-to-Noise Ratio (PSNR) is one of widely used method to measure video quality. This method is inherent with inaccuracy issue while applied to measure video quality that streamed over wireless and mobile network. Some methods have been proposed in order to tackle the issue. Modified-PSNR (MPSNR) is one of the proposed methods that show better performance than conventional PSNR. However, MPSNR is inappropriate if there are more than 5 consecutive frame losses due to its fixed window concept. This paper proposes a concept of dynamic window size, called Aligned-PSNR (APSNR). Experiment has been conducted to evaluate performance of the APSNR toward the conventional PSNR. Results from these two methods are benchmarked with Subjective Video Quality Assessment (SVQA) by utilizing Pearson product-Moment Correlation Coefficient (PMCC). Based on PMCC calculation, APSNR shows better performance (0.92) than conventional PSNR (0.89) in representing SVQA result.

Keywords: Objective Video Quality Measurement (OVQM), Peak Signal-to-Noise Ratio (PSNR), video streaming, wireless and mobile network, Pearson product-Moment Correlation Coefficient (PMCC).

Mathematics Subject Classification:

Computing Classification System:

1. INTRODUCTION

Emerging development of mobile communication technology has encouraged many new and creative mobile services. The trend is towards streaming-based service that enables customer to experience multimedia content, instantly as they requested. The service itself is evolved into variety of services such as video call, video teleconference, IP television (IPTV), and Video on Demand (VoD). As more and more request for the service, quality of the service is essential aspect that needs to be considered.

Focused on VoD service, the service quality is determined by quality of three service components. There are quality of video that will be streamed over the network, quality of network that streams the video, and capability of device that receives the streamed video. However, among these service components, quality of the streamed video is believed as the most influencing factor (Winkler and

Mohandas, 2009). Therefore, the essentiality of quality in the service is also mean the essentiality of streamed video quality.

Due to this essentiality, quality of streamed video can be defined as Quality of Service (QoS) and Quality of user Experience (QoE). QoS is defined as how well the streamed video is deliveredto mobile device. It is influenced by video encoding, network performance, and capability of mobile device. Good encoding process, excellent network performance, and advanced mobile device capability result in high QoS. On the other hand, QoE deals with how well the service quality is perceived by user. It is influenced by the QoS and subjective factors, either intrinsic or extrinsic (Herman et al., 2011). In line with QoS and QoE, evaluation on streamed video quality can be carried out using two approaches. They are Subjective Video Quality Assessment (SVQA) and Objective Video Quality Measurement (OVQM). SVQA assesses video quality based on customer’s perceived quality that is gathered through survey. It provides high accuracy result since it is straightly derived from the customer’s opinion. However, due to time and resources required for the survey, SVQA is inefficient (Piamrat, 2009; Reiter and Korhonen, 2009; Khan, 2009). To address these problems, OVQM is introduced as a method for predicting the customer’s opinion. OVQM measuresvideo quality parameters such as error of pixel and video frame quantitatively.

This paper proposes Aligned Peak Signal-to-Noise Ratio (APSNR) as one of OVQM methods for measuring video quality that streamed over wireless and mobile network. To discuss the method, this paper is organized into seven sections including this first introduction section. Problem background and related work is discussed in the next section. Section three explains the proposed APSNR method. Experiment and result are provided in the fourth section followed by result analysis in fifth section. Finally, conclusions and future work are summarized in the sixth section and closed by list of reference in the seventh section.

2. BACKGROUND AND RELATED WORKS

Video Quality Expert Group (VQEG) categorizes PSNR method as Full Reference (FR) methodto differentiate it with Reduced Reference (RR) and No Reference (NR) method. It requires existence of reference and distorted video. Therefore, application of PSNR is in out-of-service measurement.PSNR is formulated as (Vranjes, et al., 2008),

$$PSNR = 10log_{10}\frac{(2^{M}-1)^2}{MSE}. \qquad (1)$$

Where 2^M-1 is maximum pixel for M-bit video frame while MSE is abbreviation of Mean Square Error. For a video with frame size of (x, y) pixels,

$$MSE = \frac{1}{XYT}\sum_{t=1}^{T}\sum_{y=1}^{Y}\sum_{x=1}^{X}[p(x,y,t)-p'(x,y,t)]^2. \qquad (2)$$

The p(x, y, t) represents the reference video frame and the p'(x, y, t) represents the distorted video frame. Substituting Eq. (2) into Eq. (1) results on PSNR value for the distorted video. High value of PSNR in decibel indicates high video quality while the low value represents low video quality. Table 1 maps PSNR interpretation towards video quality.

