Comparative Image Quality Assessment using Free Energy Minimization

Guangtao Zhai*† and Andre Kaup†
*Institute of Image Communication and Information Processing,Shanghai Jiao Tong University, Shanghai, 200240, China
Email:zhaiguangtao@gmail.com
†Chair of Multimedia Communications and Signal Processing, Friedrich-Alexander-University, 91058 Erlangen, Germany
Email: kaup@lnt.de

***Abstract*—It is a straightforward task for human observers to judge the relative quality of two visual signals of the same content, but subject to different type/level of distortions. However, this comparative image quality assessment (C-IQA) problem remains a difficult challenge for the current research of image quality assessment (IQA). In this paper, we propose a C-IQA approach to predict the relative perceptual quality of a pair of images that are possibly subject to different artifact types/levels. The C-IQA algorithm is designed to emulate the process of comparing the relative quality of two visual stimuli as performed by the human visual system (HVS) within the framework of free energy minimization. The brain's internal generative models initialized on the inputs are used to explain the two images. And their relative quality can then be determined through comparing the free energy level of this model-data fitting process. In the existing work of IQA, the full-reference (FR) and reduced-reference (RR) methods need the prior knowledge of the original images while the no-reference (NR) algorithms usually work with a single input image. The C-IQA approach is inherently different from those existing methods in that it takes as input an image pair and predicts their relative quality without using any knowledge about the original image. A computationally efficient solution to the proposed C-IQA scheme based on a linear autoregressive image model is also introduced. Experimental results show that the proposed method achieves about 98% accuracy in line with the subjective ratings when applied on over 300,000 image pairs sampled from the LIVE database, outperforming the FR metrics such as PSNR, SSIM, and some of the most advanced NR IQA algorithms.**

I. Introduction

It is our daily experience to justify the comparative quality of two visual signals, e.g. choosing from cameras, cable companies or TV sets that provide better imaging, transmission or representation qualities for the visual signals. During those comparative image quality assessment (C-IQA) tasks, despite of the fact that we usually have no access to the original pristine image, it is fairly easy for our human visual system (HVS) to tell the relative qualities. In the current research of image quality assessment (IQA), image quality metrics are usually classified into full-reference (FR), reduced-reference (RR) and no-reference (NR) approaches depending on the accessibility to the original reference points [1]. Clearly, the C-IQA problem does not fall into the FR and RR categories because there is no prior knowledge about the original image. Yet, the C-IQA problem is also not the same as the currently practices of NR-IQA.

Existing NR-IQA methods can be classified into two categories according to whether the image distortion process or the image formation process is modeled. Most of the early attempts on NR-IQA assume certain types of distortion, e.g. image blur [2] , blockiness [3], [4] or ringing artifact [5], and therefore convert the problem of IQA to distortion measures. Recent works along this line of research include [6], [7], [8], However, by design, such methods only work well for specific types of images. Hence, difficulty arises for this type of methods to assess the relative quality of two images subject to different distortions. Although multiple distortion measures can work in tandem towards a holistic quality measure [9], accuracy of this heuristic solution is usually low due to the lack of a unified framework and because of the fact that the distortion measures can interfere with each other. Another emerging type of NR-IQA algorithms is based on the natural scene statistics (NSS) model [10], [11], [12]. The underlying reasoning is that our visual system has evolved in the natural environment for millions of years and is therefore well adapted to natural scenes. A prior statistical model derived from a large set of natural images is therefore expected to capture the "naturalness" of images. So by measuring its departure from the ideal "naturalness", we can predict the perceptual quality loss of a distorted image.

