SUST Journal of Engineering and Computer Sciences (JECS), Vol. 20, No. 3, 2019

Performance Evaluation of Natural Scenes Features to create Opinion Unaware-Distortion Unaware IQA Metric

Dr. Saifeldeen Abdalmajeed Mahmood

Assistant Prof. , Dep. of Electrical and Electronics Engineering, Faculty of Engineering and Technology, Nile Valley University, Atbara, Sudan,

Saifm11@yahoo.com,

Received: 12/10/2019

Accepted: 11/11/2019

ABSTRACT-There are many challenges facing image quality assessment (IQA) task. The greatest one which has been treated by this research is the difficulty of quantifying and evaluating distorted images quality blindly with no existence of the original (reference) image or partially from it. Choosing the appropriate features plays a significant role in measuring image quality. This study evaluates the efficiency of a set of features in quantifying image quality. The features have been gathered in spatial domain using the techniques of both rich edges and sharper regions of pristine natural images. The performance efficiency of these features examined through comparing them with both features gathered from reference and distorted images. These techniques employed to build two IQA metrics. Results clearly show the proposed pristine natural features competes reference features in assessing the distorted image quality. This proves the validity of these features in creating a robust metrics for evaluating distorted images. When testing the proposed metrics on LIVE database, experiment results show extracting features by means of rich edges is better than extracting it using sharper regions when assess the prediction monotonicity and applying the prediction accuracy evaluation. Besides they show the average outcome of the two techniques not only competes the popular full-reference peak signal-to-noise ratio (PSNR), the structural similarity (SSIM), and the developed NR natural image quality evaluator (NIQE) model but also outperform them.

Key words: *natural features, image, metric.*

المستخلص- *هناك عدة تحديات تواجه تقييم جودة الصوره أهمها -والذى تمت معالجته بهذا البحث-صعوبة معرفة جودة الصوره المشوهة فى غياب الصوره الاصلية (المرجعية) أو جزء منها. يلعب أختيار الميزا ت المناسبة دورا هاما فى تقييم جودة الصورة. هذه الدراسه تُقيم كفاءة مجموعة من الميزات فى تقدير جودة الصورة. الميزات تم استخلاصها باستخدام تقنيتى كل من الحواف والمناطق الاكثر وضوحا فى صور المشاهد الطبيعيه. هاتين التقنيتين وظفا لبناء مقياسين لتقييم جودة الصورة. لمعرفة كفاءة أداء هذه الميزات، تمت مقارنتها بكل من الميزا ت المأخوذه من الصور الاصلية والمشوهة. أظهرت النتائج بوضوح منافسة ميزات المشاهد الطبيعية للميزات الاصلية فى تقييم جودة الصور المشوهة. هذا يثبت بجلاء صحة الميزات المقترحة في إنشاء مقاييس قوية لتقييم الصور المشوهة. عند اختبار المقياسيين المقترحتين في قاعدة بيانات LIVE ،تُظهر نتائج التجربة أن استخلاص الميزات عن طريق المناطق الاغنى حوافا أفضل من أخذ الميزا ت عبر المناطق الاكثر وضوحا وذلك عند ايجاد كل من تقييم رتابة التنبؤ prediction monotonicity وتطبيق تقييم دقة التنبؤ prediction accuracy. إلى جانب ذلك تظهر التجربة أن متوسط نتائج التقنيتين لاينافس فقط نماذج نسبة إشارة الذروة إلى الضوضاء كاملة المرجع (PSNR) ، والتشابه الهيكلى (SSIM) ، ومقيم جودة الصورة الطبيعيه (NIQE) بل أيضا تتفوق عليهم.*

INTRODUCTION

The techniques using in image processing such as: A acquisition, transmission, compression, restoration and enhancement are image processing techniques are in focus of current research. Therefore, there is a demand to develop novel methods to quantify and assess images. In spite that Humans are the ultimate judge of image quality; their judgment is time consuming, subjective and at times, impractical. Thus the innovated techniques used to measure images, however, should be effective and correlate well with subjective humans opinion scores. One of the main issues of getting correlated metrics are the gathered features and the way that can be gathered.

Because full or even partial of the reference image may not be available, no-reference (NR) methods provide a useful and effective way to measure the quality of distorted images [1,2]. In addition, NR IQA techniques are the only available choice since the purification of reference images can be also uncertain. This is confidence in situation when perfect noise-free image is not available when assessing the quality of a de-noising algorithm on a real-world database.

