# USING JPEG AND JPEG2000 COMPRESSIONS FOR FAST IMAGE QUALITY METRICS BASED ON FREE ENERGY THEORY

*Ke Gu, Guangtao Zhai, Min Liu, Xiaokang Yang, and Wenjun Zhang*

Insti. of Image Commu. & Infor. Proce., Shanghai Jiao Tong Univ., Shanghai, China, 200240
Shanghai Key Laboratory of Digital Media Processing and Transmissions
{gukesjtuee, zhaiguangtao, liumin_merry, xkyang, zhangwenjun}@sjtu.edu.cn

## ABSTRACT

In this paper, we propose a fast, effective and practical algorithm for image quality assessment (IQA). Recently, a new free energy theory was revealed in the field of brain science, which illustrates that the human visual system (HVS) always strives to comprehend the input visual signal by reducing the undetermined portions. Inspired by this, our previous work recently designed a valid reduced-reference (RR) free energy based distortion metric (FEDM) using linear autoregressive model. Despite of fairly well performance, the FEDM is yet difficult to work in real-time applications owing to its weak portability and considerable computational load. Using an alternative way, this paper approaches the free energy related mechanism in human brains with JPEG and JPEG2000 compressions. The proposed metric can work quickly and practically in that it is realized with highly developed and widely employed image compression methods. Results of experiments on the most popular publicly-available LIVE database demonstrate the effectiveness of the proposed RR algorithm over classical full-reference PSNR and SSIM as well as state-of-the-art RR IQA metrics.

***Index Terms*—** Image quality assessment (IQA), reduced-reference (RR), free energy theory, human visual system (HVS)

## 1. INTRODUCTION

Nowadays, hundreds of thousands of images or video frames are being broadcasted to human consumers every moment. An accompanying problem is that an enormous amount of workers are also required to guaranty the visual quality of those images or frames, which are easily deteriorated by disturbing noise, a Rayleigh fast-fading channel and etc. So the work to monitor and control the image or video frame quality yet hire very few employees is extremely desired in the present time [1]. The image quality assessment (IQA), due to its outstanding ability to approximate human visual perception to image quality, is usually considered to be a good choice to cope with this problem.

Existing IQA methods are broadly categorized into subjective assessment and objective assessment. The former plays a crucial role because it provides the testing images and their real human quality ratings (e.g. the famous LIVE [2]) to testify the performance of objective IQA algorithms, although subjective assessment methods are expensive, laborious, time-consuming and not practical for real-time applications. As a result, more and more objective IQA metrics have been developed to fast and accurately evaluate image quality using various kinds of mathematical models. According to the availability of the original references, we can further divide the objective assessment into three classes. The first class is full-reference (FR) IQA methods, (e.g. mean squared error (MSE) and structural similarity index (SSIM)), which suppose the original image signals are entirely known [3]-[16]. However, we cannot acquire the original image in most applications, and this makes the FR approaches unable to work. Consequently, the studies of reduced-reference (RR) IQA methods, which only utilize part of the original image as features, are recently in the stage of booming evolution for high accuracy [17]-[22]. For instance, the free energy based distortion metric (FEDM) [17] was inspired by the free energy principle to approximate the internal generative model of the human brain [23]-[24] .

Despite of the emergence of a vast majority of IQA algorithms, very few of them (e.g. MSE and SSIM) have been integrated into existing image/video image processing systems, and this is always resulted from the limitations of computational complexity, portability and the requirement of the entire original image. A natural question to ask is whether we can design an effective IQA method to overcome those limitations stated above. So this paper makes a good effort to propose the Free Energy based Metric with JPEG/JPEG2000 (FEMJ). The FEMJ is motivated by the recent revealed free energy principle [23]-[24], similar to our early exploited FEDM. Instead of the linear autoregressive (AR) model used in FEDM [17], the proposed FEMJ adopts widely employed and highly developed JPEG and JPEG2000 (JP2K) compression methods, leading to less computational load and strong portability. Furthermore, the FEMJ only needs one number as the RR features, and this makes it even act as a blind IQA metric since we can encode that number precisely with very few bits in the header's file.