For the problem of C-IQA, the distortion detection type of NR-IQA algorithm is clearly not applicable because of the fact those two visual signals may undergo difference type of distortions. The NSS type of holistic NR-IQA algorithm is applicable to the problem but it always tries to quantify the qualities of the two images independently. This is in contrast to the working mechanism of the HVS to accomplish the task of assessing the relative quality of a pair of images. According to our experience, when being asked to compare the relative quality of two images, we usually alternate our focus within and between both images, making and validating assumptions, before coming to the final conclusion. In other words, we use information from both input images to fulfill the C-IQA task. As an example, try to quantify the relative quality of the two distorted images in Fig. 1 and verify the existence of cross-reference between the images. On the other hand, however, NSS type of holistic NR-IQA algorithms work with a single image input and generate an "absolute" quality prediction. To compare the relative quality of two images, we need to run the

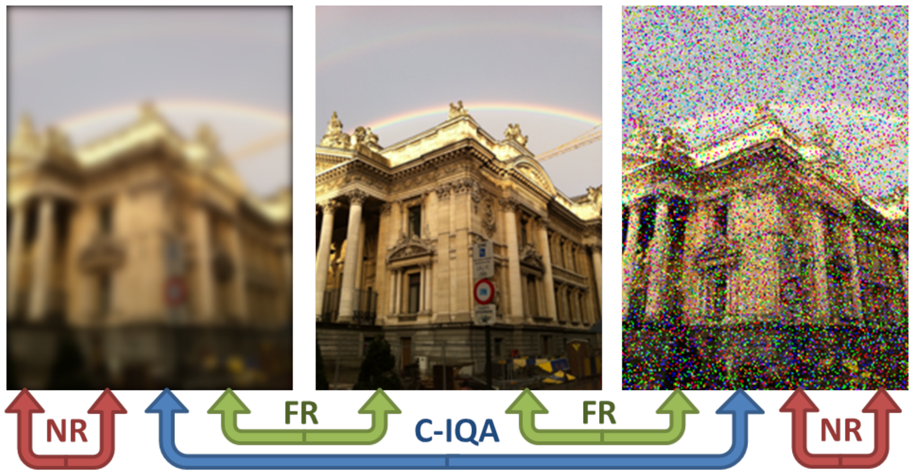

Fig. 1: An example to show the inherent difference between C-IQA and FR-/NR-IQA. From left to right: blurred image, the original image, and image with salt & pepper noise. Please cover the middle reference image to verify that the HVS extracts information from both distorted images to quantify their relative visual quality. In this example, if we define the left image as I_1 and the right image as I_2, we have estimates of the free energy levels $F_{I_1 \to I_1}(\boldsymbol{\theta}) = 1.13$, $F_{I_1 \to I_2}(\boldsymbol{\theta}) = 7.85$, $F_{I_2 \to I_2}(\boldsymbol{\theta}) = 8.01$ and $F_{I_2 \to I_1}(\boldsymbol{\theta}) = 4.39$. Since $F_{I_1 \to I_2}(\boldsymbol{\theta}) < F_{I_2 \to I_2}(\boldsymbol{\theta})$ and $F_{I_1 \to I_1}(\boldsymbol{\theta}) < F_{I_2 \to I_1}(\boldsymbol{\theta})$, so image I_1 (left) is believed to have better visual quality than image I_2 (right).

NR algorithm twice and match the scores. During this process, the quality of an image is determined without whatsoever the information from the other. The quality predicton of NR-IQA is solely based on an appearance model (e.g. the NSS model). Clearly, this observation suggests a significant departure of existing NR-IQA algorithms from the way that the HVS works when solving the C-IQA problem.

In this paper, we propose a C-IQA algorithm to simulate the way how the HVS judges the relative quality between a pair of images. The C-IQA problem is modeled as the estimation and comparison of the posterior probability of one image given the other. We resort to a free-energy theory [13] induced structural model for images. The internal generative models optimized on the inputs are used to explain the image itself as well as its counterpart (model transition). The discrepancies between one image and its internal-model and transition-model explainable parts are computed and compared. And the image that can better explain the other one is believed to have superior quality because the underlying model has higher description power. In will be shown in the paper that this free energy estimation and comparison based C-IQA algorithm achieves a remarkable 98% accuracy rate (i.e. conformance to the subjective ratings) when quantifying the relative quality of over 300,000 image pairs sampled from the LIVE database [14], outperforming FR-IQA method PSNR, SSIM and some state of the art NSS based NR-IQA algorithms.

