Fidelity and Robustness Analysis of SVD-based Image Watermarking Schemes

Franco Del Colle ⋆ and Juan Carlos Gómez

Laboratory for System Dynamics and Signal Processing
FCEIA, Universidad Nacional de Rosario
CIFASIS, CONICET
Riobamba 245 bis, 2000 Rosario, Argentina
{delcolle, jcgomez}@fceia.unr.edu.ar

Abstract. In this paper, an SVD-based perceptual fidelity metric is reviewed and used in the context of objective quantification of image watermarking fidelity for two state-of-the-art Image Adaptive SVD-based Watermarking Schemes. The proposed metric makes use of a Human Visual System perceptual model in the wavelet domain. The robustness of the watermarking schemes against JPEG compression and resizing is also tested on an image dataset of natural color images.

Keywords: Image Watermarking, Perceptual Quality Metrics, Singular Value Decomposition.

1 Introduction

Digital Watermarking has become the most efficient and widely used technique to protect copyright data in an imperceptible and robust way. The embedded information (watermark) should always be present and detectable but users should not be aware of its existence. The main requirements that any watermarking technique should met are:

– *Perceptual transparency (Fidelity)*: Property of the watermark of being imperceptible in the sense that humans can not distinguish the watermarked images from the original ones by simple inspection.
– *Robustness*: Capacity of the watermark to remain detectable after alterations due to processing techniques or intentional attacks.
– *Payload of the watermark*: Amount of information stored in the watermark, which in general depends on the application.

Good overviews on the state of the art of classical watermarking techniques can be found in the textbooks [1] and [2], and in [3], [4], [5] and the references therein.

Among the different approaches that have been proposed in the literature for the watermarking of still images, the ones in the transform domain, which are adapted to the particular image, have proved to deliver better results regarding transparency and robustness. In these methods, the length, location and strength of the watermark is adapted to the image characteristics, [6], [5], [7].

Several watermarking methods based on Singular Value Decomposition (SVD) have been proposed in the literature in recent years, see for instance [8], [9], [10] and [11]. These methods are based on the modification of the singular values of the image. The SVD of an image is an optimal decomposition, in the sense that most of the signal energy is concentrated in

⋆ Author to whom all correspondence should be addressed.

few coefficients (singular values). In addition, the SVD has properties of stability, proportion invariance and rotation invariance.

More recently, the SVD has been used in combination with the Discrete Wavelet Transform (DWT) decomposition in different watermarking schemes. In [12], a semi-blind reference watermarking scheme for copyright protection using gray scale logos as watermarks is presented. For the watermark embedding, the original image is transformed using DWT, a reference image is obtained, and then the watermark is embedded into the reference image by modifying its Singular Values (SVs), using the SVs of the watermark. In [13], a hybrid DWT-SVD watermarking scheme that takes into account the Human Visual characteristics is introduced. The embedding is done by DWT decomposing the host image into four subbands, applying SVD to each subband, and then modifying the SVs using the SVs of the watermark. The watermark strength is determined by the Human Visual System (HVS) model proposed in [14].

Regarding image watermarking fidelity assessment, two different approaches can be distinguished: subjective evaluation and objective evaluation. In the subjective assessment, a number of observers are asked to rank the distortion of the images in a given scale and a Mean Opinion Score (MOS) is computed. This type of evaluation is time consuming and it can be influenced by experimental conditions (such as lighting, monitor characteristics, etc.), and lack of motivation and mood of the participants. On the other hand, in the objective assessment approach, a distortion metric is mathematically defined and computed from the original and watermarked images, and it is then used to quantify the watermarked image fidelity in an automatic way, without the involvement of human beings. This classification of the different approaches for image watermarking fidelity assessment is considered within the framework of the so-called full reference quality evaluation techniques, where both the original and the distorted images are assumed to be available.

