Objective quality assessment for stereoscopic images based on structure-texture decomposition

Kemeng Li, Shanshan Wang, Qiuping Jiang, Feng Shao*
Faculty of Information Science and Engineering
Ningbo University
Ningbo 315211
CHINA
KemngLi@gmail.com

Abstract: - We present a novel quality assessment index for stereoscopic images based on structure-texture decomposition. To be more specific, we decompose a stereoscopic image pair into its structure and texture components. Then, gradient magnitude similarity (GMS) and luminance-contrast similarity (LCS) indexes are used to measure the qualities of structure components and texture components, respectively. Finally, the quality score is obtained by combining the above quality scores in a non-fixed manner. Experimental results on two publicly available 3D image quality assessment databases demonstrate that, in comparison with the related existing methods, the proposed technique achieves high consistency alignment with subjective assessment.

Key-Words: - Quality assessment; structure-texture decomposition; gradient magnitude similarity; luminance-contrast similarity.

1 Introduction

Currently, high-quality assessment for 3D media has a huge market demand. However, assessing the 3D image/video quality is still a challenging issue because it is affected by the factors from 2D image quality, depth perception and visual experience [1,2]. Therefore, understanding of the binocular/3D vision perception is particularly important in designing 3D image quality assessment (3D-IQA) metric.

Many well-known 2D image quality assessment (2D-IQA) approaches has been proposed over the last several decades, such as Structural SIMilarity (SSIM) [3], multi-scale SSIM (MS-SSIM) [4], UQI (universal quality index) [5], VIF (visual information fidelity) [6], etc. Among those metrics, structure degradation has been employed for IQA in different ways. Gu et al. combined structural degradation model into no-reference IQA metric [7]. Li et al. proposed an IQA metric that separately evaluates detail losses and additive impairments [8]. Wu et al. measured image quality by considering spatial distribution of structure [9]. Many such metrics can be found in literatures [10-11]. However, 3D-IQA is still a less investigated problem due to lack of understanding of 3D visual perception.

The most direct way of applying state-of-the-art 2D-IQA methods to 3D-IQA is to evaluate the two views of the stereoscopic images, disparity/depth image separately using the existing 2D-IQA metrics, and then combined into an overall score. Boev et al. [12] combined monoscopic and stereoscopic quality components from the 'Cyclopean' image and disparity map respectively. Benoit et al. [13] computed quality scores of both stereo-pair and the disparity map by 2D quality metrics, and then combined them to produce a final score. You et al. [14] investigated ten common 2D quality evaluators on a stereo-pair and on its disparity map, and found the optimal combination which can yield the best performance. Hewage et al. [15] investigated the effectiveness of three 2D metrics (PSNR, VQM and SSIM) to predict the perceived quality of compressed color plus depth 3D video. However, for effective 3D evaluation, we cannot assess the perceived depth perception directly using 2D–IQA methods.

For measuring the perceived quality of stereoscopic images, several metrics have been proposed by more closely with 3D perceptual properties. Hwang et al. [16] devised a visual attention and depth assisted stereo image quality assessment by fusing the impact of stereo attention predictor, depth variation and stereo distortion predictor. Bensalma et al. [17] devised a Binocular Energy Quality Metric (BEQM) by modeling the complex cells responsible for the construction of the binocular energy, and evaluated the similarity between the binocular energy maps of the original
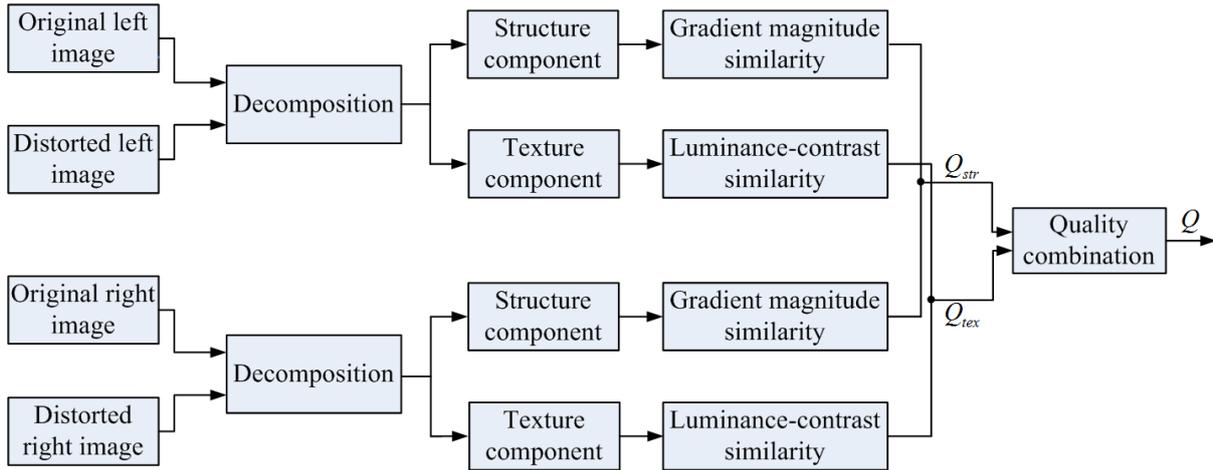
Fig.1. The framework of the proposed quality assessment metric.

and the distorted stereo-pairs. Chen et al. [18] constructed a "Cyclopean" view from the stereo-pair and depth information by modeling the influence of binocular rivalry, and evaluated the 'Cyclopean' view by 2D quality metrics. De Silva et al. [19] proposed a stereoscopic video quality metric to measure the perceptual quality of symmetrically and asymmetrically compressed stereoscopic video by extracting features that quantify the compression artifacts.

