LoG acts as a good feature in the task of image quality assessment

Xuanqin Mou[1], Wufeng Xue[1], Congmin Chen[1], and Lei Zhang[2]
[1]Institute of Image Processing and Pattern Recognition, Xi'an Jiaotong University, China
[2]Department of Computing, The Hong Kong Polytechnic University, Hong Kong
xqmou@mail.xjtu.edu.cn

ABSTRACT

In the previous work, the LoG (Laplacian of Gaussian) signal that is the earliest stage output of human visual neural system was suggested to be useful in image quality assessment (IQA) model design. This work considered that LoG signal carried crucial structural information of IQA in the position of its zero-crossing and proposed a Non-shift Edge (NSE) based IQA model. In this study, we focus on another aspect of the properties of the LoG signal, i.e., LoG whitens the power spectrum of natural images. Here our interest is that: when exposed to unnatural images, specifically distorted images, how does the HVS whitening this type of signals? In this paper, we first investigate the whitening filter for natural image and distorted image respectively, and then suggest that the LoG is also a whitening filter for distorted images to some extent. Based on this fact, we deploy the LOG signal in the task of IQA model design by applying two very simple distance metrics, i.e., the MSE (mean square error) and the correlation. The proposed models are analyzed according to the evaluation performance on three subjective databases. The experimental results validate the usability of the LoG signal in IQA model design and that the proposed models stay in the state-of-the-art IQA models.

Keywords: Laplacian of Gaussian, whitening, Principal Component Analysis (PCA), image quality

1. INTRODUCTION

Image quality assessment (IQA) models mimic functionally the ability of human vision system (HVS) to perceive image quality. They have been widely used for various image related applications including communication, restoration, processing, compression, etc. In IQA model design, some known properties of HVS, such as the contrast sensitivity functions (CSF), the masking effect and the just noticeable difference (JND), have been introduced. However most of them describe the properties of entire HVS, while it should be more interesting if we can relate the property of some specific visual neurons with the IQA model design. In our previous work [8, 9], we proposed that IQA models can be designed by using the earliest stage output of visual neural system, i.e., the ganglion cells in the retina and the neurons in the lateral geniculate nucleus (LGN), which have been functionally modeled as a Laplacian of Gaussian (LoG) function. In the IQA model design, the LoG signal was considered to carry crucial structural information of the image on the position of its zero-crossing according to Marr's theory, and thus the Non-shift Edge (NSE) based FR IQA model was developed.

In the present study, we explore another aspect of the properties of LoG that it acts as a whitening filter for natural images and whether it does also the same for distorted images. Researchers have validated that the receptive field of the ganglion cells, which resembles LoG, is a whitening filter for natural images which remove the first order and the second order statistical redundancy in natural images [15, 17]. Known that HVS is insensitive to the first order (intensity) and the second order (contrast) statistical information, and the structural information is carried by the high-order statistics of the image, LoG would be a good feature to represent image structural information. Here an interesting question is that: Does HVS process the distorted image with the same mechanism as to process the pristine natural image? More specifically, given the LoG filter can whiten the power spectrum of the natural images, will it still be appropriate to whiten the distorted images in the same way? In this study, we initiate an experiment to apply principle component analysis (PCA) and PCA based whitening [1] on distorted images to explore what property of the PCA based whitening filter is for the distorted images. The preliminary experimental results show that the obtained filter is quite analogous to a

Digital Photography X, edited by Nitin Sampat, Radka Tezaur, et. al. Proc. of SPIE-IS&T
Electronic Imaging, SPIE Vol. 9023, 902313 · © 2014 SPIE-IS&T
CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2038982

LoG function. This observation suggests that LoG filtering can remove the first-order and the second-order statistical redundancy for both the natural image and the distorted image. We thus further propose that LoG is a good feature in the task of IQA model design. To validate that, we apply two simple distance metrics, i.e., the MSE (mean square error) and the correlation, to measure the error between the LoG signals of the natural images and their distorted counterparts, hence two full-reference (FR) IQA models, i.e., LoG-MSE and LoG-COR, are proposed. Experimental results show that the proposed models work with high performance and stay in the state-of-the-art IQA models.

