Proceedings of the 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation, Madrid, Spain, February 15-17, 2006 (pp351-355)

Objective Video Quality Measurement Using Various Degradation Factors

Wonseok Ahn and Chulhee Lee
Dept. Electrical and Electronic Engineering, Yonsei University
134 Shinchon-dong Seodaemoon-gu, Seoul 120-749, Korea

Abstract: - In this paper, we propose a new video quality measurement method which is based on several degradation factors. The proposed method assumes that reference videos are available. In the proposed algorithm, the VQM (Video Quality Metric) is calculated as a weighted sum of seven features which represent sources of various video degradations. In particular, the degradation factors include PSNR in the lowest frequency band of the wavelet transform, PSNR in edge areas, edge error ratio, degradation ratio of edge images, a blocking coefficient, PSNR in visible chrominance signals and a contrast attenuation ratio. Experimental results show that the proposed method consistently provides satisfactory performances.

Key-Words: - video quality, objective video quality measurement, degradation factors.

1 Introduction

There is an increasing need for digital video quality measurement in the areas of multimedia services and various video processing techniques. In particular, as multimedia services such as VOD and video phones become widely available, quality monitoring emerges as an important topic. There have been numerous efforts to develop objective methods for video quality measurement, which can replace subjective video quality testing. In general, the performance of an objective method for video quality measurement is evaluated in the following three aspects: prediction accuracy, monotonicity and consistency [1].

In order to develop a video quality measurement model that provides consistently good performance, we investigate several degradation factors which can affect video quality. Furthermore, in order to overcome weakness of error-based methods, structural distortion-based methods are more desirable [2].

In the proposed video quality model, we investigate various degradation factors and their effects in digitally processed videos and selected most promising degradation factors. We compute the final VQM (video quality metric) as a linear combination of these degradation factors. Then, we find the optimal weight for these degradation factors so that the VQM has the maximum correlation with DMOS obtained from training data set. Fig.1 illustrates the proposed method. Experimental results show the effectiveness of the proposed method.

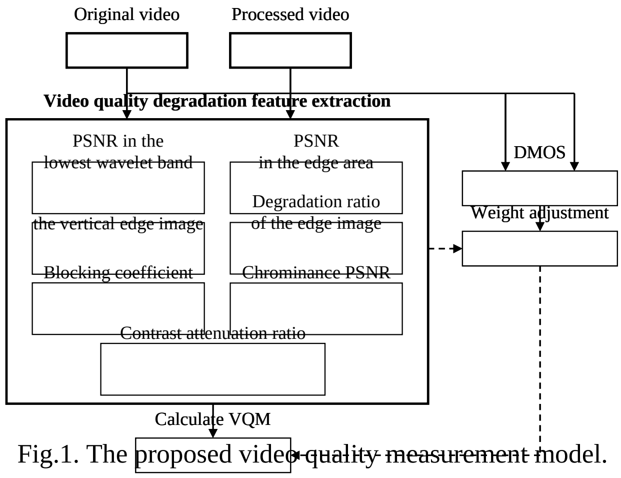

Fig.1. The proposed video quality measurement model.

2. Video Quality Degradation Factors

2.1. PSNR in the lowest frequency band of the wavelet transform

In general, low frequency components of images contain structural information of objects in the pictures. Therefore, the PSNR of the low frequency would provide information on overall degradation of videos. Thus, we propose to use this PNSR of the low frequency as a degradation factor and to use the lowest frequency band of the wavelet transform to compute the low frequency PSNR. In particular, we use the Haar filters and the low frequency PSNR is computed as follows:

$$PSNR_{LF} = 10\log(\frac{255^2}{MSE_{LF}}) \qquad (1)$$

where MSE_{LF} is the mean square error in the lowest frequency band of the wavelet transform.

2.2. PSNR in the edge area [4]

Distortion of in edge areas is a major factor of image degradation. It is reported that the degradation in edge areas of video is highly correlated with subjective scores [4]. In [4], edge areas are found using a modified Sobel filtering method and the EPSNR (edge PSNR) computed as follows:

$$SE^l{}_e = \sum_{j=0}^{N-1}\sum_{i=0}^{M-1}\left\{S^l(i,j) - P^l(i,j)\right\}^2 \quad if \quad \left|Q^l(i,j)\right| \geq t_e \qquad (2)$$

$$mse_e = \frac{1}{K}\sum_{l=0}^{L-1} se^l{}_s \qquad (3)$$

$$EPSNR = 10\log\left(\frac{p^2}{mse_e}\right) \qquad (4)$$

where $S(i,j)$ is the pixel value of the original image at the location of (i,j), $P(i,j)$ is the pixel value of the processed image and $Q(i,j)$ is the pixel value of the edge image obtained by the modified Sobel filtering method.