Among the objective image quality metrics, two different classes can be distinguished: *Pixel-based Metrics* and *Perceptual Quality Metrics*. *Pixel-based Metrics* are based only on the characteristics of the image. Within this class, the widely used root mean squared error (RMSE), the peak signal to noise ratio (PSNR), the Universal Image Quality Index (UQI) proposed in [15], and the metric based on SVD introduced in [16], can be mentioned. *Perceptual Quality Metrics* take also into account perceptual characteristics of the HVS. Within this group, the structural similarity metric (SSIM) introduced in [17] and the Komparator metric proposed in [18] can be mentioned.

As pointed out in [19] and [20], pixel-based metrics do not correlate well with human visual distortion perception. The same conclusion is drawn in [21], where a comparison of several perceptual and non perceptual metrics in the framework of image watermarking is carried out.

In this paper, the watermark fidelity metric based on DWT and Singular Value Decomposition introduced by the authors in [22] is reviewed. The metric benefits from the advantages of the Discrete Wavelet Transform Decomposition regarding space-frequency resolution, and of the Singular Value Decomposition of an image regarding the compactness of the representation of the signal energy in a few coefficients. The metric is perceptually aware in the sense that it takes into account a perceptual model of the HVS, namely the one introduced in [23]. The proposed metric is used in this paper to evaluate the fidelity of two SVD-based watermarking schemes introduced in [11]. In contrast to the work in [11], where a single image is used, a set of natural color images is considered in this paper for the testing and validation of the proposed metric, and for the study of the degradation of the watermark detectability after JPEG compression and resizing. The fidelity of other state-of-the-art Image Adaptive DWT (IADWT) watermarking techniques was also studied in [22] using the proposed metric, showing good correlation results with the subjective tests.

The rest of the paper is organized as follows. In section 2, both SVD-based watermarking techniques are briefly described. In section 3, the perceptual metric used for the evaluation of the fidelity performance is presented. The metric used to evaluate robustness of the watermarking schemes is presented in section 4. The experimental results regarding fidelity and robustness, for the whole image dataset, are shown in section 5. Finally, some concluding remarks are given in section 6.

2 SVD-based Watermarking

The watermark embedding schemes introduced in [11] are considered in this paper. The authors in [11] present two algorithms based on the SVD: the first one (called hereafter Global SVD algorithm) assumes the size of the watermark image W to be equal to the size of the original image A, and the second (called hereafter Block SVD algorithm) partitions the original images into $M \times M$ blocks and embeds one bit of the watermark in each block. The algorithms are briefly described in subsections 2.1 and 2.2, respectively.

2.1 Global SVD technique

The algorithm can be summarized in the following steps:

- Watermark Embedding
  1. Perform SVD on the original image A:
  $$A = U\Sigma V^T \tag{1}$$
  2. Add the watermark image W to Σ, with a scale factor α that controls the strength of the watermark as:
  $$\Sigma_w = \Sigma + \alpha W \tag{2}$$
  3. Obtain the watermarked image A_W:
  $$A_W = U\Sigma_w V^T \tag{3}$$
- Watermark Extraction
  1. Given the SVD components of the original image $A = U\Sigma V^T$ and a possibly corrupted watermarked image A_W^*, obtain the possibly corrupted matrix Σ_w^* as:
  $$\Sigma_w^* = U^T A_W^* V \tag{4}$$
  2. Obtain the extracted watermark as:
  $$W^* = \frac{1}{\alpha}(\Sigma_w^* - \Sigma) \tag{5}$$
  3. Compute the normalized cross-correlation between W^* and W.

2.2 Block SVD technique

In this algorithm, the original image is partitioned into smaller blocks A_{ij} of size $M \times M$, and one bit of the watermark sequence is embedded in each block. A method to generate a watermark sequence from a meaningful watermark image is presented in subsection 2.3. The algorithm can be summarized as follows:

- Watermark Embedding

1. Partition the original image A into $M \times M$ blocks A_{ij}
2. Perform SVD on each block as

$$A_{ij} = U_{ij} \Sigma_{ij} V_{ij}^T \tag{6}$$

3. Embed one bit of the watermark in each block as:

$$A_{ijW} = U_{ij} [\Sigma_{ij} (1 + \beta b_{ijW})] V_{ij}^T \tag{7}$$

where b_{ijW} is the watermark bit embedded into block A_{ij}, and β is a scale factor that determines the strength of the embedded signal. Note that the bits b_{ijW} take only the values zero or one.

