ANALYSIS AND COMPARISON OF QUALITY METRICS WITH REFERENCE BASED ON UNIFORM COLOUR SPACES

G. Dauphin[1], P. Viaris de Lesegno[1]

[1]Laboratoire de Traitement et Transport de l'Information

Institut Galilée, Université Paris 13

{gabriel.dauphin,patrick.viaris}@univ-paris13.fr

ABSTRACT

Image quality assessment is becoming increasingly used in many applications. In most of the existing image quality assessment approaches, the main objective is to develop measures that are consistent with the subjective evaluation. Therefore, the performance of a given image quality metric is evaluated against the MOS determined from a series of subjective tests performed on a database. A plethora of image quality metrics has been developed. However, a few studies have been reported on the analysis and comparison of these metrics. This study attempts to provide a new framework for analysing and comparing some of the most common image quality metrics. Three quality representative metrics of the most known approaches have been chosen for this study. The Peak Signal to Noise Ratio (PSNR), the Visible Differences Predictor (VDP) and the Mean Structural SIMilarity index (SSIM). The two latter are found to be consistent with Weber's law. However, subjective testing in literature and computations derived from uniform color spaces such as CIE-L*a*b* suggest a different photometric invariance law. In this paper, we establish this photometric invariance law and show through numerical simulations how to check whether a given quality metric is compliant with this law.

Index terms: Image Fidelity, Quality Measure, Uniform Colour Space, Weber's Law, Gamma Correction.

1. INTRODUCTION

Image quality assessment is becoming increasingly prevalent nowadays in many applications. In most of the existing approaches the interaction with the observers seems to be the only way to assess the performance of the results. Many objective image quality metrics have been proposed in the last decades. However, due to the wide variety of picture types and applications, image quality assessment could not be fully automatic and subjective approaches are still predominant. The image quality measures in the literature can be classified into two categories: subjective and objective. The subjective evaluation is the most accepted and reliable method. However, the subjective tests should be done in a well defined environment and experimental setting standardized by the ITU [8]. However, it is time-consuming, complex and useless in real-time application. This motivates the development of objective image quality assessment methods that aim at predicting the perceived image quality. Objective image quality assessment methods could be classified into three classes: Full-reference (FR), No-Reference (NR) and Reduced-Reference (RR) methods. For FR methods the image quality metric is derived from a direct comparison between the original image and the distorted image. In NR methods an estimate of the image quality is computed from the observed image without referring to the original signal. RR methods could be used when some features or a set of local or global information extracted from the original and the processed images are available. In the last three decades, a significant amount of research effort has been directed towards the development of FR quality assessment methods inspired from the Human Visual System (HVS).

Many psycho-visual experiments have been performed in order to better understand the visual perception mechanisms [14]. The objective efficiency of any HVS-inspired methods is evaluated in terms of correlation with subjective tests ([3], [13], [11]). The intent of this paper is to propose a new approach for analysing and comparing IQMs using some photometric properties based on the perceptually uniform color space. This approach is introduced and illustrated on three representative image quality metrics, namely PSNR, VDP and SSIM. The Peak Signal to Noise Ratio (PSNR) is the simplest and most used quality metric. The Visible Differences Predictor (VDP) is a representative example of models based on Human Visual System, [5]. The Mean Structural SIMilarity index (SSIM) is simple to implement and has attracted considerable attention in the literature, see [15] for applications on image and video processing and [6] for a statistical and mathematical assessment. It is worth to notice that the proposed approach could be extended to any FR image quality metric. The intent of this work is not to provide a complete study on all the known IQM's but rather to open a new direction for image quality analysis and comparison. Section 2 describes Weber's law and the three quality metrics considered in this study. Section 3 discusses the relationship between Weber's law and uniform color spaces. Section 4 states a photometric invariance law and shows numerical simulations that assess whether a given quality metric is compliant with

this photometric invariance law. The last section is devoted to concluding remarks and conclusion.