2.3. Edge error ratio

Edge differences between source video sequences and processed video sequences represent a major degradation which indicates that blurring or blocking occurs. The edge differences are computed by taking absolute edge differences between source video sequences and processed video sequences. This difference value takes the distortion of the edges of the objects and the blurring effects into account. We found that vertical edge differences are more useful and have better correlation with DMOS. The error ratio is defined as follows:

$$E_s = \left\{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1}\left|S^o{}_V(i,j) - S^p{}_V(i,j)\right|\right\} / \left\{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1} S^o{}_V(i,j)\right\} \qquad (5)$$

where $S^o{}_V(i,j)$ is a pixel value at the location of (i,j) of the original vertical edge image and $S^p{}_V(i,j)$ a pixel value of the processed vertical edge image. The edge image is obtained by applying the Sobel filter. It is noted that the error ratio is normalized.

2.4. Degradation ratio of the edge distribution

When blurring occurs, edges of processed video sequences will be severely blurred or weakened. In order to measure this degradation, the standard deviations of edge images of source and processed video sequences are compared. Then the ration of the standard deviations may roughly describe how much blurring had occurred in the processed video sequence. This ratio is obtained as follows:

$$D_e = \frac{stdev(S^p)}{stdev(S^o)} \qquad (6)$$

where S_o is the edge image of the source video sequence, S_p is the edge image of the processed video sequence, and $stdev(x)$ means the standard deviation of x.

2.5. Blocking coefficient

Blocking artifacts are dominant at low bit-rate coding when the block DCT transformation is used. Block is a major source in video quality degradation. There are a number of methods proposed to detect blocking artifacts [5]. In this paper, we propose a new method to detect blocking artifacts using the Sobel filter.

First, we obtain a horizontal edge image by applying the horizontal Sobel filter. Then we project the horizontal edge image and calculate a row image SH as follows:

$$SH(i) = \sum_{j=0}^{N-1} S_h(i,j)/N \qquad (7)$$

where $S_h(i,j)$ represents a pixel value at the location of (i,j) of the horizontal edge image of the processed video sequence. The image size is M by N. If the processed video sequence suffers from blocking artifacts, $SH(x)$ tends to have periodic impulse signals. Since the source video sequence may have vertical edges, we use a normalized blocking coefficient, which is computed as follows:

$$SH_r(i) = SH_h(i)/SH_o(i) \qquad (8)$$

$$SH_r(i) \approx \sum_{k=0}^{[M/B]-1} \delta[i - Bk] + n \qquad (9)$$

where $SH_h(i)$ is calculated from the processed video sequence and $SH_o(i)$ is calculated from the source video sequence. If blocking artifacts occur, $SH_r(i)$ can be approximated as an impulse train which has the period of the corresponding DCT block size B.
The period can be detected in the frequency domain by taking the Fourier transform of (9) as follows:

$$F(SH_r(i)) \approx \frac{M}{B} \sum_{k=0}^{[M/B]-1} \delta[i - \frac{M}{B}k] + F(n) \qquad (10)$$

It is noted that DFT of $SH_r(i)$ has impulses at the location of $M/2B$ and M/B at the same time if there are heavy blocking errors, and has impulse at the location of M/B if there is moderate blocking errors. Based on this characteristic, we can determine the degree of blocking artifacts. For this purpose, we define a blocking coefficient, which is computed as follows:

$$C_b(N_B) = \frac{1}{N_B}(1 \times B_H - 0.5 \times B_N) \qquad (11)$$

where N_B is the number of the separated image blocks, B_H is the number of blocks with heavy blocking, B_N is the number of blocks with no blocking.

2.6. PSNR in visible chrominance signals

It is known that small error at the chrominance signal is not perceivable. Especially if luminance signals have dominant errors, it is not necessary considering errors in chrominance signals. Thus we define the following rules when we use the chrominance PSNR as an image degradation factor:

Rule1. The chrominance signal is not perceivable with small errors in comparison with luminance signals.

