Perceptual Weights Based On Local Energy For Image Quality Assessment

Sudhakar Nagalla   *sudhakar.nagalla@becbapatla.ac.in*
Department of Computer Science and Engineering
Bapatla Engineering College
Bapatla, 522102, India

Ramesh Babu Inampudi   *rinampudi@hotmail.com*
Department of Computer Science and Engineering
Acharya Nagarjuna University
Guntur, 522510, India

Abstract

This paper proposes an image quality metric that can effectively measure the quality of an image that correlates well with human judgment on the appearance of the image. The present work adds a new dimension to the structural approach based full-reference image quality assessment for gray scale images. The proposed method assigns more weight to the distortions present in the visual regions of interest of the reference (original) image than to the distortions present in the other regions of the image, referred to as perceptual weights. The perceptual features and their weights are computed based on the local energy modeling of the original image. The proposed model is validated using the image database provided by LIVE (Laboratory for Image & Video Engineering, The University of Texas at Austin) based on the evaluation metrics as suggested in the video quality experts group (VQEG) Phase I FR-TV test.

Keywords: Image Quality, HVS, Full-reference Quality Assessment, Perceptual Weights.

1. INTRODUCTION

Any image processing system should be aware of the impacts of processing on the visual quality of the resulting image. Numerous algorithms for image quality assessment (IQA) have been investigated and developed over the last several decades. The objective image quality measurement seeks to measure the quality of images algorithmically. Objective image quality metrics can be classified as full-reference in which the algorithm has access to the original (considered to be distortion free) image, no-reference in which the algorithm has access only to the distorted image and reduced-reference in which the algorithm has partial information regarding the original image. A comprehensive review of research and challenges in image quality assessment is presented in [1].

In [2], a number of simple statistical image quality metrics based on numerical errors are compared for gray scale image compression. These metrics include average difference, maximum difference, absolute error, mean square error (MSE), peak MSE, Laplacian MSE, histogram and Hosaka plot. It is observed that although some numerical measures correlate well with the human response for a specific compression technique, they are not found to be reliable for evaluation across various methods of compression. The most widely adopted statistical feature is the Mean Squared Error (MSE). However, MSE and its variants may not correlate well with subjective quality measures because human perception of image distortions and artifacts is unaccounted for. A detailed discussion on MSE is presented by Girod [3].

Most HVS based quality assessment metrics share an error-sensitivity based paradigm [4], which aims to quantify the strength of the errors between the reference and the distorted signals in a

perceptually meaningful way. A well-known method, Visible Differences Predictor (VDP) [5], Lubin's algorithm [6], Teo and Heeger's metric [7], a perceptual image quality metric named information mean square error (IMSE) proposed by David Tompa et al. [8], a measure of perceptual image quality of Westen et al. [9], a comprehensive distortion metric for digital color images presented by Stefan Winkler [10], an image quality metric using contrast signal-to-noise ratio (CSNR) by Susu Yao et al. [11], image quality metric named visual information fidelity (VIF) introduced by Sheikh and Bovik [12] belong to this category. The rest of the paper is organized as follows. Section 2 explains structural similarity measure, Section 3 presents local energy model for detecting image features, Section 4 explains weighting of structural similarity indices and formulation of Perceptual Structural Similarity index. Section 5 presents the results followed by conclusions.

2. STRUCTURAL SIMILARITY MEASURE

One distinct feature that makes natural image signals different from a "typical" image randomly picked from the image space is that they are highly structured and the signal samples exhibit strong dependencies amongst themselves. These dependencies carry important in-formation about the structures of objects in the visual scene. An image quality metric that ignores such dependencies may fail to provide effective predictions of image quality. Structural similarity based methods [13, 14] of image quality assessment claim to account for such dependencies in assessing the image quality. In [14] a more generalized and stable version of the universal quality index was proposed named as Structural SIMilarity quality measure ($SSIM$).

