Journal of Knowledge Information Technology and Systems(JKITS), Vol. 9, No. 5, pp. 579~587, October 2014

Journal of Knowledge Information Technology and Systems

ISSN 1975-7700

http://www.kkits.or.kr

QoE Analysis based on Inverse Filtering For Video Quality Assessment

Yogiraj S. Jadhav[1], **Jungdae Kim**[2], **Jinsul Kim**[*1]

[1]*School of Electronics and Computer Engineering, Chonnam National University*
[2]*Korea Nazarene University*

ABSTRACT

The quality of experience (QoE) can assess the video quality with the service parameters and their relative importance at the network layer. The service providers are also examining the level of multimedia service quality to a particular aspect hence, internet service provider is also willing to provide all the services through a single line. Considering the above facts the factors associated with the user perceived experiences are responsible for the development of such digital systems, For analyzing the user perceived experience the QoE parameters play an important role. In this paper, considering QoE factor we have determined an approach that helps to estimate the user's QoE, which uses a full reference model. This method uses an image restoration technique like inverse filtering, which improves the overall appearance of the video. It also helps to evaluate factors that measure the video quality effectively.

KEYWORDS: QoE, Metrics, Modules, Refine, Estimations

ARTICLE INFO: Received 10 August 2014, Revised 10 October 2014, Accepted 10 October 2014.

1. Introduction

The multimedia services like IPTV, VoD (Video-on-Demand), Social TV, Smart TV are merged into internet since they have come into existence. In day to day life the use of video streaming services with different systems has become a key factor, as the technology is tending more towards the development of digital systems.

[*]Corresponding author is with the School of Electronics& Computer Engineering, Chonnam National University, Gwangju, 500-757 Republic of Korea.
E-mail address: jsworld@jnu.ac.kr

Among such systems or devices, IPTV, smart TV, Social TV systems are more popular now days, because of its prime factors like voice calling, video calling, SMS service, Internet, entertainment, e-commerce service, email service, a social network service, etc. Use of such digital systems is becoming more and more common now days. The heavy usage of such systems has also increased the demand of bandwidth, especially, in the real time environment. The widespread use of video streaming services is a challenging proposition due to the use of such systems with various applications. These multimedia services are accessed with an integrated network environment, which have introduced the concept of quality of experience (QoE) in the network area. The measurement of a quality for subscriber's satisfaction in real time environment still faces an issue of certain network parameters. The QoS parameters like delay, jitter, loss and bandwidth affect the end user's quality needs. Also, service quality parameters like cost, reliability, fidelity, availability, usability , etc. affects the performance of end user's video experience.

In such a context, the user's QoE considered the most important parameter to initiate an interaction with such digital systems. There are many approaches that determine the user's QoE by estimating the performance of the device or network from the user experience. These approaches are based on either multicast or unicast technologies, generally the live broadcast service uses the multicast technology and video-on-demand service uses the unicast technology. The digital systems are more focused on estimating the video performance. The video quality primarily focuses on the mean opinion score, i.e. MOS, peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) values.

In this paper, the approach uses a full reference model that uses the reference video content to compare with the transmitted video content over the network. The transmitted video contents are analyzed that determines the difference between it and reference video contents, In turn to calculate the QoE factors like peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and MOS values to determine the video quality from the user's point of view. The method uses an image restoration technique, i.e. Inverse filtering, which is a form of high pass filtering, the video frames are blurred with a low pass filter, then to get the improved version of it they are imposed with a high pass filter. This approach gains an extensive video quality assessment.

2. Related Work

Early, digital systems were with limited success as the creation of the shared connections with a remote control and the user interface design made interaction disruptive to the TV experience [1] [2]. The video streaming services have made the digital systems very feasible, because of its consistent quality of experience for the users over the network. There are many applications developed later on with different functionalities for the users.

There are two existing video quality assessment technologies subjective quality assessment and objective quality assessment. The subjective assessment is carried by a certain selected group of viewers, which are gathered together for rating the video quality. It is most convenient and reliable way of video assessment. The disadvantage of this method is that the video rated by these group members will have different scores as per the perception of individual regardless of perfection. Also, this method requires a special kind of environment and equipment, with the need of crowd of people to participate in the test. On the other hand, the objective method uses the estimation software to analyze the quality of video signals and produce relevant results based on it. The advantage of this method is the results obtained are close to standard value marks and there is no need of special environment or bunch of mass of people to assess the video quality. The peak signal-to-noise ratio (PSNR) is widely used for the objective video quality assessment. For numerical prediction of objective metrics the structural similarity index (SSIM) is used that correlates the subjective measurement with the objective ones.

