REDUCED REFERENCE IMAGE QUALITY ASSESSMENT BASED ON WEIBULL STATISTICS

Wufeng Xue and Xuanqin Mou

Institute of Image Processing and Pattern Recognition, Xi'an Jiaotong University, China
xqmou@mail.xjtu.edu.cn

ABSTRACT

Theories in fragmentation have proved that the statistics of image gradient magnitude followed a Weibull distribution, with β (scale) and γ (shape) as free parameters, which are demonstrated to be strongly correlated with brain response. In this paper, we chose β extracted from the proposed strongest component map (SCM) in scale space, as the reduced reference (RR) feature, and developed a novel method for reduced reference image quality assessment (RRIQA) named βW-SCM. For each scale, the SCM was constructed by assembling coefficients with maximum amplitude among different orientations into a single map. The Weibull parameters were then estimated from the SCM. The final image quality was computed by summing the geometric mean of the defined absolute and relative deviations of β. Performance evaluation on the well-known LIVE database demonstrated an outstanding advantage of low RR feature data rate with nearly the same prediction accuracy and consistency.

Index Terms—Natural image statistics, Weibull distribution, image quality assessment, reduced-reference

1. INTRODUCTION

Visual information or image plays an important role in human's daily life and working environment. However, degradation often occurs during various image processing procedures such as compression and communication. To make the information effective and easy to use, many efforts have been devoted to the degradation estimation or image quality assessment (IQA).

The existing IQA methods can be classified into three categories according to how much information about the perfect reference image is employed: 1) Full Reference (FR) method, 2) Reduced Reference (RR) method and 3) No Reference (NR) method. For the reason of available access to the full perfect image, many inspiring FR algorithms have been proposed and perform well across a wide range distortion types [1-5]. On the contrary, exploring a generic NR method is hard. Only when dealing with a specific distortion type, there are algorithms that perform adequately but still need improvement [6-8].

As a compromise, the RR method provides a flexibility solution for the condition when the reference image is not fully accessible. Partial information, or the so-called reduced reference (RR) features, are extracted at the sender side and transferred to the receiver side. Then the distortion of the received image is computed from the difference between the RR features of the original and the distorted images. Under the framework described by Wang et al. in [9], the appropriate RR features are desirable to: 1) provide an efficient summary of the reference image; 2) be sensitive to a variety of image distortions; 3) be relevant to the visual perception of image quality. With these criterions, several RRIQA methods have been developed. In [10], grouplet transform was used to efficiently characterize the image features and rational sensitivity thresholding was performed to obtain the sensitivity coefficients. The quality of the test image was estimated by comparing its sensitivity coefficients with the reference ones. Experiments showed high prediction accuracy on JPEG and JPEG2000 compressed images. Local harmonic amplitude information computed from an edge-detected picture was exploited as the reduced reference information to deal with the typical degradations of blackness and blur [11]. A model, called C4 [12], extracted visual features from an image presented in a perceptual space. It computed similarity between the features of the test image and the reference one to predict the distortion and worked well with JPEG and JPEG2000 compressed images on LIVE database. Two general-purpose RRIQA methods were purposed in [13] and [14]. Within these works, the image was represented in a wavelet domain and a divisive normalization domain, respectively. Then the marginal distributions of subbands coefficients of the reference image were summarized by a generalized Gaussian model and Gaussian model respectively. Finally the Kullback-Leibler distance between the distribution of the reference and distorted images was computed as the image quality index. Along with this framework, we aimed to search for a more efficient and biologically depended model and parameters to feature the nature image quality.

In this paper, we proposed a novel RRIQA method inspired by the recently discovery that brain response