Let x and y be two discrete non-negative signals where $x = \{x_i \mid i = 1, 2...N\}$ and $y = \{y_i \mid i = 1, 2...N\}$ are aligned with each other (e.g. two image patches extracted from the same spatial location of original image and distorted image being compared). Let $\mu_x, \mu_y, \sigma_x, \sigma y, \sigma_{xy}$ represent mean intensity of signal x, mean intensity of signal y, standard deviation of x, standard deviation of y, and covariance between x and y respectively. The Structural Similarity measure between the image patches is defined in (1), where C_1 and C_2 are small constants introduced to avoid instability when the denominator is close to zero.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{1}$$

Let X and Y be the two images being compared. A local moving window approach is followed, to compute $SSIM(X, Y)$. The window moves pixel-by-pixel from the top left corner to the bottom right corner of the image. In each step, the local statistics and $SSIM(x_j, y_j)$ index are calculated using (1) within the local window j. The $SSIM$ index between X and Y is defined in (2) where N_S is the number of local windows in the image, and $W_j(x_j, y_j)$ is the weight given to the j-th window of the image. If all the local regions in the image are equally weighted, then $W_j(x_j, y_j) = 1$. This results in the mean $SSIM(MSSIM)$ measure employed in [14].

$$SSIM(X, Y) = \frac{\sum_{j=1}^{N_s} W_j(x_j, y_j).SSIM(x_j, y_j)}{\sum_{j=1}^{N_s} W_j(x_j, y_j)} \tag{2}$$

It may be noted that the $MSSIM$ algorithm gives equal importance to distortions for all local regions of the image. Wang et al. [14] suggest that the performance of $SSIM$ can be improvised by weighting the local $SSIM$ indices. They also suggest that the prior knowledge about the

importance of different regions in the image if available can be converted into a weighting function. A variety of such approaches can be found in [15]-[20].

Studies of visual attention and eye movements [6, 21, 22] have shown that humans attend to few areas in the image. Even though unlimited viewing time is provided, subjects will continue to focus on few areas rather than scan the whole image. These areas are often highly correlated amongst different subjects, when viewed in the same context. In order to automatically determine the parts of an image that a human is likely to attend to, we need to understand the operation of human visual attention and eye movements. In [23], many algorithms for defining Visual regions of interest were evaluated in comparison with eye fixations. The present work adopts the local energy model to identify feature rich local regions which are normally attended to by humans and to define a weighting function that is proportional to feature richness of the region. The weighting function is used in (2) to define the Perceptual Structural SIMilarity index $PSSIM^e$ proposed in this paper.

3. LOCAL ENERGY MODELING FOR FEATURE DETECTION

The local energy model of feature detection postulates that features are perceived at points of maximum phase congruency in an image. Venkatesh and Owens [24] show that points of maximum phase congruency can be calculated equivalently by searching for peaks in the local energy function. The calculation of energy from spatial filters in quadrature pairs has been central to the models of human visual perception proposed by Heeger [25], Adelson and Bergen [26]. Local frequency and, in particular, phase information in signals are of importance in calculating local energy. To preserve phase information, linear-phase filters must be used. That is, one must use non orthogonal filters that are in symmetric/antisymmetric quadrature pairs. In this work, the approach of Morlet et al. [27] is followed with a modification in the usage of filters. Logarithmic Gabor functions [28, 29] are used instead of Gabor filters as the maximum bandwidth of a Gabor filter is limited to approximately one octave and Gabor filters are not optimal if one needs broad spectral information with maximal spatial localization.

Field [28] suggests that natural images are better coded by filters that have Gaussian transfer functions when viewed on the *logarithmic* frequency scale. Firstly, log-Gabor functions, by definition, always have no DC component, and secondly, the transfer function of the log Gabor function has an extended tail at the high frequency end. Field's studies of the statistics of natural images indicate that natural images have amplitude spectra that fall off at approximately inverse of the frequency. To encode images having such spectral characteristics one should use filters having spectra that are similar. Field suggests that log Gabor functions, having extended tails, should be able to encode natural images more efficiently than, say, ordinary Gabor functions, which would over-represent the low frequency components and under-represent the high frequency components in any encoding. Another point in support of the log Gabor function is that it is consistent with measurements on mammalian visual systems which indicate we have cell responses that are symmetric on the log frequency scale.

