OPEN ACCESS
ISSN: 2249-6645

International
Journal
Of Modern Engineering Research (IJMER)

A New Quality Evaluation for Tone Mapped Images

G.Divya[1], N.L.Aravinda[2]
* (Electronics and communication Engineering, CMR Engineering College, JNTUH, India
Email: dha.grs@gmail.com)
** (Electronics and communication Engineering, CMR Engineering College, JNTUH, India
Email: aravindanl@gmail.com)

ABSTRACT: *Conversion of High dynamic range images to Low dynamic range images is still area of concern in Image processing domain. The Tone mapped Operators are one such operators used to convert HDR to LDR for better visualization in practical Standard LDR displays. There is wide variety of Tone mapped images for wide variety of practical applications, so which Tone mapped operator (TMO) is best, so without an appropriate analysis and quality measurement, comparison of Tone mapped operators cannot be done. In order to perform this comparison, subjective rating approach is best assessment method, though its performance is outstanding its quite expensive and more time consuming one, moreover, precisely when come to optimization frameworks its quite difficult to embedded .Here, the proposed Algorithm proposes a novel Quality measurement Algorithm for TMOs i.e. Objective quality assessment algorithm by combining the Modified structural similarity index values and natural images statistics values. The proposed tone-mapped image quality index (TMQI) has good correlation as in subjective ranking score. Every time we cannot process TMO's on after visualization, for example in medical images, images are often taken in HDR images, so in that we process before visualization.*
Keywords: *High dynamic range images, Tone mapped operator (TMO)*

I. INTRODUCTION

There has been a growing interest in recent years in high dynamic range (HDR) images, where the range of intensity levels could be on the order of 10,000 to 1. This allows for accurate representations of the luminance variations in real scenes, ranging from direct sunlight to faint starlight. A common problem that is often encountered in practice is how to visualize HDR images on standard display devices that are designed to display low dynamic range (LDR) images. To overcome this problem, an increasing number of tone mapping operators (TMOs) that convert HDR to LDR images have been developed. A common problem that is often encountered in practice is how to visualize HDR images on standard display devices that are designed to display low dynamic range (LDR) images. To overcome this problem, an increasing number of tone mapping operators (TMOs) that convert HDR to LDR images have been developed.

The purpose of the current work is to develop an objective IQA model for tone mapped LDR images using their corresponding HDR images as references. Our work is inspired by the success of two design principles in IQA literature. The first is the structural similarity (SSIM) approach and its multi-scale derivations, which asserts that the main purpose of vision is to extract structural information from the visual scene and thus structural fidelity is a good predictor of perceptual quality.

II. STRUCTURAL SIMILARITY (SSIM)

This project is related to structural similarity index (SSIM) which represents perceptual image quality based on the structural information. SSIM is an objective image quality metric and is superior to traditional quantitative measures such as MSE and PSNR. This project demonstrates the SSIM based image quality assessment and illustrates its validity in terms of human visual perception. It will be very helpful to understand SSIM and its applications by reviewing the paper listed below.

A general form of SSIM is

$$SSIM(x,y) = [l(x,y)]^{\alpha}.[c(x,y)]^{\beta}.[s(x,y)]^{\gamma} \quad (1)$$

[Note that $\alpha>0$, $\beta>0$ and $\gamma>0$ are parameters used to adjust the relative importance of the three components].

where x, y are image patches and

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \; s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$ and $l(x,y)$ is luminance

comparison, $c(x,y)$ is contrast comparison, and $s(x,y)$ is structural comparison. C_1, C_2, C_3 are constants.

Gaussian weighting function has the form $w(n_1,n_2) = \exp\left(-\frac{n_1^2+n_2^2}{2\sigma^2}\right)$, $n_1, n_2 = 1, 2, \cdots, 11$

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - x')^2, \; \sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - y')^2 \quad (2)$$

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - x')(y_i - y'). \quad (3)$$

A. Structural Similarity (SSIM) Index

The structural similarity (SSIM) method is a recently proposed approach for image quality assessment. By “structured signal”, mean that the signal sample exhibit strong dependencies amongst themselves, especially when they are spatially proximate. The goal of image quality assessment research is to design method that quantifies the strength of the perceptual similarity between the test and the reference images.

