

# Research on No-reference Video Quality Assessment Based on H.264/AVC Bitstream

Zhou Zheng and Feng Weidong\*

Information and Telecommunications Branch, Hubei Provincial Electric Power Company of State Grid, Hubei Wuhan 430077, China

**Abstract:** It is an important part to efficiently evaluate the video quality for user experience in video systems. Traditional video evaluation criterion is strict for testing environment, and is not appropriate for real-time evaluation on network video. No-reference video quality assessment doesn't require any information of original video, which makes it meet the characteristics of high real-time evaluation on network video image quality and small resource reduction. By extracting key parameters distorted by characteristic video in H.264/AVC video stream flow including quantification parameter, encoding rate and bits of each macro block of coding video, the paper establishes no-reference video quality assessment model from the space complexity and time complexity of video. For the model, not only the model is not complicated, and can meet the requirement of real-time quality assessment of video, the simulation test but also proves that the correlation between the algorithm and subjective assessment value is high, which meets visual perception feature of human eye.

**Keywords:** Bitstream, H.264/AVC, no-reference, video quality assessment.

## 1. INTRODUCTION

In recent years, with rapid development of internet and mobile communication, and rapid development of integrated circuit technique towards super-large scale, and the emergence of high-speed and dedicated multimedia chip, there are more and more video techniques such as video communication, video compression and video processing. Video not only is applied to TV broadcasting service, and has been applied to various fields in life such as business application of video supervision and video conference, video service of telemedicine and web based education, and recreational activities including video on demand [1, 2]. However, compared with huge size of video data, network bandwidth and storage device is limited, which needs to compress the video for transmission or storage [3]. Compressing video inevitably makes the video distort after compressing and reconstructing the video, which reduces the quality of video. In order to ensure that the users have subjective feeling on the video, how to effectively assess video quality is more and more important in various video systems.

There are many scholars proposing assessment methods for video quality. The existing assessment for video quality is divided into subjective quality assessment and objective quality assessment [4]. And objective quality assessment includes full-reference type [5], reduced-reference type [6] and no-reference type [7]. Subjective quality assessment

[8, 9] can accurately reflect the subjective feeling of human eyes watching video. The existing objective quality assessment mathematical models are used to simulate the results of video subjective assessment. The advantages and disadvantages of objective quality assessment mathematical model are reflected by subjective quality assessment results and objective quality assessment results. The higher the correlation is, the better the objective quality assessment mathematical model.

The paper firstly introduces typical subjective quality assessment method and the disadvantages and advantages of the assessment methods. Then, the paper introduces no-reference objective quality assessment. And the objective quality assessment mathematical model researched in the paper needs to apply the results of subjective assessment.

## 2. OVERVIEW OF VIDEO QUALITY ASSESSMENT METHODS

### 2.1. Typical Subjective Quality Assessment Method

According to the test purpose and test environment, the existing typical subjective quality assessment methods are divided into the following types.

(1) Single stimulus continuous quality evaluation [10]

Single stimulus continuous quality evaluation can rapidly trace the change of video image quality, and it is appropriate for the compression system with great difference of video image quality change and with time-variant characteristic. But selecting the sequence of testing video has great influence on experiment evaluation results.

### (2) Single stimulus methods

The distorted video images are randomly combined with the original video images, and each group of video images are broadcast independently, and the quality is evaluated by the subjects. The order of the video is difference. In order to improve the accuracy of the score, SSM extends the five-point quality ranking in video quality assessment into 9-point and 11-point scale.

### (3) Double stimulus impairment scale [10]

The grade uses 5-point quality ranking. According to the mode of firstly playing the original reference video and then playing the distorted video, video groups are played for the subjects. And the subjects evaluate the quality of the distorted video by comparing the original reference video and distorted video.

