Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2013) 21: 603 – 612
© TÜBİTAK
doi:10.3906/elk-1111-11

A new color image quality measure based on YUV transformation and PSNR for human vision system

Yıldıray YALMAN*, **İsmail ERTÜRK**
Department of Computer Engineering, Turgut Özal University, 06010 Ankara, Turkey

Received: 03.11.2011 • **Accepted:** 22.02.2012 • **Published Online:** 22.03.2013 • **Printed:** 22.04.2013

Abstract: Various methods for measuring perceptual image quality attempt to quantify the visibility of differences between an original digital image and its distorted version using a variety of known properties of the human vision system (HVS). In this paper, we propose a simple and effective full-reference color image quality measure (CQM) based on reversible luminance and chrominance (YUV) color transformation and peak signal-to-noise ratio (PSNR) measure. The main motivation of this new measure relies on a unique feature of the human eye response to the luminance and color. Experimental studies about the applicability of the CQM on a well-known test image under 6 different distortions, both perceivable by the human vision system and with the same PSNR value (i.e. 27.67), are presented. The CQM results are obtained as 39.56, 38.93, 38.08, 37.43, 37.10, and 36.79 dB for each distorted image, showing that image quality of the first image is noticeably higher than the others with respect to the same PSNR value. This conclusion attests that using the CQM together with the traditional PSNR approach provides distinguished results.

Key words: Image quality measure, human vision system, color transformation, PSNR, YUV

1. Introduction

A fundamental task in many image processing applications is the visual evaluation of a distorted image. There are many measures for examining image quality, such as the mean structural similarity, mean absolute error, mean square error (MSE), and peak signal-to-noise ratio (PSNR). The simplest and the most widely used full-reference quality measure is the MSE. It is computed by averaging the squared intensity differences of distorted and original image pixels, along with the related quantity of the PSNR. Although the PSNR is mostly used in the literature, it is not very well matched with the perceived visual quality [1–4]. In the last decade, much effort has gone into the development of image quality measures that take advantage of well-known characteristics of the human vision system (HVS). Sparking from these facts, the proposed color image quality measure (CQM) follows a new strategy of changing the implementation method of the PSNR.

The proposed CQM is based on 2 major parts. First of all, a reversible color transformation is realized from red, green, and blue (RGB) to luminance and chrominance (YUV) using an original image and its distorted version. For color images, a color transformation is originally used as a preprocess before intracomponent coding in any image compression application. Here, Y is the luminance component, while U and V are the blue-difference and red-difference components of the YUV, respectively. In addition, the number of color sensors (cones) is different from the number of luminance sensors (rods) in the human eye (rods > cones) [5,6]. Thus, the light sensitivity and color sensitivity of the HVS are different from each other. The first part of the CQM

*Correspondence: yyalman@turgutozal.edu.tr

correlates this natural fact to the YUV transformation. Secondly, thanks to the YUV transformation and our innovative idea of computing the distortion on the luminance component and the distortion on the color components in an image using the classical PSNR formula, it is separately possible to take into account the number of rods and cones weightedly.

The rest of the paper is organized as follows. Section 2 summarizes digital image fundamentals and YUV color transformation and its details. Section 3 describes the new paradigm for the proposed image quality measure based on error calculation between the original and distorted images, employing PSNR and natural properties of the human eye. Section 4 presents the experimental results of the proposed CQM and a detailed evaluation study, followed by final remarks in the last section.

2. Fundamentals of the digital image and YUV color transformation

A digital image is represented with a 2-dimensional array composed of M rows and N columns. In general, line and column indexes are shown as y, x, or c. Each element of that serial is called a pixel (Figure 1). Basically, pixels are valued as either 0 or 1. Images that are formed through such pixels are called binary images. The '1' and '0' values represent light and dark areas or objects and backgrounds (the environmental background in front of or on which an object is situated), respectively. Digital image files are used in the form of 16 or 24 bits as color images while gray-level images are used as 8 bits, and they are used in many different areas, such as remote sensing or medical applications [7].