In this paper, we proposed an objective quality assessment index for stereoscopic images based on structure-texture decomposition. The motivation for the proposed structure-texture decomposition is that distortions on the two components have different impacts toward the perceptual quality. The main contributions of this paper are as follows: 1) we decompose a stereoscopic pair into its structure and texture components by region covariance matrices; 2) we measure the quality of structure components by using gradient magnitude similarity (GMS) index; 3) we measure the quality of texture components by using luminance-contrast similarity (LCS) index; 4) We conduct a non-fixed quality combination for the above scores to get a total quality score. The rest of the paper is organized as follows. Section II presents the proposed IQA for stereoscopic images. The experimental results are given and discussed in Section III, and finally conclusions are drawn in Section IV.

2 Proposed stereoscopic image quality assessment metric

The framework of the proposed quality assessment metric is illustrated in Fig.1. Given the original and distorted stereoscopic images (case of left and right images), they are firstly decomposed into their structure and texture components, respectively, and GMS and LCS indexes are used to measure the qualities of these structure and texture components, respectively. Finally, the two quality scores are combined to get a total quality score.

2.1 Structure-texture decomposition

Covariance of features is an effective means to describe the compactness of regions [20]. A region R containing n pixels can be represented with a $d \times d$ covariance matrix $\mathbf{C}_R$ of the feature points

$$\mathbf{C}_R = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{f}_i - \boldsymbol{\mu})(\mathbf{f}_i - \boldsymbol{\mu})^{\mathrm{T}} \qquad (1)$$

where $\{\mathbf{f}_i\}_{i=1,2,\ldots,n}$ denotes the d-dimensional feature points inside R, and $\boldsymbol{\mu}$ being the mean vector of these points.

In our implementation, we use simple features, namely intensity, orientation and coordinate so that a pixel can be represented with a 7-dimensional feature vector

$$\mathbf{f}(x,y) = \left[I(x,y) \quad \left|\frac{\partial I}{\partial x}\right| \quad \left|\frac{\partial I}{\partial y}\right| \quad \left|\frac{\partial^2 I}{\partial x^2}\right| \quad \left|\frac{\partial^2 I}{\partial y^2}\right| \quad x \quad y \right]^{T} \qquad (2)$$