The local energy function is computed using log-Gabor filters in 4 scales with center frequencies of 1/3 cycles/pixel, 1/6 cycles/pixel, 1/12 cycles/pixel and 1/24 cycles/pixel and 6 orientations at 0^0(horizontal), 30^0, 60^0, 90^0(vertical), 120^0, and 150^0. The following discussion [29] is made for a specific orientation θ of the filter. If $I(x)$ denotes the image signal and M_e and M_o denote the even-symmetric (cosine) and odd-symmetric (sine) filters at a scale n, the respective responses e_n and o_n of each quadrature pair of filters can be represented by the vector,

$$\left[e_n\left(x\right), o_n\left(x\right) \right] = I\left(x\right) * M_n^e, I\left(x\right) * M_n^o \tag{3}$$

The amplitude $A_n(x)$ and phase $\phi_n(x)$ of the transform at any given scale is given by

$$A_n(x) = \sqrt{e_n(x)^2 + o_n(x)^2} \tag{4}$$

$$\phi_n(x) = atan2\left(e_n(x), o_n(x)\right) \tag{5}$$

At each point x in a signal, an array of these response vectors is obtained, one vector for each scale of filter of the chosen orientation. These response vectors form the basis of localized representation of the signal, and they can be used in calculating the resultant local energy vector at point x. The design of the filter bank needs to be such that the transfer function of each filter overlaps sufficiently with its neighbors so that the sum of all the transfer functions forms a relatively uniform coverage of the spectrum. If the local energy should accurately represent the feature strength at point x, then a broad range of frequencies in the signal are to be retained. The local energy at point x of the image $E_\theta(x)$ for a given orientation θ can be calculated from $F(x)$ which can be formed by summing the even filter convolutions over all scales and $H(x)$ which can be estimated by summing the odd filter convolutions over all scales given by

$$F(x) = \sum_n e_n(x) \tag{6}$$

$$H(x) = \sum_n o_n(x) \tag{7}$$

$$E_\theta(x) = \sqrt{F(x)^2 + H(x)^2} \tag{8}$$

Figure 1(a) and Figure 1(b) show the normalized maps of the local energy function of the *Lena* image considering different ranges of frequencies. Figure 1(a) is the result of considering frequencies larger than 0.2*pixels/cycle*. Figure 1(b) is the result of considering the complete set of frequencies. One can observe that the latter makes a clear distinction among the significance of features than the former. The former shows that the majority of features are equally important while the latter shows a broad scale distinction.

(a) Energy function considering 2 scales   **(b)** Energy function considering 4 scales

FIGURE 1: Energy function of *Lena* Image.

At each location in the image, the weighted local oriented energy $E_\theta(x)$ in each orientation is calculated, and the sum over all orientations $E(x)$ is computed. The following algorithm illustrates the above steps.

Let $I(x)$ be the original image;

$c \leftarrow 0.5$;

$\varepsilon \leftarrow .0001$;

$\gamma \leftarrow 10$;

$E(x) \leftarrow [\ ]$;

for each orientation θ do

    $sum\_e(x) \leftarrow [\ ]$;

    $sum\_o(x) \leftarrow [\ ]$;

    $sum\_A_n(x) \leftarrow [\ ]$;

    for each scale n do

        compute $e_n(x), o_n(x)$ as in (3);

        compute $A_n(x)$ using as in (4);

        $sum\_e(x) \leftarrow sum\_e(x) + e_n(x)$;

        $sum\_o(x) \leftarrow sum\_o(x) + o_n(x)$;

        $sum\_A_n(x) \leftarrow sum\_A_n(x) + A_n(x)$;

        if first scale then

            $A_{max}(x) \leftarrow A_n(x)$;

        else

            $A_{max}(x) \leftarrow max(A_n(x), A_{max}(x))$;

        end if

    end for

    compute $E_\theta(x)$ as in (8);

    $E(x) \leftarrow E(x) + E_\theta(x)$;

end for

Figure 2 shows the perceptual map of *Lena* image based on local energy. The perceptual importance map assigned more weights (more bright) to image features in face, hair, hat and background. It can also be observed that the perceptual weights assigned to the features are well distributed.

FIGURE 2: Perceptual map of Lena image.

4. LOCAL ENERGY WEIGHTED STRUCTURAL SIMILARITY

We assume that the width w and the height h of the original image X and the distorted image Y are exact multiples of 9. If the size does not conform to these dimensions the images are cropped on all sides so that minimum amount of details is lost. This requirement comes from the fact that $SSIM$ indices are computed in non-overlapping 9×9 regions. These regions in X and Y are referred to as x_{ij} and y_{ij} respectively.