The existing NR IQA methods mostly are based on prior knowledge of the type of distortion. These are called “distortion-specific NR IQA” [3-8]. In such algorithms, this specification limits their applications. In contrast, the “general distortion” NR IQA algorithms are non-distortion-specific. Those algorithms collect distorted images with co-registering human scores and are called opinion aware (OA) [9-11]. On the other hand, there are algorithms do not need training on databases of human judgments of distorted images. These are opinion unaware (OU) [12]. distorted images may not be available during IQA model construction or training. So among OU models, there are ones that do not require knowledge about anticipated distortions which are distortion unaware (DU) [13, 14]. The proposed features used to build OU-DU NR IQA approaches that do not need training on databases of human judgments or even prior knowledge about expected distortions. From a practical point of view, predicting NR IQA should not depend on prior knowledge about anticipated distortions or their corresponding human opinion scores. This is the case in most general no-reference IQA. The model used for testing IQA should be generic and should have been created in such a way that it does not expect any specific distortion type as the one used in this study.

In [13] asymmetric generalized Gaussian distribution (AGGD) was used to gather features. Punit and Damon [15] claimed the sharper an image the better is its quality. Moreover, humans give more heavily weight judgments of image quality from the sharp image regions [13]. The model used in this paper applied two sharpness functions considering mentioned knowledge. The output parameters of these functions represent the proposed extracted features.

The Research Motivation and Aim

Images processing techniques mentioned in section (1) cause the images to be subject to distortion. The processing systems should be able to identify and quantify image degradations in order to maintain, control, and enhanced the quality. Moreover, these techniques are growing interest in current research and therefore quality assessment methods have increased demand as well.

Choosing the appropriate features plays a significant role in measuring image quality. Since the performance behavior of IQA evaluators depends on the way they collect their measuring features, the model of this research collects pristine natural features as measuring features and investigates their validity to create a robust metric for assessing distorted images. The created metric used two feature gathering techniques ; rich edges and sharper regions.

Materials and methods

A set of natural low level features are used. Those features composed of locally normalized luminance and contrast values. The way that modeled the features is point wise statistics for single pixels. Besides, the relation of adjacent pixels also have been taken in a count. This done using pair wise based log-derivative statistics. Then extracted features should have to be fitted in such a way that it be suitable to treat. So Multivariate Gaussian Model (MVG) is then used. The features are distinguished and classified. The features related to patches with more edges and sharper regions have been gatheredseperatly. The generation of the features was due to two sharpness functions. This will be explained and in the model in the flowing sections.

Normalized Luminance and Contrast Coefficients and their Log-derivatives

Firstly the image $I(i,j)$ will be divides into patches according to the model. These patches of the size 96×96 . For each of the patches (for both distorted and natural images), contrast and normalized luminance computed. The last parameter, denoted by $\hat{I}(i,j)$ for both distorted and natural images. Normalized luminance computed through local mean subtraction and contrast divisive normalization (MSCN) [16]. This defined as:

$$\hat{I}(i,j)=\frac{I(i,j)-\mu(i,j)}{\sigma(i,j)+1} \quad (1)$$

Where $i \in \{1,2,\dots,M\}$ and $j \in \{1,2,\dots,N\}$ } are spatial domain indices, M and N are the dimensions of the image, and

$$\mu(i,j)=\sum_{k=-K}^{K}\sum_{l=-L}^{L} w_{k,l} I(i+k,j+l) \quad (2)$$

$$\sigma(i,j)=\sqrt{\sum_{k=-K}^{K}\sum_{l=-L}^{L} w_{k,l}\left[I(i+k,j+l)-\mu(i,j)\right]^2} \quad (3)$$

are the estimated local mean and local contrast respectively and $w=\{w_{k,l}|k=-K,\dots,K, l=-L,\dots,L\}$is a 2D circularly-symmetric Gaussian weighting function sampled out to three standard deviations $(K=L=3)$ and rescaled to unit volume. After computing MSCN (1) and contrast (3) coefficients, features are calculated through these coefficients for each patch. Features are extracted by means of log-derivative statistics [14].