The remainder of this paper is processed as follows. Section 2 first reviews the free energy based human theory before proposing the FEMJ method. In Section 3, experimental results on the most famous LIVE image quality database justify and compare the performance of the FEMJ with some mainstream FR IQA algorithms and existing RR methods, and furthermore, we also analyze the reason why the proposed FEMJ can work fast and practically. In the final, we conclude this paper in Section 4.

## 2. THE FEMJ METRIC

Most existing IQA algorithms mainly target to low level features, such as structural information, image gradient, and phase congruency. However, we believe that the visual quality of images should also highly connected to the psychological mechanism of perception in the human brain. Recently, Friston *et al.* revealed the free energy based human theory to explain and unify several brain theories in biological and physical sciences about human action, perception and learning [23]-[24]. Similar to the Bayesian brain hypothesis [25] that has been widely used in ensemble learning, the free energy principle makes a basic premise that the cognitive process is controlled by an internal generative model in the human brain. With this generative model, the human brain can predict those encountered scenes in a constructive manner.

In essential, this constructive model is a probabilistic model that can be separated into a likelihood term and a prior term. Visual perception is then to inverting this likelihood term, in order to infer the posterior possibilities of the given scene. It is very natural that there always exists a gap between the encountered scene and brain's prediction, because the internal generative model cannot be universal. We believe that the gap between the external input and its generative-model-explainable part is very closely related to the visual quality of perceptions, and thereby can measure the image quality.

Specifically, we assume that the internal generative model $G$ is parametric for visual perception, and the perceived scenes can be explained by adjusting the vector $\boldsymbol{\theta}$ of parameters. Given an image $I$, its 'surprise' (determined by entropy) is evaluated by integrating the joint distribution $P(I, \boldsymbol{\theta}|G)$ over the space of model parameters $\boldsymbol{\theta}$

$$-\log P(I|G) = -\log \int P(I, \boldsymbol{\theta}|G) d\boldsymbol{\theta}. \tag{1}$$

We then introduce an auxiliary term $Q(\boldsymbol{\theta}|I)$ into both the denominator and numerator in Eq. (1) to derive:

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

Using the Jensen's inequality, we can easily obtain the following relationship from Eq. (2):

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

The right hand side of Eq. (3) is the upper bound by a term called 'free energy', which is defined by

$$J(\boldsymbol{\theta}) = -\int Q(\boldsymbol{\theta}|I) \log \frac{P(I, \boldsymbol{\theta})}{Q(\boldsymbol{\theta}|I)} d\boldsymbol{\theta}. \tag{4}$$

It is clear that the free energy is a discrepancy measure between the input image and its best explanation inferred by the internal generative model, and it thereby presents itself as a natural proxy for psychically quality of images. We accordingly define a perceptual distance between the reference image $I_r$ and its distorted counterpart $I_d$ as the absolute difference of the two images in free energy as

$$\text{FEDM}(I_r, I_d) = \left| J(\hat{\boldsymbol{\theta}}_r) - J(\hat{\boldsymbol{\theta}}_d) \right| \tag{5}$$

with

$$\hat{\boldsymbol{\theta}}_r = \arg\min_{\boldsymbol{\theta}_r} J(\boldsymbol{\theta}|G, I_r),$$
$$\hat{\boldsymbol{\theta}}_d = \arg\min_{\boldsymbol{\theta}_d} J(\boldsymbol{\theta}|G, I_d).$$

The $G$ was chosen to be the linear AR model in our previous work for its ability to approximate a wide range of natural scenes by varying its parameters and for its simplicity. The AR model is defined as

$$x_n = \boldsymbol{\chi}^k(x_n)\boldsymbol{\lambda} + \varepsilon_n \tag{6}$$

where $x_n$ is a pixel in question, $\boldsymbol{\chi}^k(x_n)$ is a vector consisting of $k$ nearest neighbors of $x_n$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_k)^T$ is a vector of AR coefficients, and $\varepsilon_n$ is additive Gaussian noise term with zero mean. So the free energy of the reference image $I_r$ is quantified by the entropy between itself and the predicted version $I_p$ by