2. WHITENINGIMAGES

2.1 PCA Whitening

Principle component analysis (PCA) was used in the investigations of whitening natural images [1, 15]. Here we use the same technique for distorted images. Let **x** denote the data vector, which is formed by rearranging the image patch into a column wise manner. We want to find the filter that could whiten the frequency spectrum of **x**. We denote by **U** the matrix with the vectors defining the principles components as its columns, denote by **y** the coefficients of the principal components of **x**, denote by **z** the whitened version of **x** with zero phase, and denote by Λ a diagonal matrix with the inverse of the square root of the variance of **y** on its diagonal:

$$\Lambda = \begin{bmatrix} \frac{1}{\sqrt{\mathrm{var}(y_1)}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{\mathrm{var}(y_2)}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{\mathrm{var}(y_K)}} \end{bmatrix} \tag{1}$$

Then the zero-phase whitening equation is as follows:

$$\mathbf{y}=\mathbf{U}^{\mathrm{T}}\mathbf{x} \tag{2}$$

$$\mathbf{z}=\mathbf{U}\Lambda\mathbf{y}=\mathbf{U}\Lambda\mathbf{U}^{\mathrm{T}}\mathbf{x}=\mathbf{W}\mathbf{x}. \tag{3}$$

Three steps constitute the whole story: $\mathbf{U}^{\mathrm{T}}$ transforms the original data x into the principle component space**y**, Λ equalizes the spectrum power of **y**, and finally one more **U** make the representation go back from the principle space to the original coordinate. All of them can be completed in one matrix transformation, which we denote as **W**. The final step makes **W** a zero-phase filter. In the implementation of the process to image data, a low-pass filter is firstly applied to make sure that noise in image is not amplified.

2.2 Whitening filters for natural image and distorted image

Taking advantages of existing databases [4, 5, 6] on IQA which include both pristine images and distorted images, we investigate the results of the above mentioned PCA based whiten procedure. The reference images existed in the three databases are with various content, ranging from indoor scene to outdoor scene, from people to animal and from wildlife to urban building. Besides, the distortion procedures exposed to these images include JPEG compression, JPEG2000 compression, Gaussian blur, and kinds of noise contamination. With all these images, similar results can be obtained, therefore we just show as an example the results of two of them.

When applied with the PCA whitening procedure, for pristine natural images, the center-surround filter **W** is obtained, as shown in the first row of Fig.1. Although the two image's content differ greatly from each other, the resulted whiten filters are both center-surround Laplacian filter. The magnitude difference can be attributed to the content: "parrots" can be clearly segmented into foreground and blurry background while "bikes" contains a rich set of textures and structures. Combined with the pre-filtering for noise suppression, the final whiten filter is actually a Laplacian of Gaussian (LoG).