Currently, developed digital systems are built using the priority based queuing designs which are prone to traffic over the network. Another, problem we can consider is that these systems use lots of bandwidth that are unable to balance the load on the system. The current digital systems are based on the non reference model, the video quality is computed without comparing with the original source. The problem here is that the system cannot identify the video quality produced by transmitted source is better or not. As these types of systems are helpful in real time video streaming. Also, the previous approach was not able to produce the video quality as there was some blockiness and low contrast level in it [2]. Thus, the proposed approach uses the full reference model with Inverse filtering that improves the video quality by enhancing the contrast levels and reducing blockiness using a high pass filter to determine the characteristics of a transmitting source with a reference source to produce better results. Also, after getting the enhanced video frames through inverse filtering it computes the PSNR, SSIM, MOS values for determining the video quality.

3. System architecture

The system uses a full reference model, which is a bilateral approach that compares the source video with the transmitted video one frame by another frame to determine the video quality. The video quality assessment through it produces higher accuracy. The figure shows the architecture of the proposed system. The system uses a reference source and transmitted source as input to it. The transmitted video source is converted in several video frames then, these frames are further processed. The temporal and spatial features are extracted from these processed frames and the information related to each frame is maintained like frame size, frame rate, pixel values, etc. After extracting the features related to

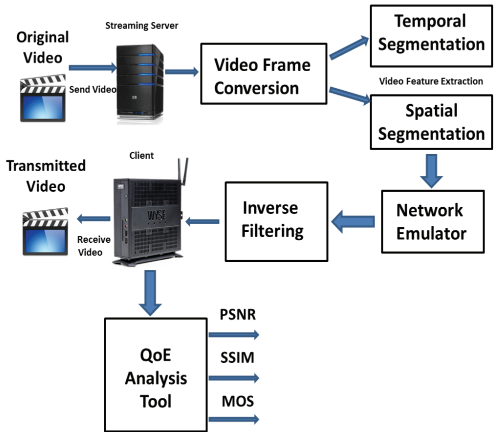

그림 1. 시스템 아키텍처

Figure 1. System Architecture

video frames the inverse filter is applied to each of these frames to enhance the contrast levels and reduce the blockiness, this happens because the lower values tend to become lower and the higher values tend to become higher. The QoE analysis tool uses the reference video to compare with the transmitted video for measuring the video quality. By computing the metrics like PSNR, SSIM, MOS values the video quality of the transmitted source can be determined.

4. System functionality

The figure 2 shows the data flow of the proposed system, Firstly, the choice variable is initialized, then the videos are converted to frames. These video frames are extracted for the temporal and spatial features to obtain the details of it. Thereafter, The inverse filter, i.e. a form of high pass filter is imposed on each frame one by one. The frames are blurred with a low pass

CREATE a Variable Choice=0

//Inverse Filtering

READ Video Frames V_{F1} and V_{F2} of a reference and a transmitted sources.

APPLY Blurring with a Low Pass Filter b(V_{F1} ,V_{F2}).

//Restoring the video frames

TAKE Convolution with a High Pass Filter h(V_{F1} ,V_{F2}).

IF (Choice == '1') THEN

READ Video Frames V_{F1} and V_{F2} of a reference and a transmitted sources.

//Compute Peak signal-to-noise ratio (PSNR)

Calculate the MSE = $(V_{F1} - V_{F2})^2$ and take a square of it.

Take a logarithm of computed MSE

Calculate PSNR.

END IF

IF (Choice == '2') THEN

READ Video Frames V_{F1} and V_{F2} of a reference and a transmitted sources.

//Compute Structural Similarity Index (SSIM)

SET variables $K_1 = 0.01$, $K_2 = 0.03$, $L = 255$

Calculate the average of video frames V_{F1}, V_{F2} as μ_{F1} and μ_{F2} respectively

Calculate the variance σ_{F1}^2 and σ_{F2}^2 of video frames V_{F1}, V_{F2} respectively

Calculate the covariance σ_{F12} of video frames V_{F1} and V_{F2}.

Calculate SSIM.

END IF

Determine MOS values with respect to computed PSNR and SSIM values.