2. THE THREE QUALITY METRICS STUDIED

Three image quality metrics are considered in this study.

2.1. Peak Signal to Noise Ratio (PSNR)

The *Peak Signal to Noise Ratio* (PSNR) is the simplest and widely used metric. It is given by:

$$PSNR = 10\log_{10}\left(\frac{(g_{\max})^2}{\frac{1}{MN}\sum_{mn}\left(g^R_{mn} - g^D_{mn}\right)^2}\right) \quad (1)$$

where g^R_{mn} and g^D_{mn} stand for the pixel intensity in the reference and distorted image, respectively ranging from 0 to $g_{\max}$. It is found that this metric is invariant under some additive changes, we shall refer to such property as *additive invariance*.

2.2 Visible Differences Predictor (VDP)

The second image quality metric used here, is the *Visible Differences Predictor* (VDP). It is considered as one of the most representative of HVS-inspired image quality metrics. Let us recall some relevant quantities and notions related to this metric. One of the most important perceptual notion is the contrast. There are many definitions and there is no unified theory for defining this measure. Michelson contrast is one of the most known and is defined as:

$$C = \frac{L_{\max} - L_{mean}}{L_{mean}} \quad (2)$$

where $L_{\max}$ and L_{mean} refer respectively to the maximum and mean luminance of a sinusoidal stimulus pattern. Note that for a given image this contrast measure provides a unique and global value.

Another property exploited in the VDP is related to the nonlinear response of the HVS. In VDP it is expressed as a nonlinear point transform relating the visual sensitivity R to the luminance L. For digital images it is defined as a pixel-wise transform. For pixel (m,n), the visual response is related to the luminance by: $\frac{R_{mn}}{R_{\max}} = \frac{L_{mn}}{L_{mn} + (c_1 L_{mn})^b}$ where c_1 =12.6cd.m^{-2} and b=0.63.

Another HVS property used in VDP and related to the visual acuity is the Contrast Sensitivity Function. The original and distorted mages are then filtered by the CSF.

A Cortex transformation is then applied. It is a filter bank with 6 orientations and 5 spatial frequencies with equal log bandwidth. The lower-frequency channel has no orientation. For each subband (k,l) and for both images, the contrast $C_{kl}[m,n]$ is given by

$$C_{kl}[m,n] = \frac{B_{kl}[m,n]}{B_K} \quad (4)$$

where $B_{kl}[m,n]$ is the value of the filtered image and B_K is either the output of the lower-frequency channel (and hence a function of location m,n) or the spatial average of the image. This contrast is slightly different from Weber's contrast in that it is a function of pixel coordinate and subband, it is computed separately for the reference image and for the distorted image. It remains consistent with Weber's law. Indeed the reference and the distorted image have close background luminance and a just-noticeable luminance change is still proportional to that background luminance.

Subband thresholds are computed to model the masking effect. These thresholds may alter slightly the consistency with Weber's law.

$$\begin{aligned} T^{kl}_i[m,n] &= \left(1 + \left(k_1\left(k_2\left|B^{kl}_i[m,n]\right|\right)^s\right)^b\right)^{1/b} \\ T^{kl}_m[m,n] &= \min\left(T^{kl}_R[m,n], T^{kl}_D[m,n]\right) \end{aligned} \quad (5)$$

where $i \in \{R,D\}$ refers either to the Reference image or the Distorted image and $k_1 = W^{1-\frac{1}{1-Q}}$ $k_2 = W^{\frac{1}{1-Q}}$ W=6, Q=0.70, b=4 and $s \in [0.65,1]$, (this parameter decreases with increased learning). These subband thresholds are used in the psychometric function which produces a probability-of-detection map.

$$\begin{aligned} \Delta C_{kl}[m,n] &= \left|C^R_{kl}[m,n] - C^D_{kl}[m,n]\right| \\ P_{mn} &= 1 - \prod_{k,l}\left(1 - e^{-\left(\frac{\Delta C_{kl}[m,n]}{T^{kl}_m[m,n]}\right)^\beta}\right) \end{aligned} \quad (6)$$

2.3. Structural SIMilarity Index (SSIM)

The *Structural SIMilarity (SSIM) Index* is based on the assumption that the human visual system is highly adapted to extract structural information from the viewing field. It follows that a measure of structural information change can provide a good approximation to perceived image distortion.