Table 1: PSNR interpretation of video quality (Gross et al., 2004).

PSNR Value (dB)	Video Quality
> 37	Excellent
31 – 37	Good
25 – 31	Fair
20 – 25	Poor
< 20	Bad

PSNR has been used to evaluate performance of video encoder by comparing encoded video towards original video. While applied in video stream service over wireless and mobile network, PSNR is used to measure degradation of streamed video quality caused by streaming process. It also can be applied in video stream simulation to predict customer's opinion towards streamed video quality.

Data transmission using wireless and mobile network is characterized by high possibility of packet loss. If the packet loss occurs in packet that contains frame pixel, then spatial error will occur that bring about blank pixel or incorrect pixel. In other case, if the packet loss occurs in packet that contains header of video frame, then temporal error will occurs as frame loss.

The frame loss is serious problem while the video is encoded with inter frame coding such as in H.264. Inter frame coding generates video frame based on one or more neighboring frames. If one of the neighboring frames is lost, video frame that depends on the lost frame also will be lost. Thus, it generates consecutive frame losses. In relation to video content characteristic, video with less motion such as news report video is encoded with less key frame. The key frame has many depended neighboring frames. If the packet loss hits packet that contain header of the key frame, then the video will suffer with more consecutive frame losses. In consequence, the customer will see jerky news video.

Due to the frame losses, total frame in streamed video will be less than in original video. As consequence, while PSNR is applied, PSNR compares non-corresponding frames between streamed video and original video. Fig. 1 illustrates this circumstance.

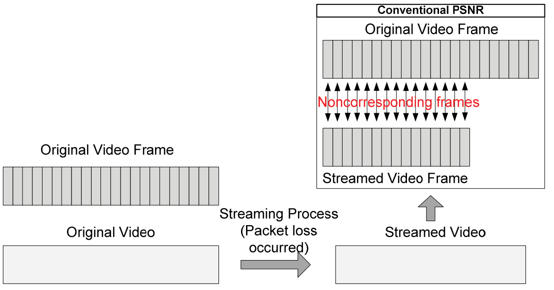

Figure 1. Illustration of conventional PSNR.

This inaccuracy issue has been studied by An (Jack) Chan et al. (2010). They proposed MPSNR to cope with non-corresponding frame issue. MPSNR is inspired from research by Wolf and Pinson (2009). It is based on assumption that the corresponding frame pair has higher PSNR value than non-corresponding frame pairs.MPSNR is also similar to sequence alignment problem in bio informatics that aligns DNA or RNA sequence (Mount, 2004).

In its process, MPSNR defines searching boundary called *window* as shown in Fig. 2. It consists of streamed video frame (Stre-frame) and original video frame (Ori-frame). In the *window*, MPSNRmatchesStre-frame against Ori-frames until it findshighest PSNR value among the pairs.MPSNR can optimize its frame searching only on definite *window*. Therefore, without searching on entire original video frames, it will require less measurement time. In its experiments, An (Jack) Chan et al. (2010) defines fixed *window* with size that consists of five frames.

The fixed *window*comes up with an issue as demonstratedin Fig. 2. The figure shows case of frame loss where frame number three to seven in Ori-frame do not have their pair in Stre-frame (five sequential frame loss in Stre-frame). In other word, Stre-frame number x should be paired with Ori-frame number eight. By defining *window* size of five, MPSNR fails to include Ori-frame number eight in its corresponding frame searching. In consequences, the highest PSNR value will not be obtained from the actual corresponding pair. This issue also will impact the next Stre-frame.

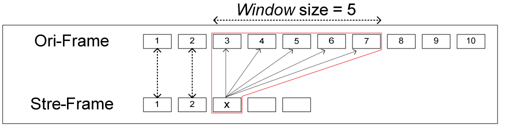

Figure 2. Possibility issue with fixed*window* size.

3. PROPOSED ALGORITHM

This paper proposes APSNR to overcome the MPSNR issue. APSNR uses dynamic*window* size, instead of fixed *window* size as in MPSNR. APSNR defines the dynamic *window* size as

$$w = sumFL + 1. \quad (3)$$

Where, w is*window* size and sumFL is total of frame loss. sumFL is obtained by subtracting total of Stre-frame from total of Ori-frame. Fig. 3 shows *window* size that defined by APSNR for previous MPSNR sample case. The *window* size is defined as six frames since there are total of five frame losses. Therefore, the correspondences between Stre-frame number x and Ori-frame number eight will be found.

