

Research Article

Visual Perception Based Objective Stereo Image Quality Assessment for 3D Video Communication

^{1,2}Gangyi jiang, ¹Xiangyin Mao, ^{1,2}Mei Yu, ¹Zongju Peng and ¹Feng Shao

¹Faculty of Science Information and Engineering, Ningbo University, Ningbo, 315211, China

²National Key Lab of Software New Technology, Nanjing University, Nanjing, China

Abstract: Stereo image quality assessment is a crucial and challenging issue in 3D video communication. One of major difficulties is how to weigh binocular masking effect. In order to establish the assessment mode more in line with the human visual system, Watson model is adopted, which defines visibility threshold under no distortion composed of contrast sensitivity, masking effect and error in this study. As a result, we propose an Objective Stereo Image Quality Assessment method (OSIQA), organically combining a new Left-Right view Image Quality Assessment (LR-IQA) metric and Depth Perception Image Quality Assessment (DP-IQA) metric. The new LR-IQA metric is first given to calculate the changes of perception coefficients in each sub-band utilizing Watson model and human visual system after wavelet decomposition of left and right images in stereo image pair, respectively. Then, a concept of absolute difference map is defined to describe abstract differential value between the left and right view images and the DP-IQA metric is presented to measure structure distortion of the original and distorted abstract difference maps through luminance function, error sensitivity and contrast function. Finally, an OSIQA metric is generated by using multiplicative fitting of the LR-IQA and DP-IQA metrics based on weighting. Experimental results shows that the proposed method are highly correlated with human visual judgments (Mean Opinion Score) and the correlation coefficient and monotony are more than 0.92 under five types of distortions such as Gaussian blur, Gaussian noise, JP2K compression, JPEG compression and H.264 compression.

Keywords: 3D video communication, human visual system, image quality assessment, stereo image processing, visual perception

INTRODUCTION

With the rapid development of internet, communication, multimedia technology, applications of 3D video communication technologies will be more extensive nowadays, such as three dimensional television, stereo video conferencing systems and consumer electronics (Fan *et al.*, 2011; Shao *et al.*, 2012; Gotchev *et al.*, 2011). However, due to the tremendous data and the limit of signal transmit channel, stereo images often suffer distortion in the process of compression, transmission, reconstruction etc. Thus, Stereo Image Quality Assessment (SIQA) is a very important issue.

SIQA measures are generally divided into two categories: subjective and objective. Subjective testes have been standardized by the ITU (ITU-R Recommendation BT.500-11, 2002; ITU-T Recommendation P.910, 2008). As human eyes are the ultimate receivers of any visual information, the results of subjective approaches are reasonable and reliable. Some researchers adopted the results of subjective approaches to analyze the effects on the perceived

quality of stereoscopic image. Analyzed the camera parameters, the display duration, quality-asymmetric coding and other factors on the perceived quality of stereoscopic image (Ijsselsteijn *et al.*, 2000; Lambooi *et al.*, 2009; Tam, 2007; Wang *et al.*, 2009, 2012). However, subjective methods are time-consuming, expensive and impractical in real systems.

Up to now, the research on the latter is relatively few, which may be easy to be real-time implemented and can be used to benchmark stereo image coding. Yang *et al.* (2008) proposed a SIQA metric based on Peak Signal to Noise Ratio (PSNR). Owing to different visual effects for different signals, it was not well consistency with HVS. Shen *et al.* (2009) proposed a SIQA metric based on Structural Similarity (SSIM). With the limit of the image pair number of metric validation, it did not well reflect metric performance. Horita *et al.* (2000) proposed a color SIQA metric. Based on the statistical characteristics, it is vulnerable to the impact of image content. Benoit *et al.* (2008 a, b) proposed SIQA metric based on single view image quality assessment SSIM and C4, which mainly discussed the correlation between subjective perception

Corresponding Author: Gangyi Jiang, Faculty of Science Information and Engineering, Ningbo University, Ningbo, 315211, China

This work is licensed under a Creative Commons Attribution 4.0 International License (URL: <http://creativecommons.org/licenses/by/4.0/>).

and proposed SIQA under different combination ways of Left-Right view Image Quality Assessment (LR-IQA) and Depth Perception Image Quality Assessment (DP-IQA). In additional, Benoit *et al.* (2008b) also verified the significance of SIQA on disparity map computed through different algorithms. Experimental results show that SIQA is enhanced when adding the disparity distortion contribution. The performances of SIQA based on SSIM and C4 has subtle differences and has certain influence on disparity obtained by different algorithms. Due to hard to obtain the accurate disparity, DP-IQA is hard to assess accurately the depth and realism. However, the combination of LR-IQA and DP-IQA is worth to be deliberate. Sazzad *et al.* (2009) and Akhter *et al.* (2010) proposed SIQA metrics under the hypothesis that LR-IQA and DP-IQA depended on the local features such as edge, texture and flat, which is combined of LR-IQA and DP-IQA. The disparity proposed is obtained through estimating on a certain range. Akhter *et al.* (2010) verified the effect on performance under disparity estimated and absolute different map and the experimental result showed it has little influence and disparity estimation need spend more time and increase the computational complexity significantly. At the same time, the parameters are more and were formed in terms of the statistical characteristics, which make reliability reduced and make applicability restrict. The disparity is hard to accurately reflect depth perception through disparity estimation.

In consideration of inconformity between human judgments (Mean Opinion Score) and the above proposed metrics and to build an assessment mode which is agreed with the human visual system, we propose a new Objective Stereo Image Quality Assessment (OSIQA) method in this study. The proposed OSIQA method includes two parts, new LR-IQA and DP-IQA metrics. The former is comprised of Left view Image Quality Assessment (L-IQA) and Right view Image Quality Assessment (R-IQA) and takes into account Watson model, which is a classic and comprehensive human visual model with contrast sensitivity, masking and error and HVS features such as contrast sensitivity. The latter takes into account structure distortion of original and distorted abstract difference map.