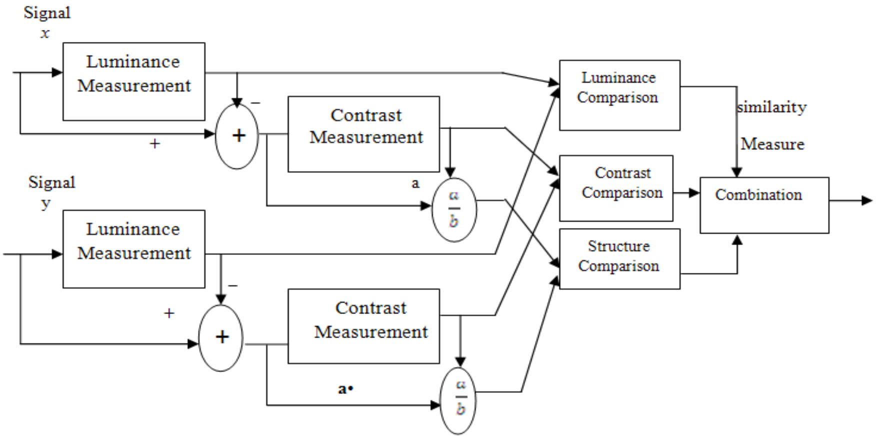

Fig 1**:** Structural similarity (SSIM) measurement system

The system diagram of the SSIM quality assessment system is shown in Fig 1. Suppose and are two nonnegative image signals, which have been aligned with each other. The purpose of the system is to provide a similarity between them. The system separates the task of similarity measurement into three comparisons: luminance, contrast and structure. First, the luminance of each signal is compared.

Use the standard deviation (the square root of variance) as an estimate of the signal contrast. An unbiased estimate in discrete form is given by

$$\sigma_x = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2]^{1/2} \quad (4)$$

The contrast comparison c(x, y) is then the comparison of σ_x and σ_y:

c(**x**, **y**) = c(σ_x, σ_y) (5)

Third, the signal is normalized (divided) by its own standard deviation; so that the two signals are being compared have unit standard deviation. The structure comparison s(x,y) is conducted on these normalized signals:

s(**x**, **y**) = $(x- \mu_x)/\sigma_x$ and $(y- \mu_y)/\sigma_y$. (6)

Finally, the three components are combined to yield an overall similarity measure:

S (**x**, **y**) = f (ℓ(**x**, **y**), c(**x**, **y**), s(**x**, **y**)) (7)

An important point is that the three components are relatively independent. The similarity measure needs to satisfy the following conditions.

1 .Symmetry: S(x,y) = S(y,x). Since the purpose is to quantify the similarity between two signals, exchanging the order of the input signals should not affect the resulting similarity measurement.

2 .Bounded ness: S(x,y) <= 1. Boundness is a useful property for a similarity metrics since an upper bound can serve as an indication how close the two signals are to being perfectly identical. This is in contrast with most signal-to-noise ratio type of measurements, which are typically unbounded.

3 .Unique maximum: S(x, y) = 1 if and only if x = y (in discrete representations xi = yi, for all i= 1,2…….,N). In other words, the similarity measure should quantify any variations that may exist between the input signals. The perfect score is achieved only when the signals being compared are exactly the same.
For luminance comparison, we define

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (8)$$

The constant C_1 is included to avoid instability when $\mu_x^2 + \mu_y^2$ is very close to zero. Specifically, we choose

$$C_1 = (K_1 L)^2 \quad (9)$$

Where L is the dynamic range of the pixel values
The contrast comparison function takes a similar form

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (10)$$

Where C_2 is a non-negative constant

$$C_2 = (K_2 L)^2 \quad (11)$$

Structure comparison is conducted after luminance subtraction and variance normalization. Specifically, the direction of two unit vectors $(x- \mu_x)/\sigma_x$ and $(y- \mu_y)/\sigma_y$, each lying in the hyper plane defined by (3.4), with the structure of the two images. The correlation (inner product) between these is a simple and effective measure to quantify the structural similarity. The correlation between $(x- \mu_x)/\sigma_x$ and $(y- \mu_y)/\sigma_y$ is equivalent to the correlation coefficient between **x** and **y**. Thus, we define the structure comparison function as follows:

