

A Novel Approach for Quality Evaluation of Digital Images By Auto Regressive Approach Based On Internal Generative Mechanism

*D. Rihana Bhanu*¹ *Trupthi gadhale (Ph.D)*²

¹(PG Scholar) Department of ECE, KV Subbareddy College Of Engineering for Women

²(Ph.D) Head of the Department, Department of ECE, KV Subbareddy College Of Engineering for Women

Abstract

The main aim of Objective image quality assessment (IQA) is to evaluate image quality consistently with human perception. We have different types of perceptual IQA metrics but they cannot accurately represent the degradations from different types of distortions, e.g., existing structural similarity metrics perform well on content dependent distortions and give the better peak signal-to-noise ratio (PSNR) but it is not well on content-independent distortions. In this paper, we integrate the merits of the existing IQA metrics with the guide of the recently revealed internal generative mechanism (IGM). The IGM indicates that the human visual system actively predicts sensory information and tries to avoid residual uncertainty for image perception and understanding. Motivated by the IGM theory, here we assume an autoregressive prediction algorithm to decompose an input scene into two portions, the predicted portion with the predicted visual content and the disorderly portion with the residual content. Distortions on the predicted portion cause to degrade the primary visual information, and structural similarity procedures are employed to measure its degradation; distortions on the disorderly portion mainly change the uncertain information and the PSNR is employed for it. Based on the noise energy deployment on the two portions, finally we mix the two evaluation results to acquire the overall quality score. Simulation results show better performance comparable with the state-of-the-art quality metrics.

KEYWORDS

Internal generative mechanism, Human perception, PSNR

INTRODUCTION

The human visual system (HVS) is the ultimate receiver of sensory information, perceptual image quality assessment (IQA) is useful for many

image and video systems, e.g., for information acquisition, compression, transmission and restoration, to make them HVS oriented. Therefore, an objective visual quality metric

consistent with the subjective perception is in demand. In order to develop an accurate IQA metric in accord with the subjective perception, researchers turn to investigate the HVS characteristics to seek for image features which affect quality assessment, such as brightness, contrast, frequency content, structure and statistical information [4]. Many HVS oriented IQA metrics are proposed such as noise quality measure (NQM) [5], structural similarity (SSIM) [6], visual information fidelity (VIF) [7], the PSNR-HVS-M [8], visual signal-to-noise ratio (VSNR) [9], and the recently proposed most apparent distortion (MAD) [10] and feature similarity (FSIM) [11].

The SSIM index is the most popular one among all of these IQA metrics. This index is based on the assumption that the HVS is highly adapted for extracting structural information from the input scene [6]. In [12], [13], SSIM is improved by using edge/gradient feature of the image so the edge conveys important visual information for understanding. And another high-level HVS property based and well accepted metric, the VIF index computes the mutual information between the reference and test images for visual information fidelity evaluation [7]. These HVS oriented IQA metrics promote our understanding on sensory signal processing and perceptual quality assessment.

Different types of distortion cause different degradation. However, these existing HVS oriented IQA metrics gives better by considering the content-independent. But this proposed HVS oriented IQA metrics perform well on content-dependent distortions (e.g., blur and compression

noise) but not well enough on content-independent distortions (e.g., white noise and impulse noise) [3]. While PSNR/MSE performs the opposite way. Recently, Larson and Chandler [10] advocated that the HVS uses multiple strategies to determine image quality. And near-threshold and clearly visible (suprathreshold) distortions are measured separately in their model. This model mainly considers the distinctions of different energy levels rather than the different effects of distortions. In [14], Li et al. introduced an ad hoc procedure to decouple the original distortion into detail loss and additive impairment for discriminative measurement.