To acquire the log-derivatives, the logarithm of $\hat{I}(i,j)$is computed using (4) to create new image sub-band J .

$$J(i,j)=\log\left(\hat{I}(i,j)+\varepsilon\right) \quad (4)$$

The small constant ε is taken to be 0.1 to prevent $I(i,j)$ from being zero. The five types of log-derivatives are then computed. These include horizontal, vertical, main-diagonal, secondary-diagonal, and combined-diagonal as given in (5-9).

$$J_x(i,j)=J(i,j+1)-J(i,j) \quad (5)$$

$$J_y(i,j)=J(i+1,j)-J(i,j) \quad (6)$$

$$J_{xy}(i,j)=J(i+1,j+1)-J(i,j) \quad (7)$$

$$J_{yx}(i,j)=J(i+1,j-1)-J(i,j) \quad (8)$$

$$J_{xandy}(i,j)=J(i,j)+J(i+1,j+1)-J(i,j+1)-J(i+1,j) \quad (9)$$

In the spatial domain, the MSCN coefficients and their log-derivatives statistics significantly change in the presence of some distortion [9, 16]. The effectiveness of these statistics in modeling natural images and their variations due to different types of distortions has been examined in this study.

The two sharpness functions based extracted features:

The MSCN coefficients in (1) and the log-derivatives (5-9) are modeled following two sharpness function: grey level “amplitude” and grey level "variance" (10) [17]. The MSCN and the five log-derivatives used with each sharpness function come up with 12 model features as outputs of these two functions. These features are computed at two scales to represent multi-scale behavior. This achieved through low pass filtering and down sampling by a factor of two, this process leads to a set of 24 features overall. All features are extracted in the spatial domain.

$$graylevel\text{"}amplitude\text{"}, \frac{1}{a\times b}\sum_{k=-K}^{K}\sum_{l=-L}^{L}|\hat{I}(i,j)-\bar{\hat{I}}(i,j)|, and$$
$$graylevel\text{"}\mathrm{var}iance\text{"}, \frac{1}{a\times b}\sum_{k=-K}^{K}\sum_{l=-L}^{L}(\hat{I}(i,j)-\bar{\hat{I}}(i,j))^2$$

(10)

Where $\bar{\hat{I}}(i,j)$ and 'a' and 'b' is mean and dimensions of a patch respectively.

The features obtained by (10) for image patches were fitted with MVG density (11), to give their rich representation [13].

$$f_X\left(x_1,\ldots,x_k\right)=\frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}}\times\exp(-\frac{1}{2}\left(x-\upsilon\right)^T\Sigma^{-1}\left(x-\upsilon\right))$$

(11)

Where, $x_1,\ldots,x_k$, are the features. The mean and covariance matrix of the MVG model are υ and $\sum$ respectively.

Edges and sharper image based natural scene statistic model

The natural scene statistic (NSS) model computed from 125 natural images, which were selected from Flickr data and from the Berkeley image segmentation database [18]. The features corresponding to patches with both rich of edges and sharper are selected. Each patch is divided to sub-patches of 6×6 size and only sub-patches those are rich in edges (effective sub-patches) and sharper are contributed into their main patches. Then the effective sub-patches of each patch were computed. Patches that had an effective sub-patch greater than 75% of the peak patch effective sub-patches over the image are selected. The features corresponding to the selected patches were gathered. These features were then fitted to MVG model (11).

To compute the quality according to the procedure mentioned above, (12) is used.

$$D\left(\upsilon_1,\upsilon_2,\Sigma_1,\Sigma_2\right)=\sqrt{\left(\upsilon_1-\upsilon_2\right)^T\left(\frac{\Sigma_1+\Sigma_2}{2}\right)^{-1}\left(\upsilon_1-\upsilon_2\right)} \quad (12)$$

The mean vectors and covariance matrices of the NSS MVG and the tested image MVG models are υ_1,υ_2 and $\sum_1,\sum_2$ respectively.

The results and discussion

The effectiveness of the proposed features in modeling pristine natural scenes for giving perfect image quality assessment will be investigated and examined.To do this statistics of these natural features are compared with statistics of features extracted from both reference and distorted images, as shown by figure (1). The reference and distorted features which plotted in figure (1) are belong to lighthouse image of LIVE (Laboratory for Image and Video Engineering) database. Lighthouse image and its five distorted versions are displayed in figure (3).

LIVE IQA database [19] is used to calibrate the proposed features and test their performance. LIVE database contains 29 reference images and 779 distorted images. These are classified into five different types of distortions and can be a result of JPEG and JPEG2000 (JPEG2K) compression or introduced as Gaussian blur (Gblur). The image transmission through a Rayleigh channel also corrupts the image and is termed as fast fading (FF) distortion. One of the common types of distortion is the additive white Gaussian noise (WN).