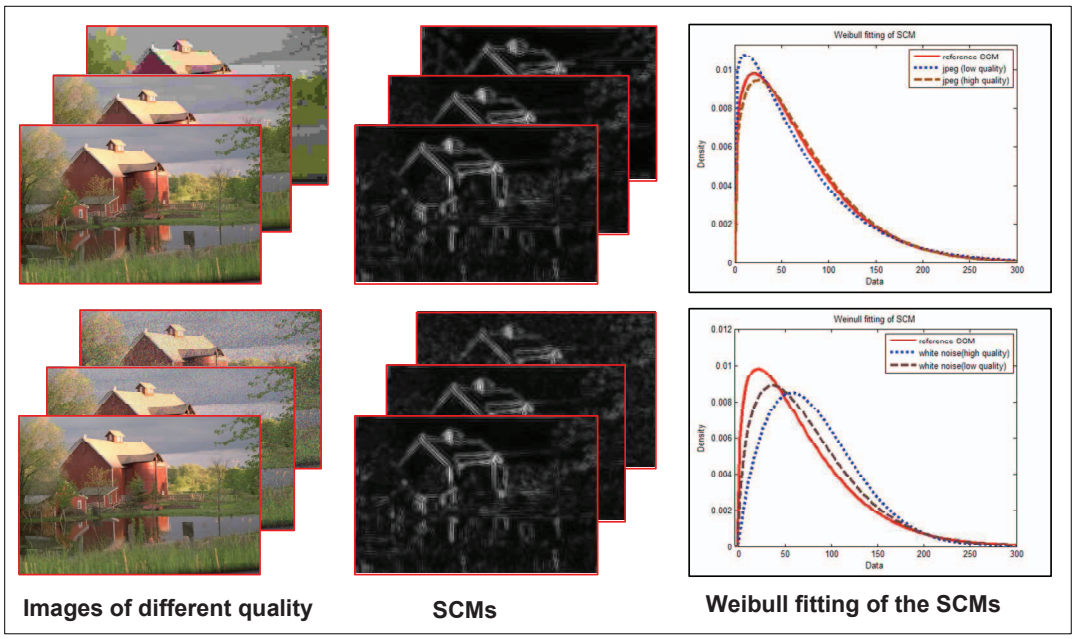

Figure 1. Shift of the Weibull distribution of the SCM for JPEG compressed images (top row) and white noise contaminated ones (bottom row) from the LIVE database. Each group contains the source image and two distorted versions with low quality and high quality. The SCMs are illustrated in the middle column and the fitted Weibull curves of the SCMs in the right column (red solid line for the source image, brown dashed line for the distorted image with high quality and blue dotted line for that with low quality).

strongly correlates with Weibull statistics when processing natural images [15]. We evaluated the new RRIQA model on LIVE database by comparing it with the published models [13, 14].

2. METHODS

2.1. Weibull distribution and brain response

In the long term of evolution, humans are imposed to various surroundings, and may have found the best way to communicate with nature. It's widespread agreement that neural processing must be influenced by the intrinsic laws of natural scene. Previous research provides insight into the natural scene statistics (NSS) [16-18]. Scale space theory imposes that natural image statistics to be of random multiplicative process for correlated intensity fluctuations combined with the additive stochastic laws for uncorrelated fluctuations cause image statistics to follow the laws of sequential fragmentation [19]. Based on these theoretical results, the gradient magnitude is proved to follow a Weibull distribution, with β and γ as scale parameter and shape parameter, respectively [20].

$$f_{wbl}(x;\beta,\gamma)=\frac{\gamma}{\beta}(\frac{x}{\beta})^{\gamma-1}\exp(-(\frac{x}{\beta})^{\gamma}),\ x>0 \qquad (1)$$

As to the aspect of brain response, experiments showed that β and γ of Weibull distribution explains up to 71% variance of the early EEG signal and can be easily estimated from the outputs of X-cells and Y-cells [15]. In addition, the two parameters can accurately describe the spatial coherence, and complexity of an image. These suggested that the brain is strongly tuned to the image's beta and gamma values. In this paper, we employed this property in our RRIQA method.

2.2. The steerable pyramid decomposition

In a wide range of image processing problems, differential algorithms and multi-scale decomposition are mostly used. Among them, the steerable pyramid decomposition is widely used for its steerability and shiftability. Here we give an outline of the steerable pyramid decomposition following [21, 22]. It divides an image into a collection of subbands localized in both scale and orientation. When expressed in the Fourier domain, it is implemented using a set of oriented complex analytic filters that are polar-separable:

$$F_{n,k}(\rho,\theta)=B_n(\rho)G_k(\theta),\ n\in[0,N],k\in[0,K-1] \qquad (2)$$

$$B(\rho)=\begin{cases}\cos(\frac{\pi}{2}\log_2(\frac{2^n\rho}{\pi})), & \frac{2^n\rho}{\pi}\in[\frac{1}{2},2]\\ 0, & \text{otherwise}\end{cases} \qquad (3)$$

