Evaluation of Image Color Metrics for Watermarked Color Images

André Rodrigues, Fabiana Fernandes, Zélia Peixoto e Flávia Freitas

***Abstract*— This paper presents a comparative study of performance metrics applied to color images with invisible digital watermark. The imperceptibility of the watermark and the distortion that can cause in the image are important factors in the evaluation of the invisible watermark insertion algorithms and usually this is done through the performance metrics. Thus, the aim of this article is to analyze and evaluate the results obtained through different performance metrics. The metrics chosen through literature were: PSNR, MPSNR, MSSIM, ΔE_{ab} and ΔE_{00}.**

Keywords—Image Quality, Color Images, Color Difference, Digital Watermark, Evaluation of Imperceptibility, PSNR, MPSNR, MSSIM, ΔE_{ab}, ΔE_{00}.

I. INTRODUCTION

A digital watermark applied to the authenticity verification of documents is an issue of great relevance and topic of current research. For example, watermarks are used in personal documents such as passports, driving licenses and other identity documents, because they are widely used and easy targets for fakes [1]. Unauthorized reproduction of documents, books and photographs (photocopy) was already a problem before the popularization of the Internet. However, with this popularity, the illegal appropriation of documents, books and digital photography became a bigger problem [2], coupled not only with the increased processing power of computers and monitors with higher resolution, size and color reproduction capability, but also with high-speed Internet connections.

The visible watermark is designed to be noticeable to people and typically contains a visual message or logo of a company indicating the copyright of the digital media used. This type of watermark is convenient for an immediate declaration of intellectual property and usually eliminates the commercial value of the document to an illegal use, for it assumes that copyright laws protect this document. Invisible watermark is designed to be imperceptive, without sacrificing the quality of digital media. However, it can be detected and extracted when needed [3].

Regarding color digital images, the use of watermarks is useful to identify the source, author, creator and/or owner of a digital media even if its free distribution has been authorized. Thus, the imperceptibility of a watermark and, consequently, some distortion included by it in the image are the most important factors in the evaluation of the invisible watermark insertion algorithms.

In this context, this paper presents an evaluation of the most usual image color metrics, taking into account the same insertion algorithm and the comparison between original and watermarked images to provide an assessment of the imperceptibility of the watermark.

This paper is organized as follows. Section II describes the watermark insertion algorithm and the performance metrics applied to its evaluation and Section III presents the experimental results. In Section IV one can find the conclusion of this work.

II. EQUATIONS AND THEOREMS

A. Watermark insertion

In this paper, a print-and-scan resilient watermarking technique operating in the discrete Fourier Transform (DFT) is applied. The invisible watermark insertion algorithm is based on the Tang and Hang [4] algorithm where by embedding a set of peaks values as a watermark into the selected DFT coefficients in the middle frequency region of the red channel of the original image. The selection of the DFT coefficients is based on the observation that DFT coefficients with large values have more chance to survive the print-scan attack and the selection of the embedding positions is generated by a secret key that increase the security of the system.

The selected coefficients are modified by the following equation:

$$\mathrm{p'}_k(\mathrm{x,y}) = \mathrm{p}_k(\mathrm{x,y}) + \mathrm{M} \tag{1}$$

where, $M = max(v)c$ and c are the watermark strength parameter that can be altered to achieve the best imperceptible watermarked image. For the purpose of this evaluation the parameter c was varied from 3 to 9 in steps of 1. This range of values represents a good compromise between imperceptibility and robustness (print-and-scan resilient).

B. Performance Metrics

Based on the current literature, the metrics chosen for evaluation of the image quality were: Peak Signal to Noise Ratio (PSNR), Masking-based Peak Signal to Noise Ratio (MPSNR) and Mean Measure Structural Similarity (MSSIM) while for the the CIELAB metrics ΔE_{ab} and ΔE_{00} were selected for the evaluation of color difference.

ΔE_{ab} and ΔE_{00} are used to measure color differences between corresponding pixels of two different images, i.e., they are not adequate for measuring the overall degradation of color between these images. Its application is for overall evaluation of reliability of colors, or the invariance regarding the color difference between the original and watermarked images.

