

ADAPTIVE HIGH-FREQUENCY CLIPPING FOR IMPROVED IMAGE QUALITY ASSESSMENT

Ke Gu, Guangtao Zhai, Min Liu, Qi Xu, Xiaokang Yang, Jun Zhou, and Wenjun Zhang

Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, China
Shanghai Key Laboratory of Digital Media Processing and Transmissions

ABSTRACT

It is widely known that the human visual system (HVS) applies multi-resolution analysis to the scenes we see. In fact, many of the best image quality metrics, e.g. MS-SSIM and IW-PSNR/SSIM are based on multi-scale models. However, in existing multi-scale type of image quality assessment (IQA) methods, the resolution levels are fixed. In this paper, we examine the problem of selecting optimal levels in the multi-resolution analysis to preprocess the image for perceptual quality assessment. According to the contrast sensitivity function (CSF) of the HVS, the sampling of visual information by the human eyes approximates a low-pass process. For images, the amount of information we can extract depends on the size of the image (or the object(s) inside) as well as the viewing distance. Therefore, we proposed a wavelet transform based adaptive high-frequency clipping (AHC) model to approximate the effective visual information that enters the HVS. After the high-frequency clipping, rather than processing separately on each level, we transform the filtered images back to their original resolutions for quality assessment. Extensive experimental results show that on various databases (LIVE, IVC, and Toyama-MICT), performance of existing image quality algorithms (PSNR and SSIM) can be substantially improved by applying the metrics to those AHC model processed images.

Index Terms— Image quality assessment (IQA), image size, viewing distance, scale transform, high-frequency clipping

1. INTRODUCTION

Image quality assessment (IQA) is an important research area due to its possible application in the design and optimization of various image processing algorithms. Generally, IQA refers to both subjective assessment and objective assessment. The subjective assessment aims to obtain mean opinion scores (MOSs) from subjective viewing test. Although the subjective assessment is always known as the ultimate image quality gauge, it is often too complicated and expensive for practical applications. Last decades have witnessed the rise of a large number of objective image quality metrics (IQMs),

among which PSNR and SSIM [1] are perhaps the most popular methods that have been integrated into many practical image processing systems. In the current IQA research, the so called multi-scale type of methods, e.g. MS-SSIM [2], IFC [3], VIF [4], IW-PSNR/SSIM [5] and MIS-SSIM [6], as inspired by the multi-resolution property of the HVS, achieved some of the best performance.

In the traditional study of IQA, the influence of subjective viewing environments in different databases on the prediction accuracies of objective IQA methods has been largely overlooked. Lin *et al* pointed out in [7] that objective image quality metrics should take into account some significant external factors, such as ambient illumination, display resolution, and viewing distance. However, as a matter of fact, existing image databases are quite diverse in terms of image sizes and viewing distances used in the subjective experiments [7]. In ITU-R BT.500 [8], although the ratio between viewing distance and image height is suggested, there is no recommendation of appropriate image size, or resolution (DPI).

Realizing the impact of viewing distance and image size on perceptual quality assessment, in our former work, we proposed a self-adaptive scale transform (SAST) model [9] for IQA metrics. The SAST model estimates the best scale parameter from the image size and viewing distance. And the images are then resized (by low-pass filtering and sampling) accordingly for better IQA performance. This approach is effective because as the viewing distance increases, the viewing angle shrinks and less image details can be noticed. On the other hand, it is obvious that by resizing the images, we essentially discard part of the high resolution information. Therefore, in this paper, we try to directly remove part of image details using adaptive high-frequency clipping in wavelet subbands. The AHC model filtered subband coefficients are then synthesized back to an image at its original resolution to be used by IQA methods.

The rest of this paper is organized as follows. In Section 2, we first review some related researches and then propose the AHC model. In Section 3, the AHC model based PSNR/SSIM are compared to some competitive IQA algorithms on the LIVE database [10], IVC database [11], and Toyama-MICT database [12]. Finally, Section 4 concludes this paper.

2. THE PROPOSED AHC MODEL

It is suggested in [7] that the external factors, including image size and viewing distance, can have considerable impacts on the prediction performance of IQA algorithms. A simple, empirical method [7] has been exploited for SSIM to determine the downsampling scale S for evaluating images viewed from a typical distance:

$$S_I = \max(1, \text{round}(H/256)) \quad (1)$$

with H being the image height.