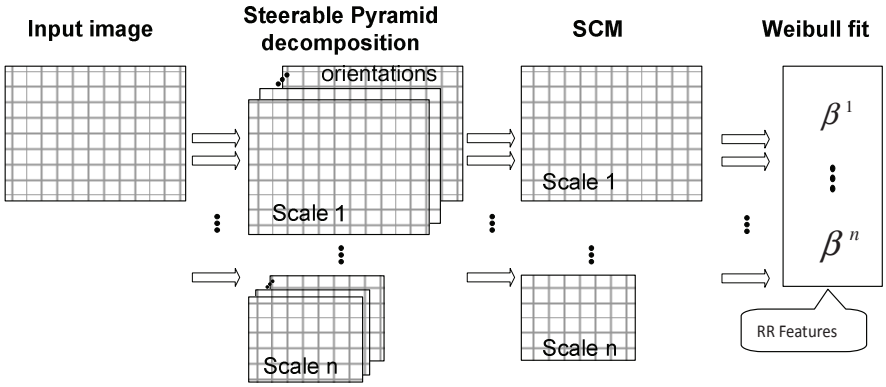

Figure 2 The building blocks of RR feature extraction.

where $G_k(\theta) = \begin{cases} [\cos(\theta - \frac{\pi k}{K})]^{K-1}, & \left|\theta - \frac{\pi k}{K}\right| < \frac{\pi}{2} \\ 0, & \text{otherwise,} \end{cases}$ (4)

where ρ and θ are polar frequency coordinates. Subbands are subsampled by a factor of 2^n along both axes. In addition, the non-oriented high pass and low pass residual bands must be retained in the following way:

$$H(\rho) = \begin{cases} \cos(\frac{\pi}{2}\log_2(\frac{\rho}{\pi})), & \rho \in [\frac{\pi}{2}, \pi] \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

$$L(\rho) = \begin{cases} \cos(\frac{\pi}{2}\log_2(\frac{2^{(N+1)}\rho}{\pi})), & \rho \in [\frac{\pi}{2^{N+1}}, \frac{\pi}{2^N}] \\ 1 & \rho < \frac{\pi}{2^{N+1}} \\ 0 & \rho > \frac{\pi}{2^N}. \end{cases} \quad (6)$$

In space domain, we denoted the subband in scale *n* and orientation *k* as $f_{n,k}$.

2.3. The Strongest Component Map

After the decomposition, details of image are distributed into subbands with corresponding orientation. However, we found that this representation may be too redundant for some applications. For example, a strong edge localized in one orientation may appear as a weaker one in another orientation. For reduction of redundancy, we constructed the strongest component map in each scale. We define it here as SCM. This map was composed of the coefficients with maximum amplitude among all orientations in each separate scale (eq.7). So in each location, only the strongest component was reserved while other weaker ones were suppressed.

$$S_n = \max_{k \in \{0,\ldots K-1\}} (\left|f_{n,k}\right|). \quad (7)$$

where S_n stands for the SCM of scale n.

Indeed, if we decompose the image into orientations as many as possible, the coefficient of the SCM acts as the gradient magnitude and the orientation which it belongs to acts as the gradient direction. In our experiments showed by figure.1, we found that the coefficients of SCM also followed a Weibull distribution, and the probability density curve shifts as the distortion increases, for both JPEG compression and additive white contamination. The distance between the distorted curve and the reference one gives a good indication of the quality degradation. So, in consideration of the multi-scale property of human vision, we adopted the SCM instead of the gradient magnitude.

2.4. RR features extraction

Based on the framework of Wang et al. [9], we extracted RR features as follows (see figure.2):

1) The input image was divided into subbands localized in both orientations and scales using the steerable pyramid decomposition;
2) The SCM was then assembled for each scale from the coefficients with the maximum amplitude across different orientations;
3) β and γ in all scales were estimated by fitting the Weibull function to the coefficients of the SCM with maximum likelihood estimation (eq.8), and β was chosen as the RR feature. The likelihood function was showed here:

$$L(\beta_n, \gamma_n \mid S_n) = \prod_{s_n(i,j) \in S_n} f_{wbl}(s_n(i,j); \beta_n, \gamma_n), \quad (8)$$

In the following sections, β_r^n and β_d^n denoted the beta value of the *n*-th scale of the reference and distorted image,

respectively. We called the proposed method βW-SCM for the reason of the way by which RR features were extracted.