where $I(x,y)$ is the intensity at pixel (x,y). Thus, the covariance descriptor of an image patch is computed as a 7×7 matrix. The design of the covariance matrix is based on the work in [21]. However, even though the covariance matrix can effectively represent the second-order statistics, first-order statistics also play an important role in feature description. To remedy this issue, the covariance matrix $\mathbf{C}_R$ is first decomposed using the Cholesky decomposition, and first-order statistics is incorporated to this representation. Thus, a $(2d+1)$-dimensional feature
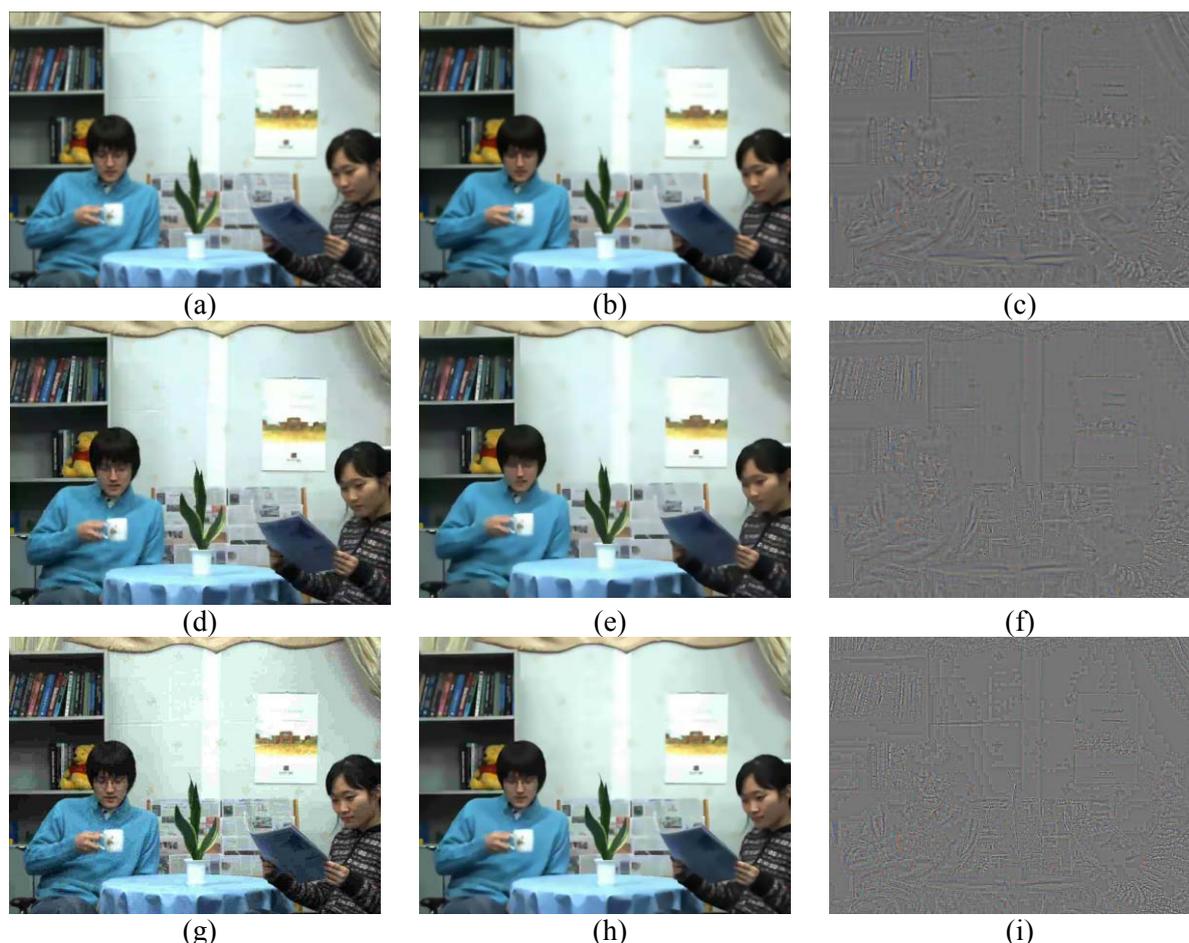
Fig.2. Examples of quality degraded left images and the corresponding structure and texture components of 'Balloons' test sequence. (a)~(d): (a) Gaussian blurred image; (b) structure component of (a); (c) texture component of (a); (d) H.264 compressed image; (e) structure component of (d); (f) texture component of (d); (g) JPEG compressed image; (h) structure component of (g); (i) texture component of (g).

vector $\mathbf{S}(\mathbf{C}_R)$ is obtained. The details of the procedure can refer to [21].

Then, we compute the structure component of a pixel $\mathbf{p}$ as

$$S(\mathbf{p}) = \frac{\sum_{\mathbf{q} \in N(\mathbf{p})} \omega_{\mathbf{pq}} I(\mathbf{q})}{\sum_{\mathbf{q} \in N(\mathbf{p})} \omega_{\mathbf{pq}}} \qquad (3)$$

where $I(\mathbf{p})$ is the intensity at pixel $\mathbf{p}$, ω_{pq} represents the similarity between two pixels $\mathbf{p}$ and $\mathbf{q}$ based on the similarity between patches centered on these pixels, and $N(\mathbf{p})$ denotes the neighborhood centered at p. The ω_{pq} is defined as

$$\omega_{\mathbf{pq}} = \exp\left(-\frac{\left\|\mathbf{S}(\mathbf{C}_{\mathbf{p}}) - \mathbf{S}(\mathbf{C}_{\mathbf{q}})\right\|^2}{2\sigma^2}\right) \qquad (4)$$

where $\mathbf{C}_p$ and $\mathbf{C}_q$ denote the covariance matrixes extracted from the patches centered at pixel $\mathbf{p}$ and $\mathbf{q}$, respectively, and σ denotes standard deviation of Gaussian function. In this work, the size of the patch is set to 9×9, the size of the neighborhood is set to 21×21 centered at a pixel, and σ is set to 0.06.

The texture component can be easily obtained by subtracting the structure component from the image

$$T(\mathbf{p}) = I(\mathbf{p}) - S(\mathbf{p}) \qquad (5)$$

Here, we present one example to illustrate the structure-texture decomposition results. The first row of Fig.2 shows: (a) Gaussian blurred left image of 'Newspaper' test sequences from NBU 3D IQA database, and the corresponding structure and texture components in (b)~(c). The second row of Fig.2 shows the H.264 compressed image in (d) and the corresponding structure and texture components in (e)~(f). The third row of Fig.2 shows the JPEG compressed image in (g) and the corresponding structure and texture components in (h)~(i). The difference mean opinion scores (DMOS) values for the Gaussian blurred, H.264 compressed and JPEG compressed stereoscopic images are 20.739, 18.391 and 21.217, respectively, that is, the subjective measures for these distorted stereoscopic images are

Table 1 Performance of the proposed method and the other schemes on the two databases.