As mentioned, human perception of images details is limited by both image size and viewing distance. The SAST model [9] was recently proposed using the concept of human visual angle and angle of gaze. The scale transform coefficient is defined as follows:

$$S_S = \sqrt{\frac{1}{4 \tan(\frac{\theta_H}{2}) \cdot \tan(\frac{\theta_W}{2})} \cdot (\frac{H}{d})^2 \cdot \frac{W}{H}} \quad (2)$$

where d is the viewing distance between the viewers and the image X , and W is the image width. θ_H and θ_W represent the human visual angles for horizontal and vertical directions, respectively.

Accordingly, we can compute the scale transformed image from X by low-pass filtering and sampling. We denote the results as X'_Z with Z ($= I$ or S) depending on the scale parameter S_I or S_S .

When the ratio between image height and viewing distance (H/d) is small, image details cannot be observed. This research aims to find the correspondence among the viewing distance, image size, and the visible spatial frequency. To that end, we first apply wavelet transform to decompose the image X , as exemplified in Fig. 1 on the “lighthouse” image. A weighting function is then introduced to assign different weights to all the LH, HL and HH subbands:

$$w(i, d, l) = \frac{b \cdot k^{t(L-l)}}{a^{(\frac{d}{d_0})}} \quad (3)$$

where L indicates the decomposition layer, and we set $L = 4$ in this paper. Namely, the number of total subbands is 12 excluding the LL subband. The coefficient i ($= [\text{LH}, \text{HL}, \text{HH}]$) represents the subband being processed. The model parameters a ($= 10$), k ($= 10$), t ($= 2$), and d_0 ($= 512$) are empirically assigned. The coefficient b favors the retaining of LH and HL subbands over the HH subband on the same level:

$$b = \begin{cases} 2 & \text{if } i \in \text{LH or HL} \\ 1 & \text{if } i \in \text{HH} \end{cases} \quad (4)$$

Then, each computed weight is compared with a threshold ($thr = 1$ in this work). If the weight is smaller than the threshold, the corresponding subband is clipped out. Wavelet reconstruction is used to get final result X'_A . Note that although

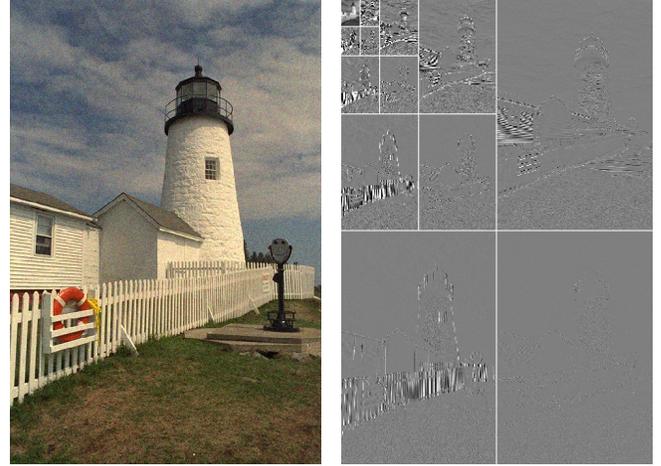


Fig. 1. Illustration of the “lighthouse” image and its wavelet decomposition with four levels.

multi-resolution analysis is used inside the AHC model, the final output is a single scale image with part of high-frequency details properly erased.

Eventually, three pairs of improved $PSNR_Z$ and $SSIM_Z$ ($Z = I, S$, and A) are computed as:

$$PSNR_Z = PSNR(X'_Z, Y'_Z) \quad (5)$$

$$SSIM_Z = SSIM(X'_Z, Y'_Z). \quad (6)$$

where X'_Z and Y'_Z are the preprocessed original and distorted images, respectively.

3. EXPERIMENTAL RESULTS

The experiments are conducted on LIVE [10], IVC [11] and Toyama-MICT [12] databases, because quite different viewing distances and image sizes were used in the subjective experiments, as highlighted in Table 1.

We adopt the four-parameter logistic function suggested by VQEG [13] as the nonlinear regression between the sub-

Table 1. Description of LIVE, IVC, and Toyama-MICT databases.