Local statistics μ^R_j, μ^D_j, σ^R_j, σ^D_j and $\sigma^{R,D}_j$ are computed within local 8×8 square window, which move

pixel-by-pixel over the entire reference and distorted image.

$$\mu_j = \sum_{m,n\in W_j} w_{mn} g_{mn}$$
$$\sigma_j = \sqrt{\sum_{m,n\in W_j} w_{mn}\left(g_{mn}-\mu_j\right)^2} \quad (7)$$
$$\sigma_j^{R,D} = \sum_{m,n\in W_j} w_{mn}\left(g_{mn}^R-\mu_j^R\right)\left(g_{mn}^D-\mu_j^D\right)$$

where w_{mn} is either identity or an 11×11 circular symmetric Gaussian weighting function with standard deviation of 1.5 samples, normalized to unit sum.

The mean luminance distortion, the contrast distortion and the structural distortion are given by

$$l_j = \frac{2\mu_j^R\mu_j^D + C_1}{\left(\mu_j^R\right)^2+\left(\mu_j^D\right)^2+C_1}$$
$$c_j = \frac{2\sigma_j^R\sigma_j^D + C_2}{\left(\sigma_j^R\right)^2+\left(\sigma_j^D\right)^2+C_2} \quad (8)$$
$$s_j = \frac{\sigma_j^{R,D}+C_3}{\sigma_j^R\sigma_j^D+C_3}$$

where $C_1=(K_1 g_{\max})^2$, $C_2=(K_2 g_{\max})^2$, usually K_1=0.01, $K_2=0.03$ and for simplification $C_3=C_2/2$. The three distortion measurements are collapsed into an index

$$SSIM_j = l_j^\alpha c_j^\beta s_j^\gamma \quad MSSIM = \frac{1}{M}\sum_M SSIM_j \quad (9)$$

where $\alpha=\beta=\gamma=1$ for simplification.

SSIM is designed so as to be consistent with Weber's law, see [16] and note that the constants C_1, C_2 and C_3 may alter this property.

3. DISCUSSION ON WEBER'S CONTRAST

3.1 Psycho-visual tests

There has been extensive work into psycho-visual tests on Weber's contrast with sinusoidal gratings of different frequencies [3], but far less with different luminance backgrounds.

Figure 4 is reproduced from [9], it shows contrast sensitivity curves for gratings of different frequencies, orientations and luminance backgrounds, it tells us that if the average luminance of an image is 10 times higher, the contrast threshold is 2.25 times lower ($65/29\approx147/65\approx2.25$). Very similar results can be found in [1] and [17].

In [7], Weber's contrast is replaced with a power-law contrast model.

$$C = L^{1/3} - L_B^{1/3} \approx \frac{1}{3}L_B^{1/3}C_W \quad (10)$$

This contrast is consistent with data shown on figure 4: when L_B is 10 times higher and C_W is 2.25 times lower,