The dynamic *window* has time inefficiency issue. If there are large frame difference between Ori-frame and Stre-frame, APSNR will need more time to find highest PSNR value among the pairs. To deal with this issue, APSNR decreases *window* size as much as number of skipped Ori-frame. As shown in Fig.3, Ori-frame number three to seven will be skipped. Then, *window* size of one frame will be used by next processed Stre-frame.

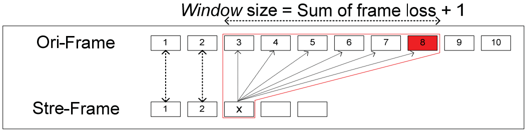

Figure 3.*Window* size in APSNR.

APSNR algorithm is divided into three phases. There are searching, shifting, and calculating. Searching phase is started by determining size of *window*. As previously mentioned, the size is obtained from difference between total frame in the original video and total frame in streamed video. Then, it starts with one of streamed video frame. The frame is paired to every original video frame. After that, the searching phases will calculate PSNR value for all the pairs. Once the calculation is finished, APSNR searches for highest PSNR among the results. If the highest PSNR is founded in the first pair of the *window* then it means no frame loss in current searching *window*. Otherwise, if the highest PSNR is not founded in the first pair then APSNR will do the shifting phase. The shifting phase will skip number of original video frame and mention the skipped frames as frame loss position of streamed video. The process is continued to the next streamed video frame and these searching and shifting phasesare repeated. This first and second phase is illustrated in the Fig. 4. When the entire streamed video frames have been processed, APSNR do its calculating phase. It will calculate average value of PSNR result and mentioned as streamed video quality.

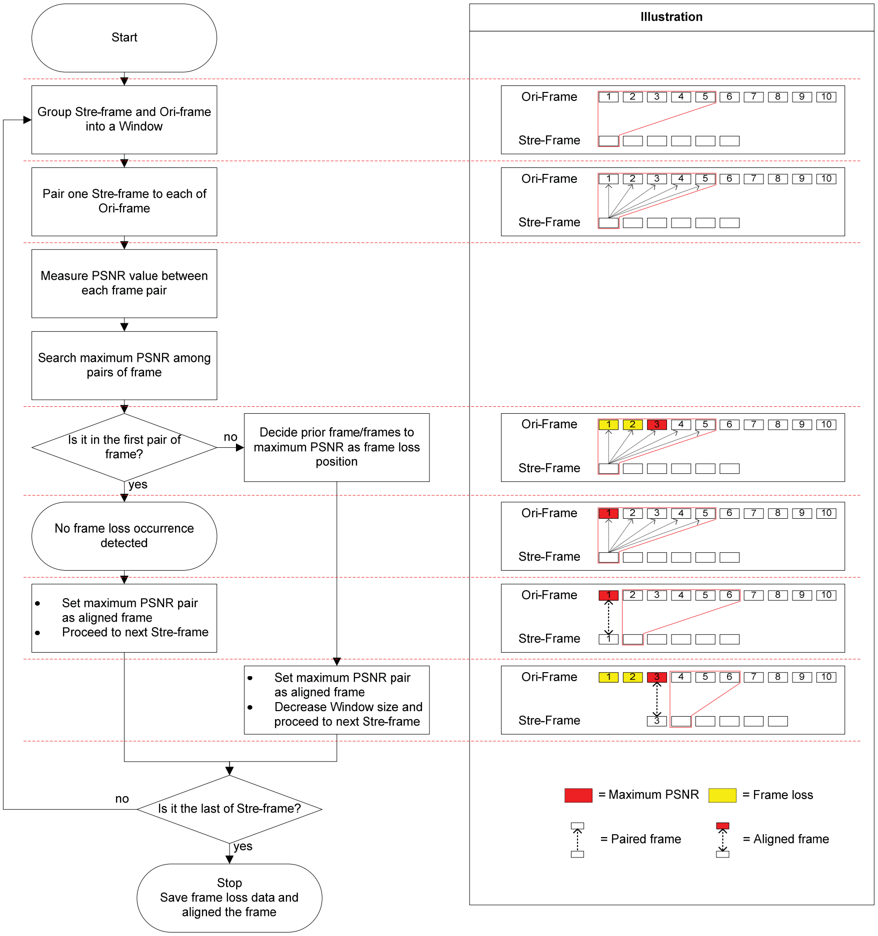

Figure 4. First and second phase of APSNR algorithm.