2.5. Geometric mean of absolute deviation and relative deviation

To describe the change of RR features when the image was distorted, absolute deviation d_A^n and relative deviation d_R^n were defined in the following way:

$$d_A^n = \left|\beta_r^n - \beta_d^n\right| \tag{9}$$

$$d_R^n = \frac{\left|\beta_r^n - \beta_d^n\right|}{\beta_r^n} \tag{10}$$

Then the summation of the geometric mean of d_A^n and d_R^n in all scales was derived to predict the perceptual distortion of the test image.

$$D_{\beta W-SCM} = \sum_{n=1}^{N} \sqrt{d_A^n \times d_R^n}\,, \tag{11}$$

Where N is the total number of scales and $D_{\beta W-SCM}$ is the finally perceptual image distortion.

3. EXPERIMENT SETUP

3.1. Orientations and scales

Previous work [23] on steerable pyramid filters indicated that a set of eight oriented filters, or an orientation bandwidth of 45°, were sufficient to span the 360° of orientation space with a high accuracy. So we divided the image into 8 orientations in the scale space. We here used 6 scales for the representation of the image. The final perceptual distortion of the tested image was then computed as follow:

$$\begin{aligned} D_{\beta W-SCM} &= \sum_{n=1}^{6} \sqrt{d_A^n \times d_R^n} \\ &= \sum_{n=1}^{6} \frac{\left|\beta_r^n - \beta_d^n\right|}{\sqrt{\beta_r^n}} \end{aligned} \tag{12}$$

3.2. Test database

We used the LIVE database to test this new RRIQA method [24]. The database contains 982 subject-rated images created from 29 high-resolution 24-bits/pixel RGB color images with five types of distortions at different levels. The distortion types include 1) JPG: JPEG compression (2 sets), 2) JP2: JPEG2000 compression (2 sets), 3) WN: white noise contamination (1 set), 4) GB: Gaussian blur (1 set), and 5) FF: fast fading channel distortion of JPEG2000 compressed bit stream (1 set). A Difference Mean Opinion Score (DMOS) value for each distorted image was computed from the observers' perception results. The proposed method was applied on the luminance component of the image.

3.3. Evaluation indices

According to the VQEG report [25], the logistic function[1] was employed to provide a nonlinear mapping between the predictive scores and the subject scores:

$$f(x) = a_1\left(\frac{1}{2} - \frac{1}{1+\exp(a_2(x-a_3))}\right) + a_4 x + a_5 \tag{12}$$

where a_1, a_2, a_3, a_4 and a_5 are model parameters and found with the MATLAB optimization toolbox.

Two indices are used to evaluate the proposed algorithm's performance: 1) Correlation coefficient (CC) and 2) Spearman rank-order correlation coefficient (SROCC).

4. RESULTS AND DISCUSSION

The performance of the proposed RRIQA method is shown in Table 1. We also list in Table 1 the performance of the RRIQA methods developed in [13, 14] as a comparison. To the best of our knowledge, among the few existing RRIQA methods, these two methods are designed for a general purpose with a low RR data rate. Note that the results of reference [14] were acquired after the training procedure to adapt well to LIVE database, while the results of our proposed method and reference [13] need no training. It can be drawn from table 1 that the proposed method performs quite well for a wide range of distortion types. Specifically, for six of the seven subsets, it gives the better prediction accuracy and monotonicity compared to the method in [13]. And it also shows nearly the same performance for all the subsets with the method in [14] which performs noticeable well among the existing RRIQA method. When applied on the overall database, that is all the images with different distortion types are mixed together, the proposed method performs nearly the same with the method in [13]. But both of them fall behind the method in [14]. However, considering that the RR features are usually transmitted in a restricted reliable communication channel, the number of used bytes is important. In this case, the proposed method using only 6 scalar features has an outstanding advantage of low RR data rate compared to 48 scalar features in [14] and 18 in [13]. Figure 3 shows the scatter plots of DMOS versus model prediction. For each test image, the point closer to the blue diagonal indicates the better accuracy in predication. The scatters can roughly show the method's performance. In this paper, we just exploit the Weibull distribution parameter

[1] The matlab code was provided by H. R. Sheikh, the primary author of the reference [3].

