A Reduced-reference Quality Assessment Scheme for Blurred Images

Zongxi Han[1], Guangtao Zhai[1], Yutao Liu[2], Ke Gu[3] and Xinfeng Zhang[4]

[1]*Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, China*
[2]*Department of Computer Science, Harbin Institute of Technology, Harbin, China*
[3]*School of Computer Science and Engineering, Nanyang Technological University, Singapore*
[4]*the Rapid-Rich Object Search (ROSE) Lab, Nanyang Technological University, Singapore*

zongxihan@sjtu.edu.cn, {zhaiguangtao, liuyutao2008, gukesjtuee}@gmail.com, xfzhang@ntu.edu.sg

***Abstract*—In this paper, we propose a reduced-reference scheme for evaluating the quality of blurred images under the theory of free-energy principle. Specifically, the free-energy principle indicates that the brain tries to account for the input image with an internal generative model and the discrepancy between the image and its model-explained version, which can be measured by free energy, is related to the image's perceptual quality. Accordingly, we define a visual distance between the blurred image and its original image in free energy to evaluate the quality of the blurred image. Therefore, the proposed quality scheme belongs to reduced-reference methods, which needs some information from the original image for quality assessment. Experimental results on public databases, LIVE, TID2013 and C-SIQ, demonstrate the proposed method works in high consistency with subjective assessment results and outperforms representative image quality assessment approaches.**

***Index Terms*—Image quality assessment, image blurriness, reduced-reference, free energy, sparse representation**

I. Introduction

Nowadays, image quality becomes the more and more important issue to be addressed in reality. The reason may be summarized into two aspects, one is that as the development of our life, the requirements of the image quality from people become higher and higher. On the other hand, the image quality can't be guaranteed during image processing or transmission. For example, the image is always compressed due to the bandwidth restrictions for transmission. Therefore, image quality assessment (IQA) is always used to monitor the image quality or employed as the performance measure for image processing algorithms [1] [2].

Generally, objective IQA methods are categorized into three classes on the basis of the access to the reference image for quality estimation, which are full reference (FR), reduced reference (RR) and no reference (NR) respectively. For FR IQA, the reference or distortion-free image is fully referred when we assess the image quality. It's not difficult to find that the reference image is absent or hard to obtain in most cases. Therefore, the FR condition is rather ideal in practice. Then researchers develop RR methods that require partial information of the reference image for quality assessment. Further, NR IQA can directly evaluate the image quality without any information of the reference image, which is highly desirable in applications. The representative works of IQA can be referred to [3] [4] etc.

In this paper, our attention is focused on RR IQA and the investigation of the quality assessment problem of a special kind of images, which refers to the blurred images. There are many causes that lead to image blur, e.g. camera out-of-focus, object motion or excessive compression etc. The intuitive observation of the blurred images is the textures or edges are always destroyed. Therefore, the image blurriness estimation can be concentrated on the image edges. In [5], the perceptual assessment model was constructed based on the pair edges detectors. Additionally, the authors proposed an objective image sharpness metric through the measure of just noticeable blur (JNB) around the image edges [6]. From another point of view, the texture degradation due to image blur can also be characterised by the attenuation of high frequency coefficients in the frequency domain. For instance, a fast wavelet-based image sharpness estimation method (FISH) [7] was developed by measuring the energy of the coefficients in the wavelet transform domain. Certainly, spatial and spectral combined analysis can assess the image blurriness complementarily [8].

For the quality assessment of the blurred images, the aforementioned blurriness / sharpness assessment methods can be applied to this question directly as blurriness is the main factor that affects the image quality. While in this paper, we view this question from a new perspective of the exploration of the brain's activities when perceiving the images. Specifically, the exploration of the brain is revealed by the recent developments in brain theory and neuroscience, particularly the free-energy principle that the perception and understanding of an image are modeled as an active inference process, in which the brain tries to predict the input image through an internal generative model. While the discrepancy, measured by free energy, between the image and its model-predicted version is closely related to

978-1-5090-5316-2/16/$31.00 © 2016 IEEE VCIP 2016, Nov. 27 – 30, 2016, Chengdu, China

the perceptual quality of the image [9] [10] [11] [12]. In other words, free energy of the image can be viewed as an quality-connected feature and we can estimate the image quality by inspect the variation of free energy. Accordingly, we design a quality metric to measure the quality of the blurred images. Experiments on public databases, LIVE, TID2013 and CSIQ, prove that the proposed metric works in high consistency with subjective assessment results and outperforms representative IQA and blurriness assessment approaches at the meantime.

