GLCM based no-reference perceptual blur metric for underwater blur image

En Cheng, Xizhou Lin, Ye Chen, Fei Yuan*, and Weidi Yang

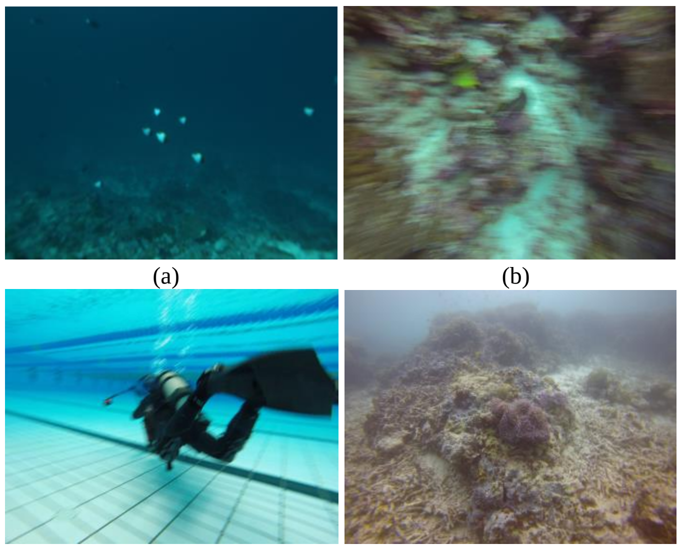
***Abstract*—**Due to the uncertainties of the ocean water environment, scattering and absorption effects of light, lack of light, and unstable imaging platform, the underwater images are usually distorted, especially been blurry. In this paper, we present a no-reference perceptual blur metric for underwater image. Due to the fact that human eye is sensitive to image edges and details, which is usually attenuated by blur. Firstly, we calculate the gradient image and then extract several features on the gray level co-occurrence matrix of gradient image to capture the image details. Then these features are sent to SVR to train and predict the objective assessment score. Experimental results show that the proposed metric outperforms other metrics both for the open datasets and real underwater blur images.

***Keywords*—**blur metric, GLCM, no-reference, underwater image

I. Introduction

THE underwater image processing area has received more and more attention in the last decades. Underwater images are essentially characterized by their poor visibility because light is exponentially attenuated in the water. Light attenuation process is caused by absorption and scattering. The poor light require longer exposure time than in air, thus more probability of blur effect happened[1]. The scattering effect, especially in the turbid water, causes the effect similar to Gaussian blur. Sample underwater image are shown in Fig.1. We could see that these images are distorted in the aspect of blur.

Image quality is critical for pose processing and performance evaluation of image processing algorithms. Based on the subject of evaluation, image quality assessment could be categorized into subjective assessment and objective assessment[2]. The subjective assessment is evaluated by human. It is random, non-repetitive and high cost. On the other hand, the objective assessment, aims to measure the image quality correlated well with human judgements, has been the research focus. It could be classified into three groups: full reference, reduced-reference and no-reference based on with full, partial information and without of origin image respectively[2]. In the case of underwater blur image assessment, no original image is available, therefore no-reference metrics are suitable.



(a) (b) (c) (d)

Fig.1 Sample underwater blur images. (a) out of focus (b) motion blur (c) motion blur (d) scattering blur

In recent years, so many studies about the quality assessment of blur image have been done. Reference [3] proposed a no-reference quality evaluation system using edge detection. Using Canny operator to measure gradient edge information in gradient image to estimate the blur image, this method is proved not so effective in detecting the edge completely. Wu *et al.*[4] used the edge information to evaluate line spread function (LSF) and point spread function (PSF). The radius of PSF is regarded as the measurement of the fuzzy degree. But it is difficult to detect the edge information when the image edge is too smooth. Yuan *et al.* [5] used Sobel operator to extract texture image and defined the strong edge characteristic of each edge pixel to get some strong edge pixel patches so as to get the relevant assessment algorithm, but it is not efficient. Ferzli *et al.* [6] proposed a novel concept about Just Noticeable Blur (JNB), it is used to determine the maximum of blur degree that fail to be detected by human subjective perception, and get the estimate index JNBM. Based on the fact that it is hard to

This work was supported by the National Natural Science Foundation of China (61571377, 61471308, 61501386).