THE PROPOSED OBJECTIVE STEREO IMAGE QUALITY ASSESSMENT (OSIQA) METHOD

The most difference of stereo images is depth perception to signal view images, which can make us real and involving. The lack of depth information can lead to discrepancy between plain image and stereo quality measures. For example, the left or right view in stereo images being subjected to blurring distortion can still make us get relatively well stereo experience, whereas single view image assessment quality assessment metric does not correlate with the enhanced stereo perception quality. Therefore, in this study, we

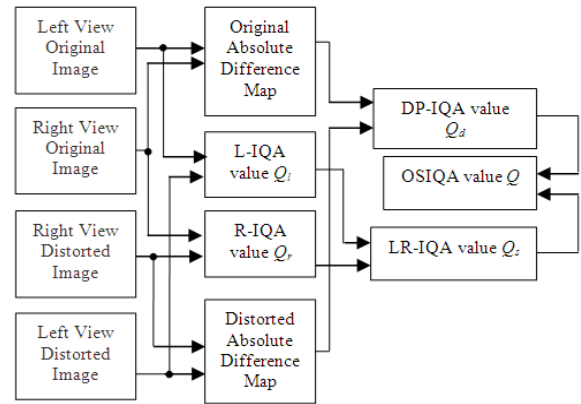


Fig. 1: The proposed Objective Stereo Image Quality Assessment (OSIQA) method

take this fact into account and propose an objective stereo image quality assessment method comprised of new LR-IQA and DP-IQA metrics, as shown in Fig. 1. Here, LR-IQA metric is mainly used to assess the circumstance of image quality distortion, while DP-IQA metric is mainly used to assess stereo perception.

In Fig. 1, the LR-IQA value Q_s decreases with increasing distortion degree, however, the DP-IQA value Q_d decreases with decreasing distortion degree. Therefore, let λ be a positive constant, then OSIQA value Q is given by:

$$Q = Q_s \times (Q_d)^{-\lambda} \quad (1)$$

The proposed LR-IQA and DP-IQA metrics will be described in detail in the succedent sections.

Left-Right view Image Quality Assessment (LR-IQA) metric based on visual perception: HVS is an extremely complex information processing system (Campisi *et al.*, 2007) and it is hard to understand deeply in the limit of many factors. An OSIQA metric, introduced visual perception characteristics such as human visual sensitivity, multi-channel characteristics, masking and stereo perception, will be able to improve the conformance between OSIQA metric and subjective perception. In the meantime, wavelet transform well presents multi-resolution decomposition characteristics of image and it can well describe multi-channel characteristics of HVS. Watson model may reflect the minimum visual error of stereo image. According to these analyses, we propose a new LR-IQA metric based on visual perception, which is comprised of Left view Quality Assessment (L-IQA) and Right view Quality Assessment (R-IQA) metrics. Its implementation includes three stages. In the first stage, L-IQA value is obtained. In the second stage, R-IQA value is obtained in the same process with L-IQA value. In the third stage, a pooling approach is employed to combine L-IQA and R-IQA metric values into a single quality score.

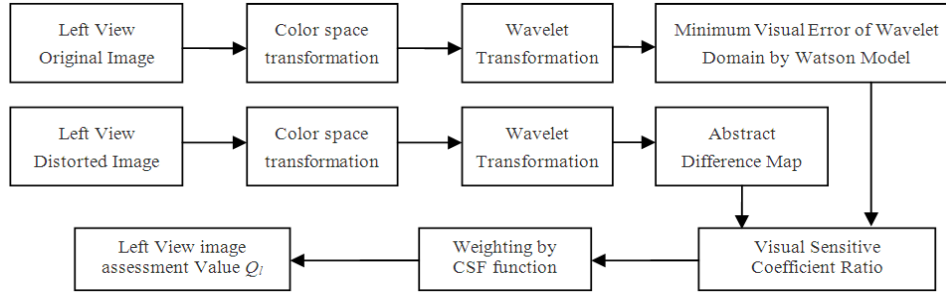


Fig. 2: Diagram of Left view Quality Assessment (L-IQA) model

Table 1: The coefficient value of $A_{l,f}$

Level direction	1	2	3	4	5	6
0	0.62171	0.34537	0.18004	0.09140	0.045943	0.023013
1	0.67234	0.41317	0.22727	0.11792	0.059758	0.030018
2	0.72709	0.49428	0.28688	0.15214	0.077727	0.039156
3	0.67234	0.41317	0.22727	0.11792	0.059158	0.030018

L-IQA metric: Figure 2 shows the diagram of L-IQA metric. Watson model is first adopted to compute the minimum visual error in each-bands after wavelet transformation. Then, the proportion of perceived coefficients in each-bands can be counted based on minimum visual error. Finally, assessment value can be gained by additive weight using visual sensitivity. YUV color space is used since it makes visually understandable and the redundancy of its components is less than that in RGB color space. Watson model (Daly, 1992) was used to describe a perception model estimating just notice difference of image. Thus, it can estimate the perception impact of image by adding distortion. Watson model is expressed by:

$$Q_{l,f} = 2 \times a \times 10^{(l \log 2^l \times f_0 \times g_f / r)^2} / A_{l,f} \quad (2)$$

where l denotes wavelet decomposition level, f expresses frequency direction, $A_{l,f}$ represents basis functions of each-bands shown in Table 1, r shows display visual resolution, a indicates the minimum threshold, g_f means direction function. When f is 0, 1, 2, or 3, the frequency direction is low-frequency, horizontal, vertical and diagonal, respectively. Table 1 shows basis functions of each-bands. Here, $a = 0.495$, $f_0 = 0.401$, g_f is given by:

$$g_f = \begin{cases} 1.501 & f = 0 \\ 1 & f = 1, 2 \\ 0.535 & f = 3 \end{cases} \quad (3)$$

According to the above discussion, $Q_{l,f}$ values of wavelet coefficients in Y component are shown in Table 2. As is well-known, the eyes are more sensibility to horizontal and vertical directions than diagonal direction. So, $Q_{l,f}$ values of horizontal and vertical directions are smaller than diagonal direction's in the same wavelet decompose level. Moreover, the

Table 2: $Q_{l,f}$ value of Y component

Level direction	1	2	3	4
0	14.05	11.11	11.36	14.50
1	23.03	14.68	12.71	14.16
2	58.76	28.41	19.54	17.86
3	23.03	16.69	12.71	14.16

sensibility is different for different wavelet decomposes level. The smaller $Q_{l,f}$ value is, the more the sensibility is.