IQA model	NBU (312 images)				LIVE I (365 images)				average			
	PLCC	SRCC	KRCC	RMSE	PLCC	SRCC	KRCC	RMSE	PLCC	SRCC	KRCC	RMSE
PSNR	0.8852	0.9032	0.7182	7.9923	0.8355	0.8341	0.6297	9.0104	0.8603	0.8687	0.6740	8.5013
SSIM	0.8111	0.9013	0.7202	10.0479	0.8766	0.8765	0.6791	7.8913	0.8439	0.8889	0.6997	8.9696
MS-SSIM	0.8623	0.9275	0.7538	8.7010	0.9335	0.9276	0.7562	5.8792	0.8979	0.9275	0.7550	7.2901
UQI	0.7678	0.7908	0.5844	11.2416	**0.9418**	**0.9373**	**0.7722**	**5.5141**	0.8548	0.8640	0.6783	8.3779
VIF	0.7925	0.8636	0.6693	10.4779	0.9250	0.9200	0.7400	6.2323	0.8587	0.8918	0.7046	8.3551
Benoit[13]	0.8645	0.8812	0.6956	8.6347	0.8927	0.8901	0.6947	7.3908	0.8786	0.8857	0.6952	8.0127
You[14]	0.7190	0.7324	0.5333	11.9396	0.9339	0.9247	0.7496	5.8638	0.8264	0.8285	0.6414	8.9017
Proposed	**0.9173**	**0.9156**	**0.7404**	**6.8404**	0.9411	0.9313	0.7644	5.5430	**0.9292**	**0.9234**	**0.7524**	**6.1917**

similar. It is seen that structure is successfully separated from texture: the structure component preserve the main edge information, while the texture component contain some randomly and irregularly structured textures. It is obvious that different prediction methods should be conducted for these components.

2.2 Quality assessment for structure component

Observed from the Fig.1, the structure components preserve the main edges while smoothing some detail information (e.g., texture). Therefore, we compute 2D gradients on horizontal and vertical directions, and use the gradient magnitudes of the original and distorted images to measure the explicit distortions around edge regions. The gradient magnitude similarity (GMS) index is defined as [22]

$$GMS = \frac{1}{N} \sum_{(x,y)} \frac{2m_o(x,y) \cdot m_d(x,y) + C_1}{m_o^2(x,y) + m_d^2(x,y) + C_1} \quad (6)$$

where the parameter C_1 is a constant to avoid the denominator being zero, N is the number of pixels in an image, $m_o(x,y)$ and $m_d(x,y)$ are the gradient magnitudes for the structure components of the original and distorted images, which is defined as the root mean square of directional gradients along two directions

$$m_o(x,y) = \sqrt{(\nabla f_x^o)^2 + (\nabla f_y^o)^2} \quad (7)$$

$$m_d(x,y) = \sqrt{(\nabla f_x^d)^2 + (\nabla f_y^d)^2} \quad (8)$$

In this paper, the GMS indexes for the structure components of the distorted left and right images are measured respectively. In order to facilitate the

following analysis, the quality scores for the left and right structure components are defined as Q_L^{str} and Q_R^{str}. Finally, the quality score for the structure component is computed as

$$Q_{str} = w_s \times Q_L^{str} + (1 - w_s) \times Q_R^{str} \quad (9)$$

where w_s is the weight assigned to the left structure component.

2.3 Quality assessment for texture component

For the texture components, the randomly and irregularly structured textures cannot be measured by the above GMS index. In this paper, the luminance-contrast similarity (LCS) index is defined as [23]

$$LCS = \frac{1}{M} \sum_{(m,n)} \frac{4 \times (\sigma_o(m,n) \times \sigma_d(m,n)) \times (\mu_o(m,n) \times \mu_d(m,n)) + C_2}{\left((\sigma_o(m,n))^2 + (\sigma_d(m,n))^2\right) + \left((\mu_o(m,n))^2 + (\mu_d(m,n))^2\right) + C_2}$$

$$(10)$$

where the parameter C_2 is a constant to avoid the denominator being zero, μ_o and μ_d are the mean intensities for the texture components of the original and distorted images respectively, and σ_o and σ_d are the corresponding standard deviations.

In this paper, the LCS indexes for the texture components of the distorted left and right images are measured respectively. In order to facilitate the following analysis, the quality scores for the left and right structure components are defined as Q_L^{tex} and Q_R^{tex}. Finally, the quality score for the texture component is computed as