### (4) Double Stimulus Continuous Quality Scale [10]

Firstly, the video pairs are combined, and each video pair should be ensured to consist of original reference video and distorted video. But the order of playing the original and distorted video in each video pair is random. Then, each video is played for the subjects, and the subjects grade the original reference video and the distorted video in the video pair. In the testing process, the video pair is firstly played for one time or many times to make the subjects have subjective acknowledge on each video pair. Then, the video is played for the subjects for one time or many times to make the subjects grade the video. The time of playing different types of video images are is different. For the video images with furious time-variant motion, each video sequence is played for 10 seconds twice. For video sequence with quiet picture, each video image is shown for three or four seconds for five times. And it is graded in the last two times.

## 2.2. Objective Quality Assessment Method

Subjective quality assessment is limited, complicated and time-consuming, and the evaluation results are not extended, so subjective quality assessment method can't be used for quality monitor with high requirement on instantaneity. Objective quality assessment overcomes the advantages of subjective quality assessment. It is easy to be implemented, and it has attracted the attention of the researcher's of video image quality. One or several parameters measuring video quality metric are selected to establish objective mathematical model to measure the quality level of video image. The advantages and disadvantages of objective quality assessment mathematical model are reflected by the correlation between subjective quality assessment results and objective quality assessment results. The higher the correlation is, the better the objective quality assessment mathematical model.

According to the condition if there is need to use original video as reference, international video quality experts group divide objective quality assessment methods into full-reference, reduced-reference and no-reference video quality assessment.

### (1) Full-reference video quality assessment

Full-reference assessment is easy to achieve good performance, the reason for which is that full-reference video quality assessment needs lots of information of original reference video. When the full-reference video quality is assessed, all information of the video original reference needs to be transmitted. The information of the original reference video and the distorted video is compared to get the quality assessment value of the distorted video. The basic principle of full-reference video quality assessment is to get the similarity of distorted video and original reference video.

### (2) Reduced-reference video quality assessment method

The full-reference video quality assessment results of HVS are consistent with subjective assessment, but the information of the original reference video is not easy to be achieved, the reason for which is that it has huge data volume, and the transmission is not realistic. Especially under the condition with higher requirement on bandwidth, the information of original reference video is not easy to be transmitted and stored. And the reduced-reference video quality assessment method emerges. The required bandwidth is small, and the computation is not complicated. And it can meet the requirement of video quality real-time assessment.

### (3) No-reference video quality assessment method

No-reference video quality assessment doesn't need to transmit any information of original video. By analyzing distorted video to be tested, and extracting characteristic parameters representing distortion of distorted video, the mathematical model is established to assess quality of distorted video. It doesn't require any reference information and there is no restriction, no-reference video quality assessment is very applicable to real-time video quality assessment.

## 2.3. Comparison of Assessment Methods

Human is the final experience on the video. Subjective quality assessment is the felling value of human eye for video, so subjective assessment result is considered to be correct. However, the process of subjective assessment is very complicated and time-consuming, and the testing results are not extended, so it can't be applied to quality monitor with high real-time requirement. Objective quality assessment simulates the results of subjective assessment. By extracting key data parameters representing video deficit, the mathematical model is established, which can get objective quality assessment value. Objective assessment can't accurately reflect the subject feeling of human vision for video like subjective evaluation, but it is rapid and easy, which can meet real-time requirement.

Full-reference video quality assessment result is consistent with subjective assessment result, but all information of the original video is not easy to be achieved, and the data is huge, and it is not easy to transmit and store all information of the original reference video under the condition that bandwidth and storage device is restricted. Reduced-reference video

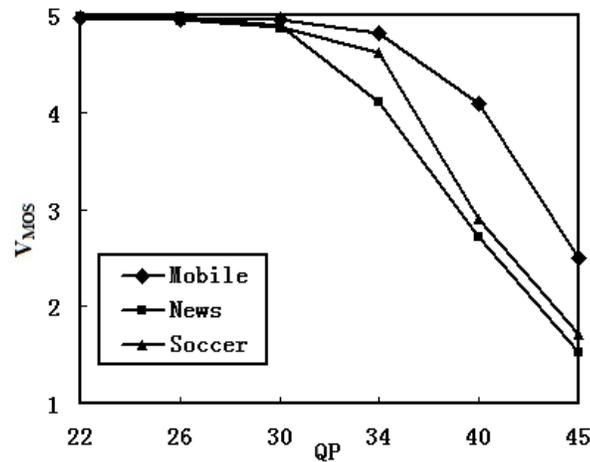


Fig. (1). Relationship between QP and  $V_{MOS}$ .