II. The Proposed Quality Metric

A. Formulation of Free-energy Principle

Our method is based on the free-energy principle. Therefore, it is needed to specify the formulation of the free-energy principle clearly. As mentioned before, the fundamental assumption in free-energy principle is that the cognitive process is governed by an internal generative model in the brain. With the internal model, the brain is able to actively infer predictions of meaningful information from the visual scenes.

For operational amenability, it is assumed that the internal generative model $\mathcal{G}$ for visual perception is parametric, which explains visual scenes by adjusting the parameter vector $\mathbf{g}$. Specifically, given an image I, its 'surprise' can be calculated by integrating the joint distribution $P(I, \mathbf{g})$ over the space of the internal model parameters $\mathbf{g}$ as:

$$-\log P(I) = -\log \int P(I, \mathbf{g}) d\mathbf{g}. \tag{1}$$

Then a dummy term $Q(\mathbf{g}|I)$ is integrated into both the denominator and numerator of the right part of equation (1) as follows:

$$-\log P(I) = -\log \int Q(\mathbf{g}|I) \frac{P(I, \mathbf{g})}{Q(\mathbf{g}|I)} d\mathbf{g}. \tag{2}$$

Here, $Q(\mathbf{g}|I)$ is an auxiliary posterior distribution of the model parameters given the image, which can be thought of as an approximate posterior to the true posterior of the model parameters $P(\mathbf{g}|I)$ calculated by the brain. The brain minimizes the discrepancy between the approximate posterior $Q(\mathbf{g}|I)$ and the true posterior $P(\mathbf{g}|I)$. Through Jensen's inequality, equation (2) can be written as:

$$-\log P(I) \leq -\int Q(\mathbf{g}|I) \log \frac{P(I, \mathbf{g})}{Q(\mathbf{g}|I)} d\mathbf{g}. \tag{3}$$

Afterwards, the right side of equation (3) is defined as the free energy as follows:

$$F(\mathbf{g}) = -\int Q(\mathbf{g}|I) \log \frac{P(I, \mathbf{g})}{Q(\mathbf{g}|I)} d\mathbf{g}. \tag{4}$$

Obviously, $F(\mathbf{g})$ defines an upper bound of 'surprise' for image I. For intuitive understanding, with $P(I, \mathbf{g}) = P(\mathbf{g}|I)P(I)$, we further derive equation (4) as:

$$\begin{aligned} F(\mathbf{g}) &= \int Q(\mathbf{g}|I) \log \frac{Q(\mathbf{g}|I)}{P(\mathbf{g}|I)P(I)} d\mathbf{g} \\ &= -\log P(I) + \int Q(\mathbf{g}|I) \log \frac{Q(\mathbf{g}|I)}{P(\mathbf{g}|I)} d\mathbf{g} \\ &= -\log P(I) + \mathbf{KL}(Q(\mathbf{g}|I) \| P(\mathbf{g}|I)), \end{aligned} \tag{5}$$

where $\mathbf{KL}(\cdot)$ refers to the Kullback-Leibler divergence between the approximate posterior and the true posterior distributions and it's nonnegative. It is clearly seen that the free energy $F(\mathbf{g})$ is greater than or equal to the image 'surprise' -log$P(I)$. The brain tries to lower the divergence $\mathbf{KL}(Q(\mathbf{g}|I) \| P(\mathbf{g}|I))$ between the approximate posterior and its true posterior distributions when perceiving image I. Interested readers can refer to more information about free energy in [9] and its application to IQA of screen content images [13] and natural scene images [14] and to visual saliency detection [15].