Recent researches on brain theory and neuroscience, such as the Bayesian brain theory [15] and the free-energy principle [16], indicate that the brain works with an internal generative mechanism (IGM) for visual information perception and understanding. Within the IGM, the brain performs as an inference system that actively predicts the visual sensation and avoids the residual uncertainty/disorder [15]–[17]. Thus, we adopt a Bayesian prediction model [15], [18] in our method, and the input scene is decomposed into predicted and disorderly portions. We suppose that distortions on the predicted content will damage the primary visual information, such as blur the edge and destroy the structure, which impact on image understanding. Therefore, edge and structure similarity [6], [12] are used for evaluation on this portion. On the other hand, distortions on the disorderly portion (predicted residual, which arouses uncomfortable sensation) is somewhat content independent.

So we take the assumption of PSNR to estimate the degradation on disorderly uncertainty since PSNR is good for content-independent noise measurement [1], [3]. Finally, we combine the results on the two portions with an adaptive

PROPOSED IQA SCHEME

AR Based Image Prediction

In image processing the decomposition is nothing but which splits an image into two or more portions for discriminately processing, e.g., to decompose an input scene into textural and cartoon parts for just noticeable difference estimation [28]. In this paper, inspired by the IGM theory about the visual perceptual process, we try to decompose an image into predicted and disorderly portions for quality evaluation. Since the Bayesian brain theory indicates that the brain performs as an active inference procedure.

We adopt a Bayesian prediction based autoregressive (AR) model [18], [32] for image content inference. The Bayesian brain theory uses Bayesian probability to imitate the inference procedure for image perception and understanding

nonlinear procedure to acquire the overall score. Simulation results on six publicly available image databases confirm that the proposed model is comparable with the state-of-the-art IQA metrics.

in the IGM [15], [16]. The key of this theory is a probabilistic model that optimizes the input scene by minimizing the prediction error. For example, with an input scene, the Bayesian brain system tries to maximize the conditional probability $p(x/X)$ between the central pixel x and its surrounding $X = \{x_1, x_2, \dots, x_N\}$ [15] for error minimization. By decomposing the conditional probability $p\left(\frac{x}{X}\right)$ and analyzing the correlation between the central pixel x and the pixels x_i in the surrounding X , it can be seen that these x_i which strongly correlated to x play dominant roles for $p(x/X)$ maximization [33]. Therefore, the mutual information ($I(x; x_i)$) between the central pixel x and its surrounding pixel x_i is adopted as the autoregressive coefficient, and an AR model is created to predict the value of the pixel x .

