Journal of Mobile Multimedia, Vol. 7, No.4 (2011) 267-278

A NOVEL FULL-REFERENCE VIDEO QUALITY ASSESSMENT METRIC FOR MULTIMEDIA BROADCASTING SYSTEMS

XINGANG LIU

School of Electronic Engineering, University of Electronic Science and Technology of China

Chengdu, China

hanksliu.xg@gmail.com

CHAO SUN*

School of Electronic Engineering, University of Electronic Science and Technology of China

Chengdu, China

ch_sun@126.com

WENJIE YANG

School of Electronic Engineering, University of Electronic Science and Technology of China

Chengdu, China

Received Dec 10, 2011
Revised March 31, 2012

In this paper, A novel full reference (FR) video quality assessment (VQA) metric for the multimedia broadcasting systems is proposed. The relationships between the VQA parameters and the quality score obtained by the subjective human visual system (HVS) model of the distorted videos are investigated. Considering the multiple effects which have relationships with the subjective VQ, Spatial frequency analysis (SFA), MSSIM, Matched PSNR (MPSNR) and Edge degradation(ED) are employed to estimate the quality of the video sequences. The simulation results show that our proposed metric can give high correlation with the subjective video quality (VQ) score and has the immersive potential to realize the application to the end-users.

Key words: FR; VQA; HVS; DMOS; correlation

1 Introduction

With the high-speed development of multimedia broadcasting system, the multimedia signal has been the requirement for ordinary customers and will have a great impact on our daily life in various applications, such as homes, workplace, public spaces and so on. For this case, the quality of the video programming is very important for the customers who would like to receive it. Therefore, the assessment for the visual quality is of fundamental importance for numerous image and video processing applications, where the goal of VQA is to automatically assess the quality of images or videos in agreement with human quality assessment. The most direct VQA way is by the subjective quality assessment. It is a psychologically based method using structured experimental designs and human participants to evaluate the quality of the video sequences [1]. Since the video quality is a subjective notion, the subjective VQA could be considered as the best method to evaluate the video quality. The VQA will give a score called Mean Opinion Score (MOS) to present its perceptual quality. It is generated by averaging the results of a set of standard, subjective tests where a number of users rate the quality on a five point scale from 1 (Bad) to 5 (Excellent). The diverse form of MOS is the Difference MOS (DMOS), which has the same theoretic basic with MOS but measures the difference between the original and distorted video frames. However, realizing the subjective quality assessment tests have to follow many recommendations, otherwise the results of these tests often become useless due to the lack of the precision on the quality scores obtained during these tests. These recommendations lead this evaluation way to be very complicated. Moreover, the subjective quality assessment is a long and expensive process [2]. Therefore, people more tend to use the objective VQA which means to establish an objective function to measure physical aspects of a video signal and consider the physical aspects and psychological issues. In general, the objective VQA metrics are divided into three categories:

- Full reference (FR): The quality assessment metric is established by making a full comparison between the original and distorted video sequences. FR metric is evidently considered as the best way to get good performances in quality assessment because it could use the maximum amount of data.
- No reference (NR): Only the information of the distorted video sequences is utilized in the quality metric algorithm. Since the appearance of the quality metric solves the problem of the requirements for the original video information, the metrics can be established only by using the distorted video information in the destination.
- Reduced reference (RR): The metrics lie between the above two metrics and established by using not all but some features of the original and distorted video sequences.

In this paper, we focus on investigating the FR VQA metric. The conventional FR VQA is based on merely pixel difference between the original and distorted video frames, e.g., peak signal-to-noise ratio (PSNR) [3] and their variations, for instance, Minkowsky Metric (MM) [4], Structural Correlation (SC) [5], Czenakowski Distance (CZD)[6] and so on. Over the years, a number of researchers have contributed significant research in the design of FR VQA metrics. Several novel VQA metrics have been built as standard such as ITU-T J.144 [7] which provides four important FR VQA metrics for both PAL and NTSC. More recently, the FR VQA metrics which utilize the spatial structural information were proposed with the representative productions named as Structural Similarity Index (SSIM) [8] and Information Fidelity Criterion (IFC) [9].