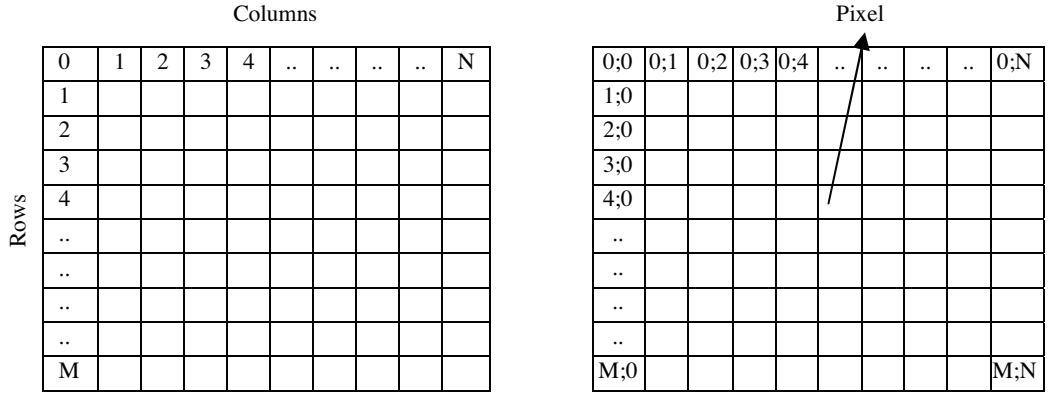

Figure 1. Digital image structure.

For a 24-bit color image (Figure 2a), a pixel is identified using 3 bytes, i.e. 8 bits for each color channel of red, green, and blue (Figures 2b, 2c, and 2d). Color images are also classified into 2 categories, as compressed and uncompressed. Uncompressed images, such as BMP and GIF formats, can host much more data than compressed images, such as JPEG format, considering the data-hiding applications [8].

Color transformation is usually used as a preprocess before the intracomponent coding in RGB color image compression [9]. A commonly preferred color transformation is from RGB (Figure 3a) to YUV (Figures 3b, 3c, and 3d). YUV was originally adopted from the JPEG and JPEG2000 standards [10]. The forward transformation and its inverse formulas (ITU-R BT.601) are given in Eqs. (1) and (2), respectively. YUV transformation not only decorrelates the original color components, but also discards some information to get lower entropy. The discarded information cannot be recovered, and thus the transformation is lossy [11].

Figure 2. An RGB image (a) and R (b), G (c), and B (d) channels.