END

그림 2. 시스템 기능

Figure 2. System functionality

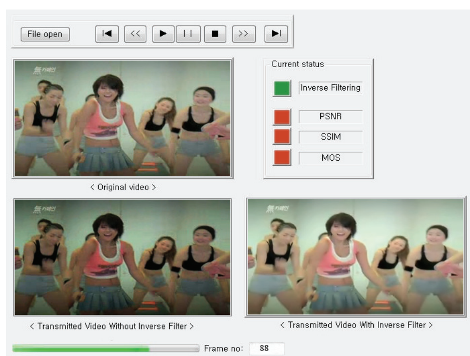

그림 3. 시스템 구현

Figure 3. Implementation of the system

filter and to obtain enhanced results the convolution of blurred function is taken with a high pass filter. The equation for the inverse filtering is as follows, Blurring function,

$$g(V_{F1},V_{F2}) = f(V_{F1},V_{F2}) * b(V_{F1},V_{F2}) \quad (1)$$

f(VF1,VF2) The original video frames that are given as input to the system, b(VF1,VF2) is a low pass filter and g(VF1,VF2) are blurred video frames. To get the enhanced results the convolution of the above equation is taken with a high pass filter h(VF1,VF2) as follows,

$$f(V_{F1},V_{F2}) = g(V_{F1},V_{F2}) * h(V_{F1},V_{F2}) \quad (2)$$

After getting the enhanced video frames, the system gets an input choice from the user to compute one of the video metric, i.e. PSNR or SSIM after that the MOS values are calculated for predicting the video quality of the transmitted source. After getting choice from the user, system reads the video frames from a reference and the transmitted sources respectively. If the user makes first choice, then the PSNR value is computed, first the mean square error value (MSE) between a reference and the transmitted video frame is computed. After getting the value of MSE, the PSNR value is computed by taking the logarithm of it. If the user makes a second choice, then the system reads video frames of a reference and the transmitted source. After reading the video frames the constant variables in the system are set to default values. The system then computes the average of these video frames. The variance and covariance of the video frames are computed. After getting all the associated values the SSIM of these frames is computed to determine the similarity between these video frames for better quality. The MOS values are determined with respect to PSNR and SSIM values.

The equation for computing the PSNR value is,

$$MSE = (V_{F1}-V_{F2})^2 \quad (3)$$

$$PSNR = 10 * \log10\ (255)^2 / MSE \quad (4)$$

The peak signal-to-noise (PSNR) is a ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation [1] [2]. For computing the PSNR value we first need to compute the MSE value. The equation (3) computes the MSE value. In the above equation VF1 is a reference video frame and VF2 is the transmitted video frame .From equation (3) by substituting the value of MSE in equation (4) the PSNR value can be computed.

The equation for computing the SSIM value is,

$$SSIM(V_{F1},V_{F2}) = ((2\mu_{F1}\mu_{F2}+C_1)\ (2\sigma_{F12} + C_2)) / ((\mu_{F1}^2 + \mu_{F2}^2 + C_1)\ (\sigma_{F1}^2 + \sigma_{F2}^2 + C_2)) \quad (5)$$

The structural similarity index (SSIM) is a method for measuring the similarity between two images, it is a full reference metric [1] [2]. In the above equation μF1 is the average of a reference video frame VF1 and μF2 is the average of a transmitted video frame VF2. C1=(K1L)2 and C2=(K2L)2 are used for stabilizing division with weak denominator [1] [2]. While, L is the dynamic range of the pixel values of the frames which is set to default value 255. The variance of a reference video frame and the transmitted video frame are given by σF12, σF22 respectively. While, σF12 determines the covariance of the video frames. From the equation (5) the SSIM value can be computed.

The figure 3 shows the implementation of the system. As per it the system gets a reference source on one side and the transmitted source is evaluated on the other side. The system evaluates these video source frame by frame, the inverse filtering is carried out for each of the frames we can see the enhanced contrast levels and reduced blockiness after this process. At the bottom of the figure we can see number of frames evaluated by the system. Based on the evaluation the system computes the PSNR, SSIM, values and determines the MOS with respect to these values.

5. Experimental Results

We had used a sample video file that was of MPEG-4 video format with 320 X 180 resolution. The frame rate of the video file was 30 frames/Sec. When the video was extracted to obtain video frames there were approximately 144 frames obtained from it. Each video frame, then was extracted for temporal and spatial features like frame size, pixel values, frame rate, etc. The inverse filtering is carried out for each of the frames, we can see the enhanced contrast levels and reduced blockiness after this process. The video frames were further processed one by one for QoE analysis, here the system computes the PSNR, SSIM values with respect to MOS values for better video quality.