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (12)$$

As in the luminance and contrast measures, we have introduced a small constant in both denominator and numerator.
Finally, we combine the three comparisons and the resulting similarity measure the SSIM index between signals **x** and **y.**

$$SSIM(x,y) = [l(x,y)]^{\alpha}.[c(x,y)]^{\beta}.[s(x,y)]^{\gamma} \quad (13)$$

Where α, β and γ are parameters used to adjust the relative importance of the three components. It is easy to verify that this definition satisfies the three conditions given above. In particular this paper α, β and γ are set to 1 and $C_3 = C_2/2$. This results in a specific form of the SSIM index:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \quad (14)$$

In practice, one usually requires a single overall quality measure of the entire image. We use a mean SSIM (MSSIM) index to evaluate the overall image quality:

$$MSSIM(X,Y) = \frac{1}{M}\sum_{j=1}^{M} SSIM(x_j, y_j) \quad (15)$$

where x and y are the reference and the distorted images, respectively; x_j and y_j are the image contents at the j^{th} local window; and M is the number of local windows of the image. The SSIM index is better in capturing poor quality regions.
A drawback of the spatial domain SSIM algorithm is that it is highly sensitive to geometrical distortions such as translation, scaling, rotation or other misalignments.

III. PROPOSED LITERATURE

A.STRUCTURAL FIDELITY

The original SSIM algorithm is applied locally and contains three comparison components − luminance, contrast and structure. Since TMOs are meant to change local intensity and contrast, direct comparisons of local and contrast are inappropriate. Let *x* and *y* be two local image patches extracted from the HDR and the tone-mapped LDR images, respectively. We define our local structural fidelity measure as

$$S_{local}(x,y) = \frac{2\sigma'_x\sigma'_y + C_1}{\sigma'^2_x + \sigma'^2_y + C_1} \cdot \frac{\sigma_{xy} + C_2}{\sigma_x\sigma_y + C_2} \quad (16)$$

where σ_x , σ_y and σ_{xy} are the local standard deviations and cross correlation between the two corresponding patches in HDR and LDR images, respectively, and C_1 and C_2 are positive stabilizing constants. Compared with

the SSIM definition, the luminance comparison component is missing, and the structure comparison component (the second term in (1)) is exactly the same.
Generally, the psychometric function resembles a sigmoid shape, and the sensory threshold is usually defined at the level of 50% of detection probability. A commonly adopted psychometric function is known as Galton's ogive, which takes the form of a cumulative normal distribution function given by

$$p(s) = \frac{1}{\sqrt{2\pi}\theta_s}\int_{-\infty}^{s} \exp\left[-\frac{(x - \tau_s)^2}{2\theta_s^2}\right]dx \qquad (17)$$

where p is the detection probability density, s is the amplitude of the sinusoidal stimulus, τs is the modulation threshold, and θs is the standard deviation of the normal distribution that controls the slope of detection probability variation. The reciprocal of the modulation threshold τs is often used to quantify visual contrast sensitivity, which is a function of spatial frequency, namely the contrast sensitivity function (CSF).

$$A(f) \approx 2.6[0.0192+0.114f]\exp[-(0.114f)^{1.1}] \qquad (18)$$

where f denotes spatial frequency. This function is normalized to have peak value 1, and thus only provides relative sensitivity across the frequency spectrum. In practice, it needs to be scaled by a constant λ to fit psychological data. In our implementation, we follow Kelly's CSF measurement.

Combining this with above, we obtain

$$\tau_s(f) = \frac{1}{\lambda A(f)}. \qquad (19)$$

This threshold value is calculated based on contrast sensitivity measurement assuming pure sinusoidal stimulus. To convert it to a signal strength threshold measured using the standard deviation of the signal, we need to take into account that signal amplitude scales with both contrast and mean signal intensity, and there is A √2 factor between the amplitude and standard deviation of a sinusoidal signal. As a result, a threshold value defined on signal standard deviation can be computed as

$$\tau_\sigma(f) = \frac{\bar{\mu}}{\sqrt{2}\lambda A(f)}. \qquad (20)$$

where μ is the mean intensity value. Based on Crozier's law, we have

$$\theta_\sigma(f) = \frac{\tau_\sigma(f)}{k} \qquad (21)$$

We can then define the mapping between σ and σ_ as

$$\sigma' = \frac{1}{\sqrt{2\pi}\theta_\sigma}\int_{-\infty}^{\sigma} \exp\left[-\frac{(x - \tau_\sigma)^2}{2\theta_\sigma^2}\right]dx \qquad (22)$$