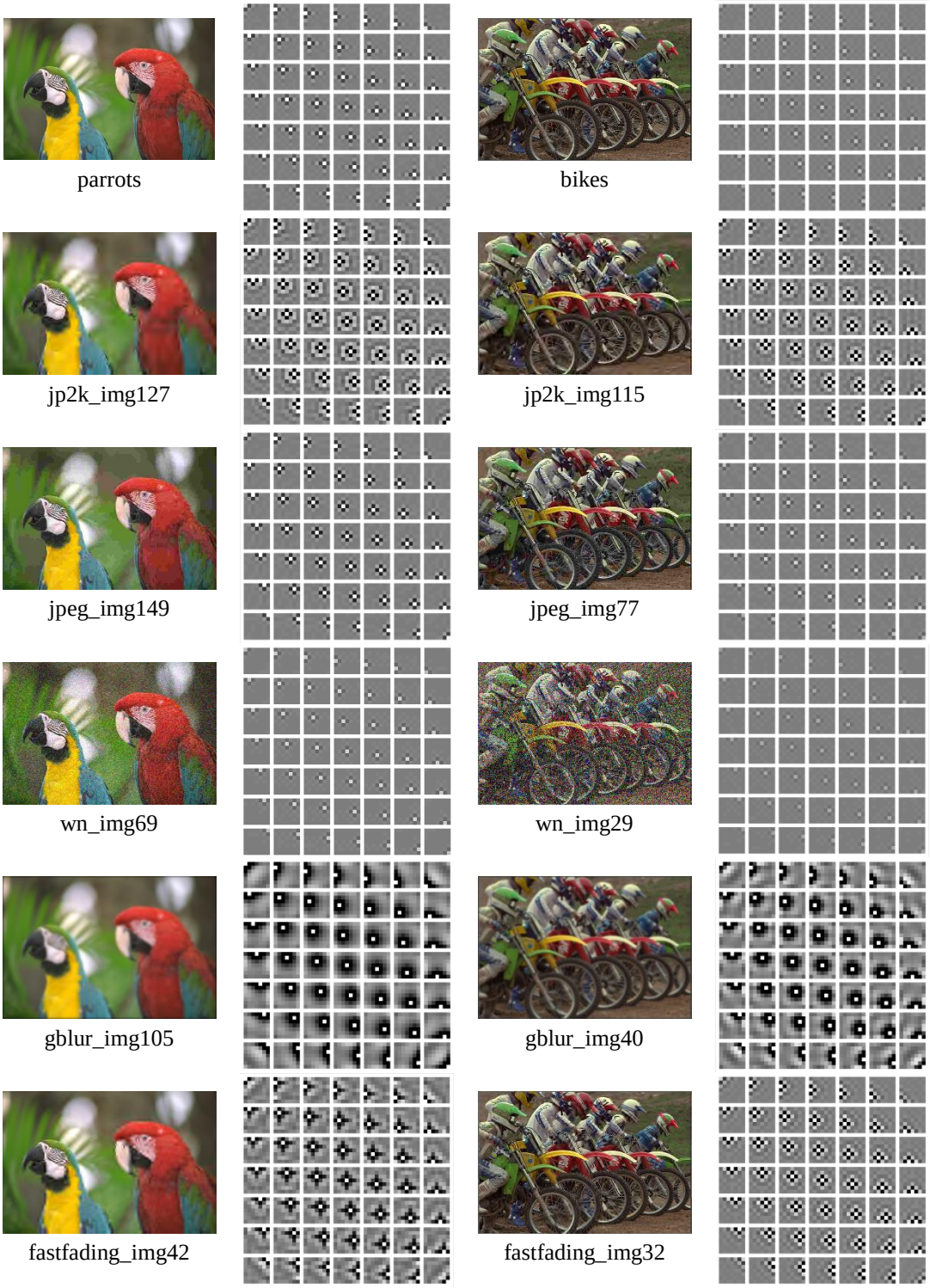

Figure1.The whiten matrix for two reference images and their distorted images. All the images are from LIVE database [4].

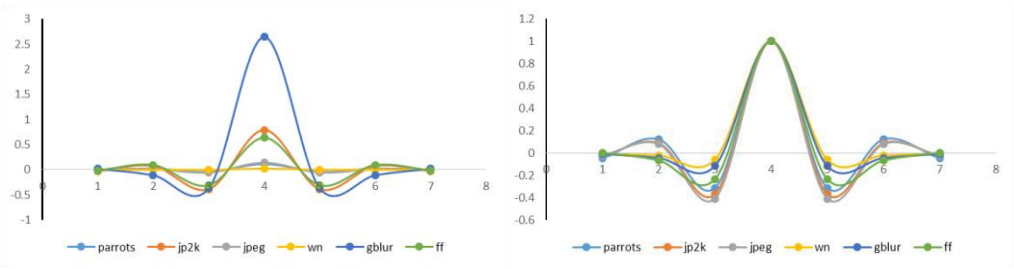

Figure 2. The profile of the PCA-whiten resulted filter for the image parrots and its distorted images. Left: the original profile; Right: the profile after normalized the maximum response to 1.