Contrast Sensitive Function (CSF) describes the relationship between HVS and frequency information. CSF is a frequency function with a characteristic of band-pass filter. If the frequency is too high or too low, CSF is decreasing sharply. In addition, CSF can make human have the same visual sensitivity in different frequencies (Wandell, 1995). Mannos gained CSF approximate curve through experiments (Mannos and Sakrison, 1974) and computed it by:

$$A(f) = 2.6 \times (0.0192 + 0.114f) \times e^{-(0.114f)^{1.1}} \quad (4)$$

Majority energy of image is concentrated in low-frequency including direct current component. So the weight of the low-frequency sub-band is set 1. The rest weights are gained by CSF integration of the corresponding frequency sub-band and is expressed as follows:

$$\omega_{l,f} = \int_a^b 2.6 \times (0.0192 + 0.114f) \times e^{-(0.114f)^{1.1}} / (b-a) \quad (5)$$

where a and b denote the lower and upper of frequency range, respectively.

Afterwards, the abstract difference of left view image in wavelet domain is given by:

$$A_{\text{left}}(l, f) = |R_{\text{left}}(l, f) - D_{\text{left}}(l, f)| \quad (6)$$

where $R_{left}(l, f)$ and $D_{left}(l, f)$ denote the original and distorted wavelet coefficient of left view in the l -th level and the f -th direction, respectively. $A_{left}(l, f)$ denotes the abstract difference in the sub-band of the l -th level and the f -th direction.

Then proportion of visual perception coefficient of abstract difference in the sub-band of the l -th level and the f -th direction can be calculated by:

$$P_{left}(l, f) = N_{left}(l, f, Q_{l,f}) / N_{left}(l, f) \quad (7)$$

where,

$N_{left}(l, f, Q_{l,f})$: The number of visual perception coefficient of abstract difference which is larger than $Q_{l,f}$ value in the l -th level and the f -th direction

$N_{left}(l, f)$: The total number of the corresponding sub-bands

In order to make the assessment metric more consistent with HVS, L-IQA metric is expressed by Eq. (8):

$$Q_l = \log_2 \left(6 + \sum_{l=1, f=0}^{l=4, f=3} \omega_{l,f} \times P_{left}(l, f) \right) / \sum_{l=1, f=0}^{l=4, f=3} \omega_{l,f} \quad (8)$$

LR-IQA metric: The assessment value Q_r of R-IQA metric can be obtained in the same way with the assessment value Q_l of the L-IQA metric described the above. According to the importance of Q_l and Q_r , the assessment value Q_s of the LR-IQA metric is given by:

$$Q_s = \omega_l \times Q_l + \omega_r \times Q_r \quad (9)$$

where ω_l and ω_r denote the importance of the L-IQA and R-LQA metrics, respectively. $\omega_l + \omega_r = 1$. The larger ω_l or ω_r is, the more the importance of L-IQA or R-IQA metric is and vice versa.

DEPTH PERCEPTION IMAGE QUALITY ASSESSMENT (DP-IQA) METRIC

Stereo perception is the capabilities of feeling the depth. Stereo perception is directly related to the stereo image quality. The literature (Akhter *et al.*, 2010) verified that it is little effect to stereo image quality using absolute difference map which is corresponding abstract differential value between left and right view images and the disparity through block matching which should spend much time in matching algorithms. Therefore, absolute difference map in the study is adopted in the DP-IQA metric. Besides, considering that it is much sensitive to edge region, the proposed DP-IQA metric is shown in Fig. 3. Its implementation includes four stages. In the first stage, absolute difference map is obtained. In the second stage, area segmentation map is obtained to original and distorted difference map. In the third stage, the assessment maps

of different areas are computed. In the fourth stage, a pooling approach is employed to combine assessment maps into a single quality score.

Abstract difference map: In order to reduce redundancy, original and distorted stereo images are converted into YUV space. Here, Y component only is considered. The original and distorted absolute difference maps are given in Eq. (9) and (10), respectively:

$$X_{org} = |I_{l_org} - I_{r_org}| \quad (10)$$

$$X_{dis} = |I_{l_dis} - I_{r_dis}| \quad (11)$$

where,

I_{l_org} and I_{r_org} : The original images of left and right view, respectively

I_{l_dis} and I_{r_dis} : The distorted images of left and right view, respectively

X_{org} and X_{dis} : The original and distorted absolute difference maps, respectively

Region segmentation: With large experiments, it is shown that the DP-IQA value of stereo image mainly depends on that of strong and weak edge regions. Edge is important feature of an image because it represent the visual separation between objects and is the most easily perceived. The importance of strong edges, weak edges, non-edges is gradually decreasing to human perception, which corresponds to their gradual decline of gradient magnitudes in the image gradient field. A simple edge detection operator, Sobel operator, is used to segment images. It has a symmetrical template. Sobel operator used has a horizontal template and a vertical template. Image segmentation is implemented as follows.

Firstly, convolute the original and distorted absolute difference map with each Sobel template individually to obtain two gradient fields, which have the same size as the original abstract difference map.

Secondly, compute the gradient magnitudes of the two gradient fields.

Thirdly, determine the thresholds $TH_1 = 0.12G_{max}$, where G_{max} is the maximal gradient magnitude of the original absolute difference map.

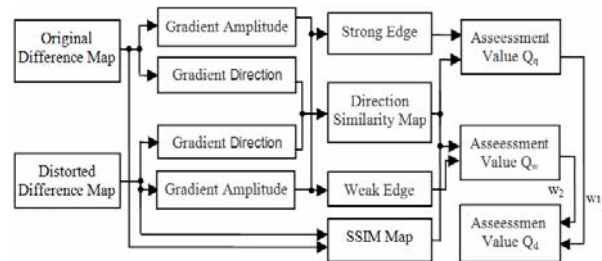


Fig. 3: Diagram of the proposed Depth Perception Image Quality Assessment (DP-IQA) metric

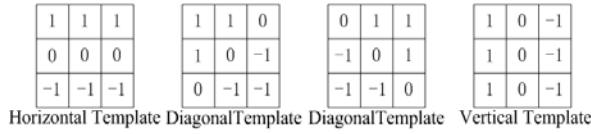


Fig. 4: The templates of four directions

Finally, determine the pixels of strong and weak edges region. Assume that the gradient of pixel at (i, j) of original absolute difference map is $P_r(i, j)$ and the gradient of distorted absolute difference map is $P_d(i, j)$. The pixel classification is carried out according to the following rules.

R1: If $P_r(i, j) > TH_1$ and $P_d(i, j) > TH_1$, then the pixel is considered as a strong edge pixel.

R2: If $P_r(i, j) > TH_1$ and $P_d(i, j) < TH_1$, or $P_r(i, j) < TH_1$ and $P_d(i, j) > TH_1$, then the pixel is considered as a weak edge pixel.