quality assessment only needs to transfer some specific features or parameters of original reference video as reference. If the transmitted parameters and features of original reference video is the same with that of distorted video to be tested, and it has great difference for human eye vision, the reduced-reference video quality assessment result is not reliable. Under the condition with high requirement on bandwidth and storage device, it is restricted to transmit the reference information parameters. No-reference video quality assessment doesn't require transferring any information of original reference video. No-reference video quality assessment doesn't need to transmit any information of original video. By analyzing distorted video to be tested, and extracting characteristic parameters representing distortion of distorted video, the mathematical model is established to assess quality of distorted video. No-reference video quality assessment needs less bandwidth, and has the advantages of low complexity, and good applicability and flexibility, so it is applied to real-time video quality assessment. And it has become one of the topics to be researched for video quality assessment method in the future.

### 3. VIDEO QUALITY ASSESSMENT BASED ON VIDEO BITSTREAM

#### 3.1. Overview of the Method

The video quality assessment method of the paper is no-reference video quality assessment based on bitstream. By extracting key parameters making video distorted from video bitstream, the appropriate mathematical model is established to assess the quality of video. The objective of the assessment method is to simulate subjective quality assessment value, so subjective video quality assessment value needs to be used to fit the coefficients of no-reference video quality mathematical model based on bitstream characteristics.

For four typical subjective quality assessment methods, and the paper uses subjective quality assessment method used by LIVE Video Quality Database [11, 12] and subjective assessment used by EPFL and Polimi [13, 14], the paper finally selects SSM to assess the subjective assessment value

of video sequence to be tested in the paper. SSM combines the distorted video image with the original video image, and the video is randomly played to the subjects, which is easy for the subjects to compare and has less difficulty. And the assessment method is as follows.

The subjective testing experiment of the paper is made in the laboratory, and the testing environment should meet the requirement of ITU-R BT50. The subjects ensure to watch video under a quiet environment, and the LCD computer monitor should be 19 inches, and the background of the monitor should be gray. Non-professional subjects without any experience of image or video quality assessment are selected, and the age is from 20 to 25 and 23 subjects are selected. In order to ensure that the evaluation results are feasible and effective, the subjects are trained in advance for simple video quality assessment. SSM is used to make 38 subjects assess quality of 10 selected video sequences. 10 selected video sequences have their own characteristics, and represent various intensity of video image movement and complexity of video image space and time (detailed description of video sequence is shown in Chapter 5). The subjective value of assessment meets 5-point grading rule, and the grading result is between 0 and 5. And the grading rule of 5 points is shown in [5]. For 23 achieved subjective assessment values from each testing video sequence, 17 surplus subjective assessment values are used as subjective assessment values of each testing video sequence.

#### 3.2. Extraction of H.264/AVC Bitstream Characteristic Parameter

The data volume of the original video is huge, and it is impossible to transfer the original video directly, so the video needs to be firstly compressed. The common encoder is H.264/AVC encoder, and the encoding statistic parameters to be used are as follows.

##### 3.2.1. Quantization Step

Quantization parameter represents index sequence of quantization step  $Qstep$ . In H.264 codec,  $Qstep$  has 52 values. QP determines the image accuracy and encoding compress-

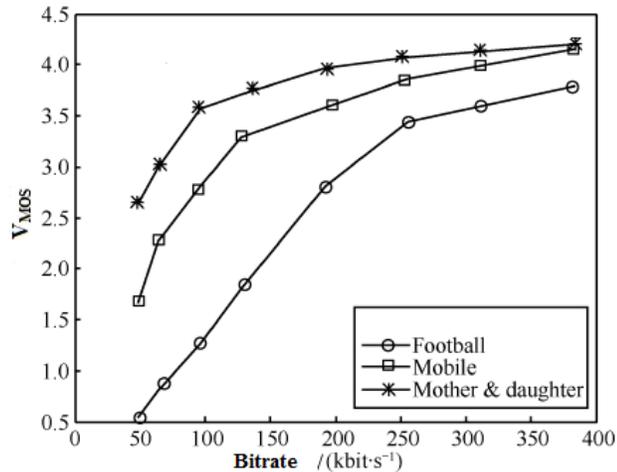


Fig. (2). Relationship between encoding rate and  $V_{MOS}$ .

