Quaternion model of video quality assessment based on structural similarity and inter-frame-different

GANG LI, YE-MING HE, XIAO-BO XU
College of Information Science and Engineering
Ningbo University
Ningbo 315211, China
ligang@nbu.edu.cn

JUN-LI LI
College of Computer Science
Sichuan Normal University
Chengdu 610068, China
li.junli@vip.163.com

***Abstract*—Recently, video quality evaluation method based on structure similarity (SSIM) has been widespread concerned. Compares with other traditional methods, it has a better performance and is simple to calculate. In this paper, we follow a new simple calculate metrics method in quaternion model, by introducing structure similarity as three parts of the quaternion, and the inter-frame-different (IFD) as the forth part to metric the structural distortion between the two frames, then we extract feature vector by using quaternion singular value decomposition. The algorithm is tested on the video quality experts group (VQEG) Phase I FR-TV test data set.**

Keywords-Video quality assessment; structure similarity; inter frame different; video quality experts group (VQEG)

I. INTRODUCTION

There has been an increasing need to develop video quality assessment methods that can predict perceived video quality automatically. These metric methods can be used in variety of video processing applications, such as compression, displaying, analysis, communication, enhancement and restoration.

Video quality assessment methods can be divided into two categories: subjective and objective assessment methods. However, subjective assessment methods are considered to be time-consuming and high cost, so objective assessment methods have gradually became a focus for researchers. The most commonly used objective video quality metrics are mean squared error (MSE) and peak signal-to-noise ratio (PSNR), because these methods are simple to calculate and have clear physical meanings. However, they have been widely criticized for not correlating well with perceived quality measurement [1][2].

There are also many HVS-based methods developed, which typically employ a frequency-based decomposition, or take into account the visual detectability of the distortion by considering human visual error sensitivities and masking effects varying in different spatial and temporal frequencies and directional channels, in spite of their complicated algorithm.

Wang et al believe that natural image signals and video signals are highly structured[3][4][5], the samples of the signals have strong dependencies between each other, especially when they are close in space, they develop a SSIM method for image quality assessment and VSSIM method for video quality assessment. Experiments show VSSIM is superior to the traditional video method, and still simple to calculate.

In this paper, we proposed a new method based on quaternion singular value decomposition, which combines structure similarity and inter-frame-difference (IFD). The results show the method is superior to MSE, PSNR and VSSIM, while still simple to calculate.

The rest of the paper is organized as follows. In section II, we give a brief introduction about singular value decomposition (SVD); in section III, we describe how to build quaternion matrix for video quality assessment; in section IV, we give out the steps to calculate; in section V, we compare the test results of different video quality assessment models tested on the video quality experts group (VQEG) Phase I FR-TV video dataset; finally, Section VI draws conclusions.

II. QUATERNION AND SINGULAR VALUE DECOMPOSITION

A quaternion matrix Q is made of one real part and three imaginary parts:

$$Q = a + bi + cj + dk \tag{1}$$

Where i , j , k obey the rules as below:

$$\begin{gathered} i^2 = j^2 = k^2 = -1 \\ i \cdot j = -j \cdot i = k, j \cdot k = -k \cdot j = i, k \cdot i = -i \cdot k = j \end{gathered} \tag{2}$$

The quaternion and quaternion matrix were introduced in [6], and can be extracted feature vector by singular value decomposition (SVD) was proved in [7]. According to the definition of SVD in [8], for any quaternion matrix A whose rank is r , there must be two quaternion unitary matrix U and V , obey:

$$A = U \begin{pmatrix} \sum_r & 0 \\ 0 & 0 \end{pmatrix} V^{\triangleleft} \tag{3}$$

Where $\sum_r = diag(\sigma_1, \sigma_2, \cdots, \sigma_r)$, $\sigma_i (1 \le i \le r)$ is the singular value of matrix A .

Quaternion matrix and singular value decomposition now is widely used in image and video research fields [14] [15].