For image with various types of distortion (from the second row to the fifth row of Fig. 1), the resulted whiten matrix **W** is also shown. For the 5 distortion types, although their whiten matrix are all of center-surround structure, they vary from each other in the magnitude of the contrast of the center-surround structure. This can be observed more clearly from Fig.2. In the left we demonstrate the 1-D profile of the center row in the whiten matrix **W** for the image parrots and its distorted versions. In the right, we show the normalized profile. The normalization is done to ensure that all the profiles have the same maximum response which equals 1. The left shows the difference in magnitude gain while the right shows the common shape. One may also see the discrepancy of the values off the center in the right plots of Fig.2. More complex model is needed to capture this difference across difference types of distortion and we leave this to the future work. In this study we still use the LoG filter to approximately model these plots because the LoG filter used here can generally normalized and decorrelate the distorted images to much extent by our experimental observations. However, when utilized in an IQA model, the disagreement in magnitude will be taken into account.

3. LOG-MSE AND LOG-COR

3.1 IQA models

From previous section, it is shown that the center-surround receptive field can whiten not only the pristine natural images, but also the distorted images. For the reason that LoG signals only contain high order information of image structure and the fact that HVS is not sensitive to the first and second order information in image, the LoG signals have been utilized in IQA model. Therefore it will be effective to investigate the LoG based whiten filters for natural images and for images with artifacts. To assess the image quality with the LoG whitened signals, we propose in this section two different IQA model, Mean squared error and correlation of LoG signals, which we denote as LoG-MSE and LoG-COR, respectively. Let R and D be the reference image and the distorted image, respectively, and l the whiten filter. Then the two models are as follows:

$$LOG-MSE=\frac{1}{N}(l\otimes R-l\otimes D)^2 \tag{4}$$

$$LOG-COR=\frac{1}{N}\frac{2(l\otimes R)(l\otimes D)+c}{(l\otimes R)^2+(l\otimes D)^2+c}. \tag{5}$$

For the whiten filter, we utilize the LoG filter:

$$l(x,y,\sigma)=\frac{1}{2\pi\sigma^2}\frac{x^2+y^2-2\sigma^2}{\sigma^4}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{6}$$

Where x and y denotes the spatial locations and σ is the scale factor of the Gaussian.

Table 1. The performance evaluation of the proposed IQA models and comparison with other state of art IQA models.

	LIVE (779)			CSIQ (866)			TID2008 (1700)			Weighted Average	
	SRC	PCC	RMSE	SRC	PCC	RMSE	SRC	PCC	RMSE	SRC	PCC
PSNR	0.876	0.872	13.36	0.806	0.751	0.173	0.553	0.523	1.144	0.694	0.664
IFC	0.926	0.927	10.26	0.767	0.837	0.144	0.568	0.203	1.314	0.703	0.537
NSER	0.942	0.939	9.362	**0.934**	**0.947**	**0.084**	0.740	0.796	0.813	0.837	0.868
SSIM	0.948	0.945	8.95	0.876	0.861	0.133	0.775	0.773	0.851	0.841	0.836
VIF	**0.964**	**0.960**	**7.61**	0.919	**0.928**	**0.098**	0.749	0.808	0.790	0.844	**0.875**
MAD	0.944	0.939	9.37	0.899	0.820	0.150	0.771	0.748	0.891	0.845	0.811
MS-SSIM	0.952	0.950	**8.56**	0.877	0.659	0.197	0.809	0.801	0.803	0.860	0.798
IW-SSIM	**0.957**	**0.952**	8.35	0.921	0.914	0.106	**0.856**	**0.858**	**0.689**	**0.896**	**0.895**
FSIM	**0.963**	**0.960**	**7.67**	**0.924**	0.912	0.108	**0.880**	**0.874**	**0.653**	**0.911**	**0.904**
LoG-MSE	0.927	0.913	11.136	0.876	0.646	0.200	0.757	0.737	0.907	0.827	0.754
LoG-COR	0.934	0.928	10.214	**0.922**	**0.916**	**0.105**	**0.835**	**0.826**	**0.756**	**0.881**	0.873

Table II: Performance comparison of the IQA models on each individual distortion type in terms of SRC.