En Cheng is with the Key Laboratory of Underwater Acoustic Communication and Marine Information Technology Ministry of Education, Xiamen University, Xiamen, Fujian, China. (e-mail: chengen@xmu.edu.cn)

Xizhou Lin is with the Key Laboratory of Underwater Acoustic Communication and Marine Information Technology Ministry of Education, Xiamen University, Xiamen, Fujian, China. (e-mail: linxizhou@gmail.com)

Ye Chen is with the Key Laboratory of Underwater Acoustic Communication and Marine Information Technology Ministry of Education, Xiamen University, Xiamen, Fujian, China. (e-mail: chenye@stu.xmu.edu.cn)

Fei Yuan is with the Key Laboratory of Underwater Acoustic Communication and Marine Information Technology Ministry of Education, Xiamen University, Xiamen, Fujian, China. (corresponding author to provide phone: 86-592-2580143; fax: 86-592-2580017;e-mail: yuanfei@xmu.edu.cn)

Weidi Yang is with the College of Ocean and Earth Sciences, Xiamen University, Xiamen, Fujian, China. (e-mail: wdyang@xmu.edu.cn).

perceive blur when the blur level is smaller than JNB, Narvekar *et al.* [7] introduced cumulative probability of blur detection (CPBD) to improve the performance of [6] greatly. Focus on the underwater image blurriness measurement, Hou and Weidemann [8] proposed an sharpness measurement for scattering blurred underwater images. The authors took the weighted average of slope of edges as the indicator of image quality.

In this paper, we proposed a no-reference underwater image quality assessment method based on the gray level co-occurrence matrix (GLCM) features and SVM. Five features are extracted on the GLCM of gradient image, to capture the texture features at which the sensitivity of human blur perception, then a predictive model based on support vector regression (SVR) is trained by using the extracted features and subjective scores. The performance is validated on the public database, as well as the real underwater images dataset.

The paper is organized as follows. Section 2 describes the proposed no-reference blur metric. Experiments results are presented in Section 3. The conclusion is given in Section 4.

II. Proposed no-reference blur metric

In this section we describe feature extraction and SVR firstly, then show draw the framework of the proposed no-reference blur metric and GUI platform.

A. Feature extraction on GLCM

The blur usually causes the attenuation or loss of high frequency components of images, thus produces the blurry edges or loss of texture details.

Haralick *et al.* [9] first introduced the use of co-occurrence probabilities using GLCM for extracting various texture features. The GLCM of an image, $p(i, j/d,\theta)$, is calculated using a displacement distance d, an orientation θ, and gray level L. The element (i, j) in GLCM stands for the number of times gray level i and j have been neighbors satisfying the condition of distance and orientation. The distance d usually takes $\{1,2,3,4\}$ and orientation θ takes $\{0^{\circ},45^{\circ},90^{\circ},135^{\circ}\}$ (Theoretically, there are 8 different orientation choices for the eight neighboring pixels of every pixel, but 0 degree is the same relationship in the GLCM as 180 degree. This concept extends to other degrees as well.)

For a 4 by 4 test image with gray level 0-3, the GLCM with distance $d=1$ and four different orientations are shown as Fig. 2.

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

(a)

4	2	1	0
2	4	0	0
1	0	6	1
0	0	1	2

(b)

6	0	2	0
0	4	2	0
2	2	2	2
0	0	2	0

(c)

4	1	0	0
1	2	2	0
0	2	4	1
0	0	1	0

(d)

2	1	3	0
1	2	1	0
3	1	0	2
0	0	2	0

(e)

Fig. 2 GLCM example (a) sample image (b) 0 degree (c) 45 degree (d) 90 degree (e) 135 degree

The GLCM itself could not been used as a feature directly. 14 features of GLCM are introduced in[9]. We choose 5 features those correlated well with blur level. For the convenience, we simply note g_{ij} rather than $p(i, j/d,\theta)$ as the $(i, j)^{th}$ element in GLCM.