B. Approximation of The Brain Generative Model

Since free energy of the image is a discrepancy measure between the image data and its explanation by the brain, we try to find its computational form for actual quality evaluation. Toward this end, the approximation of the brain generative model $\mathcal{G}$ should go first. As stated in [16], the receptive fields (RFs) of simple cells in mammalian primary visual cortex can be characterized as being spatially localized, oriented and bandpass. While sparse representation is evidenced to resemble these neural response characteristics well and the superiority of sparse representation for approximating the internal model has been verified in [17]. Inspired by this, in this paper, we approximate the generative model with sparse representation. Specifically, Sparse representation refers to representing a signal with a linear combination of a small number of atoms from a predefined or trained dictionary [18]. Mathematically, given a signal $\mathbf{y} \in \mathbb{R}^n$ with an overcomplete dictionary matrix $\mathbf{D} \in \mathbb{R}^{n \times K}$ that contains K columns, each column represents one prototype atom. Then the dictionary $\mathbf{D}$ can be denoted as $[\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3 ... \mathbf{d}_K]$. The signal $\mathbf{y}$ is represented as a sparse linear combination of the atoms in $\mathbf{D}$ as:

$$\mathbf{y} = \mathbf{D}\mathbf{x}, \tag{6}$$

or approximately represented as:

$$\mathbf{y} \approx \mathbf{D}\mathbf{x} \quad s.t. \quad \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_p \leq \xi, \tag{7}$$

where $\mathbf{x} \in \mathbb{R}^K$ represents the vector that contains the representation coefficients. $\| \cdot \|_p$ is the l^p norm. What we concerned is finding fewest number of nonzero coefficients to represent the signal $\mathbf{y}$, namely requesting for the sparsest representation:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_0 \quad s.t. \quad \mathbf{y} = \mathbf{D}\mathbf{x}, \tag{8}$$

where $\| \cdot \|_0$ is the l^0 norm, meaning the number of nonzero elements of a vector. However, l^0-minimization is an NP-hard problem, one approach is applying "pursuit algorithm" to find an approximate solution. Another alternative solution is to replace l^0 norm with l^1 norm and minimize the l^1 norm as:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1 \quad s.t. \quad \mathbf{y} = \mathbf{D}\mathbf{x}, \tag{9}$$

This equation can be further turned into an unconstrained optimization problem:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\mathrm{argmin}}\ \frac{1}{2}\|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 + \lambda\|\mathbf{x}\|_1, \tag{10}$$

where λ is a positive constant balancing the importance of the reconstruction fidelity term and the sparse constraint term. This unconstrained optimization problem can be solved by iterative shrinkage/thresholding algorithm [19]. With the obtained coefficient vector $\mathbf{x}$ and the predefined dictionary $\mathbf{D}$, we can get the sparse representation of signal $\mathbf{y}$ accordingly.

C. The Quality Metric for Blurred Images

As the free-energy principle indicates, free energy of an image which measures the discrepancy between the image and its brain prediction is closely related with the image quality. Therefore, we can estimate the image quality through checking the variation of its free energy.

Firstly, free energy of an image I can be defined as the entropy of the prediction residual, which is denoted as:

$$R = I - I_p \tag{11}$$

where R refers to the prediction residual, I_p represents the sparse representation of image I. Then free energy of I is to take the entropy of the prediction residual R as:

$$F(I) = E(R) \tag{12}$$

where $F(I)$ refers to the free energy of image I, $E(\cdot)$ is the function to calculate entropy. While as the research of image saliency prediction indicates, the human vision system (HVS) selectively pays attention to the salient areas in the image. Therefore, we redefine the free energy of I according to the saliency detection results of I in:

$$F'(I) = E(R_S) \tag{13}$$

where $F'(I)$ is the redefined free energy of I, R_S consists of the selected pixels in R whose positions are corresponding to that of $l\%$ most salient pixels in I. In implementation, the saliency of the pixel in I can be predicted by saliency prediction in advance. The value of l is empirically set to 2.

Secondly, with the specific definition of free energy and its quality proxy character, we can devise the quality metric of the blurred images through comparing the free energy between the reference image and the blurred image as follows:

$$Q = F'(I) - F'(I_{ref}) \tag{14}$$

where Q presents the quality of the blurred image, $F'(\cdot)$ is the function to calculate free energy defined in Eq. (13), I refers to the blurred image and I_{ref} is its reference image. As we want to show that the lower Q is, the closer the free energy is between the blurred image and its reference image, which means the blurred image tends to have a higher quality and vice versa. Actually, our metric belongs to RR categories which needs information (free energy) of the reference image for quality estimation. While as we defined, the free energy can be approximated by entropy of the prediction residual which is just a number and negligible to the image data.