The remainder of the paper is organized as follows: Section II outlines the C-IQA framework. Section III introduces the algorithm in detail. Experimental results and comparative studies are performed in Section IV. Concluding remarks are given in Section V.

II. Comparative Image Quality Assessment via Free Energy Minimization

A. General Formulation

Given a pair of images (I_1, I_2), we will first focus on the cognitive process of quantifying the visual quality of I_1. The Bayesian brain theory [15] and the free energy principle [13] suggest that the cognitive process is governed by an internal generative model $\mathcal{G}$ in the brain. Given a scene or image I_1, the model $\mathcal{G}$ adapts itself through varying a parameter vector $\boldsymbol{\theta}$. Then the "surprise" of observing the image I_1 using the internal model $\mathcal{G}$, is defined by integrating the joint distribution $P(I_1, \boldsymbol{\theta}|\mathcal{G})$ over the parameter space

$$-\log P(I_1|\mathcal{G}) = -\log \int P(I_1, \boldsymbol{\theta}|\mathcal{G}) d\boldsymbol{\theta}. \tag{1}$$

Note that conditional entropy of the observation I_1 given the model $\mathcal{G}$ can be computed as average of surprise $H(I_1|\mathcal{G}) = \int -P(I_1|\mathcal{G}) \log P(I_1|\mathcal{G})$ as a measure of "randomness" of the images. As in ensemble learning [16] and variational Bayesian estimation [17], we use an auxiliary posterior distribution $Q(\boldsymbol{\theta}|I_1, \mathcal{G})$ to calculate the surprise of I_1 in (1). Note that to fully utilize the information from the data, this auxiliary posterior can also be drawn from I_2 to get $Q(\boldsymbol{\theta}|I_2, \mathcal{G})$. Since the behavior of the model can be characterized by the parameter $\boldsymbol{\theta}$, from now on, we drop the latent model assumption $\mathcal{G}$ in our analysis for simplicity.

By letting the auxiliary terms $Q(\boldsymbol{\theta}|I_1)$ and $Q(\boldsymbol{\theta}|I_2)$ into (1) we have

$$-\log P(I_1) = -\log \int Q(\boldsymbol{\theta}|I_1) \frac{P(I_1, \boldsymbol{\theta})}{Q(\boldsymbol{\theta}|I_1)} d\boldsymbol{\theta}. \tag{2}$$

$$-\log P(I_2) = -\log \int Q(\boldsymbol{\theta}|I_2) \frac{P(I_1, \boldsymbol{\theta})}{Q(\boldsymbol{\theta}|I_2)} d\boldsymbol{\theta}. \tag{3}$$

Recall that for the C-IQA problem, we want to evaluate the perceptual quality of I_1 using the model derived from I_1 and I_2. The free energy of the C-IQA system $F = -P(I)$ is defined as

$$F_{I_1 \to I_1}(\boldsymbol{\theta}) = -\int Q(\boldsymbol{\theta}|I_1) \log \frac{P(I_1, \boldsymbol{\theta})}{Q(\boldsymbol{\theta}|I_1)} d\boldsymbol{\theta}. \quad (4)$$

$$F_{I_2 \to I_1}(\boldsymbol{\theta}) = -\int Q(\boldsymbol{\theta}|I_2) \log \frac{P(I_1, \boldsymbol{\theta})}{Q(\boldsymbol{\theta}|I_2)} d\boldsymbol{\theta}. \quad (5)$$

where the subscript $I_1 \to I_1$ and $I_2 \to I_1$ indicates I_1 and I_2 are respectively used to infer the quality of I_1.