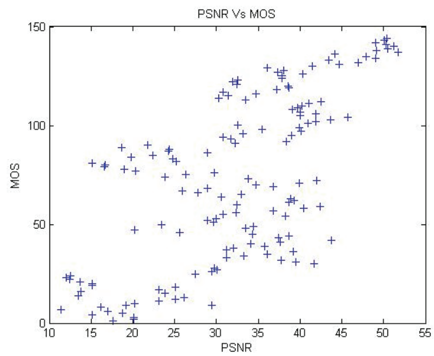

그림 4. 역 필터링이 없는 PSNR 및 MOS의 산점도

Figure 4. A scatter plot of PSNR and MOS without inverse filtering

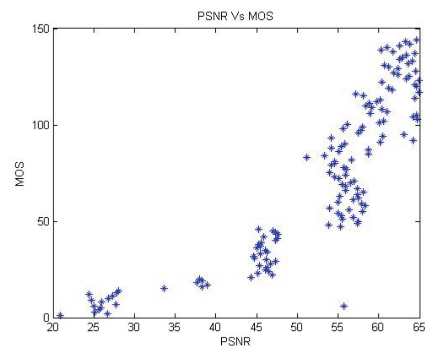

그림 5. 역 필터링이 있는 PSNR 및 MOS의 산점도

Figure 5. A scatter plot of PSNR and MOS with inverse filtering

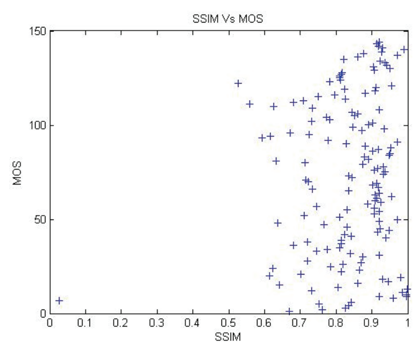

그림 6. 역 필터링이 없는 SSIM와 MOS의 산점도

Figure 6. A scatter plot of SSIM and MOS without inverse filtering

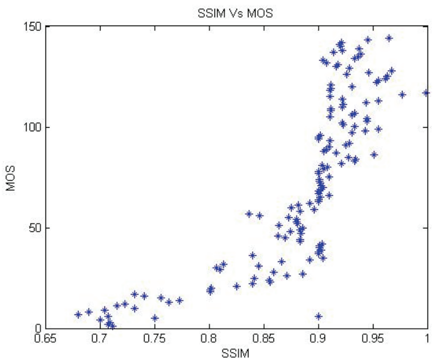

그림 7. 역 필터링이 있는 SSIM와 MOS의 산점도
Figure 7. A scatter plot of SSIM and MOS with inverse filtering

The results obtained are shown in each of the above figures. The figure 4 shows the subjective MOS values with objective PSNR values obtained, each point on the scatter plot determines an analyzed video frame. The PSNR value in the figure is obtained without inverse filtering, as the video frames are having a low contrast level and some blockiness. After applying inverse filtering the PSNR value is increased with the increase in contrast levels and reduced bockiness as shown in the figure 5. The comparison clearly determines the video quality of the transmitted source is relatively as well as a reference source with a high PSNR value which implies that the signal-to-noise ratio is higher. The figure 6 shows the subjective MOS values with objective SSIM values obtained, The SSIM value in the figure is obtained without inverse filtering, as the video quality contained some impurities thus structural similarity between these frames can be determined to some extent. After imposing the inverse filter the SSIM value is increased, we get approximately high quality as well as transmitted source as shown in the figure 7.

6. Conclusion

The approach estimates the user's QoE for the system, which uses a full reference based analysis. The system also imposes the inverse filtering technique which enhances the video quality by improving the contrast levels and reducing the blockiness. The inverse filter responds very badly to any noise that is present in the frames, which in turn provide bad results. In the future, we will implement this system with another filtering technique which doesn't have the impact of noise, but regardless of this one issue the system produces effective results. The system computes PSNR, SSIM values with respect to the MOS values that produce better video quality. As the system uses full reference based analysis, it helps to achieve higher accuracy.

References

[1] Wikipedia, *http://en.wikipedia.org/wiki.*

[2] Hsu, Wu-Hsiao, and Chi-Hsiang Lo, *QoS/QoE Mapping and adjustment model in the cloud-based multimedia infrastructure*, Systems Journal, IEEE Vol. 8, No. 1, pp. 247-255, 2014.

[3] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, *Image quality assessment: From error visibility to structural similarity*, IEEE Transactions on Image Processing, Vol. 13, No. 4, pp. 600-612, 2004.