III. QUATERNION MODEL OF VIDEO QUALITY ASSESSMENT BASED ON INTER-FRAME-DIFFERENT AND STRUCTURAL SIMILARITY

A. Inter-frame-difference (IFD) metrics

Some papers such as [9] and [10], use the difference of luminance between frames to calculate the IFD. However, in color video, the chrominance information is also important to human visual. Figure. 1 shows the IFD (c) between frame i (a) and frame i+1 (b).

(a)frame i (b)frame i+1 (c) IFD

Figure 1. Inter-frame-difference between the two frames

In this paper, We combine the difference of luminance and chrominance information between two frames. The definition of luminance difference dif_{lumi} and chrominance difference dif_{chro} of the j -th window in the i -th frame are defined:

$$dif_{lumi} = \sum_{pixel=1}^{bsize} (Y_{pixel}(i,j) - Y_{pixel}(i - i_{distance}, j)) \quad (4)$$

$$dif_{chro} = \sum_{pixel=1}^{bsize} (C_{pixel}(i,j) - C_{pixel}(i - i_{distance}, j)) \quad (5)$$

Where $bsize$ is the size of sample window, $Y_{pixel}(i,j)$ and $C_{pixel}(i,j)$ are the luminance value and chrominance value of the pixels in the j -th sampling window of the i -th frame, and $i_{distance}$ is inter-frame interval such as 3 in this paper.

The residuals difference of luminance $comp_{lumi}$ and chrominance $comp_{chro}$ between the distortion and reference video frames are defined:

$$comp_{lumi} = \frac{2(dif_{lumi})_x (dif_{lumi})_y}{(dif_{lumi})_x + (dif_{lumi})_y} \quad (6)$$

$$comp_{chro} = \frac{2(dif_{chro})_x (dif_{chro})_y}{(dif_{chro})_x + (dif_{chro})_y} \quad (7)$$

Where x and y are the distortion and reference video, respectively.

Finally, the residuals compare value $residual(x,y)$ of each window is defined:

$$residual(x,y) = (comp_{lumi} + comp_{chro})/2 \quad (8)$$

B. Structural similarity metrics

The similarity index measure in[3] is defined as:

$$SSIM(x,y) = f(l(x,y), c(x,y), s(x,y)) \quad (9)$$

Where the luminance comparison $l(x,y)$, contrast comparison $c(x,y)$ and structure comparison $s(x,y)$ are measured as follows:

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, C_1 = (K_1L)^2 \quad (10)$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, C_2 = (K_2L)^2 \quad (11)$$

$$s(x,y) = \frac{2\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}, C_3 = C_2/2 \quad (12)$$

μ_x , μ_y , σ_x^2 , σ_y^2 and σ_{xy} are the mean of x , the mean of y , the variance of x , the variance of y , and the covariance of x and y , respectively.

Constants C_1 and C_2 , are given by:

$$C_1 = (K_1L)^2, C_2 = (K_2L)^2 \quad (13)$$

where L is the dynamic range of the pixel values (for 8*8 sample window, L =255), and K_1 and K_2 are two constants whose values must be small such that C_1 or C_2 will take effect only when $(\mu_x^2 + \mu_y^2)$ or $(\sigma_x^2 + \sigma_y^2)$ is small. In this paper, we set K_1 = 0.01 and K_2 = 0.03, respectively.

In summary, we define the IFD between the two frames as one part of quaternion, and the other three parts are $l(x,y)$, $c(x,y)$, $s(x,y)$, respectively. The quaternion model can be represented as follows:

$$Q(x,y) = residual(x,y) + l(x,y)i + c(x,y)j + s(x,y)k \quad (14)$$

Where i , j , k are the unit of imaginary number. The $residual(x,y)$ part describes the structure distortion of frame sequence, and $l(x,y), c(x,y), s(x,y)$ parts describe the structure similarity of each frame.