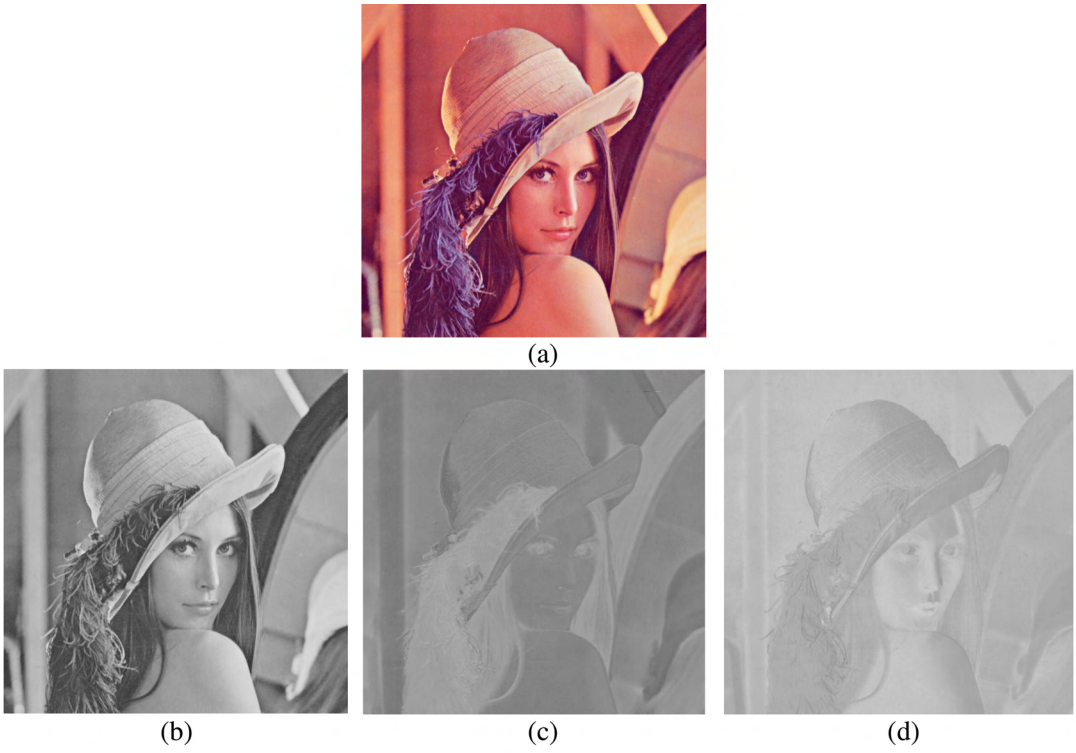

Figure 3. An RGB image (a) and Y (b), U (c), and V (d) channels.

$$\left.\begin{aligned} Y &= 0.257R + 0.504G + 0.098B + 16 \\ U &= -0.148R - 0.291G + 0.439B + 128 \\ V &= 0.439R - 0.368G - 0.071B + 128 \end{aligned}\right\} \tag{1}$$

$$\left.\begin{aligned} R &= 1.1644(Y-16) + 1.596(V-128) \\ G &= 1.164(Y-16) - 0.813(V-128) - 0.391(U-128) \\ B &= 1.164(Y-16) + 2.018(U-128) \end{aligned}\right\} \tag{2}$$

An approximated reversible YUV color transformation that originates from the JPEG2000 standard and is called the RCT in the present work is given in Eq. (3) [12,13]. The proposed CQM preferably utilizes the RCT rather than the classical YUV, as its goal is not focused on digital image compression. Therefore, data loss never occurs when the RGB to $\mathrm{YUV}_{(RCT)}$ (where the Y channel refers to the luminance, while U and V refer to the color information in an image) transformation is realized in the proposed CQM. An important feature of the human eye is its different response to the luminance and color components directly related to the natural features of the eye, which are utilized and reflected into the proposed CQM with different ratios, as detailed in the following section.

$$\begin{array}{l|l} Y = (R + 2G + B)/4 & G = Y - (U+V)/4 \\ U = R - G & R = U + G \\ V = B - G & B = V + G \end{array} \tag{3}$$

3. Description of the proposed color image quality measure

In this section, the proposed CQM, with the details of its fundamental parts, i.e. the PSNR, the human eye structure, and reversible YUV transformation ($\mathrm{YUV}_{(RCT)}$), are presented.

3.1. PSNR

The PSNR measurement is used to evaluate a distorted image quality with respect to its original version as an error measure during experimental works. A computed PSNR value indicates the quality approximation between the distorted image and the original image.

The MSE should be computed first, as given in Eq. (4) [14,15], and then the PSNR can be derived as in Eq. (5) [16,17]. Here, "O" and "D" are the original and the distorted image pixel values (binary), respectively, to be compared, and the image size is "$m \times n$". Note that Eq. (4) is specified for only monochrome images, meaning that the denominator of Eq. (5) is multiplied by a factor of 3 for color images.

$$MSE = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|O(i,j) - D(i,j)\|^2 \tag{4}$$

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right) \tag{5}$$

Here, MAX is the peak value of the pixels in an image. MAX is 255 when pixels are presented in an 8-bit format. Theoretically, the higher the PSNR value is. the better the image processing is; however, practically,

there are some problems reported in the literature about the use of the PSNR for image quality assessment [1–4]. In these studies, a test image is distorted by a wide variety of distortions, such as Gaussian, sharpen, salt & pepper, JPEG compression, median, and blurring, in order to examine the PSNR results for quality evaluations. Although the test image is exposed to different types of distortion, the obtained PSNR results can be the same, and thus its use in a genuine evaluation is not feasible. This implies that the PSNR parameter solely does not enable distinguishing results for all of the probable cases. As a result, the PSNR for image quality assessment can be used easily, but cannot be adequate alone for a statistical visual evaluation. Regarding this important fact, the proposed CQM has been developed and presented in this paper.