4. EXPERIMENTS AND RESULTS

4.1 Experiment Component

In order to simulate video stream over wireless and mobile network, the experiment is run under NS2 2.28 simulation environment. For video stream simulation, it uses MYEvalvid_RTP that is proposed by Chia-Yu Yu et al. (2007). This streaming simulation is enhanced version of EvalVid (Klaue et al., 2003)with supportfor audio streaming. The experiment also uses HSDPA network infrastructure based on European Commission 5th framework project SEACORN, EURANE. The HSDPA network infrastructure is depicted in Fig. 5. The experiment set fixed bandwidth between Radio Access Network (RNC), Serving GPRS Support Node (SGSN), Gateway GPRS Support Node (GGSN), router and video server.In order to varying the network condition, it varies channel condition between

Node B and mobile device. It uses three value of CQI, 8, 15, and 22, as experimented by Brouwer et al. (2004). These values represent channel condition that perceived by mobile device. CQI sums up influence of unpredictable factor (e.g. mobility, fading, path loss, and interference). Large CQI value represents excellent channel condition and conversely, small CQI value indicates poor channel condition.

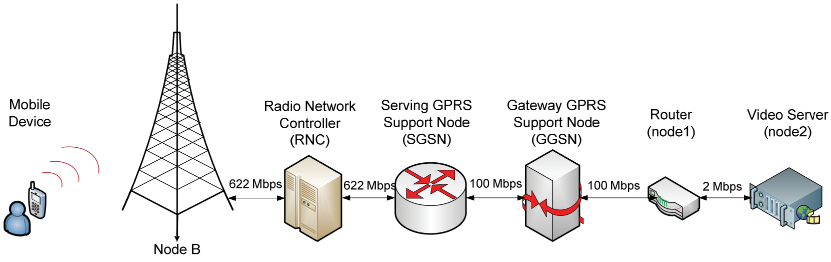

Figure 5. Network infrastructure of EURANE.

The experiment uses “ANSI T1.801.01 vtc1nw” video as master video. It depicts woman sitting and reading news report for twelve seconds. This content is characterized by low motion videonad static background. This master video is configured into twelve original videos by varying its frame rate (15 and 30 fps), video bitrate (96 and 512 kbps), and resolution (QCIF, CIF, and 4CIF). Table 2 shows these configurations in relation to Video Index (VI) that will be discussed in the following section. At the end of entire simulation process, there are total 36 streamed videos that will be measured by APSNR and PSNR.

*Table 2:*Summary of video configuration.

Video Index (VI)	Resolution	Video Rate (kbps)	Frame Rate
1	4CIF	512	30
2			15
3		96	30
4			15
5	CIF	512	30
6			15
7		96	30
8			15
9	QCIF	512	30
10			15
11		96	30
12			15

4.2 Experiment Result

Under variation of network condition as explained in previous section, the original video that streamed over the network is influenced by variety of packet loss. In short, larger CQI experiences less packet loss than smaller CQI. This packet loss influences frame loss occurrence in the streamed video. From the experiment, the packet loss occurrence results on five to twenty-one consecutive frame losses which are inappropriately if measured by MPSNR. Beside this, the packet loss also results on spatial quality degradation. The trend shows that streamed videos with high configurations suffer more quality degradation than streamed video with low configurations. Moreover, while the CQI value is lowered, the streamed video will suffer more quality degradation.

Based on this streamed video, the videos are measured by APSNR and PSNR method. Fig. 6, Fig. 7, and Fig. 8 show the measurement results.

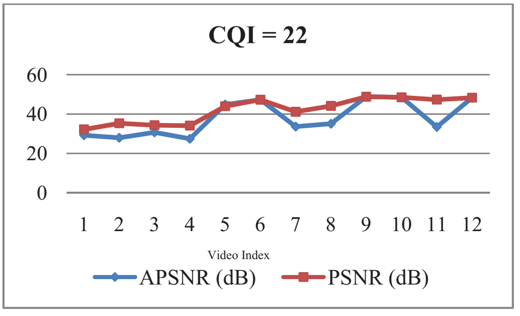

Figure 6. Experiment result for CQI value of 22.