$$Q_{tex} = w_t \times Q_L^{tex} + (1 - w_t) \times Q_R^{str} \quad (11)$$

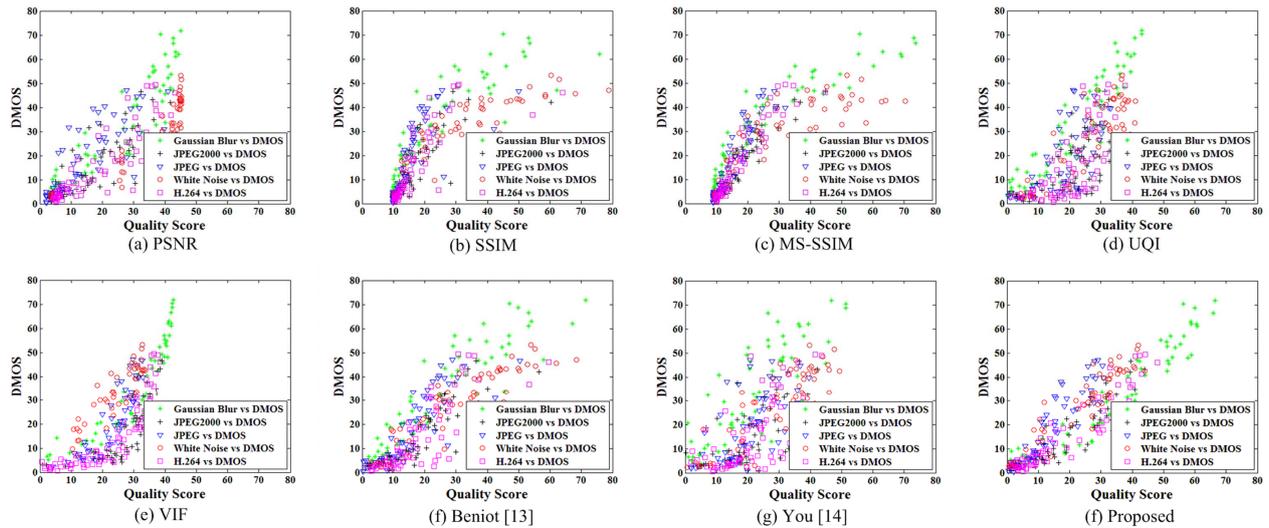where w_t is the weight assigned to the left texture component.

Fig.3. Scatter plots of predicted quality scores against the subjective scores (DMOS) of the eight methods on the NBU 3D IQA Databases.
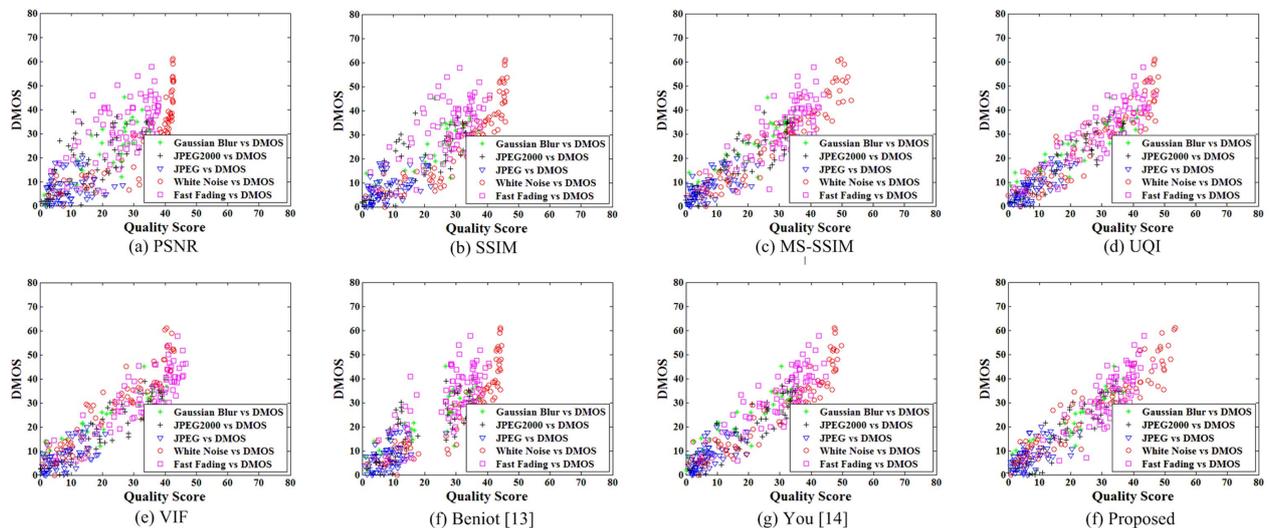

Fig.4. Scatter plots of predicted quality scores against the subjective scores (DMOS) of the eight methods on the LIVE 3D IQA Phase I Databases.

2.4 Final quality combination

After having obtained the quality scores Q_{str} and Q_{tex}, the next step is to combine the two quality scores into a final score. The direct way is to combine the quality scores Q_{str} and Q_{tex} by average weighting. In this work, we use the similar means to combine the two quality scores

$$Q = w \times Q_{str} + (1-w) \times Q_{tex} \qquad (12)$$

3 Experimental results

3.1 Databases and performance measures

In the experiment, two publicly available 3D IQA databases: NBU 3D IQA Database [24], and LIVE 3D IQA Phase I Database [25], are used to verify the performance of the proposed metric for stereoscopic images. The NBU 3D IQA Database consists of 312 distorted stereoscopic pairs generated from 12 reference stereoscopic images. Five types of distortions, JPEG, JP2K, Gblur, WN and H.264, are symmetrically applied to the left and right reference stereoscopic images at various levels. The LIVE 3D IQA Phase I Database consists of 365 distorted stereoscopic pairs generated from 20 reference stereoscopic images. Five types of distortions, JPEG, JP2K, Gblur, WN and FF, are symmetrically applied to the left and right reference

Table 2 PLCC performance comparison for the eight schemes on each individual distortion type.