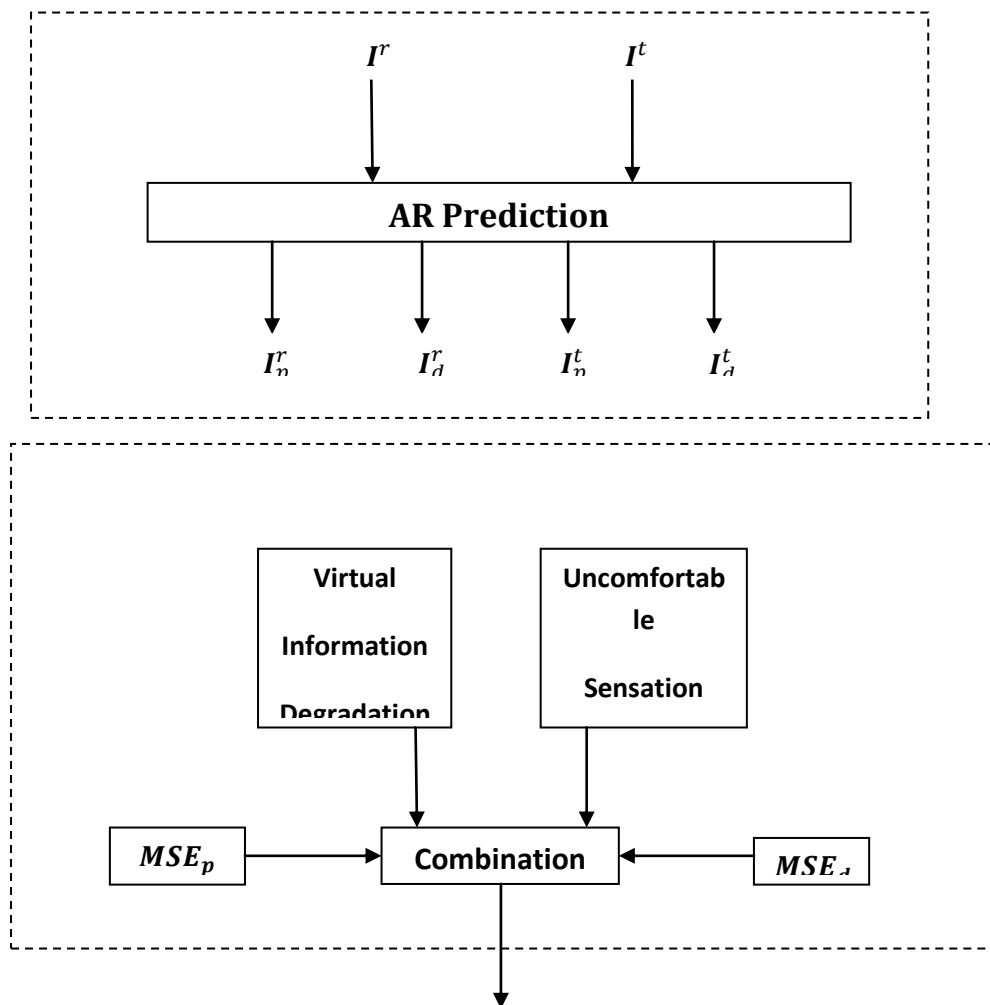


Fig. 1 Flowchart of the proposed model $I^r(I^t)$ is the **Score** (test) image, $I_p^r(I_p^t)$ and $I_d^r(I_d^t)$ are the predicted and disorderly portions of $I^r(I^t)$ respectively.

$$x' = \sum_{x_i \in \chi} C_i x_i + \varepsilon \quad (1)$$

Where x' is the predicted value of pixel x , $C_i = \frac{I(x; x_i)}{\sum_k I(x; x_k)}$ being the normalized coefficient, and ε is white noise. In this paper, we set χ as a 21×21 surrounding region. With the predicted model (1), an input image (I) is decomposed into two portions, the predicted image (I_p) and the disorderly image (I_d), as shown in Fig. 2. In the next subsections, we will evaluate the degradations on the two decomposed images, respectively, since distortions on the two portions have different impacts toward the perceptual quality.

The disorderly portion is composed of the uncertain stimuli of the original image [16]. Distortion on this portion has little effect on image understanding and mainly generates uncomfortable sensation. As a natural way to define the energy of the error signal [1], the PSNR metric presents a good match with the HVS when the error signal is independent of the original signal [3], and this point is also confirmed by the experiments in [2]. Since the distortion of the disorderly portion is independent of the original image content, the PSNR is adopted to evaluate the quality of this portion. Therefore the

B. Uncomfortable Sensation Variation

uncomfortable sensation variation is computed as follow

$$P(I_d^r, I_d^t) = \frac{1}{C_1} \text{psnr}(I_d^r, I_d^t) = \frac{1}{C_1} 10 \log_{10} \left(\frac{255^2}{\text{MSE}(I_d^r, I_d^t)} \right) \quad (2)$$

where I_d^r and I_d^t are the disorderly portions of the reference and test images, respectively; $\text{PSNR}(I_d^r, I_d^t)$ is the PSNR value between I_d^r and I_d^t , and $\text{MSE}(I_d^r, I_d^t)$ is the mean squared error between I_d^r and I_d^t (the minimal value of MSE (such as 1) is set to avoid infinite psnr); C_1 is a constant parameter which is used to normalize the PSNR value into the range [0 1], for this purpose, we set $C_1 = 10 \log_{10}^{255^2}$