3.2. The structure of the human eye

In this subsection, the main characteristics of the human eye are briefly given. The human eye consists of the sclera, the cornea, ocular muscles, the choroid, the retina, the iris, the muscular tissue, the lens, and the pupil (Figure 4) [18].

After the light enters the pupil, it directly falls onto the lens of the eye, where it is improved before passing through to the retina. The lens is a biconvex composition that is encased in a thin transparent layer. It not only refracts but also focuses the incoming light onto the retina for optic processing [18].

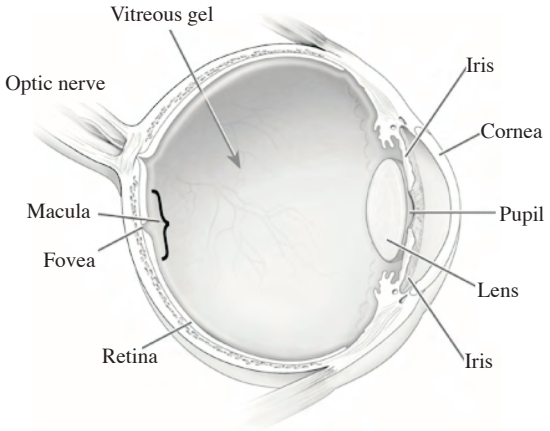

Figure 4. View of the human eye [18].

There are 2 main types of bulk photoreceptors in the human retina, named rods and cones. The former are responsible for vision at especially low light levels. They do not provide color vision and have almost no effect on spatial perception [4,17]. Unlike the rods, cones are active at higher light levels and are responsible for color vision and spatial perception. The fovea part of the human eye hosts the cones and rods. There are approximately 150,000 cones/mm^2 (7 million) and 200,000 rods/mm^2 (120 million) in the fovea [19,20]. The rods are extremely sensitive to light. On the other hand, the cones supply the eye's color sensitivity [5,6,21].

Measured density curves for the rod and cone photoreceptors show a huge density of cones in the fovea centralis that is related to both color vision and the highest visual perception (Figure 5). The visual inspection of a small detail is associated with focusing light from that detail onto the fovea centralis. On the other hand, there are no rods on the macula. At a few degrees away from the macula, the rods' density increases extremely and spreads over a large part of the fovea. These rods are what make night vision, motion detection, and peripheral vision easy [19,20].

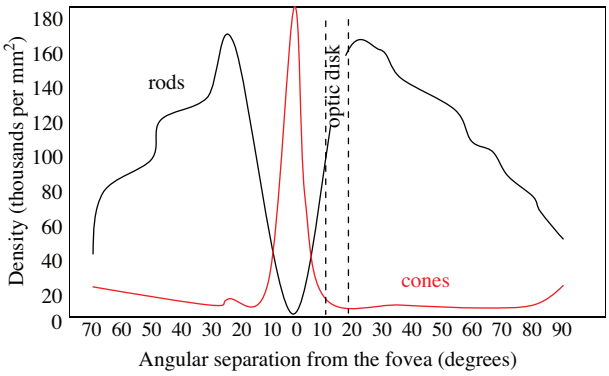

Figure 5. Rod and cone densities on the retina [6,19,21].

The weights on the human perception of these cone and rod sensors (i.e. C_W and R_W) are calculated below:

$$C_W = 7,000,000/(120,000,000 + 7,000,000) = 0.0551, \tag{6}$$

$$R_W = 120,000,000/(120,000,000 + 7,000,000) = 0.9449. \tag{7}$$

The 1st term, C_W, as seen in Eq. (6), is the weight on the human perception of the cones, and the 2nd term, R_W, as seen in Eq. (7), is the weight on the human perception of the rods. Differing from the classical PSNR approach, the proposed CQM outstandingly considers this natural fact to consequently enable the obtaining of more accurate and differentiating image quality results.