		PSNR	*IFC*	*SSIM*	*VIF*	*MS-SSIM*	*MAD*	*IW-SSIM*	*FSIM*	*LoG-MSE*	*LoG-COR*
LIVE database	*JP2K*	0.895	0.911	0.961	0.970	0.963	0.968	0.965	0.971	0.957	**0.966**
	JPEG	0.881	0.947	0.976	0.985	0.981	0.976	0.981	0.983	0.965	0.975
	AWN	0.985	0.938	0.969	0.986	0.973	0.984	0.967	0.965	**0.984**	**0.989**
	GB	0.782	0.958	0.952	0.973	0.954	0.946	0.972	0.971	0.935	**0.961**
	FF	0.891	0.963	0.956	0.965	0.947	0.957	0.944	0.950	0.845	0.849
CSIQ database	*AWN*	0.936	0.843	0.897	0.957	0.951	0.954	0.938	0.926	**0.930**	**0.967**
	JPEG	0.888	0.941	0.954	0.970	0.947	0.961	0.966	0.966	0.952	**0.956**
	JP2K	0.936	0.925	0.960	0.967	0.963	0.975	0.968	0.968	0.965	0.957
	PGN	0.934	0.826	0.892	0.951	0.968	0.957	0.906	0.923	**0.934**	**0.961**
	GB	0.929	0.953	0.961	0.974	0.933	0.968	0.978	0.972	0.958	**0.963**
	CTD	0.862	0.487	0.793	0.934	0.971	0.921	0.954	0.942	**0.894**	**0.930**
TID2008 database	*AWN*	0.907	0.581	0.811	0.880	0.953	0.839	0.787	0.857	**0.917**	**0.906**
	ANMC	0.899	0.546	0.803	0.876	0.913	0.826	0.792	0.851	**0.911**	**0.892**
	SCN	0.917	0.596	0.815	0.870	0.809	0.868	0.771	0.848	**0.917**	**0.927**
	MN	0.852	0.673	0.779	0.868	0.805	0.734	0.809	0.802	0.734	**0.782**
	HFN	0.927	0.732	0.873	0.907	0.821	0.886	0.866	0.909	**0.951**	**0.915**
	IMN	0.872	0.534	0.673	0.833	0.811	0.065	0.646	0.746	**0.878**	**0.815**
	QN	0.870	0.586	0.853	0.797	0.869	0.816	0.818	0.855	**0.902**	**0.897**
	GB	0.870	0.856	0.954	0.954	0.691	0.920	0.964	0.947	0.944	**0.960**
	DEN	0.942	0.797	0.953	0.916	0.859	0.943	0.947	0.960	**0.964**	**0.969**
	JPEG	0.872	0.818	0.925	0.917	0.956	0.927	0.918	0.928	**0.958**	**0.931**
	JP2K	0.813	0.944	0.962	0.971	0.958	0.971	0.974	0.977	**0.975**	**0.970**
	JGTE	0.752	0.791	0.868	0.859	0.932	0.866	0.859	0.871	0.779	0.814
	J2TE	0.831	0.730	0.858	0.850	0.970	0.839	0.820	0.854	0.852	0.829
	NEPN	0.581	0.842	0.711	0.762	0.868	0.829	0.772	0.749	**0.715**	**0.714**
	Block	0.619	0.677	0.846	0.832	0.861	0.797	0.762	0.849	0.795	**0.867**
	MS	0.696	0.425	0.723	0.510	0.738	0.516	0.707	0.669	**0.733**	**0.746**
	CTC	0.586	0.171	0.525	0.819	0.755	0.272	0.630	0.648	**0.565**	**0.646**