C. Visual Information Degradation

Since the predicted portion possesses the primary visual information and distortion on this portion impacts on image understanding, we should adopt some high-level HVS properties to evaluate the degradation of the visual information. In this paper, degradations on edge and structure are computed for primary visual information fidelity where ∇_K are four directional filters, as shown in Fig. 3, $\phi = 1/16$, and symbol $*$ denotes the convolution operation. However, some image regions (e.g., the feather of the parrots in Fig. 2) has no apparent edge but still represents specific structural character. In addition, the HVS is highly adapted for extracting structural information from a scene for recognition. Therefore, besides edge similarity, we need another primary visual information degradation measurement to evaluate the fidelity on image structure. Here, we adopt the structural similarity [6] to evaluate the degradation on structural information

evaluation. The HVS is highly sensitive to the edge, which conveys important visual information and is crucial for scene understanding [12], [34]. The degradation on the edge between the predicted portions of the reference image (I_p^r) and the test image (I_p^t) is computed as their edge height similarity,

$$g(x_p, y_p) = \frac{2E_p^r(x_p)E_p^t(x_p)+C_2}{E_p^r(x_p)^2+E_p^t(x_p)^2+C_2} \quad (3)$$

Where x_p and y_p are the corresponding pixels from the predicted portions of the reference and test images (I_p^r and I_p^t), respectively; $g(x_p, y_p)$ is the edge similarity between x_p and y_p , E_p^r and E_p^t are the edge height maps of I_p^r and I_p^t , respectively, C_2 is the small constant to avoid the denominator being zero and is set as $C_2 = (0.03 \times L)^2$ [6], and L is the gray level of the image. The edge height E_p^r (same for E_p^t) is computed as the maximal edge response along the four directions [27],

$$E_p^r(x) = \max_{k=1, \dots, 4} \text{Grade}_k(x_p) \quad (4)$$

$$\text{Grad}_k = |\phi \nabla_k * I_p^r| \quad (5)$$

$$s(x_p, y_p) = \frac{2\sigma_{x_p y_p} + C_3}{\sigma_{x_p}^2 + \sigma_{y_p}^2 + C_3} \quad (6)$$

where $s(x_p, y_p)$ is the structural similarity between patches ($B(x_p)$ and $B(y_p)$) centered at x_p and y_p ; $\sigma_{x_p y_p}$ is the covariance of the two patches; $\sigma_{x_p}^2$ ($\sigma_{y_p}^2$) is the variance of patch ($B(x_p)$ and $B(y_p)$); we set the patch size as 11×11 , and the constant $C_3 = \frac{C_2}{2}$ (the same as in [6]). Combining the edge and structure similarities, we deduce the degradation on primary visual information as

$$\mathfrak{V}(x_p, y_p) = g(x_p, y_p) s(x_p, y_p) \quad (7)$$

D. Overall Perceptual Quality

Distortions on the two portions codetermine the quality of the contaminated image. The distortion on the disorderly portion degrades image quality by disturbing our attention and arousing uncomfortable sensation. On the other hand, the distortion on the predicted portion changes the original visual content and affects image understanding. Therefore, we combine the evaluation of the two portions, (2) and (7), to acquire the perceptual quality score

$$Q = P^\alpha V^\beta \quad (8)$$

where V is the pooling value of the predicted portion (mean value of all $V(x_p, y_p)$); the parameters α and β are used to adjust the relative importance of the two portions. The weights of the two evaluation parts, P and V , are closely related to the noise energy level on the two decomposed portions. The more noise energy that one decomposed portion possesses, the more important role it will play. For example, if most of the noise is in the disorderly portion, the noise mainly arouses uncomfortable sensation and the uncomfortable sensation variation is dominant in the quality assessment. Thus a big value of α is required in (8) to highlight the evaluation result of the disorderly portion (P). On the contrary, when the noise is mainly in the predicted portion, the quality degradation is primarily caused by the change of the primary visual information. A big value of β is needed to highlight the evaluation result of the predicted portion (V). According to the analysis above, we compute the importance