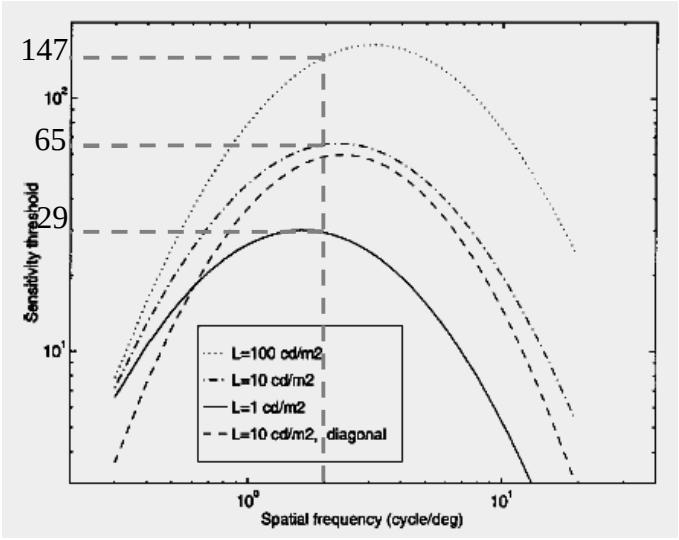

Figure 4: Contrast sensitivity curves for different gratings, reproduced from [9].

C is only 0.96 times lower, almost constant. As mentioned in [7], it is because this contrast inherits the uniform perceptual properties of the Lab system, that it is well suited.

3.2 Deriving contrast threshold from uniform colour spaces

CIE 1976 (L* a* b*) colour space is defined as

$$L^* = 116\left(\frac{Y}{Y_n}\right)^{1/3} \quad if \quad \frac{Y}{Y_n} > 0.00856$$
$$L^* = 903.3\frac{Y}{Y_n} \quad else \quad (11)$$

and

$$a^* = 500\left(f\left(\frac{X}{X_n}\right)-f\left(\frac{Y}{Y_n}\right)\right)$$
$$b^* = 200\left(f\left(\frac{Y}{Y_n}\right)-f\left(\frac{Z}{Z_n}\right)\right) \quad (12)$$

where X,Y,Z are the tristimulus values that specify the colour of a sample and $X_n, Y_n Z_n$ specify the illuminant; and

$$f(t) = t^{1/3} \quad if \quad t > 0.00856$$
$$f(t) = 7.787t + \frac{16}{116} \quad else \tag{13}$$

two colours look different if their Euclidean distance ΔE is greater than 2:

$$\Delta E = \sqrt{\left(L_1^* - L_2^*\right)^2 + \left(a_1^* - a_2^*\right)^2 + \left(b_1^* - b_2^*\right)^2} \tag{14}$$

By restricting to grey inputs and to luminance not too low, definition of CIE 1976 (L* a* b*) is modified into

$$L_1^* = 116\left(\frac{Y_1}{Y_n}\right)^{1/3} - 16 \quad L_2^* = 116\left(\frac{Y_2}{Y_n}\right)^{1/3} - 16$$
$$\Delta E = \left|L_1^* - L_2^*\right| \tag{15}$$

Substituting $Y_1 = L$ and $Y_2 = L + \Delta L$ into (15) and denoting Y_n as $L_{\mathrm{Re}\,f}$, we can approximate the new expression by its Taylor series at $\frac{\Delta L}{L} \to 0$.

$$\Delta E = L_2^* - L_1^* = 116\left(\frac{L}{L_{\mathrm{Re}\,f}}\right)^{1/3} \times \frac{1}{3}\frac{\Delta L}{L} \tag{16}$$

These formulas are now applied not only to uniform regions but also to textures and edges, L is the average luminance. Calling L the background luminance enables us to derive Weber's contrast threshold from (16).

$$C_W = \frac{\Delta L}{L} = \left(\frac{6}{116}L_{\mathrm{Re}\,f}^{1/3}\right)\frac{1}{L^{1/3}} \tag{17}$$

3.3 Gamma Correction

Gamma correction was initially a relationship between the input voltage and the output luminance L of a cathode ray tube [4]. Today the main purpose of gamma correction in video, desktop graphics is to code luminance into a perceptually uniform domain [12]. Gamma correction is defined as

$$L = L_{\max}\left(\frac{g}{g_{\max}}\right)^{\gamma} \tag{18}$$

where $\gamma \in [2.2, 2.4]$. The largest source of variation in γ is caused by varying settings of the black level control of monitors [12]. Substituting L with $L + \Delta L$ and g with $g + \Delta g$ into (18), we can approximate (18) by its Taylor series at $\frac{\Delta L}{L} \to 0$.