1) PSNR: Although the MSE (Mean Squared Error) is one of most used performance metrics in the field of signal processing, particularly, in image processing is more usual to convert the MSE in PSNR. The calculation of MSE, which precedes the calculation of PSNR returns a quantitative analysis of the similarity level or distortion between the signals x and y. In Equation 3 the parameter L is the maximum value of intensity allowable for the image.

$$\text{MSE} = \frac{\sum_{j=0}^{J-1}\sum_{k=0}^{K-1}[x(j,k) - y(j,k)]^2}{JK} \quad (2)$$

$$\text{PSNR} = 10\log\left(\frac{L^2}{\text{MSE}}\right) \quad (3)$$

However, in the case of color images, Equation 2 should be adjusted according to Equation 4, which takes into account the RGB channels. Equation 3, PSNR, remains unchanged.

The higher the value of PSNR more the image with watermark resembles the original image.

$$\text{MSE} = \frac{\sum_{c=R,G,B}\sum_{j=0}^{J-1}\sum_{k=0}^{K-1}[x(c,j,k) - y(c,j,k)]^2}{3JK} \quad (4)$$

2) *MPSNR*: The most commonly used metrics usually do not take into account the human visual system, especially regarding colors. Thus, the MPSNR is calculated in the same way as the PSNR (Equation 3), but with a different change in MSE (Equation 5): the inclusion of factor *p(c, j, k)* (Equation 5) where *m'(c, j, k)* (Equation 7) is a weight of intensity which shows how a pixel can be modified in each color channel without degrading the image.

Equation 7 is called JNCD model (Just Noticeable Differences color), and *m(c, x, y)* represents a specific pixel in a specific color channel [5].

$$\text{MSE} = \frac{\sum_{c=R,G,B}\sum_{j=0}^{J-1}\sum_{k=0}^{K-1}[x(j,k) - y(j,k)]^2\, p(c,j,k)]}{3JK} \quad (5)$$

$$p(c,j,k) = \frac{100}{1 + m'(c,j,k)} \quad (6)$$

$$m'(c,j,k) = \begin{cases} m(c,j,k)\dfrac{2}{7}, c \in R \\ m(c,j,k)\dfrac{1}{7}, c \in G \\ m(c,j,k)\dfrac{4}{7}, c \in B \end{cases} \quad (7)$$

3) *MSSIM*: The MSSIM [6], described in Equation 8, is intended to provide a measure of reliability between two images *x* and *y*, taking into account the perception of HVS (Human Visual System). This metric uses three distinct functions applied to B blocks that belong to the original image: luminance (*l*), contrast (I) and structure (*s*). In Equations 9, 10 and 11, μ_x e σ_x are respectively the mean intensity and the standard deviation of the image *x*, μ_y e σ_y correspond to the image y and σ_{xy} is the covariance between both images. *L* is the maximum value of intensity and K_1 and K_2 are very small constants used to avoid instability in the case of a division by zero.

When dealing with color images, first, the metric is calculated separately for each color channel (R, G and B), and then is averaged across the three channels. The MSSIM ranges from 0 to 1. The higher the value of the metric more the image with watermark resembles the original image.

$$\text{MSSIM}\,(x,y) = \frac{1}{B}\sum_{j=1}^{B} l(x,y).c(x,y).s(x,y) \quad (8)$$

$$l(x,y) = \frac{2\mu_x\mu_y + (K_1L)^2}{\mu_x^2 + \mu_y^2 + (K_1L)^2} \quad (9)$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + (K_2L)^2}{\sigma_x^2 + \sigma_y^2 + (K_2L)^2} \quad (10)$$

$$s(x,y) = \frac{2\sigma_{xy} + (K_2L)^2}{2\sigma_x\sigma_y + (K_2L)^2} \quad (11)$$

4) *ΔE_{ab}:* By converting the image from RGB color space to the CIELAB, one of the most important color space, we have as pixel coordinates, L*, a* and b* representing, respectively: luminance, position between red and green and position between yellow and blue. Thus, ΔE_{ab} (Equation 12) is typically used to measure color differences between two images and not the existing degradation between them [7].