$$J(I_r) = -\sum_{i=0}^{255} P_i(\tilde{I}_r) \log\ P_i(\tilde{I}_r) \tag{7}$$

where $P_i(\tilde{I}_r)$ indicates the probability density of grayscale $i$ in $\tilde{I}_r$ that is calculated by

$$\tilde{I}_r = I_r - I_p = I_r - R(I_r) \tag{8}$$

with

$$R(x_n) = \boldsymbol{\chi}^k(x_n)\boldsymbol{\lambda}_{est} \tag{9}$$

where $\boldsymbol{\lambda}_{est}$ is the optimal estimate of AR coefficients for $x_n$ using the least square method. Similarly, the free energy of the distorted image $I_d$ is accordingly defined.

We list the performance indices and the run time of the FEDM in Table 1-2. It is apparent that the FEDM achieves fairly well performance yet requires a great deal of rum time. And furthermore, one of the most important functions of IQA methods is used for instruction and optimization in real-time applications. But unfortunately, the AR prediction model is still needed to inserted into the image/video processing systems to be used, rendering the FEDM not portable. Aiming to solve this problem, we in this paper remodify the FEDM with the highly developed and widely employed JPEG and JP2K compression methods to propose the fast, effective and practical FEMJ algorithm.

More precisely, we compute the $G$ using JPEG/JP2K compression method, and redefine the free energy of the reference image $I_r$ by

$$J'(I_r) = -\sum_{i=0}^{255} P_i(\tilde{I}'_r)\log_2 P_i(\tilde{I}'_r) \qquad (10)$$

where $P_i(\tilde{I}_r)$ indicates the probability density of grayscale $i$ in $\tilde{I}_r$ that is computed by

$$\tilde{I}'_r = I_r - I'_p = I_r - F(I_r) \qquad (11)$$

with $F(I_r)$ is evaluated using the 'imwrite' command with JPEG/JP2K compression in Matlab. In a similar way, we can estimate the free energy $J'(I_d)$ of the distorted image $I_d$. At last, the FEMJ is calculated by introducing the $J'(I_r)$ and $J'(I_d)$ into Eq. (5). An important note is that both JPEG and JP2K compressions operate very quickly and have been widely integrated into most image/video processing systems so as to make the FEMJ very portable. Besides, we can adaptively choose JPEG or JPEG2000 compression to compute the FEMJ according to the applicable conditions.

## 3. EXPERIMENTAL RESULTS

We conducted five classical FR and RR IQA algorithms for comparison. They include FR PSNR and SSIM[1], RR FEDM, SDM[2], and the proposed FEMJ. We first follow the suggestion given by VQEG [26] to map the objective predictions of those five methods to subjective scores using a four-parameter logistic function based nonlinear regression:

$$Quality(q) = \frac{\rho_1 - \rho_2}{1 + \exp(-(q - \rho_3)/\rho_4)} + \rho_2 \qquad (12)$$

with $q$ and $Quality(q)$ being the input score and the mapped score. The free parameters $\rho_1$ to $\rho_4$ are to be determined during the curve fitting process. We then use three frequently employed performance measures, Pearson Linear correlation coefficient (PLCC), Spearman rank-order correlation coefficient (SROCC), and root mean-squared error (RMSE) as suggested by VQEG [26] to further evaluate the prediction accuracy of those testing five IQA metrics on the most popular LIVE database. We tabulated and compared the performance indices in Table 1, and displayed the scatter plots of the FEMJ (using JPEG and JP2K) on five various distortion categories in the LIVE database in Fig. 1-2.

It is easy to find from Table 1 that the proposed FEMJ is superior to the classical full-reference PSNR and matchable with the popular full-reference SSIM. As compared with reduced-reference FEDM and SDM methods, the FEMJ has also acquired very encouraging results. And furthermore, our FEMJ has a considerably low computation complexity and very strong portability, since both JPEG and JP2K compressions can operate quickly and have been widely integrated into most existing applications. As illustrated in Table 2, we compare the rum time of those five IQA approaches used in

**Table 1**. Pearson Linear correlation coefficient (PLCC), Spearman rank-order correlation coefficient (SROCC) and root mean-squared error (RMSE) values (after nonlinear regression) of PSNR, SSIM, FEDM, SDM and the proposed FEMJ algorithms on five image subsets of different distortion categories in the LIVE database.