With our study we found that the quality of the video signal which broadcasted through network service can be affected by multiple factors, and there are many important parameters which have definite relationships with the subjective human visual model quality score (DMOS). Each of the parameters which we consider has the ability to represent VQ, but not accurate enough only by itself. And we also find that multiple parameters can affect the observational result of HVS, and these parameters can also affect mutually. Therefore, if the useful VQA parameters can be found and combined together by the mathematics methods, we can get more accurate VQA results.

The rest of this paper is organized as follows. Several conventional FR VQA metrics are introduced in section 2. Section 3 describes our proposed FR VQA metric framework through analyzing the features of the important parameters which can affect the QA performance. Section 4 gives the simulation results which can confirm the improved performance of our proposed metric. And the conclusions are drawn in section 5.

2 Conventional FR VQA Metrics

2.1 VQA by Measuring Pixel Difference

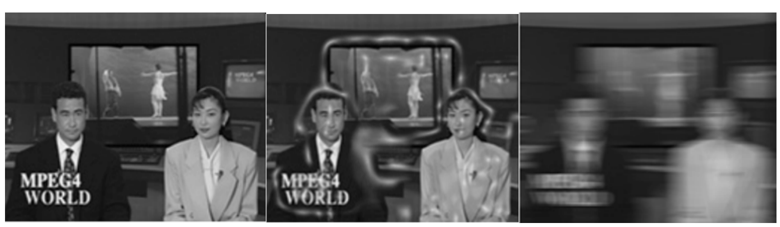
Among the existing VQA metrics, the most direct is based on making spatial domain pixel-by-pixel difference between the reference and distorted video frames, such as Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and so on.

$$PSNR=10\log_{10}\frac{\left(2^{\alpha}-1\right)^{2}}{MSE} \tag{1}$$

Eq.(1) shows the function of *PSNR*, Where α is the number of bit used to represent each pixel value, and MSE is the mean square error between the reference and distorted video frames.

$$MSE=\frac{1}{N\times M}\sum_{x=0}^{N-1}\sum_{y=0}^{M-1}\left(ref\left(x,y\right)-dis(x,y)\right)^{2} \tag{2}$$

Where *ref(.)* and *dis(.)* are the pixel value at location (x, y) of the reference and distorted video frames, respectively. *N* and *M* are the lengths of the row and column of the video frame, respectively. PSNR could response the quality degradation and has been used to measure video/image quality in wide fields. Unfortunately, numerous issues plague the PSNR. The most serious one is that it effectively considers only the global variations and neglects localized errors. In addition, it cannot detect the structured errors such as artifacts and distortions that have a particular, rather than arbitrary pattern.



(a) (b) (c)

Fig.1. PSNR for distorted image; (a) Original image; (b) PSNR=18.367; (c) PSNR=21.536

Fig.1-(a) is the original image, Fig.1-(b) and (c) are the distorted image with different PSNR values. If we use HVS subjective measurement to estimate the quality of the images, it is obvious that Figure1-(b) must have the higher score than Fig.1-(c). However, because of the limitation of PSNR, Figure1- (c) has higher PSNR value which conflicts with the subjective quality judgment.

2.2 VQA by Spatial Structural Similarity

Considering the shortcoming of the conventional VQA metrics by measuring pixel difference directly, the entire structural effect of the image should be measured. For this case, another kind of VQA which considers both the pixel difference and the structural effect was proposed. The representational one is SSIM. It separates the VQA task of the similarity measurement into three comparisons: luminance, contrast and structure. And it calculates these comparisons between the reference and distorted video frames by using the mean, variance and co-variance of the entire image, respectively. Finally, the three comparisons are combined to yield an entire similarity measurement as Eq. (3).

$$SSIM = \left(x, y\right) = \frac{\left(2\mu_x\mu_y + c1\right)\left(2\sigma_{xy} + c2\right)}{\left(\mu_x^2 + \mu_y^2 + c1\right)\left(\sigma_x^2 + \sigma_y^2 + c2\right)} \tag{3}$$