– Watermark Extraction

1. Partition the watermarked image A_W and the original image A into blocks of the same size used for the embedding stage.
2. Perform SVD on each block $A_{ij} = U_{ij} \Sigma_{ij} V_{ij}^T$
3. The embedded watermark bit could be found from (7) as:

$$I_M b_{ijW} = \frac{\Sigma_{ij}^{-1}}{\beta} [U_{ij}^T A_{ijW} V_{ij} - \Sigma_{ij}] \tag{8}$$

provided the indicated inverse exists. In a practical situation, some singular values in Σ_{ij} could be zero (this happens in most of the cases). This situation is not considered in [11]. To solve this problem, eq. (8) should be modified to

$$B_{ijW} = \frac{\Sigma_{ij}^{\dagger}}{\beta} [U_{ij}^T A_{ijW} V_{ij} - \Sigma_{ij}] \tag{9}$$

where $\Sigma_{ij}^{\dagger}$ stands for the left pseudo inverse of Σ_{ij}, and B_{ijW} is a diagonal matrix with values b_{ijW} and zeros in its diagonal.

4. Use the first element in the diagonal of B_{ij} as b_{ijW}.

2.3 Watermark sequence generation

A method to obtain a watermark sequence from a meaningful watermark image to be used in the Block SVD algorithm is proposed in this subsection. Given a watermark binary image, a binary sequence is generated by partitioning the original image into blocks which are scanned in a zigzag manner and each block is also scanned in a zigzag manner, to form the desired sequence. The procedure is depicted in Figure 1 for a 64×64 example image partitioned into 4×4 blocks, resulting in a 1024-length sequence.

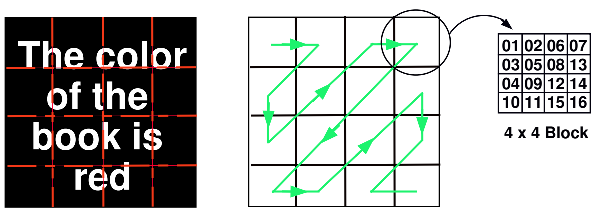

Fig. 1: Zigzag method for the generation of a binary sequence from a binary image.

3 Fidelity Evaluation

In the evaluation of image watermarking methods it is of interest to judge the fidelity of the watermarked image. Basically, the fidelity is a measure of the similarity between the images before and after the watermark insertion. For some applications, fidelity is the primary perceptual measure of concern, and in these cases the watermarked image must be indistinguishable from the original.

The metric proposed by the authors in [22] is briefly described here since it will be used in subsection 5.1 to evaluate the SVD based watermarking schemes described in section 2.

This metric resorts to a widely used perceptual model of the HVS introduced in [23], which takes into account frequency sensitivity, local luminance and contrast masking effects to determine an image-dependent quantization matrix, which provides the maximum possible quantization errors in the DWT coefficients which are not perceptible by the HVS. These values are the so-called Just Noticeable Difference (JND) thresholds. A relationship between these maximum quantization errors in the DWT domain and the maximum variation of the wavelet coefficients' singular values is also derived.

In a first stage, a 1-level DWT decomposition is performed for both the original and the watermarked images, using the biorthogonal 7/9 wavelet [24], resulting in the coefficient matrices $C_{LL}, C_{LH}, C_{HL}, C_{HH}$ for the original image and $C^w_{LL}, C^w_{LH}, C^w_{HL}, C^w_{HH}$ for the watermarked image. Here, the subindexes LL, HL, LH and HH indicate approximation, and vertical, horizontal and diagonal details, respectively.