The local structural fidelity measure S_{local} is applied to an image using a sliding window that runs across the image space. This results in a map that reflects the variation of structural fidelity across space. The visibility of image details depends on the sampling density of the image, the distance between the image and the observer, the resolution of the display, and the perceptual capability of the observer's visual system. A single scale method cannot capture such variations. Following the idea used in multi-scale and information-weighted SSIM, we adopt a multi-scale approach, where the images are iteratively low-pass filtered and down sampled to create an image pyramid structure.
At each scale, the map is pooled by averaging to provide a single score:

$$S_l = \frac{1}{N_l}\sum_{i=1}^{N_l} S_{local}(x_i, y_i) \qquad (23)$$

where xi and yi are the i -th patches in the HDR and LDR images being compared, respectively, and Nl is the number of patches in the l-th scale. In the literature, advanced pooling strategies such as information content based pooling have been shown to improve the performance of IQA algorithms

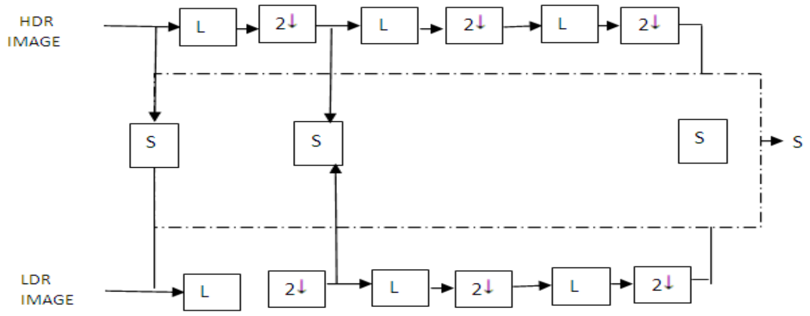

Fig 2: Framework of multi scale structural fidelity assessment

However, in our current experiment, these advanced pooling methods did not result in notable performance gain in the proposed structural fidelity measure. The overall structural fidelity is calculated by combining scale level structural fidelity scores

$$S = \prod_{l=1}^{L} S_l^{\beta_l} \quad (24)$$

where L is the total number of scales and βl is the weight assigned to the l-th scale.

B.STATISTICAL NATURALNESS

An interesting study of naturalness in the context of subjective evaluation of tone mapped images was carried out in which provided useful information regarding the correlations between image naturalness and different image attributes such as brightness, contrast, color reproduction, visibility and reproduction of details. The results showed that among all attributes being tested, brightness and contrast have more correlation with perceived naturalness.

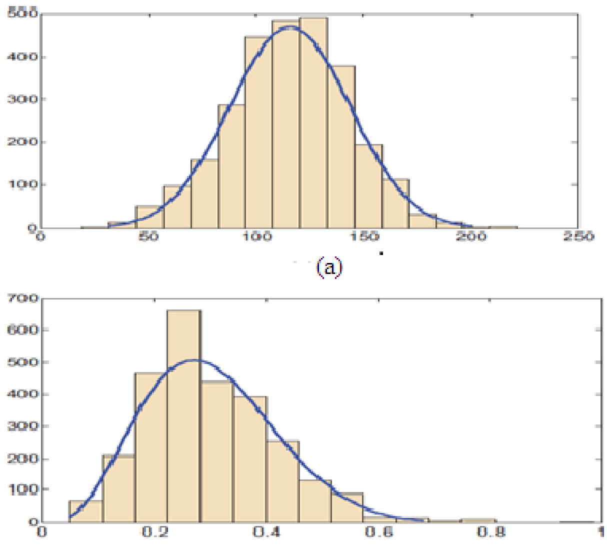

Fig 3: Histograms of (a) means (fitted by Gaussian PDF) and (b) standard deviations (fitted by Beta PDF) of natural images