Toassess the prediction monotonicity, Spearman's rank ordered correlation coefficient (SROCC) is used while Pearson's linear correlation coefficient (PLCC) is employed to evaluate the prediction accuracy of the proposed algorithm. Before PLCC is calculated, the objective scores are passed through a logistic non-linear function [20] (where its parameters are found numerically using the MATLAB function 'fminsearch' in the optimization toolbox) to maximize the correlations between subjective and objective scores.

The plots of figure (1, a-e) compare the pristine features with features of lighthouse reference image (figure 3) and its five distorted versions from LIVE database. All extracted features shown in this figure are via sharper patches. The figure shows a shift and little deformation happens to the features based natural (pristine) images in the presence of the five distortion types. But in spite of this, the distorted features are still consistent with those gathered from pristine images. This indicates the capability of these features for recognizing distortion, quantify it, and then measure image quality. Besides, figure (1) illustrates all plotted features have sharp tips. The observation from figure (1, f) is the pristine features are totally compatible with those extracted from the reference image. This compatibility makes them acts as purified reference features for measuring various distortions regardless it is classified or not.

When compare the features gathered due to sharper patches with those extracted via patches of rich edges [21], we come up with result both are consistent with features extracted from reference image, figure (2). This support the results of this research that our features can be used as alternative of the reference image to measure the distorted image quality. This is practical, where full or even partial reference image may not be available when assessing distorted image quality, as the case in de-noising techniques.

The result proved in figure (1) and figure (2) encourage us to create a robust metrics for evaluating distorted images blindly using the rich edges and sharper patches. Table (1) shows a comparison of SROCC and PLCC when extracting features using the two techniques respectively. Also it demonstrates a comparison of SROCC and PLCC for the presented algorithms versus: FR-PSNR and FR-SSIM algorithms and NR-NIQE. The table indicates that extracting features by means of rich edges better than extracting it using sharper regions when assess prediction monotonicity and applying prediction accuracy evaluation.

The table also shows the average results of each of the two techniques not only competes the popular full-reference peak signal-to-noise ratio (PSNR), the structural similarity (SSIM), and the developed NR natural image quality evaluator (NIQE) model butalso outperformthem.

CONCLUSIONS

Researchers have to think for developing perceptual no-reference models that do not train on features extracted from distorted images and human opinion scores for practical considerations. However, choosing the appropriate features and the way to collect them play a significant role in the issue of IQA. In this study, a comparison between two techniques for gathering low level features is examined. Also a performance comparison between two NR DU IQA metrics using these features is introduced. The NR OU-DU model used in this study has low computational complexity, and extracted features in the spatial domain so no transforms (e.g. DCT, wavelet, etc.) are required. The results show that the introduced method provides good performances.