By noticing that $P(I_1, \boldsymbol{\theta}) = P(I_1|\boldsymbol{\theta})P(\boldsymbol{\theta})$, we can write (4) into

$$\begin{aligned} F_{I_1 \to I_1}(\boldsymbol{\theta}) &= \int Q(\boldsymbol{\theta}|I_1) \log \frac{Q(\boldsymbol{\theta}|I_1)}{P(I_1|\boldsymbol{\theta})P(\boldsymbol{\theta})} d\boldsymbol{\theta} \quad (6) \\ &= \mathrm{KL}(Q(\boldsymbol{\theta}|I_1)||P(\boldsymbol{\theta})) + E_Q[-\log P(\mathbf{I}_1|\boldsymbol{\theta})]. \end{aligned}$$

Similarly, (5) can also be written as

$$F_{I_2 \to I_1}(\boldsymbol{\theta}) = \mathrm{KL}(Q(\boldsymbol{\theta}|I_2)||P(\boldsymbol{\theta})) + E_Q[-\log P(\mathbf{I}_1|\boldsymbol{\theta})]. \quad (7)$$

Here the terms $\mathrm{KL}(Q(\boldsymbol{\theta}|I_1)||P(\boldsymbol{\theta}))$ and $\mathrm{KL}(Q(\boldsymbol{\theta}|I_2)||P(\boldsymbol{\theta}))$ measure the distance between the recognition densities and the true prior of the model parameters. The term $E_Q[-\log P(\mathbf{I}_1|\boldsymbol{\theta})]$ is the averaged entropy of predicting I_1 over the approximating posteriors $Q(\boldsymbol{\theta}|I_1)$ and $Q(\boldsymbol{\theta}|I_2)$.

We can now formulate the dual-process of inferring the quality of I_2 with the internal model initialized with I_2 and I_1 as

$$F_{I_2 \to I_2}(\boldsymbol{\theta}) = \mathrm{KL}(Q(\boldsymbol{\theta}|I_2)||P(\boldsymbol{\theta})) + E_Q[-\log P(\mathbf{I}_2|\boldsymbol{\theta})]. \quad (8)$$

$$F_{I_1 \to I_2}(\boldsymbol{\theta}) = \mathrm{KL}(Q(\boldsymbol{\theta}|I_1)||P(\boldsymbol{\theta})) + E_Q[-\log P(\mathbf{I}_2|\boldsymbol{\theta})]. \quad (9)$$

The relative quality of I_1 and I_2 can be assessed through comparing the self free energy terms $F_{I_1 \to I_1}(\boldsymbol{\theta})$ and $F_{I_2 \to I_2}(\boldsymbol{\theta})$ and the cross free energy terms $F_{I_1 \to I_2}(\boldsymbol{\theta})$ and $F_{I_2 \to I_1}(\boldsymbol{\theta})$. First, if the free energy level can be reduced using another observation, i.e. the cross free energy is lower than the self free energy level: $F_{I_2 \to I_2}(\boldsymbol{\theta}) > F_{I_1 \to I_2}(\boldsymbol{\theta})$ or $F_{I_1 \to I_1}(\boldsymbol{\theta}) > F_{I_2 \to I_1}(\boldsymbol{\theta})$ the observations I_1 and I_2 are considered respectively to have better visual quality. In the condition that the cross free energy levels are both higher than the self free energy levels, i.e. $F_{I_1 \to I_1}(\boldsymbol{\theta}) < F_{I_2 \to I_1}(\boldsymbol{\theta})$ or $F_{I_2 \to I_2}(\boldsymbol{\theta}) < F_{I_1 \to I_2}(\boldsymbol{\theta})$, we further compare the cross free energy. The image that explains the other one better is considered to have better quality. Under the circumstances that the cross free energy levels are both lower than the self free energy levels, i.e. $F_{I_2 \to I_2}(\boldsymbol{\theta}) > F_{I_1 \to I_2}(\boldsymbol{\theta})$ and $F_{I_1 \to I_1}(\boldsymbol{\theta}) > F_{I_2 \to I_1}(\boldsymbol{\theta})$, we also compare the cross free energy terms $F_{I_1 \to I_2}(\boldsymbol{\theta})$ and $F_{I_2 \to I_1}(\boldsymbol{\theta})$, and again the image that can better explain the other one is considered to have better quality.