The computation of perceptual weights of local regions in the original image begins with the log-Gabor decomposition of the image. Log-Gabor filters with 4 scales and 6 orientations are used for this purpose. The algorithm presented earlier explains the computation of local maxima $E(x)$ of local energy function at each pixel location x for the original image X. Let the matrix $E(x)$ be divided into non-overlapping blocks of size 9×9. Each resulting block in corresponds to a non-overlapping block x_{ij} of the original image X, $1 \le i \le m = h/9$ and $1 \le j \le n = w/9$. The local maxima values present in each such block are summed up to obtain the local maxima for the block x_{ij}. The resulting matrix is normalized and these values are proposed as the perceptual weights of the corresponding to 9×9 local regions which will be indicative of the human attention the regions call for. Let the resulting matrix be E of size $m \times n$.

The structural similarity index between corresponding blocks of X and Y of is computed using (1) to obtain the matrix $SSIM$ of size $m \times n$. The Weighted Structural SIMilarity measure $PSSIM^{e}$ between X and Y is calculated using (9). $PSSIM^{e}$ indicates the quality of distorted image on a scale of 0 to 1, where a value of 1 indicates that the images are identical.

$$PSSIM^{e} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}[E][SSIM]}{\sum_{i=1}^{m}\sum_{j=1}^{n}[E]} \tag{9}$$

5. RESULTS

The proposed models are validated using the image database provided by *LIVE (Laboratory for Image & Video Engineering, The University of Texas at Austin)* [30]. The psychometric study for the development of the database contained 779 images distorted using five different distortion types and more than 25,000 human image quality evaluations.

The distorted image database consists of twenty-nine high resolution 24-bits/pixel RGB color images (typically 768 × 512). The distortions include white Gaussian noise, Gaussian blur, simulated fast fading Rayleigh (wireless) channel, JPEG compression and JPEG2000 compression and with each type the perceptual quality covered the entire quality range. Observers are asked to provide their perception of quality on a continuous linear scale that was divided into five equal regions marked with adjectives "Bad", "Poor", "Fair", "Good", and "Excellent". About 20-29 human observers rated each image. The raw scores for each subject are converted to difference scores (between test and reference images) and then converted to Z-scores , scaled back to 1-100 range, and finally a Difference Mean Opinion Score (*DMOS*) value for each distorted image is computed.

The proposed models and other models used for comparison are validated using the *LIVE* image database based on the evaluation metrics as suggested in the video quality expert's group (VQEG) Phase I FR-TV test [31]. A nonlinear regression model is fitted to the *DMOS* values in the database, and calculated image quality metric values (*IQ)* of the distorted images for each

distortion and for each quality assessment model used in comparison. The following 4 parameter logistic function is used in the present work.

$$DMOS_p = 1/\left(1 + exp\left(-b*(IQ-c)\right)\right) + d \qquad (10)$$

The nonlinear regression function is used to transform the set of *IQ* values to a set of predicted *DMOS* values, *DMOSp*, which are then compared with the actual *DMOS* values from the subjective tests. The Correlation Coefficient (*CC*), the Mean Absolute Error (*MAE*), and the Root Mean Squared Error (*RMSE*) between the subjective scores *DMOS* and predicted scores *DMOSp* are evaluated as measures of prediction accuracy. The prediction consistency is quantified using the Outlier Ratio (*OR*), which is defined as the percentage of the number of predictions outside the range of 2 times the standard deviation of errors between *DMOS* and *DMOSp*. Finally, the prediction monotonicity is measured using the Spearman Rank-Order-Correlation Coefficient (*ROCC*).

Table 1 shows the results for the proposed method to estimate the image quality index *PSSIM*[e] in comparison with the three image quality assessment models *PSNR* (Peak Signal-to-Noise Ratio) and the Structural SIMilarity quality measure *SSIM* [14]. The Correlation Coefficient (*CC*), the Mean Absolute Error (*MAE*), and the Root Mean Squared Error (*RMSE*) values for the three assessment models considered prove that the prediction accuracy of the proposed model is superior to the others. The values of the Spearman Rank-Order-Correlation Coefficient (*ROCC*) indicate that the proposed model correlates well with the human judgment. However, the values of Outlier Ratio (*OR*) are inferior marginally when compared with the other two models. This can be attributed to the fact that the human judgment is impulsive in case of images with higher levels of distortion in contrast to the computational algorithms for image quality assessment.