1) Contrast

$$F_1 = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}(i-j)^2 g_{ij} \tag{1}$$

The contrast reflect the amount of local variations present in the image.

2) Dissimilarity

$$F_2 = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\left|i-j\right| g_{ij} \tag{2}$$

This feature have similar effect as contrast to measure the sharpness.

3) Entropy

$$F_3 = -\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} g_{ij}\log_2 g_{ij} \tag{3}$$

The entropy measures the disorder of an image. Texture without blur tend to have high entropy.

4) Homogeneity

$$F_4 = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\frac{1}{1+(i-j)^2} g_{ij} \tag{4}$$

This feature is inverse to the contrast feature of being correlated in term of equivalent distribution in the pixel pairs population.

5) Energy

$$F_5 = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} g_{ij}^2 \tag{5}$$

This statistic measures the textural uniformity that is pixel pair repetitions and it detects disorders in textures.

B. Support vector regression

Support vector machine (SVM) is machine learning method proposed by Vapnik[10]. Depend on application, SVM could be divided into support vector classifier (SVC) and support vector regression (SVR).

In the SVR, it usually uses the followed linear model to estimate the mapping of input vector and output value.

$$f(x,w) = w\varphi(x) + b \tag{6}$$

where $\varphi(x)$ maps the input vector to high dimension feature space. w is the weight vector, b is the constant. Taking the ε-insensitive loss function, the object function could be modeled as following equation.

$$\min \phi(w,\xi) = \frac{1}{2}\left\|w\right\|^2 + C\sum_{i=1}^{1}(\xi_i^- + \xi_i^+) \tag{7}$$

$$s.t.\begin{cases} y_i - f(x_i) \le \varepsilon + \xi_i^- \\ f(x_i) - y_i \le \varepsilon + \xi_i^-, i = 1, ..., l \\ \xi_i^-, \xi_i^- \ge 0 \end{cases} \tag{8}$$

where C is the penalty coefficient, ξ_i^-, ξ_i^- are slack variable. It is easy to optimize the object function by using Lagrange function and transform to dual problem, then get the solution.

$$f(x) = \sum_{i=1}^{n_{sv}} (a_i - a_i^*) k(x_i, x) + b \tag{9}$$

where a_i, a_i^* are Lagrange multipliers and most of them are zeros. The associated samples are called support vector (SV), n_{sv} is the number of support vector. $k(x_i, x)$ is the kernel function. Equation (10) shows the an often used kernel function-radial basis function (RBF).

$$k(x_i, x) = \exp(-\lambda \|x - x_i\|^2) \tag{10}$$

Where λ is the kernel parameters. Model approximation capability are improved by adjusting C, λ, ε in the real application. We use the libsvm library[11] for the regression.

C. System framework

The key to assess the quality of blur images by SVR is to find the image features those capture the blurriness. Rather than directly computing the GLCM on the distorted image, we extract the gradient firstly to get the structure information, then compute GLCM and extract the 5 features to capture the blurriness, and then send the features to SVR to train the best fitting parameters. Finally we predict the blurriness of a new bur image. Our system framework is shown as Fig.3.