IV. QUATERNION SINGULAR VALUE DECOMPOSITION AND VIDEO QUALITY ASSESSMENT

A. Frame Measure

According to the definition of singular value decomposition, each quaternion matrix has a unique singular value vector. We take the Euclidean distance of two singular values vector as the measurement of frame, the feature value of the frame i is:

$$Q_i = Sqrt\left(\sum_{j=1}^{ssize}\left(s_j - \hat{s}_j\right)^2\right) \quad (15)$$

Where s_j is the j -th singular values in the singular value vector of each frame in reference video, $\hat{s}_j$ is the j -th singular values in the singular value vector of each frame in distortion video, $ssize$ is total elements number in the singular value vector of the frame.

B. Inter-Frame Correlation

As the visibility of video distortion is easy impacted by the scene change, we take the inter-frame correlation [11] as a weighting factor between frames, as follows:

$$v_i = \frac{\sum_{x=1}^{W}\sum_{y=1}^{H} P_i(x,y)P_{i-1}(x,y)}{\sum_{x=1}^{W}\sum_{y=1}^{H} P_i(x,y)^2 \sum_{x=1}^{W}\sum_{y=1}^{H} P_{i-1}(x,y)^2} \qquad (16)$$

Where $P_i(x,y)$ is the pixel value of frame i in the reference video, and $P_{i-1}(x,y)$ is the pixel value of frame $i-1$ in the reference video.

C. Video Measure

The overall quality of the entire video sequence is given by:

$$QSSIM = \frac{\sum_{i=1}^{F} v_i * Q_i}{\sum_{i=1}^{F} v_i} \qquad (17)$$

Where F is the total number of the frames.

V. Experimental Results

The VQEG Phase I test dataset [12] for FR-TV video quality assessment is used to test this quaternion model. Three metrics are employed, First, the cubic polynomial functions are used in a fitting procedure to provide a non-linear mapping between the objective/subjective scores. Metric 1 is the correlation coefficient (CC) between objective/subjective scores after non-linear regression analysis, they provide the prediction accuracy evaluation, and the large value means the better accuracy. Metric 2 is the spearman rank order correlation coefficients (SPOCC), the large value indicates the better monotonicity. Metric 3 is the outlier ratio, the small value indicates the better prediction consistency.

In Table I, we give the comparison results of the three metrics. Despite its simplicity, the CC and SPOCC metrics of proposed method provides reasonably good results compared with the other approaches except the outlier ratio.

TABLE I. Performance Comparison of Video Quality Assessment Models on VQEG Phase I Test Data Set. Metric 1: Correlation Coefficient; Metric 2: Spearman Rank Order Correlation Coefficients; Metric 3: Outlier Ratio. P0~P9: The VQEG Proponents [13], Data for P0~P9 is from [13], Data for VSSIM is from [5]

Model \ Metric	Metric 1	Metric 2	Metric 3
P0(PSNR)	0.779	0.786	0.678
P1(CPqD)	0.794	0.781	0.650
P2(Tektronix/Sarnoff)	0.805	0.792	0.656
P3(NHK/Mitsubishi)	0.751	0.718	0.725
P4(KDD)	0.624	0.645	0.703
P5(EPFL)	0.777	0.784	0.611
P6(TAPESTRIES)	0.310	0.248	0.844
P7(NASA)	0.770	0.786	0.636
P8(KPN/Swisscom CT)	0.827	0.803	0.578
P9(NTIA)	0.782	0.775	0.711
VSSIM	0.849	0.812	0.578
QSSIM	**0.856**	**0.870**	**0.700**

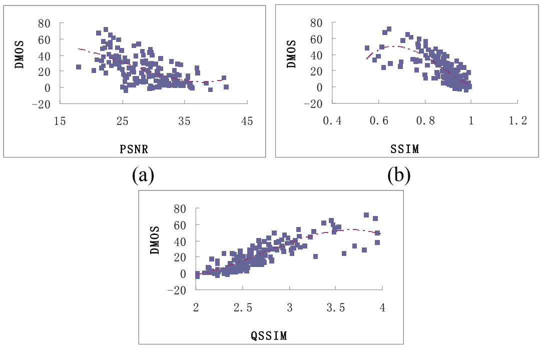

Figure 2. Scatter plot comparison of different video quality assessment models on VQEG Phase I test dataset. Vertical and horizontal axes are for subjective and objective measurements, respectively.