B. Practical Calculation of the Free Energy

For operational amenability we can assume the generative model $\mathcal{G}$ to be a 2D linear autoregressive (AR) model for its high description capability for natural images [18] and wide acceptance in image processing algorithms [19], [20], [21]. The AR model is defined as

$$x_i = \mathcal{X}^k(x_i)\boldsymbol{\alpha} + e_i \quad (10)$$

where x_i is a pixel at location i, $\mathcal{X}^k(x_n)$ defines the k member neighborhood vector of x_i and $\boldsymbol{\alpha} = (a_1, a_2, \ldots, a_k)^T$ defines the model parameter. Towards a pragmatic image quality metric, in this research we opt for straightforward solution using a piecewise AR model formulation [18]. Indeed, research shows that the free energy equals the total description length of the image I under the large sample limit [22] so the model can be estimated via minimizing the total descriptive length [23] or using the Bayesian information criterion (BIC) [24] as

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \left(-\log P(I|\boldsymbol{\alpha}) + \frac{k}{2} \log N \right) \quad (11)$$

where N is the data sample size. To further reduce the complexity, we choose a fixed model order and training set size, and thus change the complicated model comparison process into residual minimization.

Given the k-th order piecewise AR model, for a pixel x_i, we can write the linear equations as

$$x_i = \mathcal{X}^k(x_i)\boldsymbol{\alpha} + e_i, \quad x_i \in \mathcal{N}(x_i) \quad (12)$$

with $\mathcal{X}^k(x_i)$ denoting the model support and $\mathcal{N}(x_i)$ the training set in a neighborhood of x_i. To estimate $\boldsymbol{\alpha}$, the linear system can be written in matrix form as

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{x} - \mathbf{X}\boldsymbol{\alpha}\|_2 \quad (13)$$

with $\mathbf{x} = (x_1, x_2, \ldots, x_N)^T$ and $\mathbf{X}(i,:) = \mathcal{X}^k(x_i)$. The linear system can be solved easily as $\boldsymbol{\alpha} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{x}$.

With $\hat{\boldsymbol{\theta}}$, we can then get the estimation error of a local pixel as

$$\hat{e}_i = x_i - \hat{x}_i = x_i - \mathcal{X}^k(x_i)\hat{\boldsymbol{\alpha}}. \quad (14)$$

For the input image I, the point-wise error e_i can then be pooled to get an error map $E(I)$ and entropy of the errors can also be computed as $H(E) = \sum -P(e)\log P(e)$ with $P(e)$ being the probability distribution of the errors.

In our implementation, an 8-th order AR model is trained locally in a local 7×7 neighborhood. The two pseudo-inverse of the 48×8 matrix can actually be solved more efficiently using Gaussian eliminations. The algorithm written in optimized C code can operate in almost real time (i.e. 30 frames per second) for image input at VGA resolution.

III. Experimental Results and Discussions

We have tested the proposed free energy based C-IQA algorithm on the most widely used LIVE database [14] which consisted of 982 test images: 29 original reference images subject to five kinds of distortions, namely JPEG compression, JPEG2000 compression, Gaussian noise, Gaussian blur and fast fading. We first sampled all the possible image pairs from the database with the constraint that they have the same