Database name	Image size (W×H)	d / H	Number
LIVE	768×512	3~3.75	779
	480×720		
	640×512		
	632×505		
	634×505		
	618×453		
	610×488		
	627×482		
IVC	512×512	4	185
Toyama-MICT	768×512	6	168

jective scores and the prediction scores of ten metrics PSNR, PSNR_I, PSNR_S [9], IW-PSNR [5], PSNR_A, SSIM, SSIM_I, SSIM_S [9], IW-SSIM [5] and SSIM_A:

$$q(x) = \frac{\alpha_1 - \alpha_2}{1 + e^{\left(-\frac{x - \alpha_3}{\alpha_4}\right)}} + \alpha_2 \quad (7)$$

where x indicates the input score, $q(x)$ is the mapped score, and α_1 to α_4 are free parameters to be determined during the curve fitting process.

Three commonly used performance metrics, Pearson Linear Correlation Coefficient (PLCC), Spearman Rank-Order Correlation Coefficient (SROCC), and RMSE as suggested by VQEG [13], are employed to further evaluate the proposed AHC model based PSNR/SSIM metric and the other eight IQA methods on LIVE, IVC, and Toyama-MICT databases. Their performance values and directly average results are tabulated in Table 2, and all the corresponding scatter plots are displayed in Fig. 2. It can be found from those results that the proposed AHC model can significantly improve the performance of PSNR and SSIM. The improved metrics PSNR_A and SSIM_A have prediction accuracy comparable to the most sophisticated metrics of IW-PSNR and IW-SSIM [5].

4. CONCLUSION

In this paper, we propose a new adaptive high-frequency clipping (AHC) model to improve the performance of IQA metrics. The AHC model is designed within the framework of wavelet transform to remove part of the high-frequency details of the image that is believed to be indiscernible to the audiences at the viewing distance. Experimental results on LIVE, IVC, and Toyama-MICT databases are provided to demonstrate that the proposed AHC model can improve the performance of PSNR and SSIM to the level of most advanced image quality metrics, such as IW-PSNR and IW-SSIM in the literature.

In the very near future, our work will be devoted to a higher-performance scale transform model for image quality assessment by properly integrating the proposed wavelet-domain based AHC method with our early spatial-domain based SAST model.

Acknowledgment

This work was supported in part by NSERC, NSFC (61025005, 60932006, 61001145), 2013BAH54F04, 2011AA01A107, S-RFDP (20090073110022), postdoctoral foundation of China 20100480603, 201104276, postdoctoral foundation of Shanghai 11R21414200, the 111 Project (B07022), and STCSM (12DZ2272600).

Table 2. PLCC, SROCC, and RMSE values (after nonlinear regression) of PSNR, PSNR_I, PSNR_S, IW-PSNR, PSNR_A, SSIM, SSIM_I, SSIM_S, IW-SSIM, and SSIM_A on LIVE, IVC, and Toyama-MICT databases, and their directly average results.

LIVE database [10]					
<i>Metrics</i>	PSNR	PSNR _I	PSNR _S	IW-PSNR	PSNR _A
PLCC	0.8701	0.9031	0.9137	0.9329	0.9295
SROCC	0.8756	0.9056	0.9164	0.9328	0.9314
RMSE	13.468	11.735	11.104	9.8394	10.077
<i>Metrics</i>	SSIM	SSIM _I	SSIM _S	IW-SSIM	SSIM _A
PLCC	0.9014	0.9251	0.9306	0.9425	0.9321
SROCC	0.9104	0.9355	0.9446	0.9567	0.9477
RMSE	11.832	10.376	10.002	9.1317	9.8968