DP-IQA assessment map: HVS is highly adapted to extract structural information from the viewing field. SSIM proposed by Zhou *et al.* (2004) is not well conformance to subjection perception due to lack of direction information. Therefore, to improve it, a new DP-IQA metric is presented by combining SSIM and direction similarity index. SSIM includes three parts, luminance comparison $L(x, y)$, contrast comparison $C(x, y)$ and structure comparison $S(x, y)$ and is given by:

$$SSIM(x, y) = L(x, y)C(x, y)S(x, y) \tag{12}$$

$$= (2u_x u_y + C_1) \times (2\sigma_{xy} + C_2) / ((u_x^2 + u_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2))$$

where,

u_x and u_y : The (local) sample means of signal x and signal y , respectively

σ_x and σ_y : The (local) standard deviations of x and y

σ_{xy} : The (local) correlation coefficient between x and y

C_1 and C_2 : Constants

Next, in order to obtain direction similarity function D , the direction map of original absolute different map Ref and the direction map of distorted absolute different map Dis are elicited. Figure 4 shows the simple templates to extract direction information of absolute difference map. The direction similarity method is implemented as follows.

Firstly, convolute the original and distorted absolute difference maps with each direction template individually to obtain four gradient fields.

Secondly, determine the direction of pixels in original absolute difference map. Assume that the four gradient fields of the pixel at (i, j) of the original absolute difference map are $P_{r1}(i, j)$, $P_{r2}(i, j)$, $P_{r3}(i, j)$ and $P_{r4}(i, j)$. The direction classification is carried out according to the following rules:

R1: If $\max \{P_{r1}(i, j), P_{r2}(i, j), P_{r3}(i, j), P_{r4}(i, j)\} = P_{r1}(i, j)$, then $Ref(i, j) = 1$

R2: If $\max \{P_{r1}(i, j), P_{r2}(i, j), P_{r3}(i, j), P_{r4}(i, j)\} = P_{r2}(i, j)$, then $Ref(i, j) = 2$

R3: If $\max \{P_{r1}(i, j), P_{r2}(i, j), P_{r3}(i, j), P_{r4}(i, j)\} = P_{r3}(i, j)$, then $Ref(i, j) = 3$

R4: If $\max \{P_{r1}(i, j), P_{r2}(i, j), P_{r3}(i, j), P_{r4}(i, j)\} = P_{r4}(i, j)$, then $Ref(i, j) = 4$

Thirdly, determine the direction of pixels in distorted absolute difference map. Assume that the four gradient fields of the pixel at (i, j) of the distorted absolute difference map are $P_{d1}(i, j)$, $P_{d2}(i, j)$, $P_{d3}(i, j)$ and $P_{d4}(i, j)$. The direction classification is carried out according to the following rules.

R5: If $\max \{P_{d1}(i, j), P_{d2}(i, j), P_{d3}(i, j), P_{d4}(i, j)\} = P_{d1}(i, j)$, then $Dis(i, j) = 1$

R6: If $\max \{P_{d1}(i, j), P_{d2}(i, j), P_{d3}(i, j), P_{d4}(i, j)\} = P_{d2}(i, j)$, then $Dis(i, j) = 2$

R7: If $\max \{P_{d1}(i, j), P_{d2}(i, j), P_{d3}(i, j), P_{d4}(i, j)\} = P_{d3}(i, j)$, then $Dis(i, j) = 3$

R8: If $\max \{P_{d1}(i, j), P_{d2}(i, j), P_{d3}(i, j), P_{d4}(i, j)\} = P_{d4}(i, j)$, then $Dis(i, j) = 4$

Finally, compute the direction similarity of local region. Assume that $D(x, y)$ denotes the direction similarity of the $M \times N$ blocks with the center coordinate (x, y) between Ref and associated Dis . $A(x, y)$ denotes the number of the same direction of Ref and associated Dis in the corresponding block, $B(x, y)$ is the total number of Ref or Dis in the corresponding block. $D(x, y)$ is given by $D(x, y) = A(x, y) / B(x, y)$.

According to Fig. 3, strong edge assessment value Q_q and weak edge assessment value Q_w can be obtained according to the information mentioned above and Q_q and Q_w are given by:

$$Q_q = \frac{1}{N_s} \sum_{(x, y) \in s} SSIM(x, y) \times D(x, y) \tag{13}$$

$$Q_w = \frac{1}{N_w} \sum_{(x, y) \in w} SSIM(x, y) \times D(x, y) \tag{14}$$

where,

N_s = The number of strong edges in original or distorted absolute difference map

s = The strong edges

N_w = The number of weak edges in original or distorted absolute difference map

w = The weak edges

Weighted pooling: The DP-IQA metric value Q_d is gained by combining Q_q and Q_w and is given by:

$$Q_d = \omega_3 \times Q_q + \omega_4 \times Q_w \tag{15}$$

where ω_3 and ω_4 are the weights of Q_q and Q_w and the sum of Q_q and Q_w is 1. The larger ω_3 or ω_4 implies that Q_q or Q_w is more significant to DP-IQA metric.

EXPERIMENTAL RESULTS AND DISCUSSION

In order to verify the consistence of The proposed OSIQQA method with mean opinion score, the performance is evaluated over the stereo image database, which includes 12 original stereo images shown in Fig. 5, which provide by Mobile 3DTV Boev and Poikela (2010) and MPEG (1988), considering the camera spacing, wide variety of image contents, clarity of images and camera position arrangement. Five common distortion types are adopted, including Gaussian blurring, white Gaussian noise in the RGB components, JPEG, JPEG2000 and H.264 compressions. And the distortion processing referred to left and right view original stereo images is with the same distortion strength by the five distortion types. Variance σ_B of Gaussian blurring with zeros mean ranged from 0.2 to 50 pixels. Variance σ_N of white Gaussian noise with zeros mean ranged from 0.012 to 2.0. JPEG quality levels of JPEG ranged from 15 to 90.

Bit rates of JPEG 2000 ranged from 0.02 to 2 bits/pixel. Five distortion strengths for each original stereo image were adopted with each distortion proposed above. Quantization Parameter (QP) values of H.264 compressions ranged from 22 to 47 with JM 17.2 (2011). Six distortion levels were considered for each original stereo image. Eventually, 312 distorted stereo images were obtained. These distortions reflected a broad range of image impairments, from smoothing to structured distortion, image-dependent distortions and random noise. The level of distortion was varied to generate images at a broad range of quality, from imperceptible levels to high levels of impairment.