As ΔE_{ab} is calculated pixel by pixel, the result presented here is an average of all values of ΔE_{ab} of the image. So if ΔE_{ab} is less than 1, the color fidelity between the images is excellent, however if ΔE_{ab} is between 1 and 2, there is a small difference between the colors, but difficult to detect. Already if ΔE_{ab} is greater than 3 means that the colors are very different, it being perceptible to a viewer.

$$\Delta E_{ab} = \sqrt{(\Delta L*)^2 + (\Delta a*)^2 + (\Delta b*)^2} \quad (12)$$

The ΔE_{ab} can also be written in terms of luminance, chroma and hue (Equation 13) where $\Delta C_{ab}*$ is the simple difference between the pixels of the reference image and the image with the watermark.

$$\Delta E_{ab} = \sqrt{(\Delta L*)^2 + (\Delta C_{ab}^*)^2 + (\Delta H_{ab}^*)^2} \quad (13)$$

$$\Delta H_{ab}^* = \sqrt{(\Delta a*)^2 + (\Delta b*)^2 + (\Delta C_{ab}^*)^2} \quad (14)$$

In this paper, the resultant values for ΔE_{ab} consists of calculating the arithmetic mean of all values for these metrics for each image (Equation 15).

$$\Delta E_{ab} mean = \frac{\sum_{j=1}^{N}\sum_{k=1}^{M}\Delta E_{ab}(j,k)}{NM} \quad (15)$$

5) *ΔE_{00}:* As an improvement of ΔE_{ab} the ΔE_{00} (Equation 16) has a better performance to differentiate the colors blue and the colors in the scale of gray [7].

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L^*}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T\left(\frac{\Delta C'}{k_C S_C}\right)\left(\frac{\Delta H'}{k_H S_H}\right)} \qquad (16)$$

In Equation 16, the factors k_L, k_C and k_H should be adjusted according to different viewing parameters such as texture and background scene, for luminance, chroma and hue. The other factors are described in the following equations:

$$S_L = 1 + \frac{0,015(\bar{L}^* - 50)^2}{\sqrt{20 + (\bar{L}^* - 50)^2}} \qquad (17)$$

$$C' = \sqrt{a'^2 + b^{*2}} \qquad (18)$$

$$a' = \left(1 + G\right) a^* \qquad (19)$$

$$G = 0,5\left(1 - \sqrt{\frac{(\bar{C}^*_{ab})^7}{(\bar{C}^*_{ab})^7 + 25^7}}\right) \qquad (20)$$

$$S_C = 1 + 0,045\bar{C}' \qquad (21)$$

$$H' = 2sen(\Delta h'/2)\sqrt{C'_{orig}\, C'_{md}} \qquad (22)$$

$$h' = \tan^{-1}(b^*/a') \qquad (23)$$

$$S_H = 1 + 0,015\bar{C}'T \qquad (24)$$

$$T = 1 - 0,17\cos(\bar{h}' - 30^{o}) + 0,24\cos(2\bar{h}') + 0,32\cos(3\bar{h}' + 6^{o}) - 0,2\cos(4\bar{h}' - 63^{o}) \qquad (25)$$

$$R_T = -2sen(2\theta)\sqrt{\frac{C'^7}{C'^7 + 25^7}} \qquad (26)$$

$$\theta = 30\exp\left[-\left(\frac{\bar{h}' - 275^{o}}{25}\right)^2\right] \qquad (27)$$

Similar to the ΔE_{ab} the resultant values for ΔE_{00} consists of calculating the arithmetic mean for all values of these metrics for each image (Equation 28).

$$\Delta E_{00} mean = \frac{\sum_{j=1}^{N}\sum_{k=1}^{M}\Delta E_{00}(j,k)}{NM} \qquad (28)$$

III. RESULTS AND DISCUSSION

A. Image Data Base

The color test images used in this paper were acquired in the database USC-SIPI Image Database [8], which supports research in processing and analysis of images. The color images chosen from the group *Miscellaneous* are described in the Table I and shown in the Figure 1.