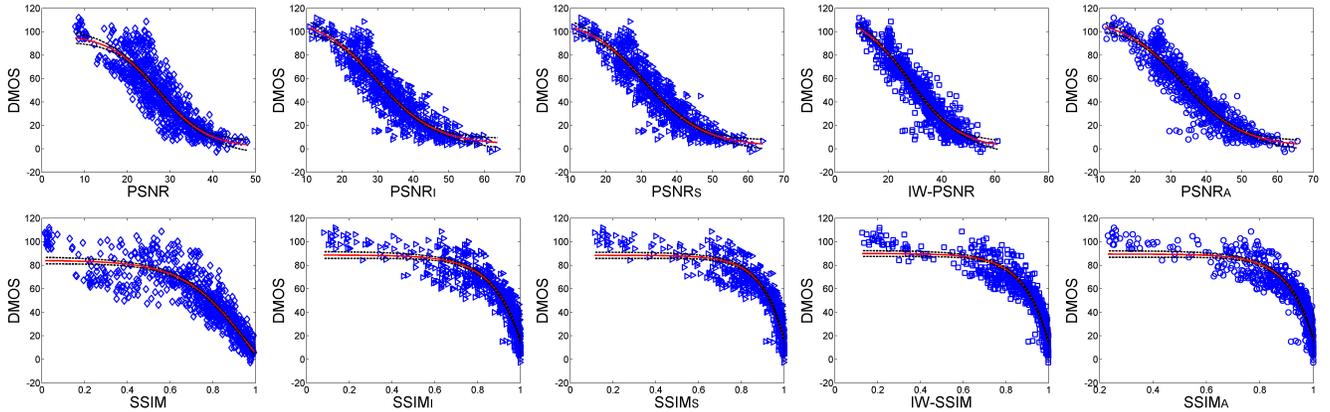
IVC database [11]					
<i>Metrics</i>	PSNR	PSNR _I	PSNR _S	IW-PSNR	PSNR _A
PLCC	0.7195	0.8912	0.8956	0.9055	0.9107
SROCC	0.6887	0.8828	0.8893	0.8999	0.9019
RMSE	0.8462	0.5527	0.5419	0.5170	0.5032
<i>Metrics</i>	SSIM	SSIM _I	SSIM _S	IW-SSIM	SSIM _A
PLCC	0.7923	0.9122	0.9046	0.9228	0.9066
SROCC	0.7785	0.9030	0.8912	0.9125	0.8957
RMSE	0.7433	0.4993	0.5195	0.4693	0.5142

Toyama-MICT database [12]					
<i>Metrics</i>	PSNR	PSNR _I	PSNR _S	IW-PSNR	PSNR _A
PLCC	0.6352	0.8003	0.8355	0.8501	0.8649
SROCC	0.6130	0.7942	0.8276	0.8475	0.8619
RMSE	0.9665	0.7504	0.6876	0.6590	0.6286
<i>Metrics</i>	SSIM	SSIM _I	SSIM _S	IW-SSIM	SSIM _A
PLCC	0.7962	0.8917	0.9079	0.9243	0.9142
SROCC	0.7865	0.8844	0.9042	0.9202	0.9117
RMSE	0.7571	0.5664	0.5247	0.4775	0.5071

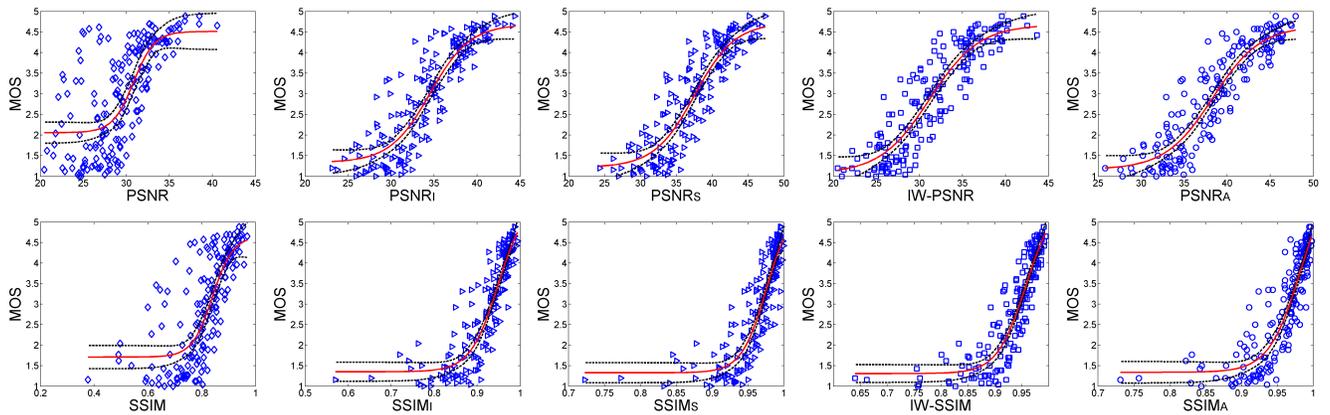
Directly average results					
<i>Metrics</i>	PSNR	PSNR _I	PSNR _S	IW-PSNR	PSNR _A
PLCC	0.7416	0.8648	0.8816	0.8962	0.9017
SROCC	0.7257	0.8609	0.8778	0.8934	0.8984
RMSE	5.0937	4.3460	4.1111	3.6718	3.7361
<i>Metrics</i>	SSIM	SSIM _I	SSIM _S	IW-SSIM	SSIM _A
PLCC	0.8300	0.9097	0.9143	0.9299	0.9176
SROCC	0.8251	0.9076	0.9133	0.9298	0.9184
RMSE	4.4443	3.8139	3.6820	3.3595	3.6394