3.3. The CQM calculation

The proposed image quality measure CQM is obtained in 3 stages (Figure 6). At first, an original RGB image and its distorted version are transformed to the $YUV_{(RCT)}$ images by employing the RCT equations explained in previous sections. Next, the PSNR quality of each $YUV_{(RCT)}$ channel (Y, U, and V) is calculated separately. At the last step, the CQM value is calculated using Eq. (8), shown below:

$$CQM = (PSNR_Y \times R_W) + \left(\frac{PSNR_U + PSNR_V}{2}\right) \times C_W, \tag{8}$$

where the CQM is composed of our new discovery of the weighted luminance quality measure $(PSNR_Y \times R_W)$ and weighted color quality measure $\left(\left(\frac{PSNR_U + PSNR_V}{2}\right) \times C_W\right)$ components. The inclusion of this new weighted approach, considering the human eye's different responses to luminance and color, leads to the superiority of the proposed CQM measure over the classical PSNR-based image quality evaluations.

3.4. Experimental results and discussion

A test image (Figure 7a) and its sample distorted versions, shown in Figure 7, are used for a detailed experimental study. They are tuned with all of the well-known distortions to yield the same PSNR values relative to the original image. The CQM and classical PSNR results for all of the Figures are given in the Table. The PSNR results for Figures 7b, 7c, 7d, 7e, 7f, and 7g are the same (i.e. 27.67 dB), meaning that the imposed different distortions have the same effect on the test image, contrary to fact based on at least the basic human vision evaluation. This implies that the PSNR parameter solely does not enable distinguishing for all of the probable

cases between the original image and the distorted images. Therefore, it exhibits a very poor performance for all kinds of numeric image quality measures. On the other hand, using the proposed CQM, it can be statistically distinguished that not only is the quality of the distorted image given in Figure 7g 36.79 dB, which is much worse than those in Figures 7b, 7c, 7d, 7e, and 7f, but they have all returned differentiating values that match well with the human vision perception results.