Rule2. Errors in chrominance signals at the pixels near the high contrast area is not perceivable.

Rule3. Errors in chrominance signals affect image degradation if errors of Cr component and errors of Cb component are large enough.

Rule4. Errors in chrominance signals are important if the processed video has a well-defined shape.

Fig. 2 shows the computational procedure of the chrominance PSNR.

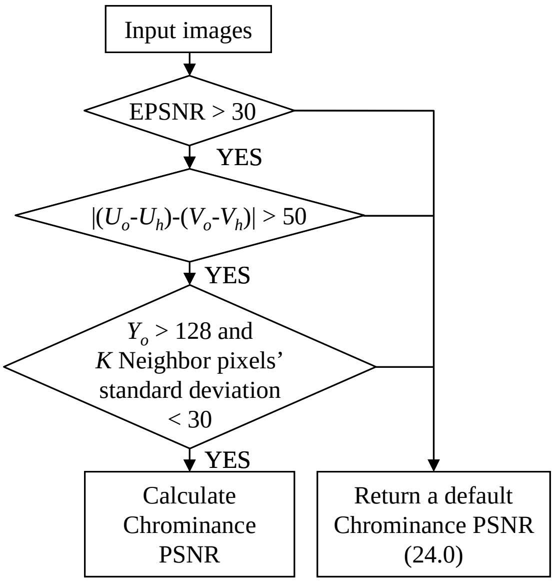

Fig.2. Calculation of chrominance PSNR.

2.7. Contrast attenuation ratio

In high quality videos, invisible high frequency noise may affect the objective score. We use the contrast attenuation ratio to take into account this phenomenon. The contrast attenuation ratio is defined as follows:

$$A_c = stdev(I_p) / stdev(I_o) \qquad (12)$$

where I_p is the processed image and I_o is the original image. In (12), $stdev(x)$ represents the standard deviation of x, which is typically used as the contrast of image x [6].

3. Computation of VQM

The final VQM is computed as a weighed sum of the seven degradation factors as follows:

$$vqm = \sum_{n=1}^{7} w_n f_n \qquad (13)$$

where w_n is the weight for the n-th factor, f_n is the n-th factor value. The video quality degradation factor value f_n is computed as follows:

$$f_1 = PSNR_w / 50 \quad (14)$$

$$f_2 = EPSNR / 50 \quad (15)$$

$$f_3 = E_s \quad (16)$$

$$f_4 = \begin{cases} 1 - \dfrac{C(I_p)}{C(I_o)} & if\ C(I_o)/C(I_p) \le 1 \\ 0 & elsewhere \end{cases} \quad (17)$$

$$f_5 = C_b(8) \quad (18)$$

$$f_6 = uvPSNR / 24 \quad (19)$$

$$f_7 = A_c \quad (20)$$

The vector form of (13) is given as follows:

$$vqm = W^T F = [w_1, w_2, ... w_7] \bullet [f_1, f_2, ... f_7] \quad (21)$$

where W is a weight vector and F a degradation factor vector.

4. Weight Optimization

The weights for seven degradation factors are calculated to maximize the Pearson correlation of DMOS and VQM. If there is v training sets of videos, the correlation of VQM and DMOS can be expressed as follows:

$$\rho = \frac{Cov(D, VQM)}{\sqrt{Var(D)Var(VQM)}} \quad (22)$$

where D is the DMOS vector of v videos ($D = [d_1, d_2, ..., d_v]$), and VQM is the VQM vector of respective videos ($VQM = [vqm_1, vqm_2, ..., vqm_v]$).
We can find the weight vector W using the optimization procedure of [3].

5. Experiments and Results

To evaluate the proposed model, we divide test videos into a number of groups. Each group does not have processed videos of the same source video. One group of videos is used to optimize the weight vector and the other group is used to validate the optimized weight vector.
Figs. 3~6 show the experimental results. The data points in the graphs have two symbols which indicate the different groups. Black circles filled represent the training data points while cross symbols represent test data points which were not used in the optimization procedure for the weight vector.