Where, μ_x and σ_x are the mean and variance of the reference video frames, respectively; μ_y, σ_y are the mean and variance of the distorted video frame, respectively; σ_{xy} is the co-variance between the reference and distorted video frames, and *C1* and *C2* are constants. In general case, SSIM cans greatly response the image quality, but due to the whole image calculation, the local error cannot be detected correctly yet. To improve the quality assessment accuracy, the SSIM has been improved to Mean SSIM (MSSIM) [8] which is shown in Eq. (4). It divides the image into $N\times N$ small blocks, and calculates the mean of each block's SSIM as the final VQA result.

$$MSSIM\left(x, y\right) = \left(\sum_{block\_n=0}^{block\_n<N} SSIM_{block\_n}\right) / N \tag{4}$$

Where $SSIM_{block\_n}$ is the SSIM of the n-th block, and N is the total divided block number. According to the general video compression standards, the video coding block size is 8×8. To avoid the blocking artifact in block boundary, the *MSSIM* uses 11×11 as the divided block size. MSSIM can give ideal correlation result with subjective perceptual image quality, but the limitation is that it is only suitable to still image significantly. For the moving video sequences especially with large resolution, it cannot measure the quality perfectly because of the characteristic differences between still images and moving video sequences.

2.3 Standarded FR VQA Metrics

As one of the famous international standard union, ITU-T applies itself to ensure an efficient and on-time production of high quality standard covering all fields of telecommunications. J.144 [7] is the research standard of ITU-T for VQA techniques which includes four FR VQA metrics noted as Annex A-D. These models have a common structure that calculates features' values from visual factors that are related to subjective quality and yield a final quality value by aggregating them with a linear or nonlinear combination. All of the VQA algorithms in J.144 can measure VQ well. But anyone has its own shortcomings. Annex A considers much useful factors which have relationships with subjective video quality. It seems like comprehensive, but the factors it considered are not ideal at all. In addition, it has factors' redundancy. J.144 Annex B is simple to be implemented because of its convenient calculation. However, it's incomprehensive because it considers edge effect so unilateral that other affects are ignored. Annex D extracts features from the original and distorted video frame to build its VQA function. Unfortunately, as Annex C, the computational complexity is very high.

3 Proposed FR VQA Algorithm

Through our study, we found that there are many of parameters which have the abilities to represent the perceptual VQ. However, for each parameter, it is not enough to give a good quality assessment only by itself. It implies that the capability of the objective VQA is affected by multiple parameters. Therefore, to establish an efficient FR VQA algorithm, we should consider the multiple effects which have relationships with the subjective VQ, such as blurring, blocking, jerkiness, noise, and so on. In this paper, several parameters are highlighted and employed as the important features as follows.

- Spatial frequency analysis (SFA),
- MSSIM,
- Matched PSNR (MPSNR),
- Spatial gradient (SG),
- Edge degradation (ED).

The MSSIM has been introduced in previous section, and the remained 4 parameters are explained one by one in the following contents.