III. Experimental Results and Analysis

A. Experiment Configurations

In implementation, we divide the image into 8×8 non-overlapped patches. Each patch is vectorized as the signal $\mathbf{y}$ and we seek its sparsest representation according to II-B described. The overcomplete DCT dictionary is employed as the predefined dictionary $\mathbf{D}$, the size of $\mathbf{D}$ is 64×128 with totally 128 atoms available for representing each patch. We restrict the maximum number of nonzero coefficients for representing each patch to 20. The orthogonal matching pursuit (OMP) algorithm [20] is utilized to find the representation coefficients. The saliency prediction model AIM [21] is employed in the computation of the image's free energy.

B. Experimental Results and Comparison

To evaluate the prediction performance of our proposed quality metric, we test it on the subsets of blurred images in three widely-adopted image databases, which are LIVE [22], TID2013 [23] and CSIQ [24] respectively. As suggested by VQEG [25], we first map the results given by objective methods to subjective ratings through a four-parameter nonlinear regression function as:

$$f(x) = \frac{\xi_1 - \xi_2}{1 + \exp(-\frac{x-\xi_3}{\xi_4})} + \xi_2 \tag{15}$$

where x and $f(x)$ respectively stand for the objective score and the mapped score. Spearman Rank-Order Correlation Coefficient (SROCC), Pearson Linear Correlation Coefficient (PLCC) and Root Mean-Squared Error (RMSE) are employed to evaluate the objective IQA method's prediction performance. It should be pointed out that a superior objective method is expected to achieve higher values in SROCC and PLCC, while a lower value in RMSE. We tabulate the performance results in Table I. As can be found, we compare our method with some representative IQA approaches, including PSNR, SSIM [3], DIIVINE [26], BRISQUE [27], NFERM [14], CPBD [28], S3 [8], FISH [7], ARISM [29], FEDM [9]. Among them, PSNR and SSIM are the most well known FR metrics, DIIVINE, BRISQUE and NFERM are the representative NR IQA methods, CPBD, S3, FISH and ARISM are the special methods for blurriness assessment and FEDM belongs to RR IQA methods. To the blurred images, SSIM is better than PSNR as can be observed and the three NR methods have no big gap on the final prediction performance. Among the special blurriness assessment methods, the newly proposed ARISM earns the best prediction results. While ARISM doesn't perform excellently on TID2013 database. All in all, our proposed method achieves the best performance all over the three databases.

IV. Conclusion

In this paper, we have proposed a RR quality metric for the blurred images. The proposed metric is under the exploration of the free-energy principle in brain theory and neuroscience. Specifically, the internal generative model was approximated

TABLE I
PREDICTION PERFORMANCE COMPARISON ON LIVE, TID2013 AND CSIQ DATABASES. THE BEST PERFORMANCE IN EACH INDICE IS MARKED IN BOLD.

Database	Index	PSNR	SSIM	DIIVINE	BRISQUE	NFERM	CPBD	S3	FISH	ARISM	FEDM	Proposed
LIVE	SROCC	0.7823	0.8944		training images		0.9186	0.9441	0.8808	0.9517	0.7594	**0.9623**
	PLCC	0.7835	0.8743		training images		0.8953	0.9527	0.9043	**0.9562**	0.7351	0.9469
	RMSE	11.4780	8.9643		training images		8.2263	4.7769	7.8844	**4.6028**	10.6586	5.0542
TID2013	SROCC	0.9149	**0.9629**	0.8344	0.8143	0.8498	0.8515	0.8046	0.8024	0.8980	0.8897	0.9616
	PLCC	0.9137	**0.9577**	0.8472	0.8248	0.8493	0.8553	0.8432	0.8327	0.8953	0.8891	0.9542
	RMSE	0.5071	**0.3592**	0.6629	0.7057	0.6588	0.6466	0.6708	0.6910	0.5560	0.5712	0.3732
CSIQ	SROCC	0.9291	0.9245	0.8716	0.9025	0.8964	0.8853	0.8681	0.8941	0.9255	0.8522	**0.9426**
	PLCC	0.9252	0.9005	0.8979	0.9274	0.9218	0.8822	0.8833	0.9232	**0.9456**	0.8150	0.9290
	RMSE	0.1087	0.1246	0.1262	0.1072	0.1111	0.1349	0.1343	0.1102	**0.0933**	0.1660	0.1061

by sparse representation and the free energy variation between the original image and the blurred image was calculated to measure the perceptual quality of the blurred image. Through extensive experiments on LIVE, TID2013 and CSIQ, we verified the proposed metric works in high consistency with subjective human ratings.