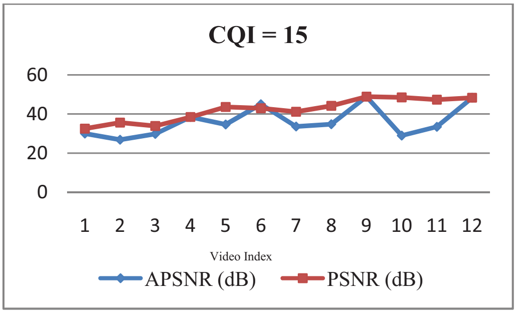

Figure 7. Experiment result for CQI value of 15.

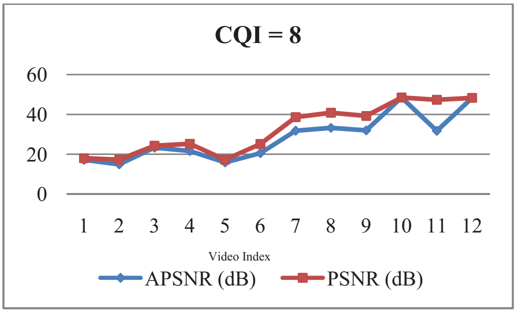

Figure 8. Experiment result for CQI value of 8.

5.RESULT ANALYSIS

Research by An (Jack) Chan et al.has mentioned approach to analyze accuracy of the result. They use SVQA resultas a benchmark. SVQA has five Mean Opinion Score (MOS) scales in representation of streamed video quality. MOS scalesare shown in Table. 3.

Table 3: MOS interpretation in terms of streamed video quality

MOS	Description
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

The analysis measures how well the OVQM represents user opinion from SVQA.It uses Pearson product-Moment Correlation (PMCC) to compare result of OVQM and result of SVQA. PMCC is formulated as

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i{}^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2}\sqrt{\sum_{i=1}^{n} y_i{}^2 - \frac{1}{n}(\sum_{i=1}^{n} y_i)^2}}. \quad (4)$$

x and y are result of OVQM and SVQA while n is total of evaluated result. r is the PMCC value that states accuracy of OVQM towardSVQA. r will be founded in range of zero to one which represent how accurate the OVQM.While the result is near to one, then the OVQM method has high accuracy in predicting the SVQA. Conversely, if the PMCC result is near to zero, the OVQM method has low accuracy.

SVQA has been conducted. Twelve people participate in the survey. The survey is conducted in two sessions namely training session and actual session. In training session, the participant is explained about the SVQA process. Then, SVQA process is conducted in actual session. Sequence of streamed video is shown to the participants and the participants are asked to determine quality of the streamed video in form of MOS scale.

Considering nonlinearity between PSNR value, MOS scale, and video quality interpretation, Table 4 shows equivalent conversion from PSNR value to MOS scale. Thus, before calculating PMCC between the experiment result (APSNR and PSNR) and SVQA result, the experiment result should be normalized into MOS scale.

Table 4: PSNR interpretation in terms of streamed video quality and MOS

PSNR Value	Video Quality	MOS
> 37	Excellent	5
31 - 37	Good	4
25 - 31	Fair	3
20 – 25	Poor	2
< 20	Bad	1

PMCC calculation between normalized APSNR result and SVQA result shows value of 0.92. As previously explained, this PMCC result shows high accuracy of APSNR in predicting the customer's opinion. Since the objective of this study is to improve PSNR accuracy in context of study, the APSNR result is also benchmarked towards the conventional PSNR result. At first, the normalized PSNR result and SVQA result is measured by PMCC calculation. It shows 0.89 as PMCC value. Then, by comparing PMCC value between APSNR and PSNR, it can be concluded that APSNR has better accuracy than the conventional PSNR.

6. CONCLUSIONS AND FUTURE WORK

This paper has proposed algorithm, APSNR, to overcome conventional PSNR limitation in measuring quality of video that is streamed over wireless and mobile network. Simulation of video stream that generates the video test material has been conducted. Then, the video test is measured by APSNR and PSNR method. To analyze accuracy of the APSNR, the video is also assessed by SVQA method. By calculating PMCC between APSNR and SVQA result, it can be concluded that APSNR has high accuracy in predicting customer's opinion. Result of APSNR method is also compared with accuracy of conventional PSNR. The result shows that APSNR has better accuracy than the conventional PSNR.