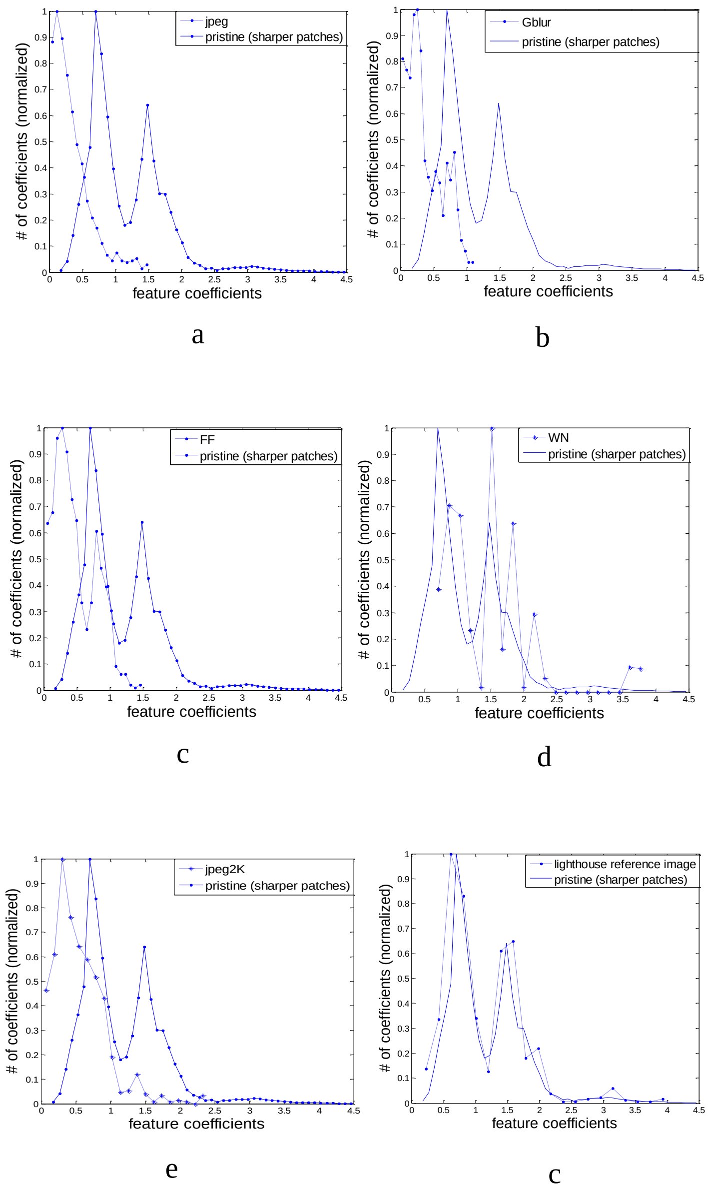

Figure (1): comparison of the pristine (natural images) features with features of lighthouse reference image and its five distorted types from LIVE database when extracting features through sharper patches.(a) JPEG, (b) Gblur, (c)FF, (d) WN, (e) JPEG2000and, (f)the reference image.

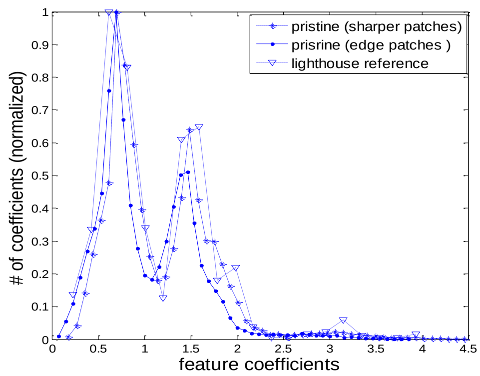

Figure (2): plot showing the pristine (natural images) features which are extracted with both sharper patches and more edge techniques are consistent with features of lighthouse reference image from LIVE database displayed in fig. (3)

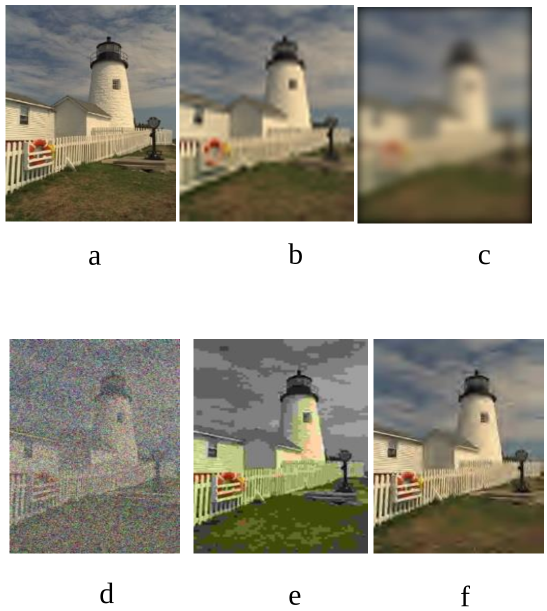

Figure (3): Lighthouse reference image and its five distorted versions in the LIVE database.(a) the reference image, (b) FF, (c) Gblur, (d) WN, (e) JPEG, and (f) JPEG2000.

Table 1: Comparison of SROCC and PLCC for the presented algorithms versus: FR-PSNR and FR-SSIM algorithms and NR-NIQE. Bold italic indicates best-performing algorithm between NR algorithms

	Metric	FF	Gblur	WN	JPEG	JP2K	average
SROCC	PSNR	0.7817	0.8086	0.6858	0.8478	0.8424	0.7933
	SSIM	0.6535	0.6536	0.6430	0.8086	0.8014	0.7120
	NIQE	0.7860	0.7602	***0.8636***	0.8669	***0.8571***	***0.8268***
	Sharper patches	***0.7882***	0.7313	0.7440	0.8860	0.8549	0.8009
	edges patches	0.7500	***0.7628***	0.78446	***0.8692***	0.8434	0.8020
PLCC	PSNR	0.7440	0.7572	0.7643	0.8271	0.8265	0.7778
	SSIM	0.7442	0.7871	0.8193	0.7840	0.7668	0.7803
	NIQE	0.8592	0.9102	0.8025	0.7520	***0.8931***	0.8434
	Sharper patches	0.8183	0.9766	0.9009	0.8712	0.8866	0.8907
	edges patches	***0.9884***	***0.9875***	***0.9797***	***0.9074***	0.8726	***0.9471***

REFERENCES

[1] SathyaVeera Reddy Dendi, ChanderDev, Narayan Kothari, Sumohana S. Channappayya, Generating Image Distortion Maps Using Convolutional Autoencoders With Application to No Reference Image Quality Assessment, Signal Processing Letters IEEE, vol. 26, no. 1, pp. 89-93, 2019.