$$L + \Delta L = L_{\max}\left(\frac{g + \Delta g}{g_{\max}}\right)^{\gamma}$$
$$\approx L_{\max}\left(\frac{g}{g_{\max}}\right)^{\gamma} \times \left(1 + \gamma\frac{\Delta g}{g}\right) \tag{19}$$

Substituting (18) into (19) yields

$$\Delta L = L\gamma\frac{\Delta g}{g} \tag{20}$$

Expressing g as a function of L using (18) and calling L the background luminance yields a slightly different Weber's contrast threshold:

$$C_W = \frac{\Delta L}{L} = \left(\gamma L_{\max}^{1/\gamma}\frac{\Delta g}{g_{\max}}\right)\frac{1}{L^{1/\gamma}} \tag{21}$$

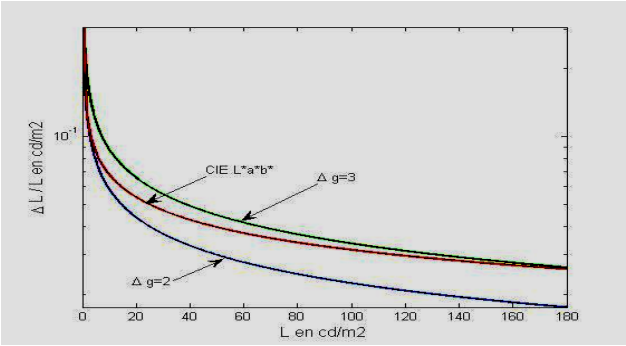

Figure 5: Weber's contrast C_W as a function of background luminance L

Figure 5 shows curves of Weber's contrast with respect to background luminance for (17) and (21), we see that these curves are very close.

4. A NEW PHOTOMETRIC INVARIANCE LAW

We propose a photometric invariance law.

Definition

A quality metric is said to follow the photometric invariant law if, for any image L_{mn}, *for any local distortion of small intensity* ΔL_{mn}, *and for all* $\lambda \in (0,1]$, *there exists* α *near 1/3 such that* $\lambda' = \lambda^{1-\alpha}$ *and*

$$Q\left(L_{mn}, L_{mn} + \Delta L_{mn}\right) = Q\left(\lambda L_{mn}, \lambda L_{mn} + \lambda'\Delta L_{mn}\right) \tag{22}$$

PSNR is additive invariant and is expected to be almost consistent with the photometric invariance law, that is with $\alpha = 1/\gamma = 0.42$ (21). By design, VDP and SSIM are consistent with Weber's law and hence the parameter α extracted is expected to be zero.

5. NUMERICAL SIMULATIONS

We have performed numerical simulations with two test images using the following methodology.

Methodology

(1) Choose a reference image (one or several images).

(2) Choose different additive distortions at different locations with intensity close to visible threshold.

(3) Compute the reference value $R = Q(L_{mn}, L_{mn} + \Delta L_{mn})$

(4) Choose a set of (λ_k) *uniformly distributed in* $(0,1]$*, and for each* λ_k*, find* λ'_k *such that* $R = Q(\lambda_k L_{mn}, \lambda_k L_{mn} + \lambda'_k \Delta L_{mn})$

(5) Compute α *by linear regression of* $\ln(\lambda_k)$ *with* $\ln(\lambda'_k)$.

(6) Check that the extracted parameter, α *do not depend on steps 1 and 2.*

The first reference image is issued from Brodatz texture (figure 6), [2]. The other reference image is the Barbara image (figure 7). The distorted images were generated by embedding three different types of local distortions (see figure 6) into the two test images:

-a luminance increase on a region having the shape of a square,

-a luminance increase on a region having the shape of a triangle,

-a luminance increase nearby a luminance decrease

This was achieved at different locations and with intensities around just-noticeable distortion.