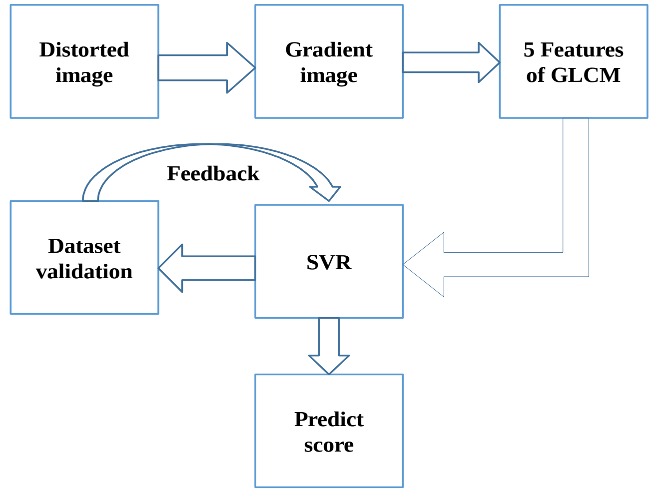


Fig.3 Framework of proposed metric

D. GUI platform

According to the framework of our model, we develop a Matlab® Graphic User Interface (GUI) platform shown as Fig.4. The platform is mainly used for single image quality prediction and dataset validation. For the image prediction, there are ‘Image Selection’, ‘Gradient Image’, ‘5 Features’ and ‘DMOS Predict’ buttons. Pushing the buttons will show the corresponding images or feature indexes. For the dataset validation operation (Bottom right corner of GUI), pushing the ‘Dataset Validation’ button after choosing dataset, the platform will pop up an window showing the scatter plot of subjective scores and predicted scores (See in performance validation Section). Performance indexes will also show in the GUI.