The first model, i.e., LOG-MSE directly uses the MSE of the LOG whitened images as the objective distortion prediction. Because much redundancy has been removed in the whitened signal, we expect that the LOG-MSE can reveal a huge

improvement upon MSE. LOG-COR computes the whitened structure similarity, i.e. high order statistical correlation between the two counterpart images. It's worthwhile to note that LOG-COR is actually a LOG-MSE normalized by the energy summation of the two LOG signals.

$$LOG-COR = 1 - \frac{(l \otimes R - l \otimes D)^2}{(l \otimes R)^2 + (l \otimes D)^2 + c} \tag{7}$$

3.2 IQA performance

To validate the performance of the two IQA models, three subjective image quality databases are employed: LIVE [4], CSIQ [5] and TID2008 [6]. Each image in the database is assigned with a quality/distortion score, which is obtained from an average observer. To evaluate the IQA models, a nonlinear regression [12] is firstly fitted to them and then three indices are computed between the subjective score and the model's score: the Spearman rank order correlation coefficient (SRC), the Pearson correlation coefficient (PCC) and the root mean squared error (RMSE). The results of the two models for the three contexts are shown in Table 1.We also compare them with PSNR and other state of art models, including IFC [13], VIF [14], MAD [5], NSER [8], MS-SSIM [7], SSIM [3], IW-SSIM [16], and FSIM [2]. It has to be noticed that IFC, VIF, MS-SSIM, NSER, and FSIM make use of the multi-scale information, and IW-SSIM and FSIM also adopt a weighted average in the final quality estimation, while the proposed two models only make use of the LoG signal in one scale and with no additional weighting information. All the computation is applied to the luminance channel only. The LoG signals are obtained with $\sigma = 3.0$ and $c = 255 \times 0.02$.

From Table 1, we can draw the following: 1) LOG-MSE outperforms the traditional PSNR with a large margin; 2) LOG-COR outperforms LOG-MSE with a large margin; 3) on CSIQ and TID2008 database, LOG-COR ranks third;4) among all the state-of-art models, LOG-COR ranks third in terms of the weighted SRC.

From all the above observations, the effectiveness of the LoG filter is clearly demonstrated. When MSE computation is deployed, the LoG-whitened signal is more effect than the original pixel based representation. The redundancy of 1st order and 2nd order statistical information, which are not sensitive to HVS, are removed and make the error signal more perceptual related. When LoG-COR is deployed in the computation, the LoG signal becomes more effective. Among all the 10 competitors, LoG-COR still performs the third best despite the fact that LoG-COR is only based on the earliest stage of HVS, and the whitening process costs much less computation complexity than the rest models.

From Section 2, we have concluded that although both the reference image and the distorted image can be whitened by the center-surround receptive field, the used filter varies in the magnitude of the contrast of the structure. This means that the whitening process may have respective magnitude gain for difference images. When MSE is utilized, this difference in magnitude gain makes the error signal be computed in an uncommon space. This will either amplify or shrink the error estimation. When the similarity function (Eq. 5) is used, this effect is removed by the fact that the magnitude gain appears simultaneously in the denominator and numerator. The discrepancy in the magnitudes is eliminated, thus better performance is then obtained. From another viewpoint, the correlation computation is equivalent a normalized MSE (Eq.7). The difference of the magnitude gain is reduced by the normalization factor and lead to the performance improvement. From all above, we can conclude that the whitened LoG signal brings more high order information into account, and make it more appropriate for IQA task.

Table 2 gives the performances of the IQA models in terms of SRC value. As mentioned above, we are not aiming to outperform the various sophisticated IQA models such as VIF, IW-SSIM, MAD and FSIM, et al. To show the effective of the LoG signals in the task of IQA, we pay more attention to the performance comparison of the proposed models and the classic structural similarity model SSIM. Shown in bold fonts in the two right-most columns of Table 2 are the distortion types where the LoG based models beats SSIM. Among all the 28 groups of difference distortions, LoG-MSE performs better than SSIM on 16 of them, while LoG-COR on 23 of them. Notice that the MSE-like computation may make LoG-MSE more appropriate for predicting the quality of noise contaminated images, such as Group 3 on LIVE, Groups 1 and 4 on CSIQ and Groups 1-7 on TID2008, LoG-MSE also shows superior SRC values on other types of distortions like DEN, JPEG, JP2K, MS, NEPN, and CTC on TID2008. The overwhelming advantages of LoG-COR for 25 groups of distortions gives evidence that LoG signals are very effective for various artifacts introduced in natural images and that the LoG based whiten filters also applies to images with these different types of distortions.