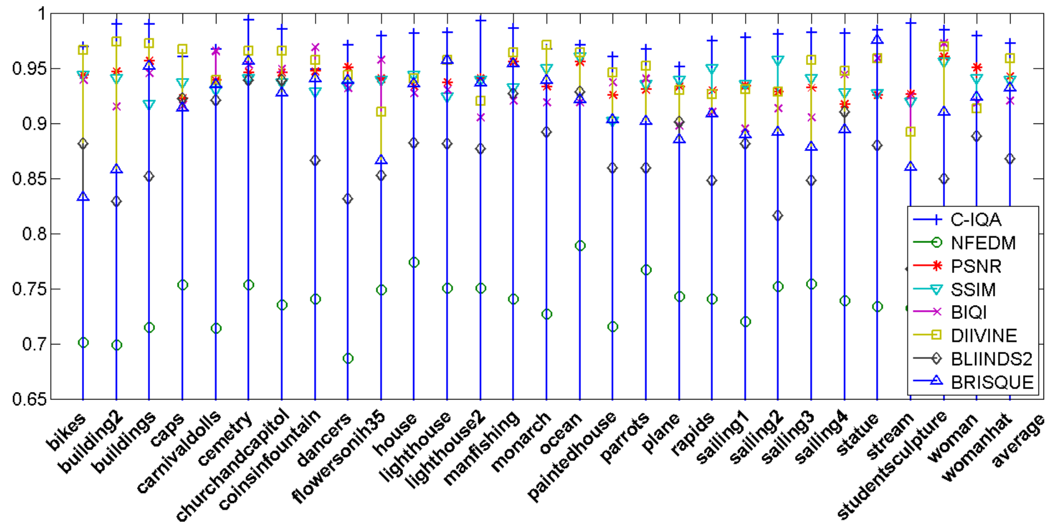

Fig. 2: Prediction accuracy of the proposed C-IQA algorithm and other test algorithms for the 29 original images in the LIVE database. The horizontal axis is the image indexes. The vertical axis is the prediction accuracy.

TABLE I: Comparisons of Prediction Accuracy of the Proposed C-IQA and some FR- and NR-IQA Algorithms

	C-IQA	NFEDM[25]	PSNR	SSIM[26]	BIQI[27]	DIIVINE[28]	BLIINDS2[29]	BRISQUE[30]
Minimum	85.7%	28.57%	61.9%	61.9%	71.4%	73.3%	76.2%	66.7%
Mean	98.11%	77.98%	94.6%	95.0%	92.9%	90.4%	94.8%	95.9%

reference image. The total number of the image pairs used in our test is 304,800.

The prediction result of the proposed C-IQA algorithm for each of the original 29 reference images in the LIVE database for the proposed C-IQA algorithm and some competing algorithms are shown in Fig. 2. We have tested the free energy based NR IQA algorithm NFEDM [25] and some recent NSS based NR IQA metrics: BIQI [27], DIIVINE [28], BLIINDS2 [29] and BRISQUE [30]. Note that those NSS training based algorithms were trained on the entire LIVE database. For the sake of completeness, we have also included the benchmark FR IQA metrics PSNR and SSIM [26]. It can be seen that the proposed C-IQA algorithm achieved the highest accuracy for almost all of the test images, outperforming the FR methods and those NR methods optimized for the LIVE database.

The overall prediction accuracy rates of the tested algorithms are summarized in Table I. It can be seen that the Proposed C-IQA algorithm achieves the highest rate of 98.1%. The BRISQUE [30] method has the second best result of 95.9%. And the results of PSNR, SSIM [26] and BLIINDS2 [29] are all around 95% while the results of BIQI [27], DIIVINE [28] are 92.9% and 90.4% respectively. Understandably, self free energy based NFEDM [25] has the worst performance. The lowest performance under the test scenarios is also shown in Table I. The Proposed C-IQA algorithm has the best "worst case" performance. Note that although some of the NR metrics have lower averaged accuracy rates than PSNR and SSIM, their minimal performance is still much better than PSNR and SSIM. This observation suggests that the performances of those NR algorithms are more consistent than PSNR and SSIM. Again, it should be emphasized that here we have compared our proposed method with the FR method SSIM and PSNR as well as the NR algorithms trained on the LIVE database itself. It is a remarkable finding that the proposed training free C-IQA method achieves the best overall performance.

IV. CONCLUSION

We have introduced in this paper the concept of comparative image quality assessment (C-IQA) that predicts the relative quality of two images subject to arbitrary distortions at different level. We then introduced a free energy minimization based C-IQA algorithm. Unlike the no-reference or blind IQA methods that treat the two input images independently, information from both images are combine to initialize posterior distributions for parameters of the internal generative model. The model learned from each image is then used to explain the image itself and the other one. And the relative quality of the two images can be precisely estimated through examining the free energy levels of the model-image fitting process. The proposed algorithms has been tested using over 300,000 image pairs sampled from the LIVE image database with a remarkable overall prediction accuracy of 98% which even outperforms the full reference methods PSNR and SSIM as well as some of the most advanced natural scene statistics based no-reference image quality metrics.