The Singular Value Decomposition of each coefficient matrix is then performed, resulting in four singular values matrices for each subband of the original image, namely Σ_{LL}, Σ_{LH}, Σ_{HL} and Σ_{LL}, and four singular values matrices for each subband of the watermarked image, namely, Σ^w_{LL}, Σ^w_{LH}, Σ^w_{HL} and Σ^w_{LL}. Then, the absolute difference of the singular values matrices for each subband is computed (element-wise) according to

$$\Delta\Sigma_i(j,k) \triangleq |\Sigma_i(j,k) - \Sigma^w_i(j,k)|, \tag{10}$$

with $i = LL, LH, HL, HH$.

The watermark in the watermarked image will be imperceptible if the variation of the wavelet coefficients associated to the singular value differences in (10) do not exceed the JND thresholds of the DWT domain HVS model. An SVD decomposition of the DWT perceptual thresholds for the ith-subband, JND_i, permits to obtain the singular value perceptual thresholds as follows,

$$JND_i = U_i \Sigma_{JND_i} V_i^T \Rightarrow \Sigma_{JND_i} = U_i^T JND_i V_i \quad , \tag{11}$$

with $i = LL, LH, HL, HH$.

A variation of the singular values of a specific subband will then be perceptible if the difference $\Delta\Sigma_i(j,k)$ in (10) exceeds the singular value perceptual thresholds $\Sigma_{JND_i}(j,k)$.

A threshold matrix $Thresh(\Delta\Sigma_i)$ can be defined from $\Delta\Sigma_i$ by zeroing the entries which are below the perceptual thresholds Σ_{JND_i}, and then, a single value of distortion for each subband can be defined as follows:

$$d_i \triangleq \frac{\|Thresh(\Delta\Sigma_i)\|_F}{\|\Sigma_i\|_F}, \qquad i = LL, LH, HL, HH \tag{12}$$

where $\|\cdot\|_F$ stands for the Frobenius norm of a matrix, and the normalization by $\|\Sigma_i\|_F$ has been performed in order for the distortions d_i to be in the range $[0,1]$.

Finally, to provide a unique parameter quantifying the distortion, a pooling of the four subband distortion measures is needed. An objective fidelity metric can then be defined as the complement of the linear combination of the four distortion measures in (12), *i.e.*,

$$f \triangleq 1 - (k_{LL}d_{LL} + k_{LH}d_{LH} + k_{HL}d_{HL} + k_{HH}d_{HH}), \quad (13)$$

where the coefficients k_{LL}, k_{LH}, k_{HL} and k_{HH} must satisfy the constraint

$$k_{LL} + k_{LH} + k_{HL} + k_{HH} = 1, \quad (14)$$

in order for f to be in the range $[0,1]$.

4 Robustness Evaluation

Another important issue when evaluating image watermarking methods is the robustness, *i.e.*, the capacity of the watermark to survive standard image processing alterations, such as lossy compression, scaling, cropping, etc..

In this paper, robustness of the watermark against JPEG compression and resizing is evaluated by computing a degradation coefficient, $\mathcal{D}$, which quantifies the degradation in the watermark detectability caused by these image processing tasks. To perform the robustness test, the watermarked image is subjected to the above mentioned attacks, and then the watermark is extracted following the procedures described in subsections 2.1 and 2.2 for the global SVD and block SVD techniques, respectively. The normalized cross-correlation, $r_{w,w_e}(k)$, between the original, $w(\ell)$, and the extracted, $w_e(\ell)$, watermarks is then computed. The *detectability degradation coefficient* is then defined as

$$\mathcal{D} \triangleq (1 - r_{w,w_e}(0)) \times 100. \quad (15)$$

5 Results

In order to compare the performance of the watermarking schemes proposed in [11], a set of fifteen (256×256) natural color images was used. Four of these images, called Image 1 to Image 4, are shown in Figure 2. The complete image dataset have not been included here due to space limitations but it can be downloaded from the authors's research group website (http://www.fceia.unr.edu.ar/lsd/mrg/watermark/).

Fig. 2: From left to right: Image 1, Image 2, Image 3 and Image 4.

For all the experiments reported in the following subsections, the values of the parameters α and β in the Global SVD and the Block SVD schemes, respectively, have been chosen so that the energy of the embedded watermark is similar for both methods (the chosen values were $\alpha = 0.08$, $\beta = 0.04$).