Figure. 2 (a), (b) and (c) show the scatter plots of the subjective/objective comparisons on test video sequences given by PSNR, the SSIM model, and the proposed model, respectively. The SSIM model here for video metrics is the mean SSIM value of all frames which is defined in [3].

VI. Conclusions

In this paper, We developed a new objective video quality assessment method using quaternion matrix. The proposed method use the inter-frame-difference (IFD) and structural similarity as the four parts of the quaternion, then the quaternion calculated by singular value decomposition. In addition, the new method is still simplicity. Experiments on VQEG FR-TV Phase I test dataset show that quaternion model has good correlation with perceived video quality.

Acknowledgment

This paper is supported by the natural science foundation of Ningbo City(2012A610047).

References

[1] Z. Wang, A. C. Bovik, and L. Lu, "Why is image quality assessment so difficult?", in Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing, May 2002.

[2] Z. Wang, H. R. Sheikh, and A. C. Bovik, "Objective video quality assessment," in The Handbook of Video Databases: Design and Applications (B. Furht and O. Marques, eds.), CRC Press, 2003.

[3] Wang Z, Bovik A C, Sheikh H R, et al, "Image Quality Assessment:From Error Visibility to Structural Similarity," IEEE Transactions on Image Processing, Vol 13, No. 04, pp. 600-612, 2004.

[4] Wang Z, Lu L G, Bovik A C, "Video Quality Assessment using Structural Distortion Measurement", 2002 International Conference on Image Processing, 3(24-28), pp. 65-68, 2002.

[5] Wang Z, Lu L G, Bovik A C, "Video Quality Assessment Based on Structural Distortion Measurement", signal processing: image communication, Vol19, No. 1, pp. 1-9, 2004.

[6] Zhang F, "Quaternion and Matrices of Quaternion," Lin. Appl. , pp. 215:21-57, 1997.

[7] D. Kahaner, C. Moler and S. Nash, "Numerical Methods and Software," Prentice-Hall, Inc, 1989.

[8] Bihan N L, Mars J, "Singular Value Decomposition of Quatemion Matrices: A New Tool for Vector-Sensor Signal Processing," Signal Processing, Vol 84, No. 7, pp. 1177-1199, 2004.

[9] Qiang Zhang, Long Wang, Huijuan Li, Zhaokun Ma, "Video fusion performance evaluation based on structural similarity and human visual perception," Signal Processing 92 (2012) pp. 912–925, 2012.

[10] Ee Ping Ong, Xiaokang Yang, Weisi Lin, et al, "Perceptual quality and objective quality measurements of compressed videos," J. Vis. Commun. Image R. 17 , pp. 717–737, 2006.

[11] A. P. Hekstra, J. G. Beerends, D. Ledermann, "PVQM-A Perceptual Video Quality Measure," Signal Processing:Image Communication, Vol 17, No. 10, pp. 781-798, 2002.

[12] VQEG, "Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment", 2003.

[13] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," Mar. 2000. http://www.vqeg.org/.

[14] Zhang Fei-yan, Xie Wei, Chen Rong-yuan, Qin Qian-qing, "Compression Image Quality Assessment Based on Human Visual Weight and Singular Value Decomposition", Journal of Electronics & Information Technology, Vol. 32, No. 5, pp. 1061-1065, 2010.

[15] Zhang Fu-qiang, Li Jun-li, Li Gang, et al, "Video Quality Assessment Based on Quaternion Singular Value Decomposition", ACTA ELECTRONICA SINICA, Vol 39, No.1, pp. 219-223, 2011.