sion rate of H.264 encoder. When QP is 0, Qstep is 0.625, and the quantization should be the most refined. But the bit rate for coding is large, and the compression ratio is smaller. With the increase of Qp, the compressed video loses more information, which the image details lose after reconstructing video bitstream. But the bit rate for coding reduces, and the compression ration increases. When QP is 51, Qstep is 224, and the quantization is coarsest. Coding needs little bit rate, but the reconstructed video image is fuzzy and even makes people boring and can't be watched.

In order to directly show the relationship between QP and VMOS, the paper selects three video sequences with the format of CIF and representative video feature, the paper gets the relationship between subjective quality assessment value (subjective assessment uses five-point grading system) and QP of three video sequence under the condition with certain OP value (QP is 22, 26, 30, 34, 40 and 4), as shown in Fig. (1).

### 3.2.2. Encoding Rate

Code rate means the transmission speed per second of video after compression into bitstream. Under the condition with enough bandwidth, for the same video, the greater the code rate is, the greater the bitstream after video is compresses, and the more clear the achieved video image is. Under the condition that bandwidth is restricted, code rate is not large, and the bitstream reduces of the achieved video image after reconstructing video. For different video sequences, the features of each video are different, which is reflected in the complexity of airspace feature and time domain feature. It indicates that encoding the same QP is different form the achieved bit rate.

In order to directly show the relationship between code rate and subjective video quality value, the paper selects three video sequences with the format of CIF with representative video feature, and draws the relationship between subjective quality assessment value (subjective assessment uses five-point grading score) of three video sequences and code rate under the condition with a series of constant code rate, as shown in Fig. (2).

### 3.2.3. Macroblocks Skipped by Coding

Macroblocks skipped by coding are the macroblocks which don't need to code, and it is a phenomenon when H.264/AVC coder predicts encoding on frames of each frame of video image. There is flat area of frame image in each video, and there is area with more details. The flat area is low-frequency area, and the area with more details is high-frequency area. In order to make the encoded code rate smaller, under the condition that the video quality after reconstructing video can be accepted, H.264 coder allows each frame of image of video to use skip block in flat area while predicting coding blocks of each frame of video image. Skip block means that it has no any data. The macroblocks which have been encoded by H.264 encoder predict the skip block, to recover skip block.

Skip block can make video frame image distorted after reconstructing video on the receiving end, which makes subjective assessment results of video reduce.

Compressing each video by H.264 encoder can generate skip blocks, which is a compromise of the reduction of subjective quality evaluation results and code rate caused by skip blocks. And it reduces code rate under the premise of not evident reduction of subjective evaluation results.

In order to directly show the relationship between macroblocks skipped by coding and VMOS, the paper selects three video sequences with the format of CIF which have different video characteristics, and draws the relationship between VMOS (subjective assessment uses 5-point grading score) of three video sequences and macroblocks skipped by coding under macroblock values skipped by a series of different codes, as shown in Fig. (3).

Above all, H.264/AVC can be divided into two predictive coding. Intra-frame predicts coding is for space complexity, and inter-frame predict coding is for time complexity. The paper uses the extracted parameters influencing video quality to establish mathematical model with space complexity and time complexity to assess video quality.