3.1 Spatial Frequency Analysis (SFA)

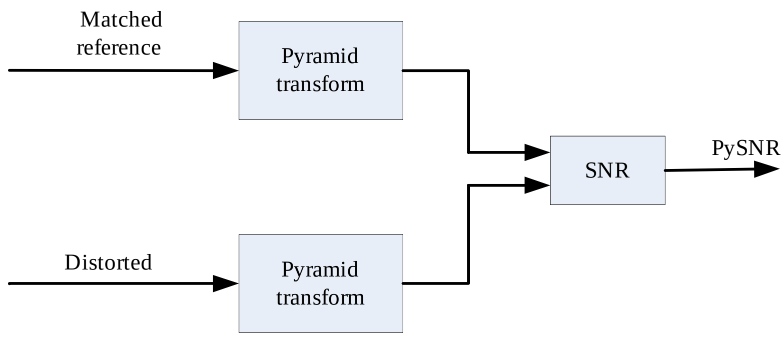


Fig.2. Spatial frequency analysis

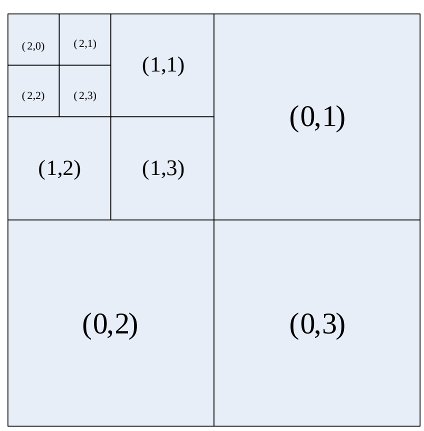


Fig.3. Pyramid transform

The SFA is based on a “pyramid” transformation of the reference and distorted video sequences. According to Fig.2, the reference and distorted video sequences are transformed to give the pyramid arrays, and the differences are calculated by using MSE as a pyramided signal to noise ratio (PySNR). The pyramid transform is based on a 3-stage pyramid transform in the spatial-domain frequency analysis which is given in Fig. 3. Horizontal and vertical low pass filter (LPF) and high pass filter (HPF) are utilized in each stage for both the reference and distorted video frames to build the pyramid transformed video frames. Eq. (5) gives the calculation of the PySNR as the VQA for the SPA, where E(s,q) is a squared error between the reference and distorted video frames; s means the pyramid stage, s = 0, 1 and 2; and q denotes the quadrant index for each stage, q = 0, 1, 2 and 3.

$$VQA_{SFA} = 10\log_{10}\left(255^2 / E(s,q)\right) \tag{5}$$

3.2 Matched PSNR (MPSNR)

As the traditional VQA, PSNR has shown its significant contribution for video/image quality measurement in numerical forms. Although it has the limitation to relate with the subjective perceptual VQA, it is still an important measurement factor for video/image quality if it could be utilized correctly. In our paper, we employ the utilization of PSNR in [7] which named as *MPSNR*. To get the *MPSNR*, two important steps are included. The first one is the matching process, and the second is to calculate the PSNR by using the matched reference and distorted video frames.

The matching signal is generated by the processing of finding the best matching area for the small blocks within each distorted video frame from a buffer of the neighboring region of the original video frame. This process yields the matched reference video sequence to be used in place of the reference sequence. The matching analysis is performed on the 9×9 pixel block of the intensity arrays which are added a field number dimension, and the pixel located on (P_x, P_y) in the reference field N can be represented as follows.

$$\mathrm{Re}\, f\left(N, Px, Py\right) = \mathrm{Re}\, fYField(Px, Py) \tag{6}$$

Where *RefYField*(.) is the intensity arrays of the reference video sequence, and *Ref* (.) denotes the matching sequence. And each 9×9 pixel block with centre pixel (P_x, P_y) within the N-th field can be represented as Eq. (7).

$$\begin{gathered} Block\,\mathrm{Re}\, f\left(N, Px, Py\right) = \mathrm{Re}\, f\left(n, x, y\right) \\ x = Px-4, ..., Px, ..., Px+4 \\ y = Py-4, ..., Py, ...Py+4 \\ n = N-4, ..., N, ..., N+5 \end{gathered} \tag{7}$$

For the distorted video frames, the *Deg*(*n, x, y*) and Block *Deg*(*n, x, y*) can be defined similarly. According to the detected blocks, the minimum matching error, $E(N, P_x, P_y)$, can be calculated by searching the neighboring reference fields by using Eq. (8).

$$E(N, Px,\ Py) = \min\left(\frac{1}{81}\sum_{j=-4}^{4}\sum_{k=-4}^{4}\left(Deg(N, Px+j, Py+k) - \mathrm{Re}\, f(n, x+j, y+k)\right)^2\right) \tag{8}$$

Where N is the index of the distorted field which contains the being matched distorted block. The matching process of the first search for the distorted block followed by the copy of the resulting block into the matched reference array is repeated for the whole of the desired analysis region. This analysis region is defined by block centre points $P_x(.)$ and $P_y(.)$ which are shown in Eq. (9).

$$\begin{aligned} P_x(h) &= 16 + 8 \times h, \ \ h = 0, \ldots, Q_x - 1 \\ P_y(v) &= 16 + 8 \times v, \ \ v = 0, \ldots, Q_y - 1 \end{aligned} \tag{9}$$

Where Q_x and Q_y define the number of horizontal and vertical analysis blocks. Until now, the matched blocks of the original and distorted video sequences can be used to calculate the VQA of MPSNR as follows.