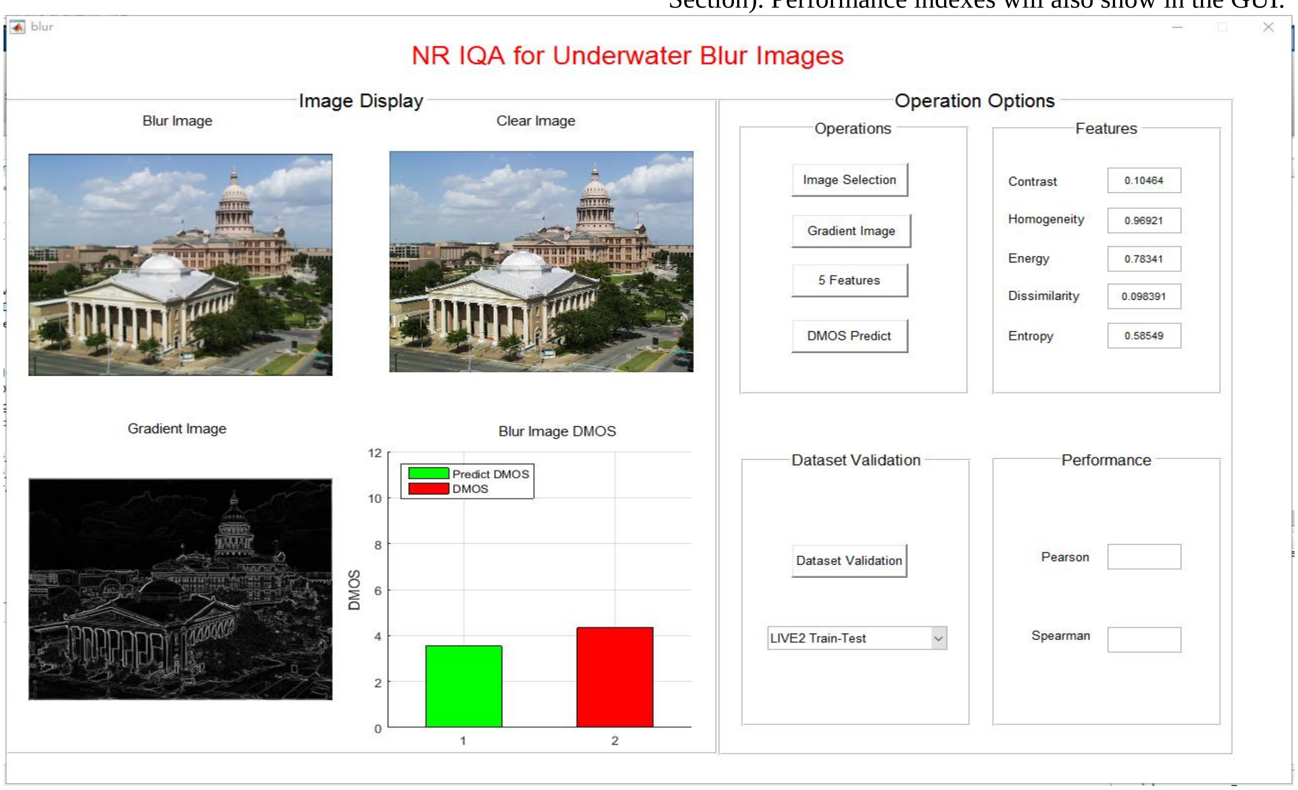


Fig.4 GUI platform for proposed no-reference blur metric

III. PERFORMANCE VALIDATION

To test the performance of the method, Gaussian-blurred images from LIVE2[12], CSIQ[13] and TID2008[14] dataset are used. Performance is measured by calculating the LCC (Pearson linear correlation coefficient, indicates the prediction accuracy), SROCC (Spearman rank order correlation coefficient, indicates the prediction monotonicity) index between the subjective scores and algorithm scores after a 5 parameter non-linear logistic regression[15].

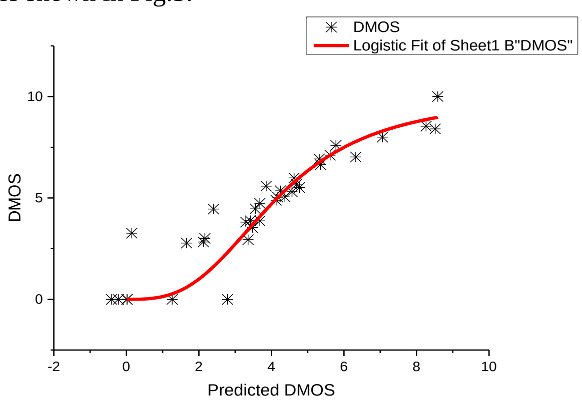
A. Validation on LIVE2

We validate the LIVE2 dataset by 5-fold cross validation methods. 174 blurred images are randomly separated into 5 distinct non-overlapped subsets, 4 of which are taken as training and the other for testing each time. So we get 5 different train test pair $C1-C5$ .Best model parameters $C=165.1, \lambda=8$, $\varepsilon=0.4$ are chose after several training and optimizing. Table 1 shows the test result. We could see that both the LCC and SROCC index are larger than 0.9, which means the predicted scores correlated well with subjective scores.

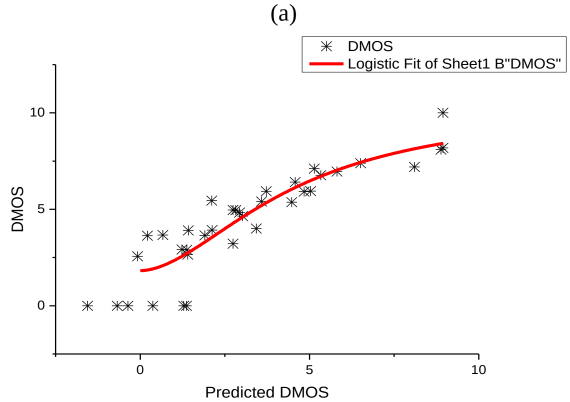
Table.1 Performance on LIVE2 dataset

Performance	C1	C2	C3	C4	C5	Average
LCC	0.93	0.90	0.92	0.94	0.90	0.92
SROCC	0.97	0.94	0.94	0.96	0.92	0.95

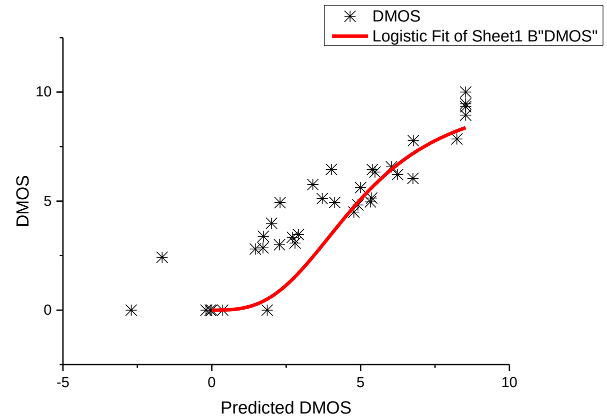
The scatter plots, where the x-axis is the predict objective score, y-axis the DMOS provided by the dataset, of five train-test pairs between the predicted scores and the subjective scores shown in Fig.5.