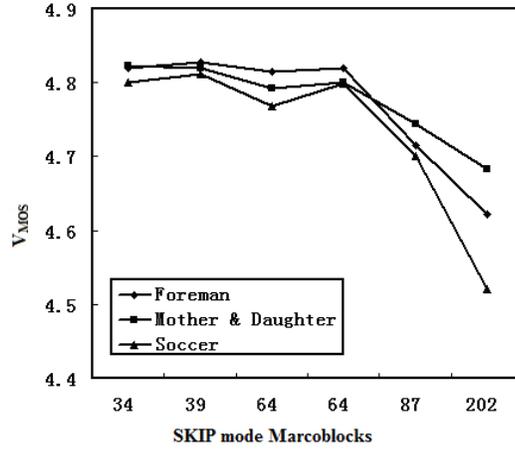


Fig. (3). Relationship between macroblocks skipped by coding and  $V_{MOS}$ .

#### 4. NO-REFERENCE ASSESSMENT MODEL BASED ON BITSTREAM STATISTICAL CHARACTERISTICS

##### 4.1. Video Quality Assessment Model with Space Complexity

QP and code rate of coding [15] is used to establish video quality assessment model with space complexity. And space complexity  $S$  meets the following formula.

$$S = f_{QP}R + g_{QP} \quad (1)$$

In the formula,  $f_{QP}$  is the slope fitting function relating to QP, and  $g_{QP}$  is the intercept fitting function relating to QP. In order to improve the quality of reconstructed video, H.264 coder allows QP not fixable, which means that the selection of QP can determine appropriate QP according to the self-adaption feature of the present blocks. If video sequence has  $N$  frames of images, each frame of image has  $M$  macroblocks, and  $QP(n,m)$  is the QP of the  $m$  macroblock of the  $n$  frame, the expression of slope fitting function  $f_{QP}$  is formula (2).

$$f_{QP} = \frac{1}{NM} \sum_1^N \sum_1^M (a_1 2^{(QP(n,m)-4)/6} + b_1) \quad (2)$$

The expression of intercept fitting function  $g_{QP}$  is formula (3).

$$g_{QP} = \frac{1}{NM} \sum_1^N \sum_1^M (a_2 QP(n,m) + b_2) \quad (3)$$

In the formula,  $a_1$ ,  $a_2$ ,  $b_1$  and  $b_2$  is undetermined constant. Substituting formula (1) and formula (2) into formula (3) can get

$$S = \frac{1}{NM} \sum_1^N \sum_1^M (a_1 2^{(QP(n,m)-4)/6} + b_1)R + \frac{1}{NM} \sum_1^N \sum_1^M (a_2 QP(n,m) + b_2) \quad (4)$$

$S$  is space complexity.  $a_1$ ,  $a_2$ ,  $b_1$  and  $b_2$  is undetermined constant, which can be achieved by fitting data. Simplifying formula (4) can get the expression of space complexity in formula (5).

$$S = \frac{1}{NM} \sum_1^N \sum_1^M [(a_{11} 2^{QP(n,m)/6} + b_{11})R + (a_{12} QP(n,m) + b_{12})] \quad (5)$$

##### 4.2. Video Quality Evaluation Model with Time Complexity

Inter-frame predictive coding means that after each frame of image of video sequence is divided into blocks or macroblocks, coding eliminates the redundancy between macroblocks, to achieve the objective of compressing video. For the area that video frame image is flat, the numbers of macroblocks which have no need to encode would be increased. only referring to the information of the adjacent encoded macroblocks can recover the image information of skip block when video is reconstructed on the receiving end. So the skip number can be used to express video quality evaluation model with time complexity. For a video sequence, under the condition with fixed encoding byte, the more the macroblocks skipped by coding, the less the bits occupied by each macroblock. And the parameter can be read from vision bitstream. Therefore, in order to simplify the complexity of the algorithm in the paper, the paper selects bits/MB to measure time complexity  $T$ , as shown in formula (6).

$$T = a \cdot \text{bits} / MB \quad (6)$$

In the formula,  $a$  is the undetermined coefficient.

##### 4.3. Assessment Model Based on H.264 Bitstream Statistical Characteristics

Video quality assessment model based on H.264 bitstream can synthesize the time complexity  $T$  and space complexity of video image. The expression of video quality assessment mathematical model based on H.264 bitstream is formula (7).

Table 1. Fitting coefficients.