	Criteria	PSNR	SSIM	MS-SSIM	UQI	VIF	Benoit[13]	You[14]	Proposed
NBU	JPEG	0.8309	0.6899	**0.9274**	0.7215	0.8210	0.8279	0.7059	**0.9307**
	JP2K	0.8310	0.7511	**0.9353**	0.6203	0.7541	0.8507	0.6527	**0.9153**
	Gblur	0.9255	0.8678	0.9009	**0.9400**	0.9139	0.9345	0.8289	**0.9686**
	WN	**0.9606**	0.8819	0.8452	0.8558	0.9445	0.9386	0.8770	**0.9486**
	H.246	0.9078	0.7576	**0.9445**	0.7028	0.8333	0.8108	0.6637	**0.9465**
LIVE I	JPEG	0.2276	0.4893	**0.6848**	**0.7737**	0.6813	0.5728	0.6333	0.6588
	JP2K	0.7878	0.8753	0.9346	**0.9510**	0.9374	0.8922	**0.9410**	0.9324
	Gblur	0.9160	0.9180	0.9446	**0.9568**	**0.9654**	0.9320	0.9549	0.9434
	WN	0.9352	**0.9421**	**0.9481**	0.9270	0.9309	0.9401	0.9351	0.9285
	FF	0.7007	0.6699	0.8102	**0.8791**	**0.8619**	0.7478	0.8592	0.8593

Table 3 SRCC performance comparison for the eight schemes on each individual distortion type.

	Criteria	PSNR	SSIM	MS-SSIM	UQI	VIF	Benoit[13]	You[14]	Proposed
NBU	JPEG	0.8490	0.8993	**0.9434**	0.7551	0.9392	0.8889	0.7606	**0.9436**
	JP2K	0.8628	0.8882	**0.9432**	0.6641	**0.9236**	0.8951	0.6754	0.9179
	Gblur	0.9529	0.9431	**0.9725**	0.9351	**0.9791**	0.9304	0.8175	0.9613
	WN	**0.9499**	**0.9492**	0.8825	0.7513	0.9137	0.9278	0.8341	0.9311
	H.246	0.9048	0.9199	**0.9436**	0.7404	**0.9257**	0.8402	0.6685	0.9172
LIVE I	JPEG	0.1212	0.4361	0.6244	**0.7374**	0.5807	0.4983	0.6008	**0.6310**
	JP2K	0.7993	0.8584	0.8947	**0.9077**	0.9015	0.8730	**0.9051**	0.9012
	Gblur	0.9020	0.8793	0.9258	0.9274	**0.9341**	0.8802	0.9300	**0.9344**
	WN	0.9316	0.9379	**0.9417**	0.9255	0.9313	0.9369	**0.9403**	0.9388
	FF	0.5875	0.5861	0.7496	**0.8329**	0.8042	0.6242	0.8030	**0.8205**

stereoscopic images at various levels for the LIVE 3D IQA Phase I Database.

In the paper, four commonly-used performance indicators are used to benchmark the proposed metric against the relevant state-of-the-art techniques: Pearson linear correlation coefficient (PLCC), Spearman rank order correlation coefficient (SRCC), Kendall rank-order correlation coefficient (KRCC), and root mean squared error (RMSE), between the objective and subjective scores. For a perfect match between the objective and subjective scores, PLCC=SRCC=KRCC=1, and

RMSE=0. For the nonlinear regression, we use the following five-parameter logistic function [26]:

$$DMOS_p = \beta_1 \cdot \left(\frac{1}{2} - \frac{1}{1+\exp(\beta_2 \cdot (x - \beta_3))} \right) + \beta_4 \cdot x + \beta_5$$

(13)

where β_1, β_2, β_3, β_4 and β_5 are determined by using the subjective scores and the objective scores.

In the experiment, the parameters w_s, w_t and w are trained by optimizing the PLCC values between the objective and subjective scores across the whole database. The parameter determination results is w_s=0.980, w_t=0.888 and w=0.882 for the NBU 3D

Table 4 Performance comparison for each component on the proposed scheme.