Figure 4 shows the scatter plots for different distortions in which each data point represents true mean opinion score (*DMOS*) versus the predicted score of one test image by the proposed method after the nonlinear mapping.

Model	CC	ROCC	MAE	RMSE	OR
White Noise					
PSNR	0.922	0.938	4.524	6.165	0.055
SSIM	0.94	0.914	4.475	5.459	0.027
PSSIM[e]	0.967	0.958	3.344	4.073	0.034
Gaussian Blur					
PSNR	0.744	0.725	8.395	10.501	0.034
SSIM	0.947	0.940	3.992	5.027	0.034
PSSIM[e]	0.971	0.966	3.02	3.711	0.041
Fast Fading					
PSNR	0.857	0.859	6.383	8.476	0.068
SSIM	0.956	0.945	3.806	4.799	0.055
PSSIM[e]	0.965	0.964	3.471	4.315	0.041
JPEG Compression					
PSNR	0.842	0.828	6.636	8.622	0.062
SSIM	0.891	0.863	5.386	7.236	0.057
PSSIM[e]	0.916	0.888	4.727	6.393	0.062
JPEG2000 Compression					
PSNR	0.859	0.851	6.454	8.269	0.059
SSIM	0.899	0.894	5.687	7.077	0.023
PSSIM[e]	0.931	0.926	4.754	5.876	0.041

.

TABLE 1: Performance comparison of image quality assessment models.

6. CONCLUSION

In this paper, we proposed a full-reference perceptual image quality metric for gray scale images based on structural approaches unified with perceptual regions humans attend to in a natural image. The local energy model was used to indentify feature rich regions in natural images and to formulate a weighting function for distortions a given image.
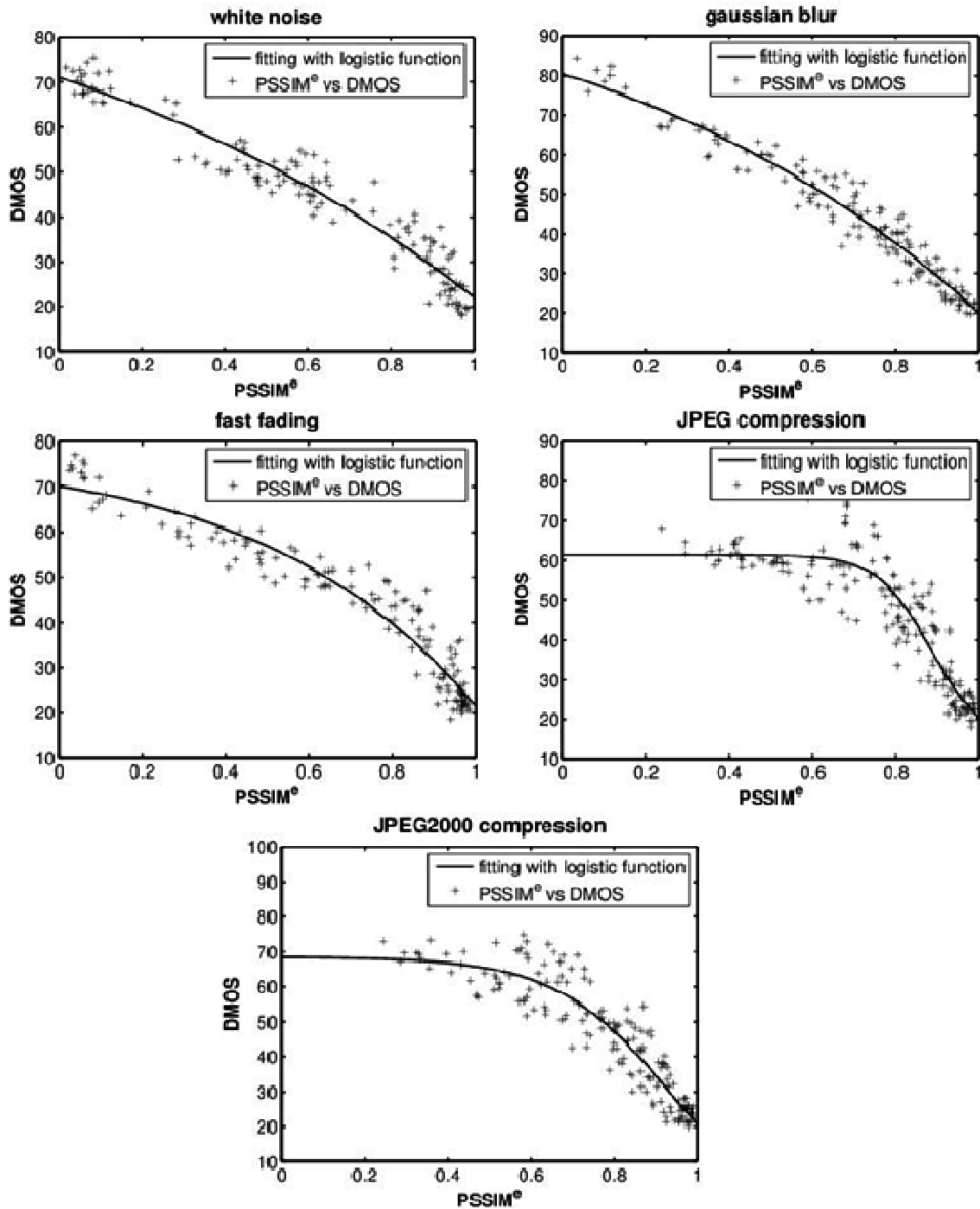