$$\begin{aligned} & if \quad \sum_{h=0}^{Q_x-1} \sum_{v=0}^{Q_y-1} E(N, P_x(h), P_y(v)) > 0 \\ & \rightarrow VQA_{MPSNR} = 10 \log_{10} \left(\frac{255^2 \times Q_x \times Q_y}{E\left(N, P_x(h), P_y(v)\right)} \right) \\ & Else\ if \ \sum_{h=0}^{Q_x-1} \sum_{v=0}^{Q_y-1} E\left(N, P_x(h), P_y(v)\right) \\ & \rightarrow VQA_{MPSNR} = 10 \log_{10} \left(255^2\right) \end{aligned}$$

Unlike the traditional PSNR, MPSNR calculates the signal to noise ratio by using the matched original and distorted video sequences. The matching operation can find the best match region for the small blocks within the limited field of distorted video frame to build a new matched reference. This operation can avoid the inaccurate measurements for the classic distortion such as jitter, jerkiness and give better correlation with subjective DMOS compared with the traditional PSNR.

3.3 Spatial Gradient (SG)

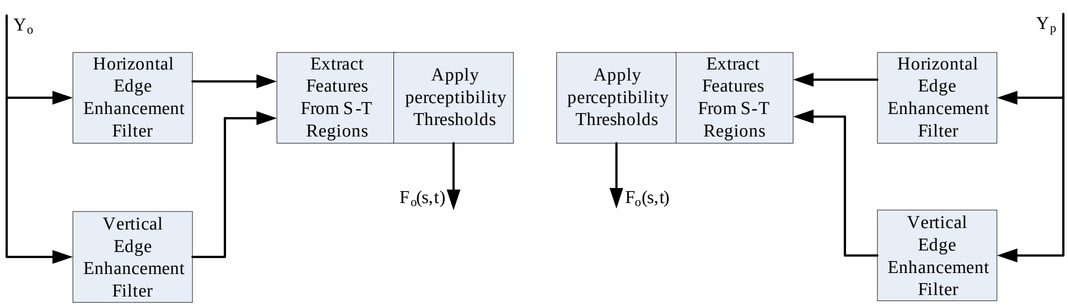


Fig.4. Extract spatial gradient features

Fig.4 presents the processing which is used to extract features based on the SG to characterize the perceptual distortions of edges. This measurement utilizes the precondition that HVS is more sensitive to the luminance components of the video frame. Firstly, the luminance of the reference and distorted video frames are processed with the horizontal and vertical edge enhancement filters that enhance edges while reducing noise. And then, the filtered video frames are divided into spatial-temporal (S-T) regions to extract two spatial activity features which can be used to detect spatial impairments such as blurring and blocking. The first feature is a measurement of overall spatial information (SI). It can be computed simply as the standard deviation over the S-T region and clipped at the perceptibility threshold. This feature is sensitive to changes in the overall amount of spatial activity within a given S-T region. The second feature is sensitive to changes in the angular distribution, of spatial activity within a given S-T region. It provides a simple means to include variations in the sensitivity of the HVS with respect to angular orientation.

3.4 Edge Degradation (*ED*)

For the subjective VQA system, HVS is very sensitive to the edge of the video frames. Therefore, the information extracted from edged video frame is necessary to establish the objective VQA. In our paper, we first apply the horizontal and vertical gradient operators by using the Sobel filter [10] to the reference and distorted video sequences to obtain the edged video frames, respectively. And then, the edged PSNR (*EPSNR*) between the reference and distorted video frames is calculated as Eq. (10).

$$EPSNR = 10\log_{10}\left(\frac{p^2}{mse_e}\right) \tag{10}$$

Where P is the peak pixel value which equals to 255 in common video frames and mse_e is the mean square error between the edged reference and distorted video frames. It is observed that when edges are severely blurred in low quality videos, evaluators tend to give lower subjective scores. In other words, if the edge areas of the processed video sequence are substantially smaller than those of the source video sequence, the evaluators give lower scores. In addition, it is observed that some video sequences have a very small number of pixels which have high frequency components. It implies that the number of pixels of edge areas is very small. In order to take into account these problems, the edge areas of the source and processed video sequences are computed and the EPSNR is reworked to modified *EPSNR* (*MEPSNR*).