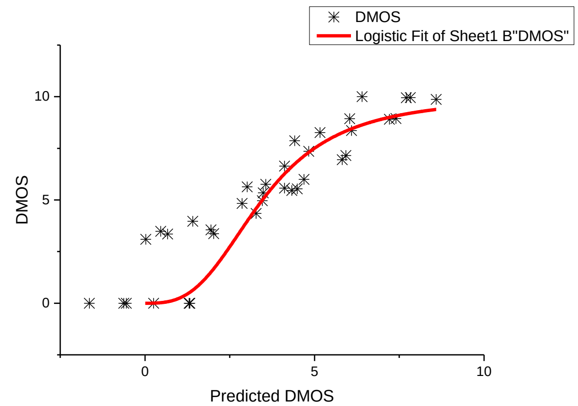
(a)



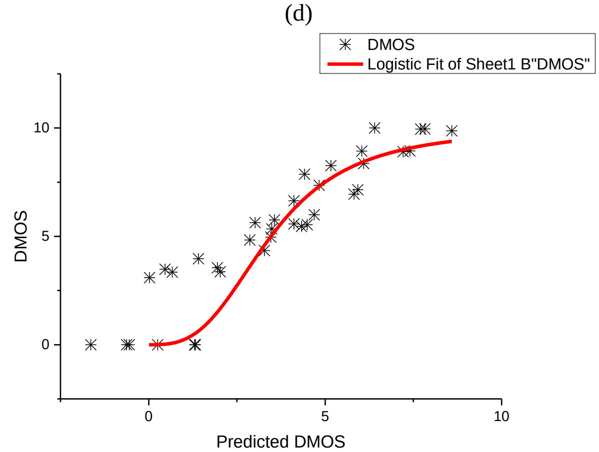
(b)



(c)



(d)



(e)

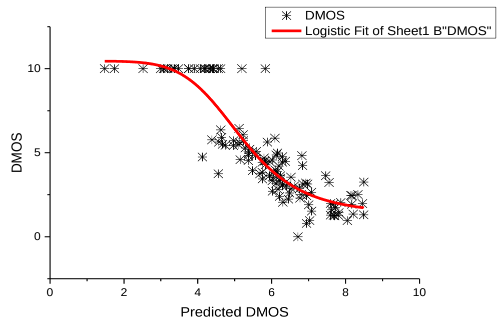
Fig.5 Scatter of LIVE2 validation (a)C1 (b)C2 (c)C3 (d)C4 (e)C5

B. Cross-validation on different datasets

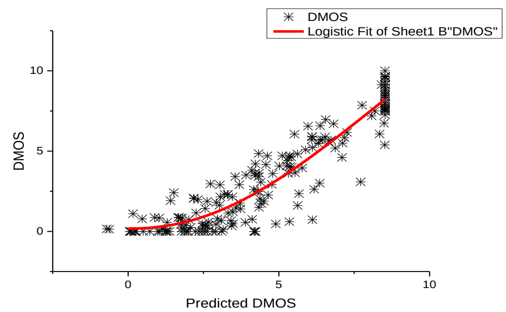
In order to demonstrate the generality of our methods, we train one of the three datasets and test the performance on the other two datasets.

1) Train LIVE2, Test CSIQ and TID2008

We train the model by using LIVE2 dataset and predict the objective scores of images from CSIQ and TID2008. Optimizing the model parameters, we get the prediction scatter plot shown as Fig. 6.