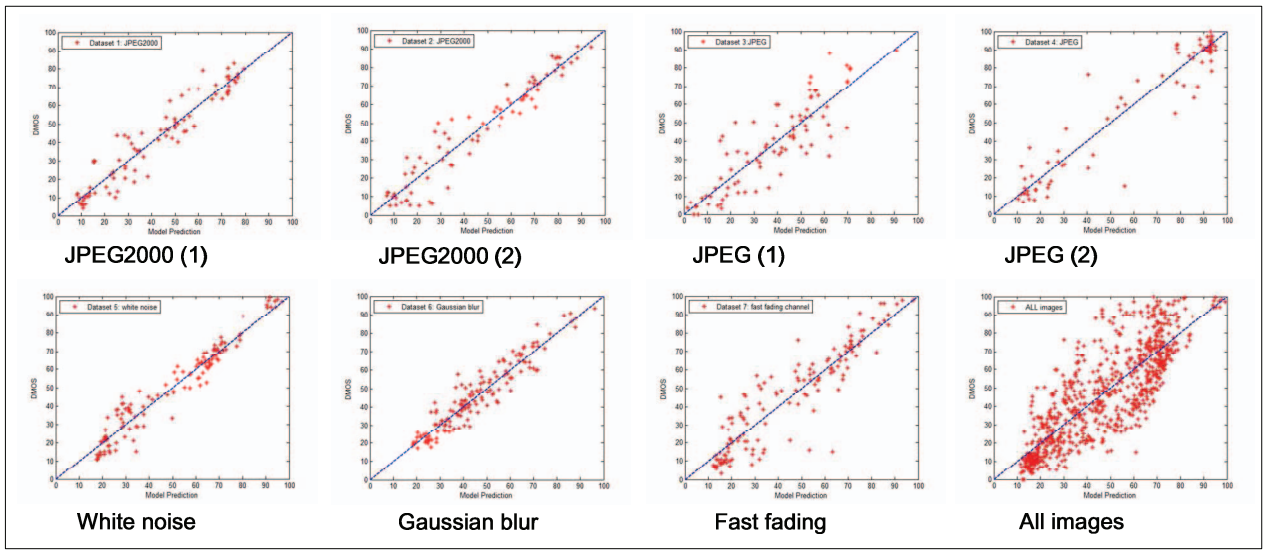

Figure 3. Scatter plots of DMOS versus model prediction for the seven image datasets respectively and all images in the LIVE database. Each star represents one image in the database.

of β to extract the RR features. The change in γ should also be incorporated for improvement in further study.

We analyze the execution time of the three methods with five pairs of reference and distorted images from the database on a Pentium IV computer with 1G RAM. Table 2 shows the actual run time of these methods. Obviously the method in [14] demands more than twice of the time used by the two others, which due to the time-demanding operation of divisive normalization. Owing to the similarity framework, it takes nearly the same time for the proposed method and the one in [13].

5. CONCLUSION

In this paper we propose a novel general purpose RRIQA algorithm. Firstly we define the SCM for redundancy reduction among differently oriented subbands in scale space to present image features. Secondly the Weibull distribution is employed to describe the statistics of the SCM and the scale parameter β is extracted as the RR feature. Finally a new method called βW-SCM is proposed to estimate the quality of the test image. The new method has the merits of less RR features and short execution time. Experiments with the LIVE database show significant correlation between the model predictions and the subjective scores, nearly the same as the state of the art metrics.

Table 1 Performance of the proposed method RRIQA method on LIVE database. Prediction accuracy and monotonicity were list for 7 subsets with different distortion type. The results from [13, 14] were also listed for comparison.

LIVE dataset	JP2(1)	JP2(2)	JPG(1)	JPG(2)	WN	GB	FF	ALL
CC (Prediction accuracy)								
Method in [14]	0.9485	**0.9655**	0.8203	0.9579	0.9654	**0.9562**	**0.9464**	**0.9173**
Method in [13]	0.9353	0.9490	0.8452	**0.9695**	0.8889	0.8872	0.9175	0.8226
βW-SCM	**0.9514**	0.9569	**0.8673**	0.9568	**0.9755**	0.9454	0.9243	0.8353
SROCC (Prediction monotonicity)								
Method in [14]	0.9478	**0.9610**	0.8143	**0.8937**	0.9559	**0.9584**	**0.9443**	**0.9287**
Method in [13]	0.9298	0.9470	0.8332	0.8908	0.8639	0.9145	0.9162	0.8437
βW-SCM	**0.9495**	0.9517	**0.8535**	0.8705	**0.9715**	0.9371	0.9258	0.8391

Table 2. Comparison of execution time of the three methods (in seconds).