FIGURE 3: Scatter plots between *DMOS* and *PSSIM^e* for different image distortions.

The proposed models and other models used for comparison are validated using the metrics suggested by VQEG. The results prove that the performance of the metric is close to the human (subjective) judgment. The metrics are also found to be superior in performance in comparison with the other models of quality assessment considered. The metrics are generic and they are applicable to a wide variety of image distortions like white noise, Gaussian blur, fast fading, and different compression artifacts.

7. FUTURE RESEARCH

The present work formulated a framework for the image quality assessment is evolved in which the model of identifying perceptual regions and the process of computing the image distortions are independent. Such a frame work facilitates a modular approach so that the individual models can be modified and optimized independently. As the framework for formulating perceptual quality metric is flexible, different combinations of distortion modeling and perceptual region modeling can be explored. In the present work, the notion of perceptual regions is used in image quality assessment. It can be extended to other possible areas of image processing like face recognition and watermarking.

8. REFERENCES

[1] Damon M. Chandler, "Seven challenges in image quality assessment: Past, present, and future research," ISRN Signal Processing, Article ID 905685, [Online]. Available:http://dx.doi.org/10.1155/2013/905685, 2013.

[2] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance", IEEE Transactions on   Communications, vol. 43, Dec. 1995, pp. 2959-2965.

[3] B. Girod, "What's wrong with mean-squared error," in Digital Images and Human Vision, A. B. Watson, Ed., MIT press, 1993, pp. 207-220.

[4] Z.Wang, H.R. Sheikh, and A.C. Bovik, Objective video quality assessment: The Handbook of Video Databases: Design and Applications, B.Furht and O. Marques, Eds., CRC press, 2003.

[5] S. Daly, "The visible differences predictor: an algorithm for the assessment of image fidelity," in Digital Images and Human Vision, A. B. Watson, Ed., MIT press, 1993, pp.197-206.

[6] J.Lubin, "A visual discrimination model for image system design and evaluation," in Visual Models for Target Detection and Recognition, E.Peli, Ed., World Scientific Publishers, Singapore, 1995, pp. 245-283.

[7] P. C. Teo and D. J. Heeger, "Perceptual image distortion," Proc. SPIE, vol. 2179, 1994, pp. 127-141,

[8] David Tompa, John Morton, and Ed Jernigan, "Perceptually based image comparison", International conference on image processing, ICIP, vol. 1, 2000, pp. 489-492.

[9] S. Western, K.L. Lagendijk, and J. Biemond, "Perceptual image quality based on a multiple channel hvs model," ICASSP, vol. 4, 1995, pp. 2351-2354..