4. CONCLUSIONS

In this study, we first analysis the PCA based whitening simultaneously for both the pristine natural images and the distorted images. Experimental results show that the whitening filter for the distorted images is also like to center-surrounded LoG except that the magnitude gain varies. Based on this, we then propose two FR IQA models by applying the distance metric, i.e., the MSE and the correlation for the LoG signals. Experimental results show that the proposed models stay in the state-of-the-art IQA models. More specifically, the correlation between the LoG signals is more perceptual relevance than the MSE computation due to the fact that the difference magnitude gain of the whiten procedure is removed by the normalized factor in the correlation computation.

ACKNOWLEDGEMENT

This work is partly supported by Natural Science Foundation of China (No.90920003 and No. 61172163).

REFERENCES

[1] Eero P.Simoncelli, and Bruno A. Olshausen. "Natural image statistics and neural representation." Annual review of neuroscience 24.1, 1193-1216 (2001).
[2] Lin Zhang, Lei Zhang, X. Mou and D. Zhang, "FSIM: A Feature Similarity Index for Image Quality Assessment," IEEE Trans. Image Process. 20(8):2378-2386 (2011).
[3] Z. Wang, A C. Bovik and H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, pp. 600-612, 2004.
[4] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik. (2005) "Live Image Quality Assessment Database Release 2." [Online]. http://live.ece.utexas.edu/research/quality.
[5] Eric C. Larson and Damon M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy", J. Electron. Imaging 19, 011006, Jan 07, (2010).
[6] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, "TID2008 - A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics", Advances of Modern Radio electronics, Vol. 10, pp. 30-45, (2009).
[7] Z. Wang, E. P. Simoncelli, and A. C. Bovik. "Multiscale structural similarity for image quality assessment." Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, Vol. 2. IEEE, (2003).
[8] M. Zhang, X. Mou, and L. Zhang. "Non-Shift Edge Based Ratio (NSER): An Image Quality Assessment Metric Based on Early Vision Features." Signal Processing Letters, IEEE 18.5, 315-318 (2011).
[9] W. Xue, and X. Mou. "An image quality assessment metric based on Non-shift Edge." 18th IEEE International Conference on Image Processing (ICIP), 2011.
[10] G. S. Brindley. "Physiology of the retina and the visual pathway." London: Edward Arnold, (1960).
[11] G.C.DeAngelis, I. Ohzawa, and R.D.Freeman, "Receptive field dynamics in the central visual pathways", Trends Neurosci.,vol. 18,pp. 451-458 (1995).
[12] VQEG, "Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment – Phase II," August 2003, available at http://www.vqeg.org/.
[13] H.R. Sheikh, A.C. Bovik and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," IEEE Transactions on Image Processing, vol.14, no.12pp. 2117- 2128 (2005).
[14] H.R. Sheikh. and A.C. Bovik, "Image information and visual quality," IEEE Transactions on Image Processing, vol.15, no.2,pp. 430- 444 (2006).
[15] J. J. Atick and A. N. Redlich, "What does the retina know about natural scenes?," Neural Computation, vol. 4, no. 2, pp. 196–210 (1992).
[16] Z. Wang, and Q. Li. "Information content weighting for perceptual image quality assessment." IEEE Transactions on Image Processing, 20.5, 1185-1198 (2011).
[17] Daniel J. Graham, D. M. Chandler, and David J. Field. "Can the theory of "whitening" explain the center-surround properties of retinal ganglion cell receptive fields?." Vision research 46, no. 18, 2901-2913 (2006).