$$\begin{aligned}
&if\ \ EPSNR < 25\,\&\&\,\frac{EP_{commom}}{EP_{src}} < 0.35\,\&\&\,t_e \ge 80 \\
&MEPSNR = EPSNR - 60\times\left(0.1225 - \left(\frac{EP_{commom}}{EP_{src}}\right)^2\right) \\
&Else\ if\ \ 35 \le EPSNR \le 40 \qquad MEPSNR = EPSNR \times 0.9 \\
&Else\ if\ \ EPSNR > 40 \qquad MEPSNR = EPSNR \times 0.8 \\
&Else \qquad MEPSNR = EPSNR
\end{aligned}$$

where *EPsrc* and *EPhrc* are the total number of the edge pixels in the source and processed video sequences, respectively. *EPcommon* is the total number of the common edge pixels in the source and processed video sequences. MEPSNR can eliminate the quality scaling error in the special cases and give the more correct VQA function. And the final objective ED scores can be rescaled as Eq. (11) so that they will be between 0 and 1.

$$VQM_{ED} = 1 - MEPSNR \times 0.02 \tag{11}$$

3.5 The entire function of the proposed VQA metric

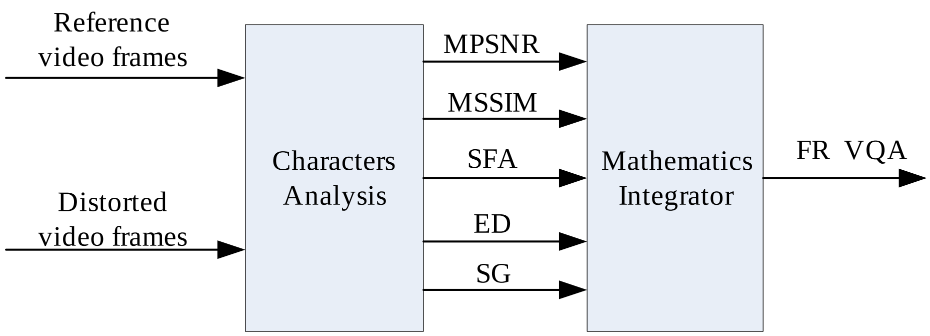


Fig.5. Proposed FR VQA structure

According to the context, the proposed VQA metric can be realized by using the five selected parameters, and Fig. 5 shows the integration structure. After the character analysis for the reference and distorted video frames, all the selected VQA parameters are input to the mathematics integrator which produces an estimation of the perceived video quality by appropriate weightings. The baseline of giving weight to each parameter in the integrator is "more essentiality more weight" which means that the bigger weight should be given to the parameter which has higher relationship with the subjective VQA to generate better correlation result with the DMOS. For instance, bigger weights should be given to MSSIM and SG, and the smaller weights should be given to the other parameters. Through the mathematic calculation, each weighted parameters can compensate to each other to work out the ideal VQA values. Eq. (12) gives the proposed FR VQA cost function in which Pi means the selected VQA parameters, w_i is the weighted value for the corresponding VQA parameters, and λ is the offset constant.

$$VQA_{proposed} = \sum_{i=1}^{N} w_i \times P_i + \lambda \tag{12}$$

The "Training Convergence Method" [11] is employed in the integrator to find the weight for each parameter. With the statistical analysis, the accurate function is given by Eq. (13).

$$VQA_{proposed} = -\begin{pmatrix} 68.11 \times VQA_{MSSIM} \\ +1.01 \times VQA_{MPSNR} \\ +26.3 \times VQA_{SPA} \\ +37.1 \times VQA_{ED} \\ +52.2 \times VQA_{SG} \end{pmatrix} + \lambda \tag{13}$$

The proposed FR VQA metric considers more aspects which can affect the quality assessment results compared with the existing metrics. Referring to Eq. (13), the five factors cover the most important aspects which can strongly response the VQA, such as the spatial domain, edge domain, matching field, and the most popular video impairments, for instance blurring, blocking, jitter jerkiness, and so on.