(a)



(b)

Fig.6 Train LIVE2, scatter plot (a) CSIQ (b) TID2008

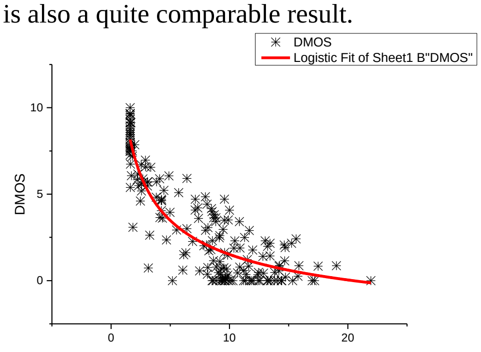
Table 2 shows the prediction performance on dataset CSIQ and TID2008. We could see that the LCC and SROCC of CSIQ are positive value, but TID2008 is negative value. Also the slope in Fig.6 (a) and Fig.6 (b) are different sign. Because the subjective scores of dataset LIVE2 and dataset CSIQ are DMOS (Differential Mean Opinion Score), and TID2008 are MOS (Mean Opinion Score). The lower value of DMOS, the higher quality the images is, but MOS is on the contrary. From the Table.2 and Fig.6, we could see that the train model predict the blurriness quite well.

Table.2 Cross validation by training LIVE2

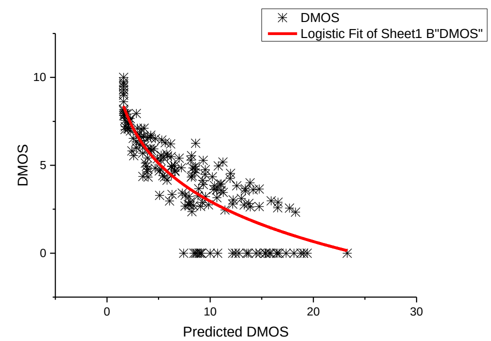
Performance	CSIQ	TID2008
LCC	0.90	-0.87
SROCC	0.89	-0.89

2) Train TID2008 and Test LIVE2 and TID2008

Similarly, Fig.7 and Table.3 show the result that trained by TID2008 and test on LIVE2 and CSIQ. Although the performance index is not as high as model trained by using LIVE2, it is also a quite comparable result.



(a)



(b)

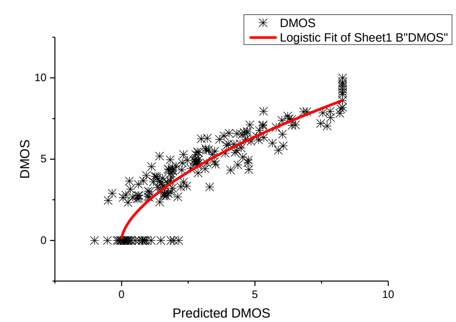
Fig.7 Train TDI2008, scatter plot (a) LIVE2 (b) TID2008

Table.3 Cross validation by training LIVE2

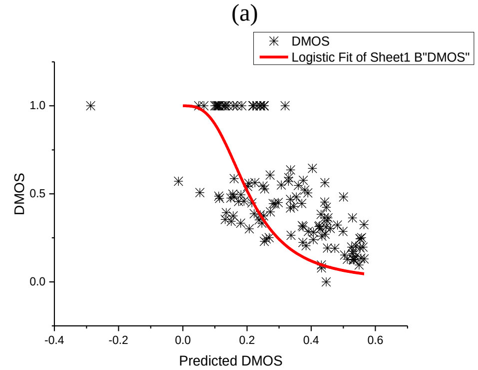
Performance	LIVE2	CSIQ
LCC	-0.82	-0.82
SROCC	-0.86	-0.80

3) Train CSIQ, Test LVIE2 and TID2008

Similarly, Fig. 8 and Table.4 show the result that trained by CSIQ and test on LIVE2 and TID2008. We can see that the performance index for TID2008 is around 0.7, the possible reason is that we add some unblurry images in TID2008 for prediction.



(a)



(b)

Fig.8 Train CSIQ, scatter plot (a) LIVE2 (b) TID2008

Table 4. Cross validation by training CSIQ

Performance	LIVE2	TID2008
LCC	0.89	-0.69
SROCC	0.92	-0.73

C. Performance comparison with other metrics

Besides testing of our own metric, we also compare the performance with JNBM[6] and CPBD[7] shown as Table.5. We could see that our results show advance on the others.