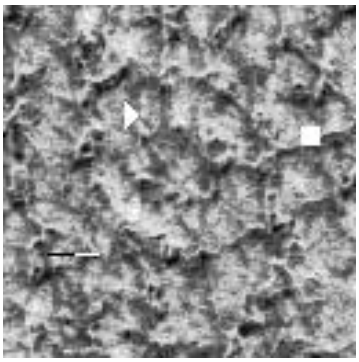

Figure 6: first test image and distortions used in the numerical simulations

Figure 7: second test image

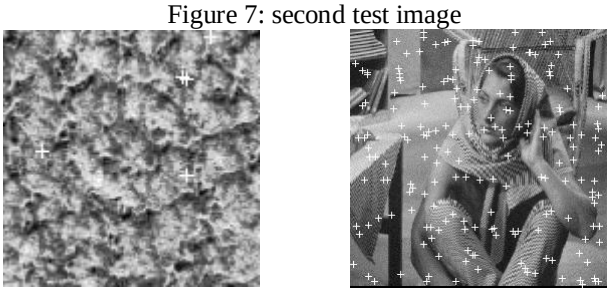

Figure 8: locations of embedded distortions

Note that in the fourth step a dichotomy algorithm is used as the mapping $\lambda' \mapsto Q(\lambda L_{mn}, \lambda L_{mn} + \lambda' L_{mn})$ is monotonic.

Results of numerical simulations[1] are shown on figure 9 and 10. They confirm the additive invariance of PSNR and Weber's law-like behaviour of VDP and SSIM with null constants. The masking effect may explain the slightly negative parameter α found for VDP. Simulations have also shown that SSIM with constants set to their default values is neither really consistent with Weber's law, nor with the photometric invariance law.

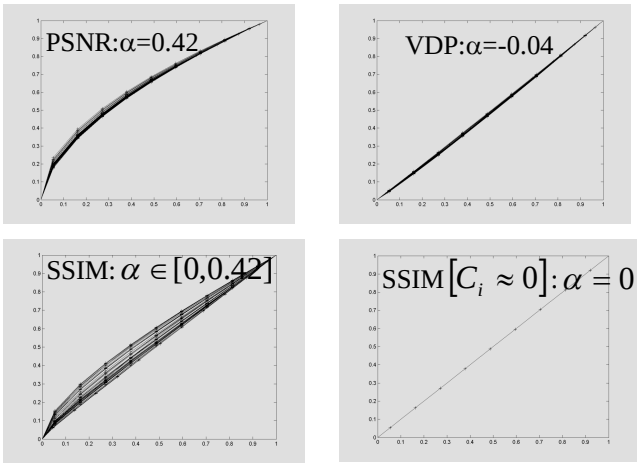

Figure 9: Relationship between λ and λ' for the three quality metrics PSNR, VDP and SSIM.

[1] We are grateful to Abdelhalim Mayache for implementing VDP in Matlab.

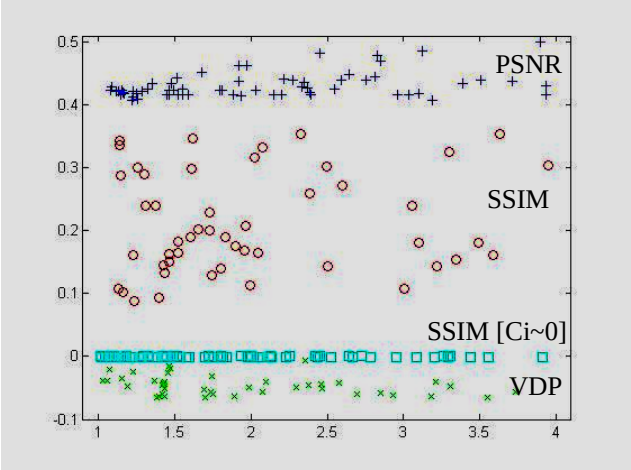

Figure 10: Parameter α extracted as a function of the intensity of the distortion measured in grey levels for PSNR, VDP and SSIM.