References

[1] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication and Image Representation*, vol. 44, no. 4, pp. 297–312, 2011.

[2] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric," *IEEE Internatinal Conference on Image Processing*, pp. 57–60, 2002.

[3] H. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding," *IEEE Signal Processing Letters*, vol. 4, no. 11, pp. 317–320, 1997.

[4] Z. Wang, A. Bovik, and B. Evan, "Blind measurement of blocking artifacts in images," *IEEE International Conference on Image Processing*, pp. 981–984, 2000.

[5] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: Application to jpeg2000," *Signal Process: Image Communication*, vol. 19, no. 2, pp. 163–172, 2004.

[6] R. Ferzli and L. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Transactions on Image Processing*, vol. 18, no. 4, pp. 717–728, 2009.

[7] H. Liu, N. Klomp, and I. Heynderickx, "A no-reference metric for perceived ringing artifacts in images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 4, pp. 529–539, 2010.

[8] X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Transaction on Image Processing*, vol. 19, no. 12, pp. 3116–3132, 2010.

[9] X. Li, "Blind image quality assessment," *IEEE International Conference on Image Processing*, pp. 449–453, 2002.

[10] E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *Annual Review of Neuroscience*, vol. 24, pp. 1193–1216, 2001.

[11] A. Srivastava, A. B. Lee, E. P. Simoncelli, and S.-C. Zhu, "On advances in statistical modeling of natural images," *Journal of Mathematical Imaging and Vision*, vol. 18, no. 1, pp. 17–33, 2003.

[12] E. Simoncelli, *The Essential Guide to Image Processing*. San Diego, CA:Adademic, 2009, ch. Capturing visual image properties with probabilistic models.

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

[14] H. Sheikh, Z.Wang, L. Cormack, and A. Bovik, "Live image quality assessment database release 2." [Online]. Available: http://live.ece.utexas.edu/research/quality

[15] D. C. Knill and A. Pouget, "The bayesian brain: the role of uncertainty in neural coding and computation," *Trends in Neurosciences*, vol. 27, pp. 712–719, 2004.

[16] D. Mackay, "Ensemble learning and evidence maximization," *Advances in Neural Information Processing Systems*, 1995.

[17] G. E. Hinton and D. von Camp, "Keeping neural networks simple by minimising the description length of weights," *Proceedings of ACM Conference on Computational Learning Theory*, pp. 5–13, 1993.

[18] X. Wu, G. Zhai, X. Yang, and W. Zhang, "Adaptive sequential prediction of multidimensional signals with applications to lossless image coding," *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 36–42, 2011.

[19] X. Li and M. Orchard, "New edge-directed interpolation," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1521–1527, 2001.

[20] X. Zhang and X. Wu, "Image interpolation by adaptive 2d autoregressive modeling and soft-decision estimation," *IEEE Transactions on Image Processing*, vol. 17, no. 6, pp. 887–896, 2008.

[21] X. Wu, X. Zhang, and X. Wang, "Low bit-rate image compression via adaptive down-sampling and constrained least squares upconversion," *IEEE Transactions on Image Processing*, vol. 18, no. 3, pp. 552–561, 2009.

[22] H. Attias, "A variational bayesian framework for graphical models," *Advances in Neural Information Processing Systems*, vol. 12, pp. 209–215, 2000.

[23] J. Rissanen and G. G. Langdon Jr., "Universal modeling and coding," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 12–23, Jan 1981.

[24] G. Schwarz, "Estimating the dimension of a model," *Annals of Statistics*, vol. 6, pp. 461–464, 1978.

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

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

[27] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 513–516, 2010.

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

[29] M. Saad and A. C. Bovik, "Blind image quality assessment: A natural scene statistics approach in the dct domain," *IEEE Transactions Image Processing*, vol. 21, no. 8, pp. 3339–3352, 2012.

[30] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *EEE Transactions on ImageProcessing*, vol. To appear, 2012.