4 Experiment Results

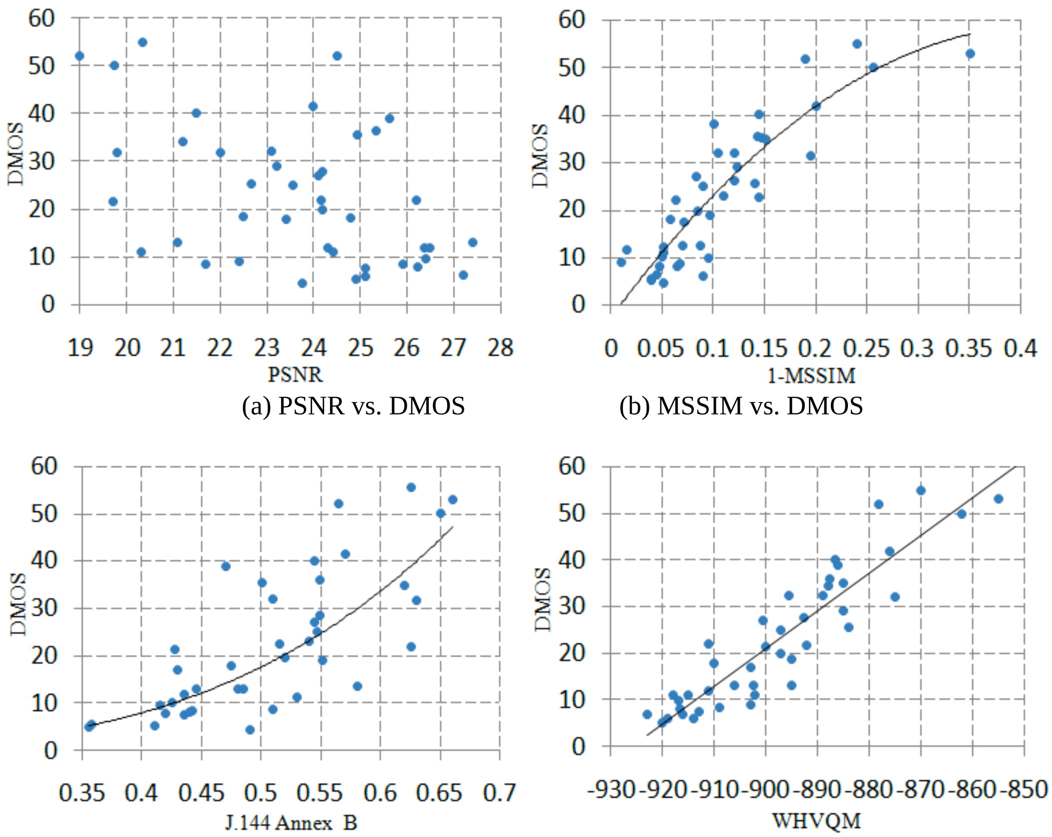
Table 1. Simulation environments

Items	Content	
Hardware	Intel® core TM2:CPU 3.2G Hz,2 GB RAM	
Software	Visual C++ 6.0	
Test video sequence	Video resource	VQEG Phase I: SRC_13~22 in HRC_1~16
	Video format	Frame size: 720 × 486
		Bit rate: 2.4 Mbps
Test VQA Metric	PSNR, MSSIM, J.144 Annex B, Proposed FR VQA	

To evaluate our proposed metric, in this section we present the simulation results on validation of the proposed metric. The SDTV sequence groups in HRC_N which are provided by VQEG Phase I [12] are utilized as the test video sequences, where HRC_N (N=1~16) means the SDTV sequence groups, and each of them stores 22 SDTV sequences noted as SRC_M (M=1~22) which are generated from the different reference video sequence. Each SRC_M in HRC_N is generated by the same reference video sequence but has different subjective quality scores. For example, SRC_1 in HRC_1 and SRC_1 in HRC_2 come from the same reference video sequence but have different DMOS values. Table 1 gives the simulation environments. To evaluate the proposed metric, we conduct a set of experiments by using the different test video sequences, and we also select several relevant FR VQA metrics, such as PSNR, MMSIM and Annex B of J.144 as the reference FR VQA metrics to compare the experimental results.