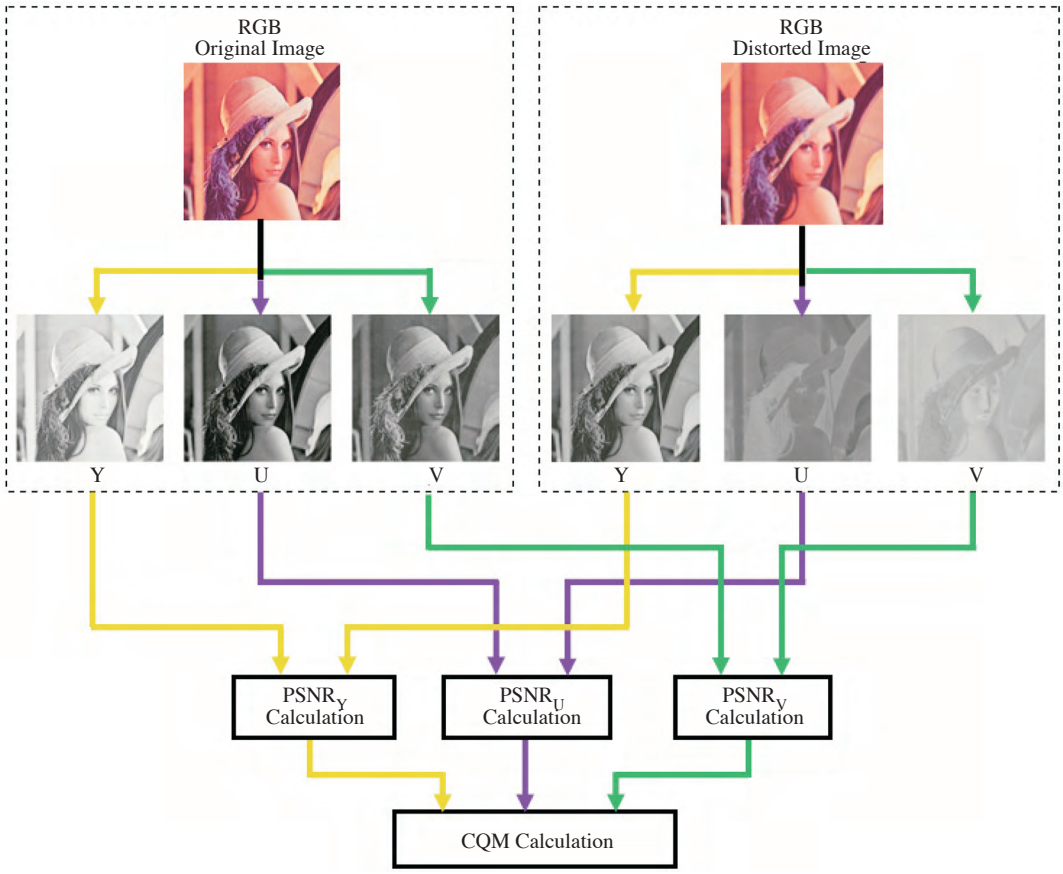

Figure 6. Block diagram of the proposed CQM calculation.

In this section, we consider another counterpart perceptual image quality measure presented in the literature, named the PSNR-HVS-modified (PSNR-HVS-M) [22], that is used for a comparative evaluation of the proposed CQM. Examining the distorted test images, the CQM, classical PSNR, and PSNR-HVS-M values are computed and presented in the Table.

Considering the PSNR-HVS-M, it can be stated that this image quality metric also produces differentiating results, like the proposed CQM, with respect to the classical PSNR. However, the PSNR-HVS-M results prove to be false, as they indicate that the image quality for Figure 7d is better than that of Figure 7c, and that of Figure 7f is better than those of Figures 7c and 7e, as opposed to a straightforward visual analysis of all of the Figures.

The experimental study presented in this section is based on the "Lena" image under different well-known distortions. These images and an easy MATLAB implementation of the proposed quality measure CQM are available online at http://www.turgutozal.edu.tr/yyalman/contents/yyalman/files/CQM.zip.

Figure 7. Evaluation of the "Lena" test image contaminated by some distortions: (a) Original "Lena" image, 512 × 512, 24 bits/pixel; (b) Gaussian noise, PSNR = 27.67, CQM = 39.56; (c) sharpen noise, PSNR = 27.67, CQM = 38.93; (d) salt & pepper noise, PSNR = 27.67, CQM = 38.08; (e) JPEG compression, PSNR = 27.67, CQM = 37.43; (f) median noise, PSNR = 27.67, CQM = 37.10; (g) blurring noise, PSNR = 27.67, CQM = 36.79.

Table. PSNR-HVS-M and CQM results of a test image under different types of distortions.

Test image (Lena)	Distortion type	PSNR (dB)	PSNR-HVS-M (dB)	CQM (dB)
Figure 7b	Gaussian	27.67	29.08	39.56
Figure 7c	Sharpen	27.67	24.75	38.93
Figure 7d	Salt & pepper	27.67	28.77	38.08
Figure 7e	JPEG compression	27.67	24.69	37.43
Figure 7f	Median	27.67	25.40	37.10
Figure 7g	Blurring	27.67	24.42	36.79

The CQM can be also applied to local regions of an image using a sliding window approach. This approach can be very valuable for examining any specific portion of a test image in order to clarify whether only a dedicated part of the image is distorted or worked out for any reason. For example, starting from the top-right corner of an image, a sliding window of size X $\times$ Y moves horizontally and vertically all around the image until the bottom-right corner is reached. At the jth step, the local quality result of the CQM_j can be computed within the sliding window, similar to the classical CQM calculation (Eq. (8)). Afterwards, it is straightforward that the lowest CQM result shows the worst distorted test image portion. In addition, the CQM can be easily utilized as an image quality measure for examining data-hiding applications and compression algorithm performance.

4. Conclusions

An integrated $\mathrm{YUV}_{(RCT)}$ and PSNR-based image quality measure named the CQM is proposed in this paper. The CQM is developed distinctively taking into account the biological characteristics of the human eye. The CQM is obtained at 3 different stages. At first, an original RGB image and its distorted version are transformed to the $\mathrm{YUV}_{(RCT)}$ images by employing the RCT equations. Next, the PSNR value of each $\mathrm{YUV}_{(RCT)}$ channel (Y, U, and V) is calculated separately. At the last step, the CQM value is computed using an important discovery of weighted luminance quality measure and weighted color quality measure components regarding the natural eye features. The inclusion of this new approach considering the human eye's different responses to luminance and color leads to the superiority of the proposed CQM measure over the classical PSNR-HVS-M–based evaluations.