	Criteria	Structure component only				Texture component only				Proposed			
		PLCC	SRCC	KRCC	RMSE	PLCC	SRCC	KRCC	RMSE	PLCC	SRCC	KRCC	RMSE
NBU	JPEG	0.8713	0.8880	0.7183	6.9911	0.7820	0.8275	0.6516	8.8783	0.9307	0.9436	0.7917	5.2084
	JP2K	0.9259	0.9245	0.7511	4.5280	0.6527	0.6752	0.4956	9.0808	0.9153	0.9179	0.7511	4.8268
	Gblur	0.9694	0.9639	0.8381	5.1760	0.8948	0.8994	0.7228	9.4156	0.9686	0.9613	0.8290	5.2412
	WN	0.9425	0.9438	0.7903	5.2236	0.7732	0.6940	0.5404	9.9093	0.9486	0.9311	0.7733	4.9440
	H.246	0.9540	0.9303	0.7766	4.2057	0.7859	0.8046	0.5988	8.6770	0.9465	0.9172	0.7523	4.5287
	All	0.8683	0.8714	0.6839	8.5222	0.6827	0.7355	0.5345	12.5523	0.9173	0.9156	0.7404	6.8404
LIVE I	JPEG	0.3359	0.3374	0.2179	6.1592	0.5801	0.5609	0.3933	5.3265	0.6588	0.6310	0.4345	4.9194
	JP2K	0.8673	0.8768	0.6873	6.4465	0.8969	0.8457	0.6639	5.7292	0.9324	0.9012	0.7316	4.6818
	Gblur	0.9195	0.9317	0.7772	6.5398	0.7986	0.8028	0.6184	10.0134	0.9285	0.9388	0.7892	6.1774
	WN	0.9425	0.9327	0.7980	4.8371	0.8830	0.8555	0.6949	6.7933	0.9434	0.9344	0.7980	4.7991
	FF	0.8353	0.8094	0.6181	6.8320	0.7449	0.7072	0.5029	8.2896	0.8593	0.8205	0.6302	6.3563
	All	0.9013	0.9042	0.7220	7.1030	0.8429	0.8430	0.6425	8.8231	0.9411	0.9313	0.7644	0.7644

IQA Database, and w_s=0.629, w_t=0.503 and w=0.838 for the LIVE 3D IQA Phase I Database. It is clear that structure component is more important than texture component in measuring the quality degradation (i.e., w>0.5).

3.2 Overall assessment performance

In Table 1, we compare the competing 2D-IQA and 3D-IQA metrics' performance on the two databases in terms of PLCC, SRCC and RMSE. For the five 2D-IQA metrics, they directly estimate the quality of each view separately and generate a weighted average score. The proposed scheme outperforms the five 2D-IQA schemes in the NBU 3D IQA Database, and outperforms the four 2D-IQA schemes in the LIVE 3D IQA Phase I Database expect for UQI scheme. For You et al.'s and Benoit et al.'s schemes, since they are the combination of 2D image quality metrics for stereoscopic images and disparity maps, the performance of the two schemes is highly dependent on both the 2D image quality and the estimated disparity maps (stereo matching algorithm [27] is used in this paper). Since UQI always deliver the best performance on the LIVE 3D IQA Phase I Database (In agreement with the conclusion in [25]), this lead to the high performance of You et al.'s scheme promoted (UQI

is used as IQM in the scheme). Fig.3 and Fig.4 show the scatter plots of predicted quality scores against subjective quality scores (in terms of DMOS) for the eight schemes on the two databases, respectively. Overall, the proposed scheme has an impressive consistency with human perception.

3.3 Performance comparison on individual distortion types

To more comprehensively evaluate the prediction performance of the proposed method, we compare the eight schemes on each type of distortion. The PLCC and SRCC results are listed in Table 2 and Table 3, where the top two metrics have been highlighted in boldface. One can see that the proposed scheme is among the top 2 metrics 5 times in terms of PLCC, the same with MS-SSIM and UQI schemes, and is among the top 2 metrics 4 times in terms of SRCC, lower than MS-SSIM (among the top 2 metrics 5 times). However, the overall performance of MS-SSIM and UQI schemes do not perform significantly better than the proposed scheme on all the two databases. This validates that structure-texture decomposition can serve as an excellent processing for quality prediction.

3.4 Impact of each components in the proposed scheme

To demonstrate the impact of structure and texture components in the proposed scheme, we design a experiment for comparison by only measuring the quality score of structure or texture component with Eq.(9) or Eq.(11). The results are presented in Table 4. From the table, we can see that the structure component always provide higher performance than the texture component. Especially, the structure component is more effective for Gblur and WN distortion types, but the overall performances of both schemes are lower than the proposed scheme. It means that the quality combination operation can highly promote the evaluation performance.

4 Conclusion

In this study, we devised quality assessment index for stereoscopic images. More specifically, we decompose stereoscopic images into structure and texture components, and evaluate them respectively by adopting gradient magnitude similarity and luminance-contrast similarity indexes. Compared with state-of-the-art 2D and 3D image quality assessment metrics, the proposed metric performs well in terms of both accuracy and efficiency on two publicly available databases. However, structure-texture decomposition and quality assessment may suffer from large computational burdens. In the future work, we will consider how to speed-up the processing using GPGPUs. Furthermore, we will further explore how to model primary visual cortex receptive field into the metric.

References:
[1] A. K. Moorthy, A. C. Bovik , "A survey on 3D quality of experience and 3D quality assessment," in Proc. of SPIE: Human Vision and Electronic Imaging XVIII, vol. 8651, no. 86510M, San Jose, CA, USA, 2013.

[2] R. Vlad, P. Ladret, A. Guérin, "Three factors that influence the overall quality of the stereoscopic 3D content: image quality, comfort and realism," in Proc. SPIE: Human Vision and Electronic Imaging XVIII, vol. 8651, San Jose, CA, USA, March 2013.

[3] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, April 2004.

[4] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment", in Proc. of IEEE Asilomar Conf. Signals, Systems, and Computers, pp. 1398-1402, 2003.