Pearson linear correlation coefficient (PLCC)

| Algorithm | JP2K | JPEG | AGWN | Blur | FF |
|---|---|---|---|---|---|
| PSNR | 0.8996 | 0.8879 | 0.9858 | 0.7835 | 0.8895 |
| SSIM | 0.9410 | 0.9504 | 0.9695 | 0.8743 | 0.9428 |
| FEDM | 0.9260 | 0.9210 | 0.9253 | 0.7355 | 0.8410 |
| SDM | 0.9447 | 0.9569 | 0.9789 | 0.9252 | 0.9316 |
| FEMJ (JPEG) | 0.9627 | 0.9659 | 0.9446 | 0.9294 | 0.8866 |
| FEMJ (JP2K) | 0.9109 | 0.9131 | 0.9529 | 0.9726 | 0.9049 |

Spearman rank-order correlation coefficient (SROCC)

| Algorithm | JP2K | JPEG | AGWN | Blur | FF |
|---|---|---|---|---|---|
| PSNR | 0.8954 | 0.8809 | 0.9854 | 0.7823 | 0.8907 |
| SSIM | 0.9355 | 0.9449 | 0.9629 | 0.8944 | 0.9413 |
| FEDM | 0.9200 | 0.9226 | 0.9152 | 0.7594 | 0.8229 |
| SDM | 0.9439 | 0.9447 | 0.9729 | 0.9342 | 0.9384 |
| FEMJ (JPEG) | 0.9581 | 0.9702 | 0.9523 | 0.9452 | 0.8744 |
| FEMJ (JP2K) | 0.9222 | 0.9164 | 0.9473 | 0.9714 | 0.9147 |

Root mean-squared error (RMSE)

| Algorithm | JP2K | JPEG | AGWN | Blur | FF |
|---|---|---|---|---|---|
| PSNR | 11.017 | 14.653 | 4.7027 | 11.478 | 13.015 |
| SSIM | 8.5349 | 9.9070 | 6.8533 | 8.9643 | 9.4963 |
| FEDM | 9.5226 | 12.409 | 10.613 | 12.516 | 15.410 |
| SDM | 8.2737 | 9.2455 | 5.7166 | 7.0095 | 10.357 |
| FEMJ (JPEG) | 6.8253 | 8.2452 | 9.1854 | 6.8192 | 13.175 |
| FEMJ (JP2K) | 10.410 | 12.987 | 8.4871 | 4.2908 | 12.122 |

[1] Other FR IQA methods are not very practical in real-time applications owing to their complicated models and large computational load.

[2] Most of other RR IQA algorithms require a great amount of RR features, rendering them unable to use in real-time applications.