The subjective experiments (Zhou *et al.*, 2011) were conducted by a linear polarization stereo display system, which consisted of a workstation with a stereo player, two projectors with polarization filters, a metal screen and a number of linear polarization glasses. Test conditions complied with those described in ITU-R Recommendation (ITU-R Recommendation BT.500-11, 2002). Twenty-six subjects were recruited from Ningbo University, whose ages ranged from 20 to 25 with the mean age of 23.6. All subjects passed the test of visual acuity, depth perception ability, color vision and contract sensitivity ability. A Double Stimulus



Fig. 5: Left images of test stereoscopic image pairs used in the database

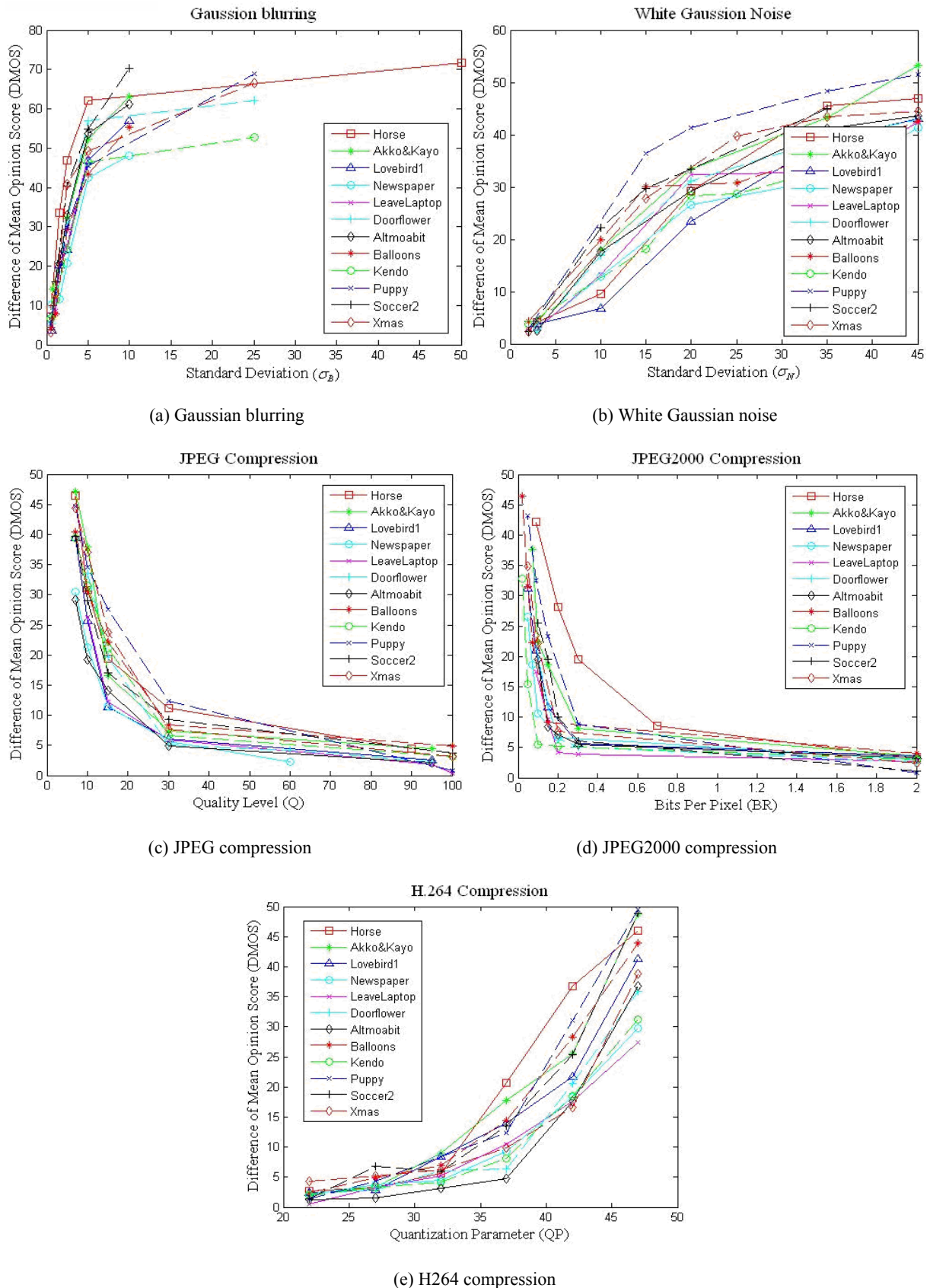


Fig. 6: Dependencies of subjective quality on distortion parameters with different distortion types (Zhou *et al.*, 2004)

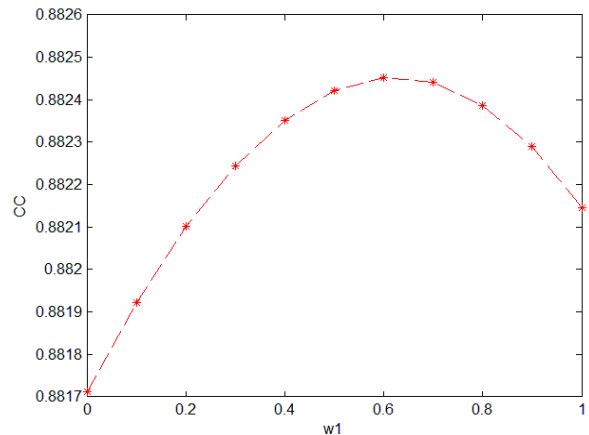
Continuous Quality Scale (DSCQS) (ITU-R Recommendation BT.500-11, 2002) test methodology was adopted. Meantime, the subjects scored up their judgments of quality and raw quality scores were generated. The subjective quality scores are given in terms of Difference Mean Opinion Score (DMOS) on a scale of 0 to 100 with larger value indicating declined visual quality through processing of raw quality scores according to Kurtosis method in which the raw data of four subjects were rejected. Figure 6 is given the dependence of DMOS and distortion parameters with different distortion types. The subjective quality of stereo images drops with the increasing strength of different distortion types. It is consistent well with the tendency of the “ground truth” quality. In fact, the results illustrate the rationality of the experiments.