$a_{11}$	-2.5609	$a_{12}$	0.0721	$a_{21}$	0.9398
$a_{22}$	1.1100	$b_{11}$	-0.0022	$b_{12}$	14.3163
$b_{22}$	0.2342	$a$	1.6579	$b$	-19.0501

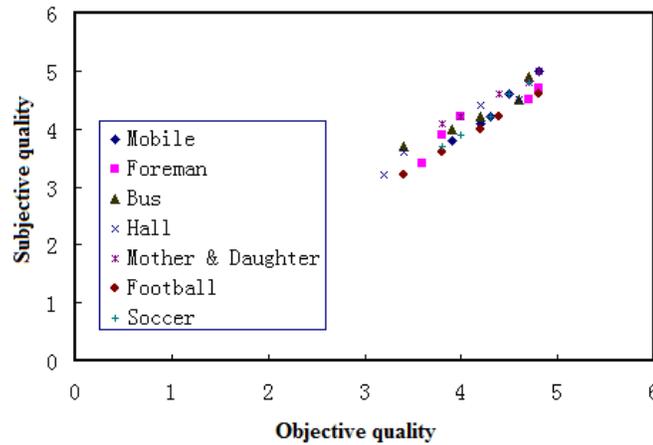


Fig. (4). Splattering figure of objective quality assessment value and subjective quality assessment value.

$$Q = a_{21}(S)^{a_{22}} + b_{21}(T)^{b_{22}} + b \tag{7}$$

In the formula, Q is the video quality assessment value based on H.264 bitstream. b, a<sub>21</sub>, a<sub>22</sub>, b<sub>21</sub> and b<sub>22</sub> are undetermined constants. Substituting formula (5) and formula (6) into formula (7) can get the expression of video quality assessment value Q based on H.264 bitstream, as follows.

$$Q = a_{21} \left( \frac{1}{NM} \sum_1^N \sum_1^M [(a_{11} 2^{QP(n,m)} + b_{11})R + a_{12}QP(n,m) + b_{12}]^{a_{22}} + a (bits / MB)^{b_{22}} + b \right) \tag{8}$$

a, b, a<sub>11</sub>, a<sub>12</sub>, a<sub>21</sub>, a<sub>22</sub>, b<sub>11</sub> and b<sub>12</sub> are undetermined coefficients, and can be achieved by using least square method to fit.

### 5. EXPERIMENTAL RESULTS AND ANALYSIS

For undetermined coefficients in formula (8), the paper selects representative video sequences with the format of CIF. JM14.2 of H.264/AVC coder software platform is used to achieve a series of different QP to test coding rate of video sequences and coding bit of each macroblock, and assessment value of subjective video quality. The coding level of JM14.2 in the paper is High level of FREXT level, the coding mode is I-B-B-P-B-B-P, and the frame rate is 25. QP in the paper is 9, 16, 25, 32, 36 and 40.

The experiment data of the paper is substituted into video quality assessment model established in formula (8) for fit, which can get undetermined coefficients in Table 1. The experiment data is achieved under JM14.2 software platform,

so for other software platform of H.264 coder, the coefficients of the model in the paper need to be adjusted.

In order to verify the reliability of the video quality assessment model proposed in the paper, the paper refers to the present video quality base. The video quality of EPFL also uses JM14.2 coder, so the paper analyzes the correlation with video quality of EPFL. For the same testing video sequence, the quality assessment value of tested video is extracted from video bitstream, and is compared with subjective quality assessment value of EPFL. In order to show the intuition of analysis results, the paper selects 8 test sequences, and draws splattering figure of objective quality assessment value and subjective assessment value, as shown in Fig. (4).

From Fig. (4), we can see that the objective quality assessment model proposed in the paper not only has better linear correlation than subjective quality assessment value, but also is easy to assess video quality in real time.

### CONCLUSION

The paper firstly compares the existing video quality assessment methods including subjective video quality assessment and objective video quality assessment. No-reference objective quality assessment doesn't require transmitting any information of the original reference video and can meet the requirement of real-time quality assessment. If the proposed no-reference video quality assessment model can accurately reflect the quality of video, taking compensation measures to improve the quality of distorted video or improving video compression algorithm is helpful to improve video compression.