[5] Z. Wang and A. C. Bovik, "A universal image quality index," IEEE Signal Processing Letters, vol. 9, no. 3, pp. 81-84, March 2002.

[6] H. R. Sheikh, and A. C. Bovik, "Image information and visual quality," IEEE Transactions on Image Processing, vol. 15, no. 2, pp. 430-444, Feb. 2006.

[7] K. Gu, G. Zhai, X. Yang, W. Zhang, L. Liang, "No-reference image quality assessment metric by combining free energy theory and structural degradation model," in Proc. of IEEE International Conference on Multimedia & Expo (ICME), San Jose, California, USA, pp. 1-6, July 2013.

[8] S. Li, F. Zhang, L. Ma, et al, "Image quality assessment by separately evaluating detail losses and additive impairments," IEEE Transactions on Multimedia, vol. 13, no. 5, pp. 935-949, Oct. 2011.

[9] J. Wu, W. Lin, G. Shi, "Image quality assessment with degradation on spatial structure," IEEE Signal Processing Letters, vol. 21, no. 4, pp. 437-440, April 2014.

[10] Z. Zhu, Y. Wang, "Perceptual distortion metric for stereo video quality evaluation," WSEAS Transactions on Signal Processing, vol. 7, no. 5, pp. 241-250, July 2009.

[11] I. Vlachos, G. Sergiadis, "Image quality indices based on fuzzy discrimination information measures," WSEAS Transactions on Circuits and Systems, vol. 3, no. 2, pp. 241-246, April 2004.

[12] A. Boev, A. Gotchev, K. Egiazarian, A. Aksay, and G. B. Akar, "Towards compound stereo-video quality metric: a specific encoder-based framework," in Proc. of IEEE Southwest Symposium on Image Analysis and Interpretation, Denver, Colorado, pp. 218-222, 2006.

[13] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," EURASIP Journal on Image and Video Processing, vol. 2008, Article ID 659024, 13 page, Jan. 2009.

[14] J. You, L. Xing, A. Perkis, and X. Wang, "Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis," in Proc. of International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, AZ, USA, 2010.

[15] C. T. E. R. Hewage, S. T. Worrall, S. Dogan, S. Villette, and A. M. Knodoz. "Prediction of stereoscopic video quality using objective quality models of 2-D video," Electronics Letters, vol. 44, no. 16, pp. 963-965, July 2008.

[16] J. J. Hwang, and H. R. Wu, "Stereo image quality assessment using visual attention and distortion predictors", KSII Trans. on Internet and Information Systems, vol.5, no.9, pp.1613-1631, September 2011.

[17] R. Bensalma, and M. C. Larabi, "A perceptual metric for stereoscopic image quality assessment based on the binocular energy," Multidimensional Systems and Signal Processing, vol.24, no.2, pp. 281-316, June 2013.

[18] M.-J. Chen, C.-C. Su, D.-K. Kwon, L. K. Cormack and A. C. Bovik, "Full-reference quality assessment of stereopairs accounting for binocular rivalry," Signal Processing: Image Communication, vol. 28, no. 9, pp. 1143-1155, Oct. 2013.

[19] V. De Silva, H. K. Arachchi, E. Ekmekcioglu, A. Kondoz, "Towards an impairment metric for stereoscopic video: A full-reference video quality metric to assess compressed stereoscopic video," IEEE Trans. on Image Processing, vol. 22, no. 9, pp. 3392-3404, Sept. 2013.

[20] O. Tuzel, F. Porikli, P. Meer, "Region covariance: A fast descriptor for detection and classification," In Proc. of European Conference of Computer Vision (ECCV), pp. 589-600, 2006.

[21] L. Karacan, E Erdem, A. Erdem, "Structure-preserving image smoothing via region covariances," ACM Transactions on Graphics, vol. 32, no. 6, pp. 176-186, November 2013.

[22] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: an highly efficient perceptual image quality index," IEEE Transactions on Image Processing, vol. 23, no. 2, pp. 684-695, Feb. 2014.

[23] M. Shlh, G. AIRegib, "MIQM: a multicamera image quality measure," IEEE Transactions on Image Processing, vol. 21, no. 9, pp. 3902-3914, Sep. 2012.

[24] F. Shao, W. Lin, S. Gu, G. Jiang, T. Srikanthan, "Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics," IEEE Transactions on Image Processing, vol. 22, no. 5, pp. 1940-1953, May 2013.

[25] A. K. Moorthy, C. C. Su, A. Mittal, A. C. Bovik, "Subjective evaluation of stereoscopic image quality," Signal Processing: Image Communication, vol. 28, no. 8, pp. 870-883, Dec. 2013.

[26] P. G. Gottschalk, and J. R. Dunn, "The five-parameter logistic: a characterization and comparison with the four-parameter logistic," Analytical Biochemistry, vol. 343, no. 1, pp. 54-65, Aug. 2005.

[27] V. Kolmogorov, and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," in Proc. of IEEE International on Computer Vision, Vancouver, Canada, vol. 2, pp. 508-515, Jul. 2001.