The performance criteria (Chou and Li, 1995) adopted in our tests includes the linear Correlation Coefficient (CC), Spearman Rank-Order Correlation Coefficient (SROCC) and Root Mean Squared Error (RMSE). All the performance criteria are computed based on DMOS and OSIQA scores. CC and RMSE are all calculated after the compensation by nonlinear mapping, while SROCC is independent of nonlinear mapping. The procedure of nonlinear mapping, adaptation of objective quality scores to fit subjective quality scores through a logistic curve, follows Eq. (16) (Zhang and Le, 2010) for all the quality measures under our tests:

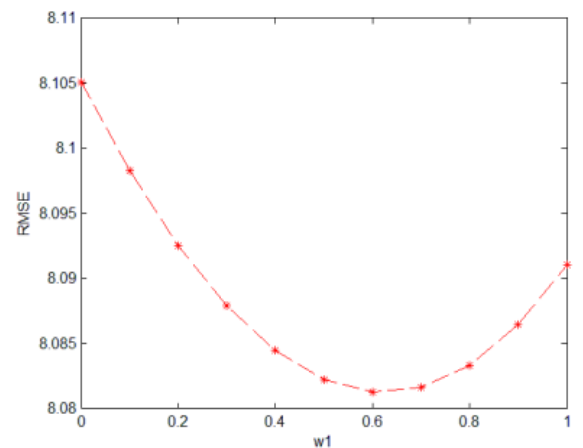
$$DMOS_p(Q) = \beta_1 / (1 + \exp[-\beta_2 \times (Q - \beta_3)]) + \beta_4 \quad (16)$$

where $DMOS_p(Q)$ denotes the predicted DMOS and β_k ($k = 1, 2, 3$ and 4) are fitting parameters. The fitting is done by the nonlinear regression over the dataset. Compared with CC, SROCC is generally considered as a less sensitive correlation measure because it operates only over data rank and ignores distance between data points. Considering SROCC as a widely used performance criterion, it is used in the following tests for completeness.

Analysis of weights: In the proposed method, wavelet basis adopts bior 4.4 and l is four. Firstly, it can make sure ω_1 and ω_2 through DMOS and the L-IQA metric value Q_l and R-IQA metric value Q_r . The importance of L-IQA value Q_l and R-IQA value Q_r is different for Different distortion types (Tam, 2007). Generally speaking, the stereo image quality depends largely on better quality between left and right view image under fuzzy distortion and the stereo image quality is about half of left and right view image quality under block effect. Figure 7 shows the relation between ω_1 value and DMOS. The performance of CC and RMSE is



(a) CC



(b) RMSE

Fig. 7: The relation between the ω_1 and DMOS

obtained through the total 312 L-IQA values Q_l , R-IQA value Q_r and $DMOS_p$. From Fig. 7, both of CC and RMSE exist a peak value when ω_1 is 0.6. Therefore, ω_1 is 0.6 and ω_2 is set to 0.4.

Secondly, it can make sure ω_3 and ω_4 value through DMOS, strong edges assessment value Q_s and weak edges assessment value Q_w . Figure 8 shows the relation between ω_3 value and DMOS. Both of CC and RMSE exist a peak value when ω_3 is 0.6. Therefore, ω_3 is 0.6 and ω_4 is set to 0.4.

Finally, the λ value is measured through LR-IQA value Q_s , DP-IQA value Q_d and DMOS. Figure 9 shows the relation between λ value and DMOS. Both of CC and RMSE exist a peak value when λ is 0.2. Therefore, λ is set to 0.2.

The performance of the new LR-IQA and DP-IQA metrics: The results of performance criteria concerning LR-IQA metric and DP-IQA metric are provided in

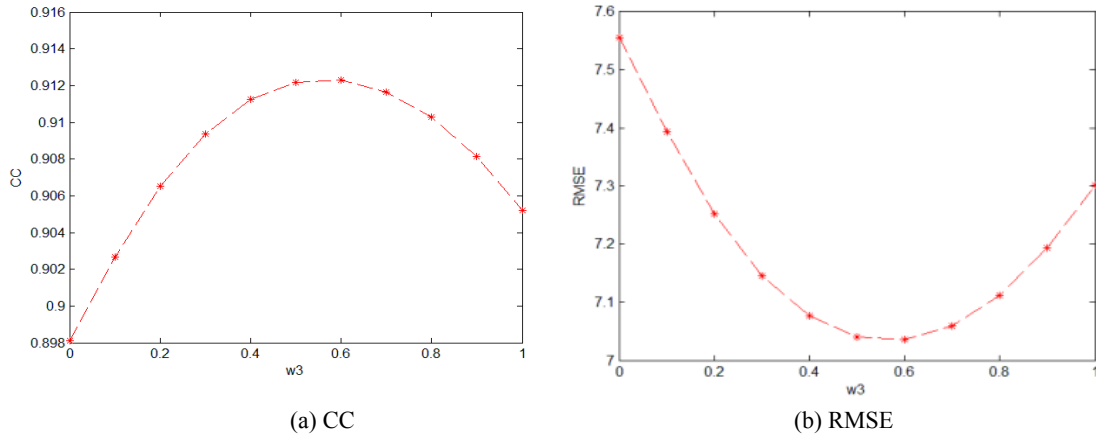


Fig. 8: The relation between the weighting of ω_3 and DMOS

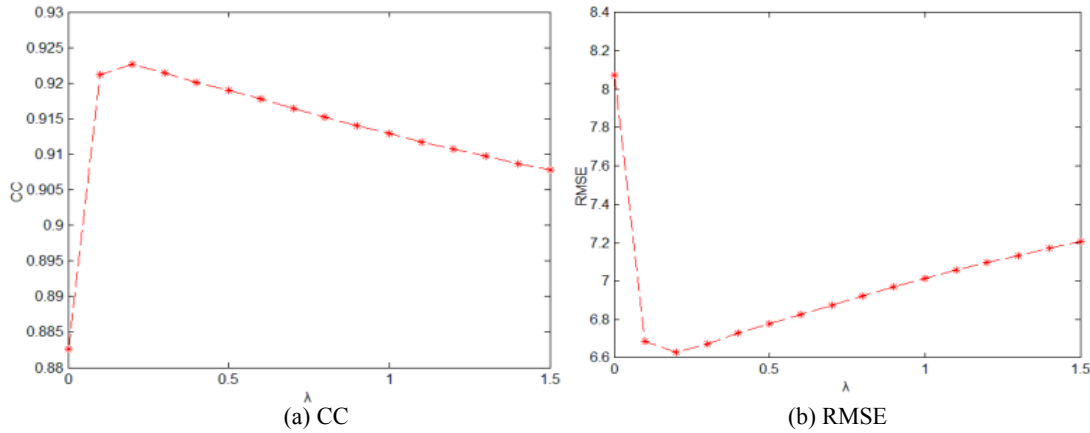


Fig. 9: The relation between the weighting of λ and DMOS

Table 3 and 4, respectively. As shown in Table 3 and 4, the performance of LR-IQA metric and DP-IQA metric are out line with DMOS. Especially, both of CC and SROCC value are lower than 0.9 under JP2K distortion in Table 3 and under JPEG distortion in Table 4.