Our statistical naturalness model is built upon statistics conducted on about 3,000 8bits/pixel gray-scale images obtained from that represent many different types of natural scenes. Above Fig. shows the histograms of the means and standard deviations of these images, which are useful measures that reflect the global intensity and contrast of images. We found that these histograms can be well fitted using a Gaussian and a Beta probability density functions given by

$$P_m(m) = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left[-\frac{m-\mu_m}{2\sigma_m^2}\right] \quad (25)$$

$$P_d(d) = \frac{(1-d)^{\beta_d-1} d^{\alpha_d-1}}{B(\alpha_d, \beta_d)} \quad (26)$$

where B(·, ·) is the Beta function. The fitting curves are shown in above Fig.3 where the model parameters are estimated by regression, and the best values we found are μm = 115.94 and σm = 27.99 in (11), and αd = 4.4 and βd = 10.1 respectively. As a result, their joint probability density function would be the product of the two. Therefore, we define our statistical naturalness measure as

$$N = \frac{1}{K} P_m P_d \qquad (27)$$

where K is a normalization factor given by $K = \max\{Pm\ Pd\}$.
This constrains the statistical naturalness measure to be bounded between 0 and 1.

Adaptive Fusion of Tone-Mapped Images:

To take the advantages of multiple TMOs, image fusion techniques may be employed to combine multiple tone mapped image and an objective quality measure can play an important role in this process. Given multiple tone mapped images created by different TMOs, we first apply Laplacian pyramid transform that decomposes these images into different scales. In the pyramid domain, this results in multiple coefficients at the same scale and the same spatial location, each corresponds to a different TMO.. A fusion strategy can then be applied to combine multiple coefficients into one at each location in each scale before an inverse. The general idea is to use the TMQI as the weighting factors in the fusion process.
Let Sj and $c\ j$ be the local structural fidelity measure and the Laplacian pyramid transform coefficient computed from the j -th tone mapped image being fused, respectively. The fused coefficient is computed as

$$c^{(fused)} = \frac{\Sigma_j S_j c_j}{\Sigma_j S_j} \qquad (28)$$

This is applied to all scales except for the coarsest scale, for which we use the statistical naturalness measure as the weighting factor:

$$c^{(fused)} = \frac{\Sigma_j N_j c_j}{\Sigma_j N_j} \qquad (29)$$

where N_j denotes the statistical naturalness score of the jth tone mapped image.
Fig.4 provides an example with natural scene, where one tone mapped image (a) better preserves structural details, and another (b) gives more natural overall appearance (but loses structural information, especially at the brightest areas). Three fused images created by three different image fusion algorithms are given in (c), (d) and (e), respectively. The image created by the proposed method achieves the best balance between structure preserving and statistical naturalness, and also results in the best quality score using TMQI.