Secondly, the paper compares the parameters of the existing no-reference objective video quality assessment evaluation model. By combining the principle of H.264 video encoder and computation complexity of evaluation model, the paper proposes to use the distorted key parameters of characteristic video (QP, coding rate and Bits/MB for each macroblock of coding video) to establish no-reference video quality assessment model from the perspective of time complexity and space complexity. And the paper proposes a no-reference video quality assessment with low complexity based on H.264/AVC bitstream statistic feature, which not only reflects video quality in real time, but also reduces computation complexity. And the experimental results indicate that the model of the paper has good linear correlation with subjective quality assessment value of typical model.

### CONFLICT OF INTEREST

The author confirms that this article content has no conflict of interest.

### ACKNOWLEDGEMENTS

Declared none.

### REFERENCES

- [1] K.R. Rao, Z.S. Bojkovic, and D.A. Milovanovic, *Multimedia Communication System: Techniques, Standards, and Networks*, Prentice Hall, London, 2002.
- [2] Y. Wang, J. Ostermann, and Y.Q. Zhang, *Video Processing and Communications*, Pearson Education, New York, 2001.
- [3] J.C. Liu, S.G. Rao, B. Li, and H. Zhang "Opportunities and challenges of peer-to-peer internet video broadcast," In: *Proceeding of the IEEE, Special Issue on Recent Advances in Distributed Multimedia Communications*, vol. 96, no. 1, pp. 11-24, 2008.
- [4] Z. Wang, and A.C. Bovik, "Mean squared error: love it or leave it? A new look at signal fidelity measures," In: *United States: IEEE Signal Process*, 2009.
- [5] ITU-T, *Recommendation J.246, Perceptual Visual Quality Measurement Techniques for Multimedia Services Over Digital Cable Television Networks in the Presence of a Reduced Bandwidth Reference*, United Nations: International Telecommunication Union, 2008.
- [6] H.R. Wu, and K.R. Rao, *Digital Video Image Quality and Perceptual Coding*, United States: CRC Pr I Llc, 2005.
- [7] C. Keimel, T. Oelbaum, and K. Diepold, "No-reference video quality evaluation for high-definition video," In: *United States: IEEE International Conference*, 2009.
- [8] ITU-T, *Recommendation J.148. Requirements for an Objective Perceptual Multimedia Quality Model*, 2003.
- [9] ITU-T, *Recommendation P.910, Subjective Video Quality Assessment Methods for Multimedia Applications*, 1999.
- [10] ITU-R Recommendation BT.500-11, *Methodology for the subjective assessment of the quality of television pictures*: Geneva, 2002.
- [11] S. Winkler, "Analysis of public image and video databases for quality assessment," *IEEE Journal on Selected Topics in Signal Processing*, vol. 6, no. 6, pp. 616-625, 2012.
- [12] S. Winkler, and R. Subramanian: "Overview of eye tracking datasets," In: *Proceedings of 5<sup>th</sup> International Workshop on Quality of Multimedia Experience (QoMEX), Klagenfurt, Austria, 2013*, pp. 212-217.
- [13] F. De Simone, M. Naccari, M. Tagliasacchi, F. Dufaux, S. Tubaro, and T. Ebrahimi, "Subjective assessment of H.264/AVC video sequences transmitted over a noisy channel", In: *First International Workshop on Quality of Multimedia Experience (QoMEX 2009)*, San Diego, California, USA, 2009, pp. 204-209.
- [14] F. De Simone, M. Tagliasacchi, M. Naccari, S. Tubaro, and T. Ebrahimi, "H.264/AVC video database for the evaluation of quality metrics", In: *35<sup>th</sup> International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010)*, Dallas, Texas, USA, 2010, pp. 2430-2433.
- [15] F. Yang, S. Wan, Q. Xie, and H.R. Wu, "No-reference quality assessment for networked video via primary analysis of bit-stream," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 11, pp. 1544-1554, 2010.

Received: September 16, 2014

Revised: December 23, 2014

Accepted: December 31, 2014

© Zheng and Weidong; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.