The experimental results indicate that the CQM significantly outperforms the PSNR-HVS-M for analyzing the quality of images under different types of distortions (salt & pepper, JPEG, blurring, etc.).

Acknowledgment

The authors would like to thank the editor and anonymous reviewers for their invaluable comments and suggestions.

References

[1] Z. Wang, A.C. Bovik, H.D. Sheikh, E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", IEEE Transactions on Image Processing, Vol. 13, pp. 600–612, 2004.

[2] B. Girod, What's Wrong With Mean-Squared Error, Digital Images and Human Vision, A.B. Watson, Ed., Cambridge, MA, USA, MIT Press, pp. 207–220, 1993.

[3] Z. Wang, "Rate scalable foveated image and video communications", PhD, Department of Electrical and Computing Engineering, University of Texas at Austin, 2001.

[4] A.M. Eskicioglu, P.S. Fisher, "Image quality measures and their performance", IEEE Transactions on Communications, Vol. 43, pp. 2959–2965, 1995.

[5] Hecht Eugene, Optics, 2nd ed., Boston, Addison Wesley, 1987.

[6] Georgia State University, The Rods and Cones of the Human Eye, available at http://hyperphysics.phy-astr.gsu.edu/hbase/vision/rodcone.html#c2.

[7] S. Candemir, Y.S. Akgül, "Statistical significance based graph cut regularization for medical image segmentation", Turkish Journal of Electrical Engineering & Computer Sciences, Vol. 19, pp. 957–972, 2011.

[8] M.D. Hassan, "Comparison for steganalysis approaches", MSc, Gazi University, Ankara, 2008.

[9] E. Elbaşı, "Robust multimedia watermarking: hidden Markov model approach for video sequences", Turkish Journal of Electrical Engineering & Computer Sciences, Vol. 18, pp. 159–170, 2010.

[10] C. Christopoulos, A. Skodras, T. Ebrahimi, "The JPEG2000 still image coding system: an overview", IEEE Transactions on Consumer Electronics, Vol. 46, pp. 1103–1127, 2000.

[11] Y. Chen, P. Hao, "Integer reversible transformation to make jpeg lossless", IEEE International Conference on Signal Processing, Vol. 1, pp. 835–838, 2004.

[12] M.J. Gormish, E.L. Schwartz, A. Keith, M. Boliek, A. Zandi, "Lossless and nearly lossless compression of high-quality images", SPIE/IS&T Conference on Very High Resolution and Imaging II, Vol. 3025, pp. 62–70, 1997.

[13] A. Mesut, "New methods in data compression", PhD, Trakya University, Edirne, 2006.

[14] H.T. Sencar, M. Ramkumar, A.N. Akansu, Data Hiding Fundamentals and Applications, New York, Elsevier Academic Press, 2004.

[15] C.C. Chang, C.C. Lin, Y.H. Chen, "Reversible data-embedding scheme using differences between original and predicted pixel values", IET Information Security, Vol. 2, pp. 35–46, 2008.

[16] A.N. Netravali, B.G. Haskell, Digital Pictures: Representation, Compression and Standards", New York, Plenum Press, 1995.

[17] M. Rabbani, P.W. Jones, Digital Image Compression Techniques, Washington, SPIE Optical Engineering Press, 1991.

[18] Freedom Scientific, Vision Loss Facts: Anatomy of the Eye, available at http://www.freedomscientific.com/resources/vision-anatomy-eye.asp.

[19] E.D. Montag, Chester F. Carlson Center for Imaging Science: Rods & Cones, available at http://www.cis.rit.edu/people/faculty/montag/vandplite/pages/chap_9/ch9p1.html.

[20] S.J. Williamson, H.Z. Cummins, Light and Color in Nature and Art, New York, Wiley, 1983.

[21] C.W. Oyster, The Human Eye: Structure and Function, Sunderland, Massachusetts, Sinauer Associates, 1999.

[22] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, V. Lukin, "On between-coefficient contrast masking of DCT basis functions", 3rd International Workshop on Video Processing and Quality Metrics for Consumer Electronics, pp. 1–4, 2007.