Table 1- Test color images description

Image	Description
'4.1.01'	Girl
'4.1.02'	Couple
'4.1.03'	Girl
'4.1.04'	Girl
'4.1.05'	House
'4.1.06'	Tree
'4.1.07'	Jelly beans
'4.1.08'	Jelly Beans
'4.2.01'	Splash
'4.2.02'	Girl (Tiffany)
'4.2.03'	Mandrill (Baboon)
'4.2.04'	Girl (Lena)
'4.2.05'	Airplane (F-16)
'4.2.06'	Sailboat on lake
'4.2.07'	Peppers

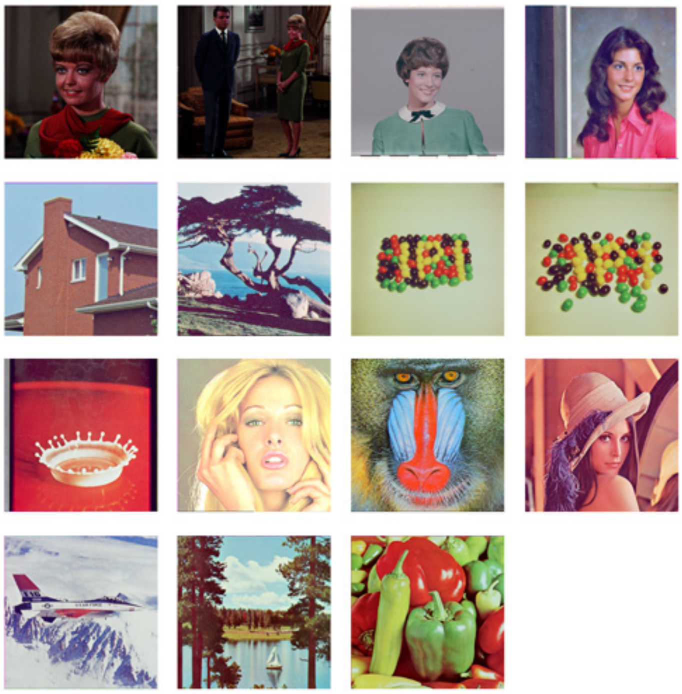

Fig 1. Test color images

B. Results

The graphs from Figures 2 to 6 refer to the values obtained with the performance metrics for the test images varying the strength parameter *c* from 3 to 9.

The PSNR metric that is the most used metric to evaluate the imperceptibility of the watermark does not take into account the perception of the HVS, especially with regard to color perception. The metric MPSNR, in turn, is calculated taking into account the JNCD model, by assigning weights to the RGB components, which have a correlation with the perception of the HVS to the three individual channels in response to the same change of intensity.

From the graphs of Figures 2 and 3, can be noticed that comparing the metrics MPSNR and PSNR both present the same response for the extensive range of values applied to the strength parameter. Therefore, the metric MPSNR do not present advantages for calibration of the watermark insertion algorithm regarding its imperceptibility.

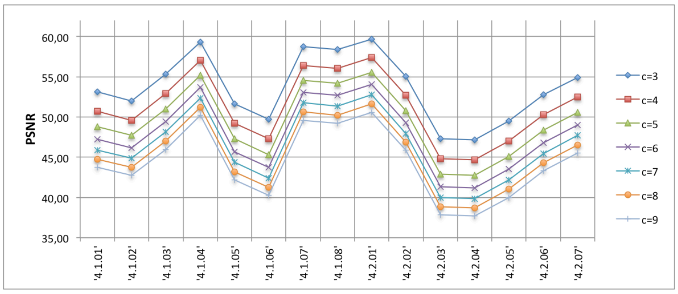

Fig 2. PSNR values calculated

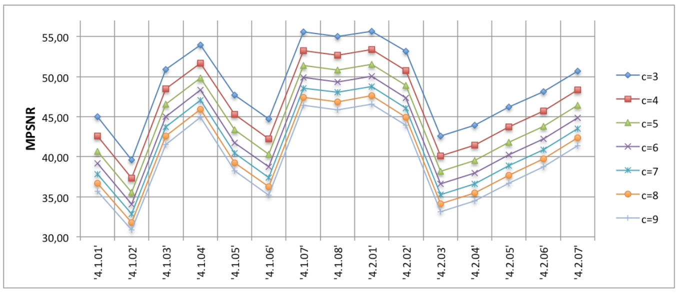

Fig 3. MPSNR values calculated

The MSSIM metric aims to provide a measure of reliability between the original and watermarked images taking into account the perception of the HVS in relation to the luminance, contrast and structure.