[2]Jomgyoo Kim, Sanghoon Lee, Deep Image Quality Asement By Employing FR-IQA, International conference on Image Processing, IEEE, 17-20, Beijing, China, Sept. 2017.

[3] Wang, Z., A.C. Bovik, H.R. Sheikh and E.P. Simoncelli. Image quality assessment: From Error visibility to structural similarity. IEEE Trans.13: 600-612, 2004.

[4] Li, X. Blind image quality assessment. Proceedings of the IEEE International Conference on Image Processing, Rochester, September 22-25, New York, USA, 2002.

[5] Gabarda, S. and G. Cristobal. No-reference image quality assessment through the von Mises distribution. J. Opt. Soc. Am. A,29: 2058-2066, 2012

[6] Wang, Z., A.C. Bovik and B.L. Evans, Blind measurement of blocking artifacts in images.Proceedings of the International Conference on Image Processing, September 10-13, Vancouver, BC, 2000.

[7] Bovik, A.C. and S. Liu. DCT-domain blind measurement of blocking artifacts in DCT-coded images.Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing May 7-11, Salt Lake City, UT., USA, 2001.

[8] Zhang, J., S.H. Ong and T.M. Le. Kurtosis-based no-reference quality assessment for JPEG2000 images. Signal Process. Image Commun., 26: 13-23, 2011.

[9] Mittal, A., A.K. Moorthy and A.C. Bovik.No-reference image quality assessment in the spatial domain. IEEE Trans. Image Process.21: 4695-4708, 2012.

[10] A.K.Moorthy, and A.C. Bovik. Blind image quality assessment: From natural scene statistics to perceptual quality. IEEE Trans. Image Proces.20: 3350-3364, 2011.

[11] M.,Saad, A.C. Bovik and C. Charrier. Blind image quality assessment: A natural scene statistics approach in the DCT domain EEE Trans. Image Process.21: 3339-3352, 2012.

[12] A.,Mittal, G.S. Muralidhar, J. Ghosh and A.C. Bovik. Blind image quality assessment without human training using latent quality factors.IEEE Signal Process.Lett. 19: 75–78, 2011.

[13] A.,Mittal, R. Soundararajan and A.C. Bovik. Making a completely blind image quality analyzer.IEEE Signal Process.Lett.20: 209-212, 2013.

[14]SaifeldeenAbdalmajeed and Jiao Shuhong, No-reference image quality assessment algorithm based on Weibull statistics of log-derivatives of natural scenes, Electronics Letters, Volume 50, Issue 8, , p. 595 – 596, DOI: 10.1049/el.2013.3585, 10 April 2014.

[15] Punit Singh and Damon M. Chandler, "F-MAD: A feature-Based Extension of the Most Apparent Distortion Algorithm for Image Assessment," in *Proc.of SPIE*, 2013.

[16] D. L. Ruderman, The statistics of natural images. Netv a : Comput. N b l Syst., 5: 517–548, 199

[17] Ng KuangChern, N., Poo AunNeow and Marcelo H. Ang Jr.”. Practical Issues in Pixel-based Autofocusing for Machine Vision”, Proceedings of the IEEE,3:2791 – 2796, 2001.

[18] D., Martin, C. Fowlkes, D. Tal and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. Proceedings of the 8th International Conference on Computer Vision, (2001) July 7-14, , California University Berkeley, CA., USA.

[19] H.R.,Sheikh, Z. Wang, L. Cormack and A.C. Bovik, LIVE image quality assessment database release 2, 2005. Laboratory for Image and Video Engineering.http://live.ece.utexas.edu/research/quality/subjective.htm.

[20] H.R., Sheikh, M.F. Sabir and A.C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans. Image Processing,.15: 3440–3451 (2006).

[21] Abdalmajeed, Saifeldeen&Shuhong, Jiao. Using the Natural Scenes’ Edges for Assessing Image Quality Blindly and Efficiently.Mathematical Problems in Engineering.1-9. 10.1155/2015/389504, 2015.