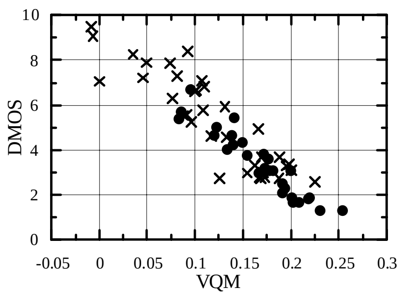

Fig.3. The model with *Weight 1*.

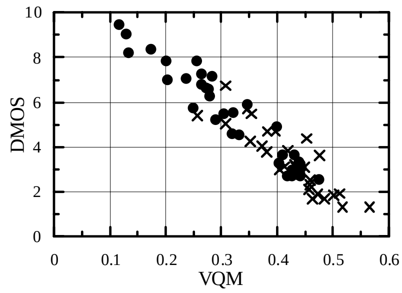

Fig.4. The model with *Weight 2*.

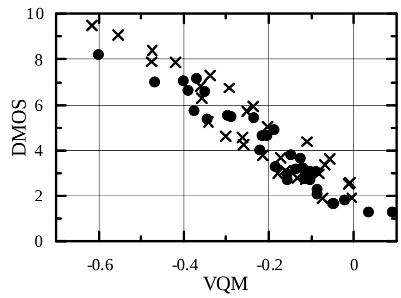

Fig.5. The model with *Weight 3*.

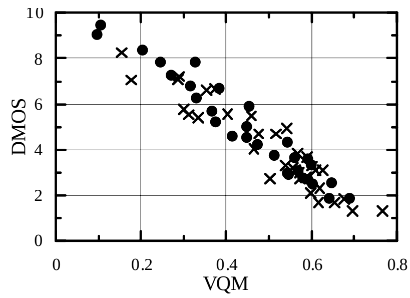

Fig.6. The model with *Weight 4*.

As can be seen in the figures, the proposed model provides consistent performances even for test videos which were not used in the optimization procedure.

However, it is observed that the performance slightly suffers if the training data are not adequate (Fig. 3). The training data of Fig. 3 have a narrow range of video quality.

The optimized weight vectors were applied to different video data sets that have completely different source and processed video sequences. Table I shows the results. Each data set has different source and processed video sequences. As can be seen in the table, the proposed method provides consistently better performances than the conventional PSNR. There is one exception (Weight 1, Data Set 2). It is found that the training data used to optimize Weight 2 were not a good representation of video quality.

6. Conclusions

In this paper, we proposed a new method for objective video quality measurement using video quality degradation factors. The proposed model consistently provides good correlations with subjective scores. However, we can improve the performance of the proposed method by developing and optimizing degradation factors. Experimental results show that the proposed method is consistent and provides robust performances even for data which were not used for optimization. However, cares should be taken in selecting training data which should be a good representation of video quality.

Table I. The correlation results by applying the different weight vectors to the various video sets.

	PSNR	The model using				
		Weight 1	*Weight 2*	*Weight 3*	*Weight 4*	*Weight 6*
Data Set 1	0.690	-0.791	-0.837	-0.850	-0.851	-0.808
Data Set 2	0.772	-0.661	-0.831	-0.868	-0.810	-0.843
Data Set 3	0.787	-0.922	-0.953	-0.948	-0.954	-0.933
Data Set 4	0.719	-0.859	-0.863	-0.856	-0.872	-0.872

References:

[1] P. Corriveau, *et al*., “Video quality experts group: Current results and future directions,” *Proc. SPEI Visual Comm. And Image Processing,* vol. 4067, Jun 2000.

[2] Z. Wang, L. Lu, and A.C. Bovik, “Video quality assessment using structural distortion measurement, ” ICIP, vol. III, pp. 65-68, Jun 2002.

[3] C. Lee, and O. Kwon, “Objective measurements of video quality using wavelet transform,” *Optical Engineering*, vol. 42, no. 1, pp. 265-272, 2003.

[4] C. Lee, “A new method for objective measurement of video quality based on edge degradation,” ITU-R Doc. 6Q/15-E, Sep 2003.

[5] Z. Wang, A.C. Bovik, B.L. Evan, “Blind measurement of blocking artifacts in images,” Image Processing, vol. 3, pp. 10-13, Sep 2000

[6] Z. Yu, and H.R Wu, “Human visual system based objective digital video quality metrics,” Proceedings of ICSP2000, Aug. 2000, pp. 1088-1095.