The most difference between single view image and stereo image is the depth perception. Therefore, OSIQA value is comprised of LR-IQA value and DP-IQA value not only LR-IQA value or DP-IQA value. Neither LR-IQA nor DP-IQA is neglected in OSIQA method.

The performance of the proposed OSIQA method:

Table 5 shows the performance between the proposed OSIQA method and DMOS. Both of CC and SROCC are more than 0.92 under all kinds of distorted types. The RMSE are lower than 7. The scatter plot of DMOS versus OSIQA scores computed by the proposed measure is shown in Fig. 10. With most points close to the fitted logistic curve in the scatter plot, the proposed measure is observed to provide satisfactory prediction of DMOS for most images. It means that the judgment

Table 3: Performance of LR-IQA and DMOS

	CC	SROCC	MSE
Gblur (60)	0.954	0.952	6.297
JP2K (60)	0.863	0.898	6.063
JPEG (60)	0.895	0.913	6.353
WN (60)	0.944	0.924	5.167
H264 (72)	0.924	0.938	5.355
All (312)	0.912	0.921	7.036

Table 4: Performance of DP-IQA and DMOS

	CC	SROCC	RMSE
Gblur (60)	0.951	0.951	6.532
JP2K (60)	0.872	0.900	5.877
JPEG (60)	0.858	0.872	7.317
WN (60)	0.955	0.932	4.735
H264 (72)	0.920	0.912	5.507
All (312)	0.892	0.909	7.749

Table 5: Performance between OSIQA model and DMOS

	CC	SROCC	MSE
Gblur (60)	0.958	0.961	6.062
JP2K (60)	0.921	0.944	4.682
JPEG (60)	0.931	0.940	5.206
WN (60)	0.954	0.933	4.689
H264 (72)	0.941	0.943	4.759
All (312)	0.923	0.929	6.627

of the proposed OSIQA method is highly consistent with DMOS.

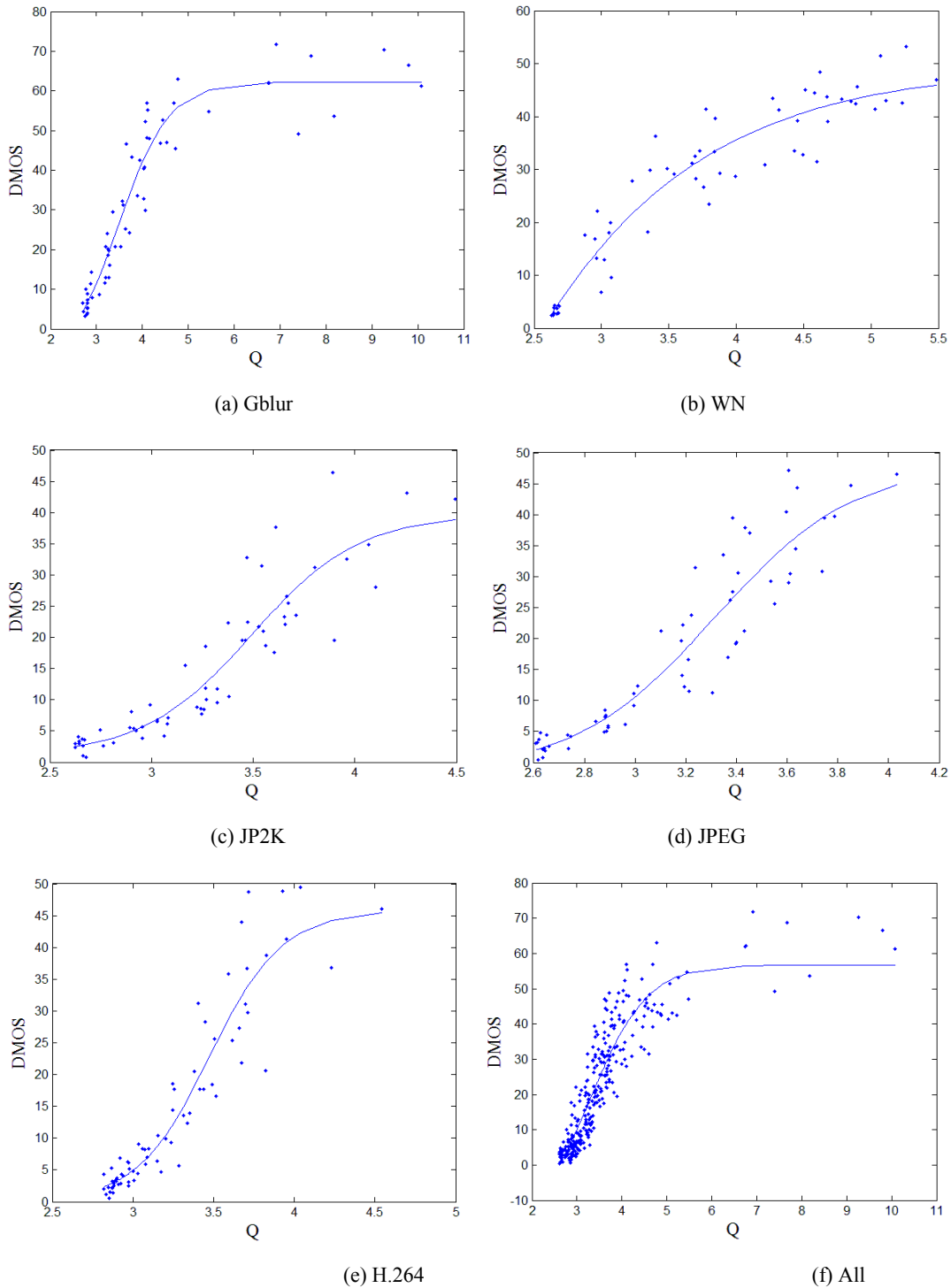


Fig. 10: Scatter plot of DMOS versus OSIQ scores computed by the proposed measure

CONCLUSION

In this study, we have proposed an Objective Stereo Image Quality Assessment (OSIQA) metric based Watson model, which consider the depth perception and some characters of Human Visual System (HVS). Experimental results show the proposed

OSIQA measure is proven to be an effective measure and is well consistent with human subjectivity. Both of CC and SROCC are more than 0.92 under five distortion types and RMSE is smaller than 6.7. However, the coefficient is only obtained by great many experiments. The coefficient needs to be improved in the future.