5. CONCLUSION

VDP and SSIM were designed to be consistent with Weber's law. However this law is not confirmed by psycho-visual testing. A new photometric invariance law inherited from uniform colour spaces is proposed and numerical simulations confirm the expectations.

Further work is necessary, to show how quality metric can be modified so as to be consistent with this photometric invariance law, and to investigate the effectiveness of the modified quality metrics.

REFERENCES

[1] P.G.J. Barten. "The effects of picture size and definition on perceived image quality". *Proceedings of the society for information display*, 30(2):67–71, 1989.

[2] P. Brodatz. "Textures: A Photographic Album for Artists and Designers". New York, Dover, 1966.

[3] T. Carney, S.A. Klein, and B. Beuttere D. Levif A.B. Watson A.J. Reevesg A.M. Norciac C. Chenc W. Makoush M.P. Ecksteini C.W. Tyler, A.D. Silversteind. "The development of an image/threshold database for designing and testing human vision models". *SPIE*, 3644, May 1999.

[4] R.G. Clapp and E.M. Creamer and S.W. Moulton and M.E. Partin and J.S. Bryan. "A New Beam Indexing Color Television Display System", *Proceedings, I.R.E.*, Vol. 44, September, 1956, Pages 1108-1114.

[5] S. Daly. "Visual Factors in Electronic Image Communications", chapter The Visible Differences Predictor: An Algorithm for the Assessment of Image Fidelity. 1993 http://vision.arc.nasa.gov/modelfest/documents/spie99.pdf.

[6] R. Dosselmann and X.D. Yang. "A comprehensive assessment of the structural similarity index". Signal, Image and Video Processing, page 9, 2009.

[7] Thomas Frese, Charles A. Bouman, and Jan P. Allebach. "A methodology for designing image similarity metrics based on human visual system models". In *Proceedings of SPIE/IS&T Conference on Human Vision and Electronic Imaging II*, pages 472–483, 1997.

[8] ITU R BT. 500-11. "Methodology for the subjective assessment of the quality of television pictures". Technical report, 2004.

[9] Y.-K. Lai and C.-C.J. Kuo. "A Haar wavelet approach to compressed image quality measurement." *Journal of Visual Communication and Image Representation*, pages 17–40, 2000.

[10] L. Mandic, S. Grgic, and M. Grgic. "Comparison of color difference equations". In 48th *International Symposium ELMAR*, Zadar, Croatia, June 2006.

[11] N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola, M. Carli, and F. Battisti. "Color image database for evaluation of image quality metrics". In *IEEE 10th Workshop on Multimedia Signal Processing*, October 2008.

[12] C. Poynton. "The rehabilitation of gamma". In *SPIE Human Vision and Electronic Imaging III*, pages 232–249, 1998.

[13] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik. "Live image quality assessment database", rel. 2, 2005.

[14] A.B. Watson. "Visual detection of spatial contrast patterns: Evaluation of five simple models". *Optics Express*, 6(1):12–33, 2000.http://www.opticsexpress.org/oearchive/source/14103.htm.

[15] Z. Wang and A.C. Bovik. "Mean-squared error: Love it or leave it. A new look at signal fidelity means". *IEEE Signal Processing Magazine*, 28(1):98–117, 2009.

[16] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. "Image quality assessment: From error measurement to structural similarity". *IEEE Transactions on Image Processing*, 13(1):14, January 2004.

[17] S.Westland, H. Owens, V. Cheung, and I.Paterson-Stephens. "Model of luminance contrast-sensitivity function for application to image assessment". *Color Research & Applications*, 31(4):5, August 2006.