Image size	480*720	768*512	618*453	618*453	640*512	Mean
Method in [14]	30.1978	30.2543	22.9196	22.7241	25.3701	26.2932
Method in [13]	12.0757	12.1505	10.4753	10.5979	11.8497	11.4298
βW-SCM	12.0387	15.4103	11.8381	11.4202	12.7671	12.6949

6. ACKNOWLEDGEMENTS

The project is partially supported by National Nature Science Fund of China through No.90920003 and No. 60472004. The authors are also grateful to H. R. Sheikh for providing the matlab code of logistic function

7. REFERENCES

[1] J. Lubin, "A visual discrimination model for imaging system design and evaluation," *Vision Models for Target Detection and Recognition: In Memory of Arthur Menendez,* p. 245, 1995.

[2] Z. Wang, A. C. Bovik and H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE transactions on image processing,* vol. 13, pp. 600-612, 2004.

[3] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing,* vol. 15, pp. 430-444, 2006.

[4] Damon M. Chandler and Sheila S. Hemami, "Vsnr: A wavelet based visual ignal-to-noise ratio for natural images, " *IEEE Trans. On Image Processing*, vol. 16, no. 9, pp. 2284–2298, 2007.

[5] 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, p. 011006, 2010.

[6] H. R. Sheikh, A. C. Bovik and L. Cormack, "No-reference quality assessment using natural scene statistics: JPEG 2000," *IEEE Transactions on Image Processing,* vol. 14, pp. 1918-1927, 2005.

[7] Z. Wang, H. R. Sheikh and A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images,", pp. 477 – 480, 2002.

[8] P. Marziliano, F. Dufaux and S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric,", *Proc. of IEEE Int. Conf. on Image Processing,* pp. 57 – 60, 2002.

[9] Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," in *Proc. of SPIE Human Vision and Electronic Imaging*, pp. 149--159, 2005.

[10] Maalouf, A. , Larabi, M.-C and Fernandez-Maloigne. C, "A grouplet-based reduced reference image quality assessment," *Quality of Multimedia Experience, 2009. International Workshop on*, pp.59-63,2009.

[11] IP Gunawan and M. Ghanbari. "Reduced reference picture quality estimation by using local harmonic amplitude information," *in Proc. London Communications Symposium*, pages 137–140, September 2003.

[12] Carnec, M., P. Le Callet and D. Barba, "Objective quality assessment of color images based on a generic perceptual reduced reference," *Signal Processing: Image Communication*, 23(4): p. 239—256, 2008.

[13] Z. Wang, G. Wu and H. R. Sheikh, E. P. Simoncelli, E. H. Yang, and A. C. Bovik, "Quality-aware images," *IEEE Transactions on Image Processing,* vol. 15, pp. 1680-1689, 2006.

[14] Q. Li and Z. Wang, "Reduced-Reference Image Quality Assessment Using Divisive Normalization-Based Image Representation," *IEEE Journal on Selected Topics in Signal Processing,* vol. 3, pp. 202-211, 2009.

[15] H. S. Scholte, S. Ghebreab, L. Waldorp, A. Smeulders, and V. Lamme, "Brain responses strongly correlate with Weibull image statistics when processing natural images," *Journal of Vision,* vol. 9, 2009.

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

[17] B. A. Olshausen and D. J. Field, "Natural image statistics and efficient coding," *Network-computation in neural Systems,* vol. 7, pp. 333-339, 1996.

[18] S. Lyu and E. P. Simoncelli, "Nonlinear image representation using divisive normalization," in *IEEE Conference on Computer Vision and Pattern Recognition, 2008*, pp. 1-8, 2008.

[19] J. M. Geusebroek, "The stochastic structure of images," *Scale space and PDE methods in computer vision, proceedings,* vol. 3459, pp. 327-338, 2005.

[20] Geusebroek, J.M. and A. Smeulders, "Fragmentation in the vision of scenes," *Ninth IEEE International Conference on Computer Vision, vols I and II, proceedings*, p. 130-135, 2003

[21] Simoncelli, E.P. and J. Portilla, "Texture Characterization via Joint Statistics of Wavelet Coefficient Magnitudes," *Fifth International Conference on Image processing, vols I, proceedings*, 1998.

[22] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *2nd IEEE International Conference on Image Processing*, pp. 444-447, 1995.

[23] H. Greenspan, S. Belongie and R. Goodman, P. Perona, S. Rakshit, and C. H. Anderson, "Overcomplete steerable pyramid filters and rotation invariance,", pp. 222-228 , 1994.

[24] H. R. Sheikh, Z. Wang and L. Cormack, and A. C. Bovik, "LIVE Image Quality Assessment Database Release 2," http://live.ece.utexas.edu/research/quality, 2005.

[25] VQEG, "Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II", http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseII/, 2003