5. REFERENCES

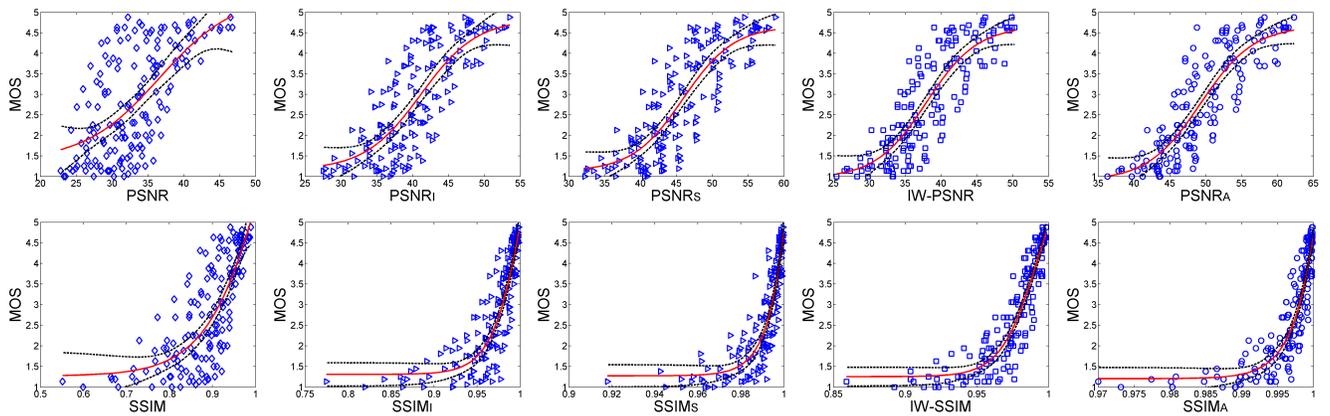
- [1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, April 2004.
- [2] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," *IEEE*



(a) LIVE database



(b) IVC database



(b) Toyama-MICT database

Fig. 2. Scatter plots of DMOS/MOS vs. PSNR, PSNR_i, PSNR_s, IW-PSNR, PSNR_A, SSIM, SSIM_i, SSIM_s, IW-SSIM and SSIM_A on LIVE, IVC, and Toyama-MICT databases. The (red) lines are curves fitted with the logistic function and the (black) dash lines are 95% confidence intervals.

Asilomar Conference Signals, Systems and Computers,
November 2003.

- [3] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117-2128, 2005.
- [4] 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.
- [5] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185-1198, 2011.
- [6] M. Liu, G. Zhai, K. Gu, Q. Xu, X. Yang, X. Sun, W. Chen, and Y. Zuo, "A new image quality metric based on mix-scale transform," *Proc. IEEE Workshop on Signal Processing Systems*, 2013.
- [7] W. Lin and C.-C. Jay Kuo, "Perceptual visual quality metrics: A survey," *J. Vis. Commun. Image R.*, vol. 22, no. 4, pp. 297-312, 2011.
- [8] ITU, "Methodology for the subjective assessment of the quality of television pictures," *Recommendation, International Telecommunication Union/ITU Radiocommunication Sector*, 2009.
- [9] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Self-adaptive scale transform for IQA metric", *Proc. IEEE International Symposium on Circuits and Systems*, 2013.
- [10] 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>.
- [11] A. Ninassi, P. Le Callet, and F. Atrousseau, "Subjective quality assessment-IVC database," [Online]. Available: <http://www2.irccyn.ec-nantes.fr/ivcdb>
- [12] Y. Horita, K. Shibata, Y. Kawayoke, and Z. M. P. Sazzad, "MICT image quality evaluation database," [Online]. Available: <http://mict.eng.u-toyama.ac.jp/mict/index2.html>
- [13] 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/>.