[10] Stefan Winkler, "A perceptual distortion metric for digital color images," Proc. International Conference on Image Processing, ICIP98, vol. 3, Oct. 1998, pp. 399-403.

[11] Susu Yao, Weisi Lin, EePing Ong, and Zhongkang Lu, "Contrast signal -to-noise ratio for image quality assessment," Proc. IEEE International Conference on Image Processing, ICIP 2005, vol. 1, Sept. 2005, pp. 397-400.

[12] Hamid Rahim Sheikh and Alan C.Bovik, "Image information and visual quality," IEEE Transactions on Image Processing, vol. 15, Feb. 2006.

[13] Z.Wang and A.C. Bovik, "A universal image quality index," IEEE Signal Processing Letters, vol. 9, pp. 81-84, Mar. 2002.

[14] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Si-moncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, pp. 600-612, Apr. 2004.

[15] D.Venkata Rao, N.Sudhakar, B.R.Babu, and L.Pratap Reddy, "An image quality assessment technique based on visual regions of interest weighted structural similarity," ICGST international journal on Graphics Vision and Image Processing, vol. 6, pp. 69-75, Sept. 2006.

[16] D.Venkata Rao, N.Sudhakar, B.R.Babu, and L.Pratap Reddy, "Image quality assessment complemented with visual regions of interest," ICCTA 2007-Proc. International Conference on Computing: Theory and Applications, IEEE Computer Society Press, vol. 2, Mar. 2007, pp. 681-687.

[17] D.Venkata Rao and L.Pratap Reddy, "Image quality assessment based on perceptual structural similarity," in Pattern Recognition and Machine Intelligence, ser. Lecture Notes in Computer Science, Springer-Verlag, vol. 4815, Dec. 2007, pp. 87-94.

[18] D.Venkata Rao and L. Pratap Reddy, "Weighted similarity based on edge strength for image quality assessment," International Journal of Computer Theory and Engineering, vol. 1, pp. 138-141, June. 2009.

[19] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," IEEE Transactions on Image Processing, vol. 20, pp. 1185-1198, 2011.

[20] Punit Singh and Damon M. Chandler, "F-mad: a feature-based extension of the most apparent distortion algorithm for image quality assessment," Proc. SPIE, Image Quality and System Performance X, SPIE Digital Library , vol. 8653, Feb. 2013.

[21] J.Findlay. "The visual stimulus for saccadic eye movement in human observers," Perception, vol.9, pp. 7-21, Sept. 1980.

[22] J. Senders. "Distribution of attention in static and dynamic scenes," Proceedings of SPIE, vol. 3016, Feb. 1997, pp. 186-194.

[23] Claudio M. Privitera and Lawrence W. Stark, "Algorithms for defining visual regions of interest: Comparison with eye fixations," IEEE Tans. on Pattern Analysis and Machine Intelligence, vol. 22, Sept. 2000.

[24] S. Venkatesh and R. Owens, "An energy feature detection scheme," Proc. of The International Conference on Image Processing, 1989, pp. 553-557.

[25] D. J. Heeger, "Normalization of cell responses in cat striate cortex," Visual Neuroscience, vol.9, pp.181-197, 1992.

[26] E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," Journal of the Optical Society of America A, vol. 2, pp. 284-299, 1985.

[27] J. Morlet, G. Arens, E. Fourgeau, and D. Giard, "Wave propagation and sampling theory - part ii: Sampling theory and complex waves," Geophysics, vol. 47, pp. 222-236, Feb. 1982.

[28] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," Journal of Optical Society of America, vol. 12, pp. 2379-2394, 1987.

[29] Peter Kovesi, "Image features from phase congruency," Videre: A Journal of Computer Vision Research, vol.1, 1999.

Sudhakar Nagalla & Ramesh Babu Inampudi

[30] H. R. Sheikh, A. C. Bovik, L. Cormack, and Z. Wang, LIVE image quality database, [Online]. Available: http://live.ece.utexas.edu/research/quality.

[31] Ann Marie Rohaly, Philip Corriveau, John Libert, Arthur Webster, Vittorio Baroncini, and John Beerends,   VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," [Online]. Available: http://www.vqeg.org/.