For further research, APSNR can be improved by considering region of interest from the customer (saliency awareness). However, it should be noted that this further improvement should keep simplicity of the measurement method. Since, there is a trade-off between feature of OVQM method in detecting many video characteristics and complexity of the OVQM method. It also can be noticed that the video quality measurement is still concerns only on the video quality. This research can be enhanced by considering audio quality in the quality measurement. It is due to customary combination of audio and video that streamed through wireless and mobile network in mobile-VoD. Therefore, there will be more accurate measurement for quality of the video stream service.

7. REFERENCES

An, C., Kai, Z., Mohapatra, P., Sung-Ju, L., Banerjee,S., 2010, *Metrics for Evaluating Video Streaming Quality in Lossy IEEE 802.11 Wireless Networks*, Proceedings IEEE INFOCOM, 1-9.

Bradeanu,O.,Munteanu,D.,Rincu, I.,Geanta, F., 2006,*Mobile Multimedia End-User Quality of Experience Modeling*, ICDT'06 International Conference on Digital Telecommunications.

Brouwer,F.,de Bruin,I., Silva,J. C.,Souto,N.,Cercas,F.,Correia,A., 2004,*Usage of Link-Level Performance Indicators for HSDPA Network-Level Simulations in E-UMTS*, Eighth International Symposium on Spread Spectrum Techniques and Applications, 844-848.

Chia-Yu, Y., et al., 2007,*MyEvalvid_RTP: A New Simulation Tool-set Toward More Realistic Simulation,* Future Generation Communication and Networking (FGCN), 90-93.

Engelke, U., Kusuma, T. M., Zepertnick, H.-J, 2006, *PerceptualQualityAssessment of Wireless Video Applications.*ITG Source and ChannelCoding.

Ennaji,Y.,Boulmalf, M.,Alaoui,C., 2009,*Experimental Analysis of Video Performance over Wireless Local Area Networks*, International Conference on Multimedia Computing and Systems (ICMCS), 488-494.

Film by NTIA/ITS (National Telecommunications and Information Administration/The Institute for Telecommunication Sciences (ITS). Available at http://www.cdvl.org/find-videos/details.php?id=18.

Gross, J.,Klaue, J., Karl, H.,Wolisz,A., 2004,*Cross-layer optimization of OFDM transmission systems for MPEG-4 video streaming*, Computer Communications, 1044–1055.

Herman, Rahman, A. A., Yoanda Alim Syahbana, Y. A., Bakar, K. A., 2011, *Nonlinearity Modelling of QoE for Video Streaming over Wireless and Mobile Network*, Second International Conference on Intelligent Systems, Modelling and Simulation, 313-317.

Huynh-Thu, Q., Ghanbari, M., 2008, *Scope of Validity of Psnr in Image/Video Quality Assessment*, Electronics Letters Vol. 44 No. 13.

Khan, Lingfen, S.,Ifeachor, E., 2009,*Content Clustering Based Video Quality Prediction Model for MPEG4 Video Streaming over Wireless Networks*, International Conference on Communications (ICC), 1-5.

Klaue,J.,Rathke, B.,Wolisz,A.,*EvalVid – A Framework for Video Transmission and Quality Evaluation*, Proc. of the 13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation, 255-272.

Kwangjin, C., Jun,K. C., Jae, H. H., Gyeong, J. M.,Jongkuk, L, 2008, *Comparison of Video Streaming Quality Measurement Methodologies*,10th International Conference on Advanced Communication Technology (ICACT), 993-996.

Mount, D. M., 2004,*Bioinformatics: Sequence and Genome Analysis*. Cold Spring Harbor Laboratory Press.

Piamrat,K., Viho, C., Bonnin, J. -M., Ksentini,A., 2009,*Quality of Experience Measurements for Video Streaming over Wireless Networks*, Sixth International Conference on Information Technology: New Generations, 1184-1189.

Reiter, U., Korhonen, J., 2009,*Comparing Apples and Oranges: Subjective Quality Assessment of Streamed Video with Different Types of Distortion*, International Workshop on Quality of Multimedia Experience (QoMEx), 127-132.

Rohaly, M., et al., 2000,*Video Quality Experts Group: Current Results and Future Directions,* SPIE Visual Communications and Image Processing. Perth, Australia, 742-753.

SEACORN project.Available at http://eurane.ti-wmc.nl/eurane/. 2004.

Winkler, S., Mohandas,P., 2008,*The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics*, IEEE Transactions on Broadcasting, 660-668.

Wolf, S., Pinson,M. H., 2009, *Reference Algorithm for Computing PeakSignal to Noise Ratio (PSNR) of a Video Sequence with a Constant Delay*.