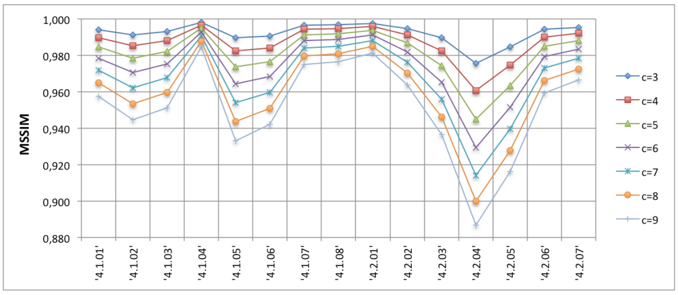

Fig 4. MSSIM values calculated

Comparing the graphs of Figures 2 and 4, it is possible to notice that the PSNR and MSSIM metrics presents a similar response curve for the extensive range of values applied to the strength parameter.

The graph from Figure 5 shows the MSSIM and components values for the strength parameter *c*=5. The luminance component close to 1 can be interpreted as follows: the insertion of watermark introduces no change in intensity values of the image. The contrast and structure components have a profile close to the metric MSSIM, which turn, also has a profile close to the PSNR metric.

Therefore, MSSIM does not present advantages for the calibration of the watermark insertion algorithm in comparison with the PSNR.

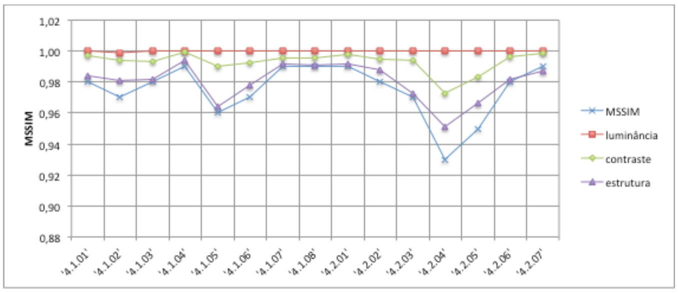

Fig 5. MSSIM and components values calculated

Likewise from the graphs of Figures 6 and 7 the metrics ΔE_{ab} and ΔE_{00} do not show significant advantages with respect to PSNR for the calibration of the algorithm.

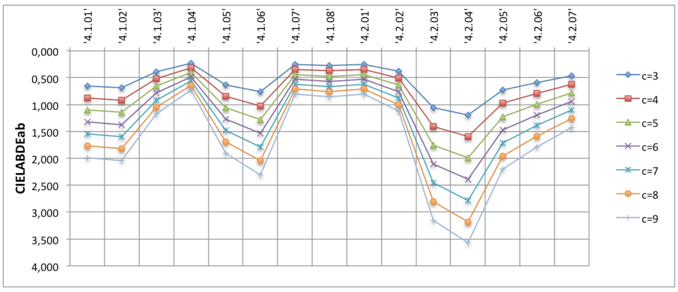

Fig 6. ΔE_{ab} values calculated

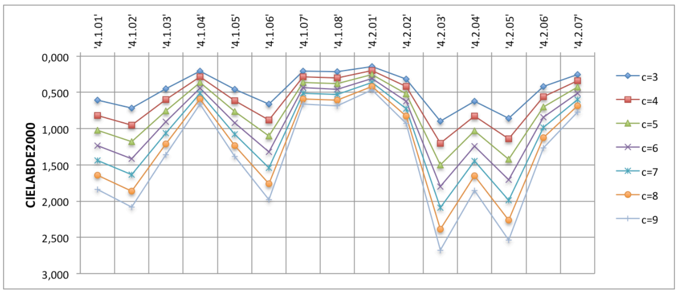

Fig 7. ΔE_{00} values calculated

It is noteworthy that the ΔE_{ab} is the metric recommended by CIE for assessing large color differences ($\Delta E_{AB} > 5$) while ΔE_{00} is used to evaluate small differences [9].

The resultant values obtained from the 4.1.04 and 4.2.0.1 images are showed in the Table II and Table III. Disregarding what metric was used, these images showed best results for the extensive range of values applied.