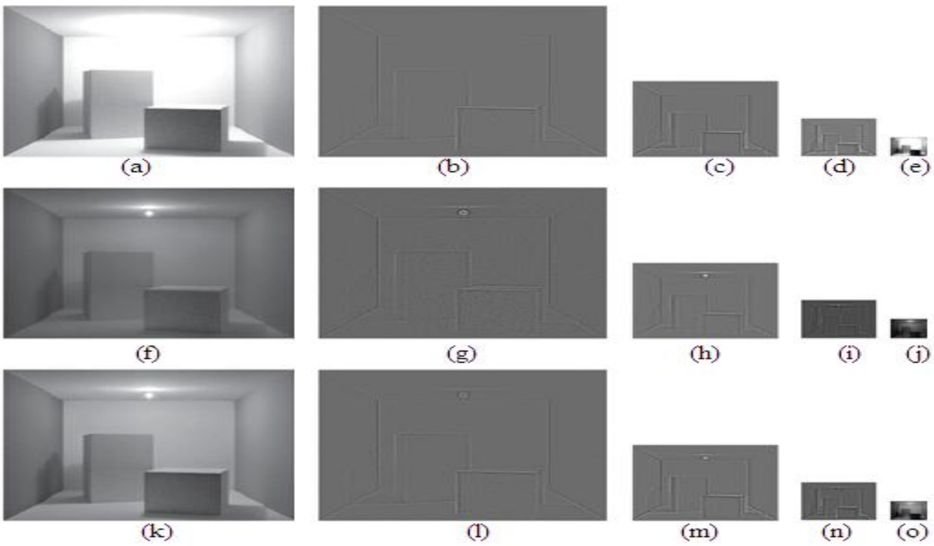

Fig 4**:** Image fusion in Laplacian pyramid domain. Top row: first tone-mapped image (a) created by TMO, and its (b)–(e) Laplacian pyramid subbands, S = 0.5034, N = 0.1263, Q = 0.6937. Middle row: second tone-mapped image (f) using "Exposure and Gamma" method in Adobe Photoshop, and its (g)–(j) Laplacian pyramid subbands, S = 0.6642, N = 0.0786, and Q = 0.7386. Bottom row: fused image by (k) the proposed method, and its (l)– (o) Laplacian pyramid domain representation

To further validate the proposed fusion scheme, we have conducted an additional subjective experiment, where ten subjects were invited to rank five sets of tone-mapped images, each of which includes eight images.
Seven of these images are generated using the TMOs employed in the third experiment in above. Two of these seven TMOs are chosen to produce the eighth image using the proposed fusion method. Compares average

subjective rankings of the source images and their corresponding fused images, where lower ranking scores correspond to better quality. It can be seen that the fused image is almost always ranked significantly higher than the two source images being fused.

IV. RESULTS

The Structural fidelity is done in multi scale approach as shown in following figures, here 'S' is total multi scale score and 's1''s2''s3''s4''s5' are number of multi scales used in Structural fidelity. 'Q' is total quality score and where as 'N' is scene naturalness. The objective approach mainly relies on structural fidelity "S" and Scene naturalness measurement "N".

'S1'

'S2'

'S3'

'S4'

'S5'

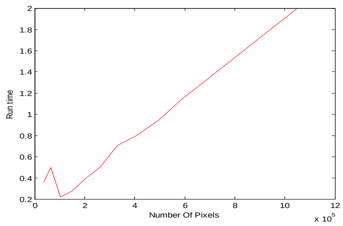

Fig. 5**:** Performance evaluation

V. CONCLUSION

The proposed method is proposed based on the objective quality assessment approach, basically objective quality approach is proposed on gray scale images but HDR images is taken in Color spaces. The objective quality assessment approach is mainly relies on two methods, first multi scale structural fidelity and second one is scene naturalness. First the structural fidelity performance is measured. Second the fidelity is checked at each stage and then by using the window overall standard deviation is obtained. Third distance of image from resultant user is measured by applying the CSF. Fourth main intensity values are attained by setting the mean of dynamic range of LDR image values. Then by combing all this measures we get overall psychological experiments.

In order to perform this on RGB image it has to convert to Y_{xy} space using the Y component. Every time we cannot process TMO's on after visualization, for example in medical images, images are often taken in HDR images, so in that we process before visualization. In order to apply TMO's in all applications, optimization methods are adapted for this type of scenarios. The proposed method gives good quality score compared to previous objective approach in grayscale images and it is more reliable than available subjective approach.

REFERENCES

[1]. Image Processing and Analysis: Variation, PDE, Wavelet, and Stochastic Methods by Tony F. Chan, Jianhong (Jackie) Shen

[2]. Objective Image and Video Quality Assessment with Applications by Qiang Li The University of Texas at Arlington, 2009

[3]. Objective Quality Assessment of Digital Video Jacqueline Suk Lai Kee Electrical and Electronic Engineering)-- University of Auckland, 1997 - Digital video

[4]. Modern Image Quality Assessment Zhou Wang, Alan Conrad Bovik Morgan & Claypool Publishers, 01-Jan-2006 - Technology & Engineering

[5]. http://www-2.cs.cmu.edu/afs/cs/project/cil/www/v-images.html.

[6]. http://www-sta_.lboro.ac.uk/ cogs/datasets/UCID/ucid.html.

[7]. Martin Cadk and Pavel Slavk. The naturalness of reproduced high dynamic range images. In IV '05: Proceedings of the Ninth International Conference on Information Visualisation, pages 920{925, Washington, DC, USA, 2005. IEEE Computer Society.

[8]. F. Durand and J. Dorsey. Fast bilateral _ltering for the display of high dynamic range images. In ACM Transactions on Graphics, volume 21, pages 257-266, 2002.

[9]. R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. In Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '02, pages 249-256, 2002.