[4] Loza, Artur, Lyudmila Mihaylova, Nishan Canagarajah, and David Bull, *Structural similarity-based object tracking in video sequences*, Proc. Of the 9th International Conf. On Information Fusion, 2006.
[5] D. Brunet, E. R. Vrscay, and Z. Wang, *On the mathematical properties of the structural similarity index*, IEEE Transactions on Image Processing, Vol. 21, No. 4, pp. 1488-1499, 2012.
[6] Agboma, Florence, and Antonio Liotta. *QoE-aware QoS management,* Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia. ACM, 2008.
[7] Kim, Hyun Jong, and Seong Gon Choi, *A study on a QoS/QoE correlation model for QoE evaluation on IPTV service*, The 12th International Conference on Advanced Communication Technology (ICACT), IEEE Vol. 2, pp. 1377-1382, 2010.
[8] Khan, Asiya, Lingfen Sun, and Emmanuel Ifeachor, *QoE prediction model and its application in video quality adaptation over UMTS networks*, Multimedia, IEEE Transactions on, Vol. 14, No. 2, pp. 431-442, 2012.
[9] Kim, Hyun Jong, and Seong Gon Choi, *QoE assessment model for multimedia streaming services using QoS parameters*, Multimedia Tools and Applications, Vol. 72, No. 3, pp. 2163-2175, 2013.
[10] Seppänen, Janne, Martín Varela, and Aggeliki Sgora, *An autonomous QoE-driven network management framework*, Journal of Visual Communication and Image Representation Vol. 25, No. 3, pp. 565-577, 2014.

역 필터링 기반의 비디오 품질 평가를 위한 QoE분석

Yogiraj S. Jadhav[1], 김정대[2], 김진술[1]
[1] *전남대학교 전자컴퓨터공학부*
[2] *나사렛대학교*

요 약

QoE(Quality of Experience)는 네트워크 계층에서 비디오 서비스 매개변수 및 품질의 상대적 비교성을 평가 할 수 있다. 서비스 제공업체는 특정형태에 멀티미디어 서비스 품질 레벨을 검사한다. 이는 인터넷 서비스 제공업체가 모든 서비스를 하나의 라인을 통해 서비스를 제공하기 때문이다. 이러한 사실을 고려하여, 사용자 경험인지를 분석하는 QoE 매개변수가 중요한 역할을 하며, 이를 위하여 사용자 경험인식과 연관된 요소는 디지털 시스템 개발에 책임을 갖는다.

본 논문에서는 참조모델을 사용하며, 이는 QoE 요인을 고려하여, 사용자의 체감품질을 예상하는데 도움이 되는 접근방식을 결정한다. 이 방법은 역 필터링 등 이미지 복원 기술을 사용하여, 비디오의 전체적인 외관을 향상시킨다. 또한, 측정 요인에 따른 비디오 품질을 효율적으로 평가하는데 도움을 준다.

Acknowledgments

This Research was supported in Research Program of Chonnam National University and Korea Nazarene University, Korea, 2014. also, Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(NRF-2013R1A1A2013740)

Yogiraj S. Jadhav received the master's degree in the Department of Computer Science from the Shivaji University, India in 2012. Currently, He is Ph.D. student in the school of Electronics and Computer Engineering in Chonnam National University. His current research interests include IPTV, Social TV, QoS/QoE measurements.

E-mail address: yogirajjadhav17@gmail.com

Jungdae Kim received the BS and M.S. degree in Industrial Engineering from Dongguk University, Korea 1983 and he completed his Ph.D. in Industrial and Management Systems Engineering from West Virginia University, USA, in 1998. Before he joined Korea Nazarene University in 2009, he worked in chemical industry for many years. His research area is mainly rehabilitation and assistive technology, quality management, and computer application.

E-mail address: jdkim@kornu.ac.kr

Jinsul Kim received the B.S. Degree in computer science from University of Utah, Salt Lake City, Utah, USA, in 2001, and the M.S. and Ph.D. degrees in digital media engineering, department of information and communications from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2005 and 2008. He worked as a researcher in IPTV Infrastructure Technology Research Laboratory, Broadcasting/ Telecommunications Convergence Research Division, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea from 2005 to 2008. He worked as a professor in Korea Nazarene University, Chon-an, Korea from 2009 to 2011. Currently, he is a professor in Chonnam National University, Gwangju, Korea. He has been invited reviewer for IEEE Trans. Multimedia since 2008. He has been invited for TPC(Technical Program Committee), IWITMA2009/2010, and PC(Program Chair), ICCCT2011 His research interests include QoS/QoE, Measurement/ Management, IPTV, Mobile IPTV, Smart TV, Multimedia Communication and Digital Media Arts.

E-mail address: jsworld@jnu.ac.kr