Table II - Performance metrics results for the 4.1.04 image (girl)

	PSNR	MPSNR	MSSIM	ΔE_{AB}	ΔE_{00}
c = 3	59,3	54	1	0,2	0,2
c = 4	57	51,7	1	0,3	0,3
c = 5	55,2	49,9	1	0,5	0,4
c = 6	53,7	48,3	1	0,5	0,4
c = 7	52,4	47,0	1	0,6	0,5
c = 8	51,2	45,9	1	0,6	0,6
c = 9	50,2	44,9	1	0,7	0,7

Table III - Performance metrics results for the 4.2.01 image (splash)

	PSNR	MPSNR	MSSIM	ΔE_{AB}	ΔE_{00}
c = 3	59,6	55,6	1	0,2	0,1
c = 4	57,4	53,4	1	0,3	0,2
c = 5	55,5	51,5	1	0,4	0,3
c = 6	54	50	1	0,6	0,4
c = 7	52,7	48,7	1	0,6	0,4
c = 8	51,6	47,6	1	0,7	0,4
c = 9	50,6	46,6	1	0,8	0,5

IV. CONCLUSIONS

This paper presented an evaluation of performance metrics for watermarked color images based on image quality and color difference metrics taking into account the same insertion algorithm and the comparison between original and watermarked images to provide an assessment of the imperceptibility aspect to be used during the calibration of the watermark insertion algorithm.

The evaluation found a response profile very similar between the metrics. Simulation results showed that the metrics commonly proposed to approximate quantitative evaluations to the perception of HVS and also regarding the reliability of colors do not have significant advantage compared to PSNR, whose implementation is much simpler.

Although is possible to find several papers in the literature that use these image quality metrics, to evaluate the imperceptibility of the watermark, few describe a comparison between these metrics, and very few describe a comparison between metrics that evaluate image quality with ones that assess color difference. Image quality metrics are normally used to evaluate the imperceptibility of the watermark and color difference metrics are normally used in cases regarding color management and rarely used to measure the watermark quality. Simulation results also show that the PSNR can be used for both assessments.

The authors believe, from the conclusion of this work, to have contributed to the research and applications of color watermarking technology.

REFERENCES

[1] G. Nian, X. Tang, D. Wang and H. Liu, Print-Scan Resilient Data Hiding Scheme Applied in Certificate Verification, 3rd International Congress on Image and Signal Processing (CISP), vol. 3, pp.1161-1165, 2010.

[2] J. Nah, J. W. Kim, J. Kim, A New Image Watermarking Using Peak Position Modulation for ID Photos, 11th IEEE International Symposium on Multimedia, pp.595-599, 2009.

[3] S. Samuel; W.T. Pentzhorn; "Digital Watermarking for Copyright Protection", 7th AFRICON Conference in Africa, 2004.

[4] Y. Tang; Y. Huang, Print-and-Scan Resilient Watermarking for Authenticating Paper-Based Certificates, First International Conference on Pervasive Computing Signal Processing and Applications (PCSPA), pp.357-361, 2010.

[5] H.M. Yang; Y.Q. Liang; X.D. Wang, S.J. Ji, "A DWT-Based Evaluation Method of Imperceptibility of Watermark in Watermarked Color Image", International Conference on Wavelet Analysis and Patter Recognition, 2007.

[6] Z. Wang; A. Bovik; H. Sheikh; E. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity", IEEE Transactions on Image Processing, v.13, 2004.

[7] L. Mandic; S. Grgic; M. Grgic; "Comparison of Color Difference Equations", 48th International Symposium ELMAR, 2006.

[8] "The USC-SIPI Image Database", Available: http://sipi.usc.edu/database/

[9] S. Westland; C. Ripamonti; V. Cheung. Computacional Colour Science using Matlab. 2nd Edition. ed. [S.l.]: Wiley, 2012.

André Rodrigues, Fabiana Fernandes, Zélia Peixoto e Flávia Freitas¸ Graduate Program on Electrical Engineering, Pontifical Catholic University of Minas Gerais, Belo Horizonte - MG, Brazil. E-mails: alsr@ieee.org, fabianapaixao@ieee.org, assiszmp@pucminas.br, flaviamagfreitas@pucminas.br.