Table.5 Performance comparison with other metrics

Performance	Metrics	LIVE2	CSIQ	TID2008
LCC	CPBD[7]	0.91	--	0.83
	JNBM[6]	0.84	0.85	0.72
	Proposed	**0.92**	**0.90**	**0.87**
SROCC	CPBD[7]	0.94	--	0.84
	JNBM[6]	0.84	0.76	0.70
	Proposed	**0.95**	**0.89**	**0.89**

D. Performance on real underwater blur image

To verify the practicability of our metric, we compare the metrics on the real underwater blur images. We capture the underwater water images in the ocean in, samples are shown as Fig. 1. And we organize a subjective image blurriness assessment with 20 naïve students without expert experience, followed the Rec.ITU-R BT.500-11T recommendation [16], to get the subjective scores MOS. The comparison results shown in Table.6 reveals that our metric get better result than the other two metrics.

Table.6 Performance comparison on real dataset

Metrics	LCC	SROCC
JNBM[6]	0.41	0.32
CPBD[7]	0.45	0.45
Proposed	**0.58**	**0.54**

IV. Conclusion

We proposed a GLCM based no-reference perceptual blur metric for underwater blur image. Compared to previous methods, instead of counting the edge numbers or measuring the edge width, we implicitly represent the image details, which is important for human blur perception, by extracting the GLCM features, and then map them to the perceived blur via SVR. Performance validations on different datasets show that our method is capable to catch the perceived image blurriness quite well. Our work outperforms other blur metrics both for the public datasets and the real underwater blur images.

However, the absolute performance index of LCC and SROCC for the underwater images is not high enough. Therefore, extracting the better and comprehensive features, which correlate well to the blurriness of underwater blur images, is the possible future work.

Acknowledgment

The authors would like to thank Dr. ChiChi Liu from City University of Hong Kong for providing us the valuable underwater images, students who participated in the subjective assessment for their works, Zhengshan Ye for the help in training the svm.

References

[1] R. Schettini and S. Corchs, "Underwater Image Processing: State of the Art of Restoration and Image Enhancement Methods," *Eurasip Journal on Advances in Signal Processing*, vol. 2010, p. 14, 2010.

[2] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, pp. 600-612, 2004.

[3] Bo. X. Zuo, D. L. Ming, Z. G. Ao, and W. J. Zheng, "No-Reference Blurred Image Based on Image Structure and the Quality Assessment Method of Its Restoration Image [J]," *Geomatics World*, vol. 1, p. 004, 2008.

[4] S. Wu, W. Lin, L. Jian, W. Xiong, and L. Chen, "An objective out-of-focus blur measurement," *in Information, Communications and Signal Processing, 2005 Fifth International Conference on*, 2005, pp. 334-338.

[5] W. Yuan and C. Li, "No-reference blur image quality metric combining HVS with SSIM," *Computer Engineering and Applications*, vol. 49, 2013.

[6] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *Image Processing, IEEE Transactions on*, vol. 18, pp. 717-728, 2009.

[7] N. D. Narvekar and L. J. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," *Image Processing, IEEE Transactions on*, vol. 20, pp. 2678-2683, 2011.

[8] W. Hou and A. D. Weidemann, "Objectively assessing underwater image quality for the purpose of automated restoration," *Visual Information Processing XVI*, 2007.

[9] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *Systems, Man and Cybernetics, IEEE Transactions on*, pp. 610-621, 1973.

[10] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, pp. 273-297, 1995.

[11] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology* (TIST), vol. 2, p. 27, 2011.

[12] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment database release 2" , 2005.

[13] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, pp. 011006-011006-21, 2010.

[14] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008-a database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, pp. 30-45, 2009.'

[15] Sheikh, H. R., Sabir, M. F., & Bovik, A. C. (2006). A statistical evaluation of recent full reference image quality assessment algorithms. *Image Processing, IEEE Transactions on*, 15(11), 3440-3451.

[16] Recommendation, I. T. U. R. B. T. (1995). 500.7, Methodology for the Subjective Assessment of the Quality of Television Pictures. *ITU-R Recommendations*, Geneva.