ACKNOWLEDGEMENT

This work was supported in part by the National Science Foundation of China under Grants 61422112, 61371146, 61521062, 61527804, and National High-tech R&D Program of China under Grant 2015AA01590.

REFERENCES

[1] X. Liu, D. Zhai, J. Zhou, X. Zhang, D. Zhao, and W. Gao, "Compressive sampling-based image coding for resource-deficient visual communication," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2844–2855, 2016.

[2] X. Liu, X. Wu, J. Zhou, and D. Zhao, "Data-driven soft decoding of compressed images in dual transform-pixel domain," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1649–1659, 2016.

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

[4] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, 2006.

[5] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric," in *Image Processing. 2002. Proceedings. 2002 International Conference on*, vol. 3. IEEE, 2002, pp. III–57.

[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, no. 4, pp. 717–728, 2009.

[7] P. V. Vu and D. M. Chandler, "A fast wavelet-based algorithm for global and local image sharpness estimation," *Signal Processing Letters, IEEE*, vol. 19, no. 7, pp. 423–426, 2012.

[8] C. T. Vu, T. D. Phan, and D. M. Chandler, "S3: A spectral and spatial measure of local perceived sharpness in natural images," *Image Processing, IEEE Transactions on*, vol. 21, no. 3, pp. 934–945, 2012.

[9] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 41–52, 2012.

[10] G. Zhai, W. Zhang, X. Yang, S. Yao, and Y. Xu, "Ges: a new image quality assessment metric based on energy features in gabor transform domain," in *2006 IEEE International Symposium on Circuits and Systems*. IEEE, 2006, pp. 1715–1718.

[11] G. Zhai, W. Zhang, Y. Xu, and W. Lin, "Lgps: Phase based image quality assessment metric," in *2007 IEEE Workshop on Signal Processing Systems*. IEEE, 2007, pp. 605–609.

[12] G. Zhai, A. Kaup, J. Wang, and X. Yang, "A dual-model approach to blind quality assessment of noisy images," *APSIPA Transactions on Signal and Information Processing*, vol. 4, p. e4, 2015.

[13] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Learning a blind quality evaluation engine of screen content images," *Neurocomputing*, 2016.

[14] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *Multimedia, IEEE Transactions on*, vol. 17, no. 1, pp. 50–63, 2015.

[15] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Visual saliency detection with free energy theory," *Signal Processing Letters, IEEE*, vol. 22, no. 10, pp. 1552–1555, 2015.

[16] B. A. Olshausen *et al.*, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.

[17] Y. Liu, G. Zhai, X. Liu, and D. Zhao, "Perceptual image quality assessment combining free-energy principle and sparse representation," in *Circuits and Systems (ISCAS), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 1–4.

[18] M. Aharon, M. Elad, and A. Bruckstein, "K-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, 2006.

[19] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, pp. 629–654, 2008.

[20] Y. C. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE, 1993, pp. 40–44.

[21] N. D. Bruce and J. K. Tsotsos, "Saliency, attention, and visual search: An information theoretic approach," *Journal of vision*, vol. 9, no. 3, p. 5, 2009.

[22] L. C. H. R. Sheikh, Z. Wang and A. C. Bovik, "Live image quality assessment database release 2," 2006.

[23] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti *et al.*, "Image database tid2013: Peculiarities, results and perspectives," *Signal Processing: Image Communication*, vol. 30, pp. 57–77, 2015.

[24] E. C. Larson and D. Chandler, "Categorical image quality (csiq) database," *Online, http://vision. okstate. edu/csiq*, 2010.

[25] A. M. Rohaly, J. Libert, P. Corriveau, A. Webster *et al.*, "Final report from the video quality experts group on the validation of objective models of video quality assessment," *ITU-T Standards Contribution COM*, pp. 9–80, 2000.

[26] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, 2011.

[27] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, 2012.

[28] 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, no. 9, pp. 2678–2683, 2011.

[29] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *Image Processing, IEEE Transactions on*, vol. 24, no. 10, pp. 3218–3231, 2015.