ACKNOWLEDGMENT

This study was by the Natural Science Foundation of China (Grant Nos. 61171163, 61071120, 61271270, 61271021), the Zhejiang Provincial Natural Science Foundation of China (Grant No. Y1101240) and Natural Science Foundation of Ningbo (2011A610200, 2011A610197).

REFERENCES

- Akhter, R., Z. Sazzad and Y. Horita, 2010. No-reference stereoscopic image quality assessment. *Proceeding of SPIE*, 7524(75240T).
- Benoit, A., P.L. Callet, P. Campisi and R. Cousseau, 2008a. Using disparity for quality assessment of stereoscopic images. *Proceeding of 15th IEEE International Conference on Image Processing (ICIP, 2008)*, pp: 389-392.
- Benoit, A., P.L. Callet, P. Campisi and R. Cousseau, 2008b. Quality assessment of stereoscopic images. *EURASIP J. Image Video Process.*, 659024. Doi: 10.1155/2008/659024.
- Boev, A. and M. Poikela, 2010. Technical Report D5.3. Retrieved from: <http://sp.cs.tut.fi/mobile3dtv/impaired-vides/>.
- Campisi, P., P.L. Callet and E. Marini, 2007. Stereoscopic images quality assessment. *Proceeding of EUSIPCO*, Poznan, Poland.
- Chou, C. and Y. Li, 1995. A perceptually tuned sub-band image coder based on the measure of just-noticeable distortion profile. *IEEE T. Circ. Syst. Vid.*, 5(6): 467-476.
- Daly, S., 1992. The visible difference predictor: An algorithm for the assessment of image fidelity. *Proceeding of SPIE*, 1616: 2-15.
- Fan, Y., Y. Kung and B. Lin, 2011. Three-dimensional auto-stereoscopic image recording, mapping and synthesis system for multiview 3D display. *IEEE T. Magn.*, 47(3): 683-686.
- Gotchev, A., G.B. Akar and T. Capin, 2011. Three-Dimensional media for mobile devices. *Proc. IEEE*, 99(4): 708-741.
- Horita, Y., Y. Kawai and Y. Minami, 2000. Quality evaluation model of coded stereoscopic color image. *Proceeding of SPIE*, 4067: 389-398.
- Ijsselstein, W.A., H.D. Ridder and J. January 03, 20132000. Subjective evaluation of stereoscopic images: Effects of camera parameters and display duration. *IEEE T. Circ. Syst. Vid.*, 10(2): 225-233.
- ITU-R Recommendation BT.500-11, 2002. Methodology for the Subjective Assessment Evaluation of the Quality of Television Pictures. ITU, Geneva, Switzerland.
- ITU-T Recommendation P.910, 2008. Subjective Video Quality Assessment Methods for Multimedia Applications. ITU, Geneva, Switzerland.
- JM 17.2, 2011. Retrieved from: <http://iphone.hhi.de/suehring/tml/download/>.
- Lambooi, M., W.A. Ijsselstein and M. Fortuin, 2009. Visual discomfort and visual fatigue of stereoscopic displays: A review. *J. Imaging Sci. Techn.*, 53(3): 0302011-03020114.
- Mannos, J. and D. Sakrison, 1974. The effects of a visual fidelity criterion on the encoding of images. *IEEE T. Inform. Theory*, 20(4): 525-536.
- MPEG, 1998. Retrieved from: <http://mpeg.chiariglione.org/>.
- Sazzad, Z.M.P., S. Yamanaka, Y. Kawayokeita and Y. Horita, 2009. Stereoscopic image quality prediction. *Proceeding of International Workshop on Quality of Multimedia Experience (QoMEX, 2009)*, pp: 180-185.
- Shao, F., G.Y. Jiang, M. Yu, K. Chen and S. Ho, 2012. Asymmetric coding of multi-view video plus depth based 3-D video for view rendering. *IEEE Multimedia*, 14(1): 157-167.
- Shen, L., J.C. Lang and Z.Y. Zhang, 2009. Quality assessment of stereo images with stereo vision. *Proceeding of 2nd International Congress on Image and Signal Processing (CISP '09)*, pp: 1-4.
- Tam, W.J., 2007. Image and depth quality of asymmetrically coded stereoscopic video for 3D-TV. Joint Video Team Document, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6. JVT-W094.
- Wandell, B., 1995. *Foundations of Vision*. 1st Edn., Sinauer Associates Inc., Sunderland, MA.
- Wang, X., M. Yu, Y. Yang and G.Y. Jiang, 2009. Research on subjective stereoscopic image quality assessment. *Proceeding of SPIE*, 7255: 725501-725509.
- Wang, X., G.Y. Jiang, J.M. Zhou, Y. Zhang, F. Shao, Z.J. Peng and M. Yu, 2012. Visibility threshold of compressed stereoscopic image: Effects of asymmetric coding. *Imag. Sci. J.*, Doi: 10.1179/1743131X11Y.0000000035.
- Yang, J., C.P. Hou, L.L. Shen and Z.Y. Zhang, 2008. Objective evaluation method for stereo image quality based on PSNR. *J. Tianjin Univ., Sci. Technol.*, 41(12): 1448-1452.
- Zhang, J. and T.M. Le, 2010. A new no-reference quality metric for JPEG 2000 images. *IEEE T. Consum. Electr.*, 56(2): 734-750.
- Zhou, W., A. Bovik, H. Sheikh and E. Simoncelli, 2004. Image quality assessment: From error visibility to structural similarity. *IEEE T. Image Process.*, 13(4): 600-612.
- Zhou, J.M., G.Y. Jiang, X.Y. Mao, M. Yu, F. Shao, Z.J. Peng and Y. Zhang, 2011. Subjective quality analyses of stereoscopic images in 3DTV system. *Proceeding of IEEE Visual Communications and Image Processing (VCIP'2011) Conference*. Tainan, pp: 1-4.