parameter based on the noise energies of the two portions, and we set

$$\alpha = \frac{MSE_d}{MSE_d + MSE_p} \quad (9)$$

where MSE_d is the energy of noise between the disorderly portions of the reference image (I_d^r) and the test image (I_d^t); MSE_p is the energy of noise between the two predicted images (I_p^r and I_p^t), and $\alpha \in [0, 1]$. Meanwhile, as same as (9), we set $\beta = MSE_p / (MSE_d + MSE_p) = 1 - \alpha$. Moreover, considering the viewing conditions [35] (i.e., the viewing distance and the display resolution), multiscale evaluation is adopted to deduce the overall quality score,

$$S_0 = \prod_{i=1}^5 Q_i^{\rho_i} \quad (10)$$

where Q_i is the perceptual quality score on the i^{th} level based on (8), the parameter ρ defines the relative importance of different scales, and its value is set as $\rho = [0.0448, 0.2856, 0.3001, 0.2363, 0.1333]$ [35], which is obtained through psychophysical experiment.

SIMULATION RESULTS

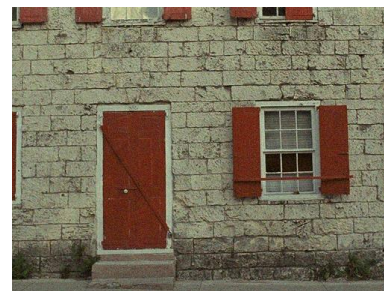


Fig: noisy image



Fig : Original Image

igm_value = 0.9787

CONCLUSION

This paper introduces a novel IQA metric by integrating the best existing IQA metrics. SSIM and GSIM perform well on content-dependent distortions but not well enough on content-independent distortions. However PSNR/MSE performs the opposite way. So we integrate the merits of these metrics by decomposing the input scene into predicted and disorderly portions, and distortions on these two portions are

REFERENCES

- [1] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it?" *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.
- [2] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electron. Lett.*, vol. 44, no. 13, pp. 800–801, Jun. 2008.
- [3] 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," *Adv. Modern Radioelectron.*, vol. 10, pp. 30–45, May 2009.
- [4] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image

discriminatively treated. The decomposition is inspired by the recent IGM theory which indicates that the HVS works with an internal inference system for sensory information perception and understanding, i.e., the IGM actively predicts the sensory information and tries to avoid the residual uncertainty/disorder. Since the predicted portion holds the primary visual information and the disorderly portion consists of uncertainty, the distortions on the two portions cause different aspects of quality degradations. Distortions on the predicted portion will affect the understanding of the visual content, and that on disorderly portion mainly arouse uncomfortable sensation. Considering the different properties of the two decomposed portions, we separately evaluate their quality degradations. Firstly, a Bayesian prediction model is adopted to decompose the reference and test images into predicted and disorderly portions, respectively.

quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.

- [5] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A. C. Bovik, "Image quality assessment based on a degradation model," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 636–650, Apr. 2000.
- [6] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [7] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans.*

Image Process., vol. 15, no. 2, pp. 430–444, Feb. 2006.

[8] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin, “On between-coefficient contrast masking of DCT basis functions,” in Proc. 3rd Int. Workshop Video

Process. Quality Metrics Consumer Electron., Jan. 2007, pp. 1–10.

[9] D. M. Chandler and S. S. Hemami, “VSNR: A wavelet-based visual signal-to-noise ratio for natural images,” IEEE Trans. Image Process., vol. 16, no. 9, pp. 2284–2298, Sep. 2007.



D. RIHANA

BHANU received the B.tech degree in Electronics and Communication Engineering from INDIRA PRIYADARSHINI College of engineering and technology Affiliated to Jawaharlal Nehru Technological University (ANANTAPUR) in 2012 and She is currently pursuing M.tech on Digital electronics and Communication Systems (DECS) at K V Subba Reddy college for women Her research Interests include Digital Image processing.



TRUPTHI GRUTTI

GAGADHALE Received the B.tech degree in ECE from Amaravathi University in 2001 and obtained the M.tech too from Amaravathi University. Presently she is doing her Ph.D from Sri Venkateswara University and working as Associate professor in K V Subba Reddy College for women. Her research Interests include Digital Image processing