Fig. 6 shows the scatter plots of the subjective quality score versus different VQA metric. And Table 2 gives the summaries of our simulation results.

From Fig. 6 we can find that the PSNR almost cannot response the DMOS of the SDTV video sequences because of its weak relationship with subjective HVS judgment. Although it can be used as the criterion in theoretical analysis, however, it is defective to represent the quality of the video programs correctly. Except for the proposed VQA metric, MSSIM can present the DMOS more correctly. But for the high DMOS region, it cannot catch the approximate logical mapping with the DMOS of the video programs sequences because the blurring impairment of the low quality video frame cannot be detected accurately. J.144 Annex B can be implemented flexible because of it simple and convenient VQA function. However, its VQA ability is limited because it only measures the VQ only considering the edge affection. Fig. 6 (d) gives nearly linear correlation between the subjective DMOS of the test video sequences and our proposed objective FR VQA metric. Since we consider multiple effects for the VQA, the proposed metric can measure the multiple impairments for the distorted video sequences and estimate the video quality more accurately. According to Table 2, we can see that the proposed metric has the highest correlation with the subjective VQ DMOS.



(a) PSNR vs. DMOS (b) MSSIM vs. DMOS

(c) J.144 Annex B vs. DMOS (d) Proposed FR VQA vs. DMOS

Fig.6. Scatter plots for the video programs quality predictions by the four metrics

Table 2 · Correlation for the VQA metrics

VQA	PSNR	MSSIM	Annex B	Proposed
Correlatio n	0.4671	0.8369	0.7583	0.9162

5 Conclusions

In this paper, we proposed an efficient FR objective VQA metric for the multimedia broadcasting systems. The relationships between the VQA parameters and the quality score obtained by the subjective human visual system (HVS) model of the distorted videos are investigated. After reviewing the features of the parameters which can affect the results of VQS, we selected several of them based on the essentiality and generation, and built a weighted hybrid VQA. According the simulation results, the proposed FR VQA can represent the quality of the video signals which broadcasted through network efficiently has higher correlation with subjective video quality DMOS than the other existing video quality assessment metrics.

References

1. H.R. Wu and K.R. Rao, “Digital video image quality and perceptive coding,” CRC Press Taylor & Francis Group Boca Raton, London, New York, 2006.
2. M.H. Pinson and S. Wolf, “A new standardized method for objectively measuring video quality,” IEEE Trans. on Broadcasting, vol. 50, no. 3, pp. 312, Sept. 2004.
3. D. Richard and X.D. Yang, “Existing and emerging image quality metrics”, in Proceeding of the IEEE Int. Conf. on Electrical and Computer Engineering, pp. 1906-1913, May 2005.
4. J.S. Pan, F.R. Mclnnes, and M.A. Jack, “Bound for Minkowski metric or quadratic metric applied to VQ codeword search”, IEEE Proceedings-Vision, Image and Signal Processing, vol.143, pp.67-71, Feb. 1996.
5. Mikhail Cherniakov, Jonathan Lo, “Structural correlation between signal and CW interference in the spread spectrum systems”, in Proceeding of Global Telecommunications Conference , vol.4, Nov. 1998.
6. Avcibas I. Sankur B. and Sayood K., “Statistical evaluation of image quality measures”, Journal of Electronic Imaging, vol.11, April 2002
7. ITU-T Recommendation J.144, “Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference”.
8. Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, “Image quality assessment: from error visibility to structural similarity”, IEEE Trans. on Image Processing, vol. 13, No. 4, Apr. 2004.
9. H.R. Sheikh, A.C. Bovik and G.D. Veciana, “An information fidelity criterion for image quality assessment using natural scene statistics”, IEEE Trans. on Image Processing, vol.14, No.12, Dec. 2005.
10. Ziou D., Tabbone S., “Edge det ection techniques: an overview”, Pattern Recognition and Image Analysis, vol.8, no.4, 1998.
11. T. Poggio and S. Smale, “The mathematics of learning: dealing with data”, Notices of the AMS 50, pp.536-544, 2003.
12. VQEG Phase I SDTV video sequences and DMOS, http://www.its.bldrdoc.gov/vqeg/