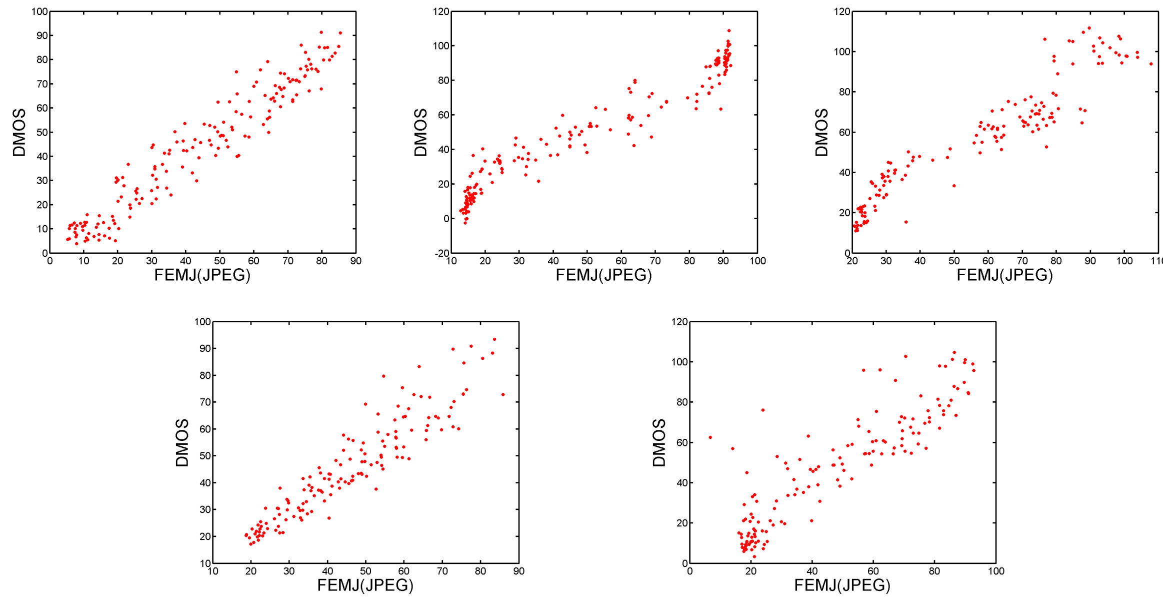


**Fig. 1**. Scatter plots of DMOS vs. the proposed FEMJ (JPEG) (after nonlinear regression) on five different distortion categories in the LIVE database.

**Table 2**. The average run time of PSNR, SSIM, FEDM, SDM and the proposed FEMJ algorithms for per image on the LIVE database.

| Algorithm | Average run time (second per image) |
|---|---|
| PSNR | 0.0070 |
| SSIM | 0.0719 |
| FEDM | 80.856 |
| SDM | 0.0973 |
| FEMJ (JPEG) | 0.0684 |
| FEMJ (JP2K) | 0.0924 |

this research, and prove the fast implementation of the proposed FEMJ metric. In addition, we want to emphasize two points: 1) the JPEG or JPEG2000 compression can be adaptively selected for the FEMJ according to different applicable conditions; 2) the FEMJ only needs one number so as to make itself even act as a blind IQA metric by encoding that number precisely with very few bits in the header's file.

## 4. CONCLUSION

In this paper, we propose a novel Free Energy based Metric with JPEG/JPEG2000 (FEMJ) based on the recent revealed free energy based brain theory, which illustrates that the HVS always attempts to perceive the input image signal by discarding the uncertain parts. In this implementation, we utilize the highly developed JPEG and JPEG2000 compression methods that are widely used in most existing image/video processing systems to realize the FEMJ, so as to render it fast and practical. We conduct comparative studies of our FEMJ and popular full-reference PSNR and SSIM methods and state-of-the-art reduced-reference FEDM and SDM approaches, and confirm the effectiveness in experimental results.


# Acknowledgment

This work was supported in part by NSFC (61025005, 61371146, 61221001), 973 Program (2010CB731401) and FANEDD (201339).


## 5. REFERENCES


[1] A. C. Bovik, "Automatic prediction of perceptual image and video quality," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 2008-2024, September 2013.

[2] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment Database Release 2," [Online]. Available: http://live.ece.utexas.edu/research/quality

[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, April 2004.

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

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

[6] 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, no. 1, March 2010.

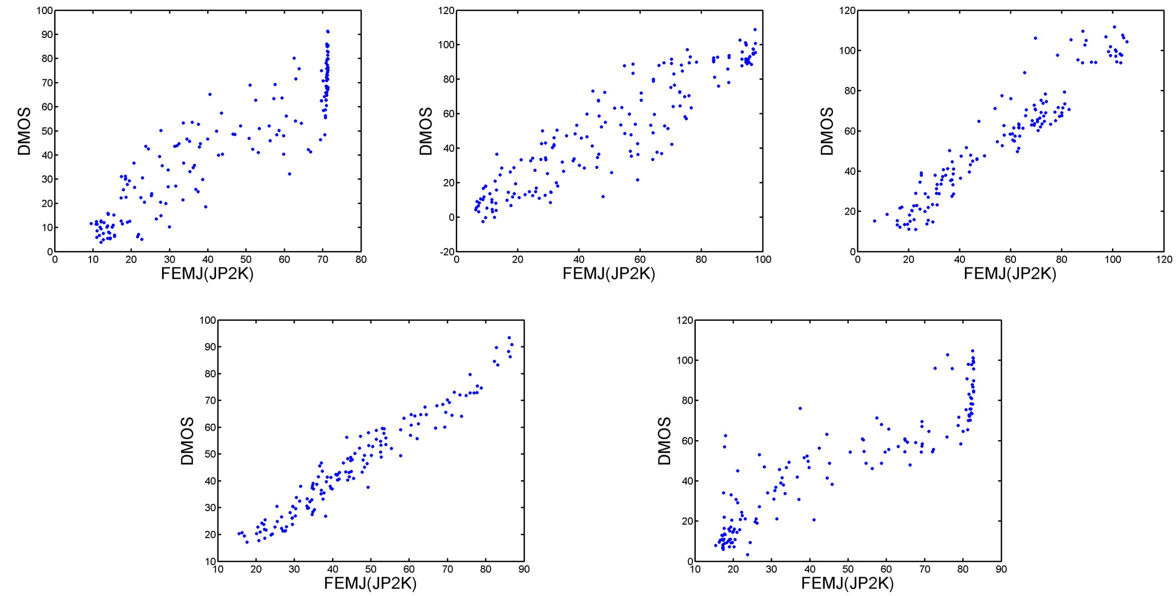


**Fig. 2**. Scatter plots of DMOS vs. the proposed FEMJ (JP2K) (after nonlinear regression) on five different distortion categories in the LIVE database.


[7] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185-1198, May 2011.

[8] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378-2386, August 2011.

[9] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500-1512, April 2012.

[10] K. Gu, G. Zhai, X. Yang, L. Chen, and W. Zhang, "Nonlinear additive model based saliency map weighting strategy for image quality assessment," *IEEE International Workshop on Multimedia Signal Processing*, pp. 313-318, September 2012.

[11] K. Gu, G. Zhai, X. Yang, and W. Zhang, "An improved full-reference image quality metric based on structure compensation," *APSIPA ASC*, December 2012.

[12] J. Wu, W. Lin, G. Shi, and A. Liu, "Perceptual quality metric with internal generative mechanism," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 43-54, January 2013.

[13] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Self-adaptive scale transform for IQA metric," *Proc. IEEE Int. Symp. Circuits and Syst.*, pp. 2365-2368, May 2013.

[14] K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new psychovisual paradigm for image quality assessment: From differentiating distortion types to discriminating quality conditions," *Signal, Image and Video Processing*, vol. 7, no. 3, pp. 423-436, May 2013.

[15] K. Gu, G. Zhai, X. Yang, W. Zhang, and M. Liu, "Structural similarity weighting for image quality assessment," *Proc. IEEE Int. Conf. Multimedia and Expo Workshops*, pp. 1-6, July 2013.

[16] K. Gu, G. Zhai, M. Liu, Q. Xu, X. Yang, and W. Zhang, "Adaptive high-frequency clipping for improved image quality assessment," *Proc. IEEE Int. Conf. VCIP*, November 2013.

[17] 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, January 2012.

[18] R. Soundararajan and A. C. Bovik, "RRED indices: Reduced-reference entropic differencing for image quality assessment," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 517-526, February 2012.

[19] M. Narwaria, W. Lin, I. V. McLoughlin, S. Emmanuel, and L. T. Chia, "Fourier transform-based scalable image quality measure," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3364-3377, August 2012.

[20] A. Rehman and Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3378-3389, August 2012.

[21] K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new reduced-reference image quality assessment using structural degradation model," *Proc. IEEE Int. Symp. Circuits and Syst.*, pp. 1095-1098, May 2013.

[22] K. Gu, G. Zhai, X. Yang, W. Zhang, and M. Liu, "Subjective and objective quality assessment for images with contrast change," *Proc. IEEE Int. Conf. Image Process.*, pp. 383-387, September 2013.

[23] K. Friston, J. Kilner, and L. Harrison, "A free energy principle for the brain," *Journal of Physiology Paris*, vol. 100, pp. 70-87, 2006.

[24] K. Friston, "The free-energy principle: A unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, pp. 127-138, 2010.

[25] D. C. Knill and A. Pouget, "The Bayesian brain: The role of uncertainty in neural coding and computation," *Trends Neurosci.*, vol. 27, no. 12, pp. 712-719, 2004.

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