5.1 Fidelity Evaluation results

In this section, the proposed metric described in section 3 is used to compare the SVD-based watermarking schemes described in section 2.

The results for all the images in the dataset are shown in Figure 3, where the values of the SVD fidelity metric f in (13) have been normalized to lie in the range $[1,5]$. In this scale, the values correspond to the level of distortion of the watermarked images compared to the original images, according to 5=Imperceptible, 4=Perceptible but not annoying, 3=Slightly annoying, 2=Annoying, and 1=Very annoying. This is usually the scale used for the subjective validation of fidelity metrics. The SVD-based metric described in section 3 has been validated through subjective tests in [22] for two IADWT watermarking schemes. A similar test was carried out to validate the metric for the SVD-based watermarking schemes described in this paper showing a good correlation between the subjective and the objective assessments. The results have not been included here due to space limitations.

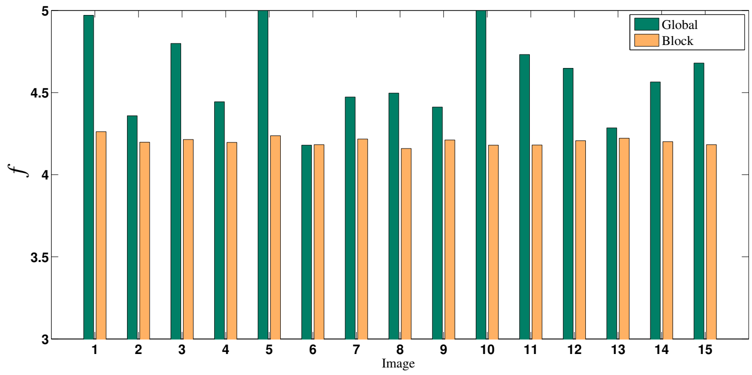

Fig. 3: Objective Assessment based on the SVD fidelity metric f, for Global SVD Watermarking Method (green) and for Block SVD Watermarking Method (orange).

As it can be observed from Figure 3, the Global SVD scheme outperforms the Block SVD one regarding fidelity, for all the images in the dataset. One possible explanation for this is the fact that the block processing introduces noticeable artifacts in the watermarked image since each block is processed independently of its neighboring blocks.

For comparison purposes, the state-of-the-art Komparator metric introduced in [18] is also used on the same image dataset. The results are shown in Figure 4. As it can be observed, also for this metric the Global SVD scheme outperforms the Block SVD one regarding fidelity.

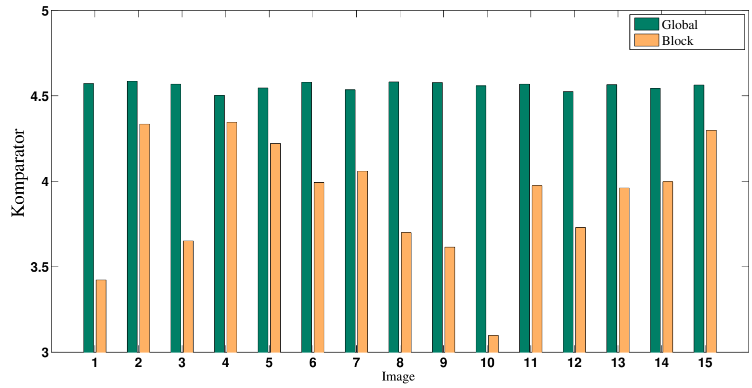

Fig. 4: Objective Assessment based on the Komparator metric [18], for Global SVD Watermarking Method (green) and for Block SVD Watermarking Method (orange).

5.2 Robustness Evaluation Results

In this subsection the robustness of the watermarked images against JPEG compression and resizing is evaluated, for both SVD-based watermarking schemes. The detectability degradation coefficient $\mathcal{D}$, as defined in (15), is used to quantify this robustness.

The robustness results for JPEG compression with compression rates of 90%, 80% and 70%, for the complete dataset, are shown in Figure 5.

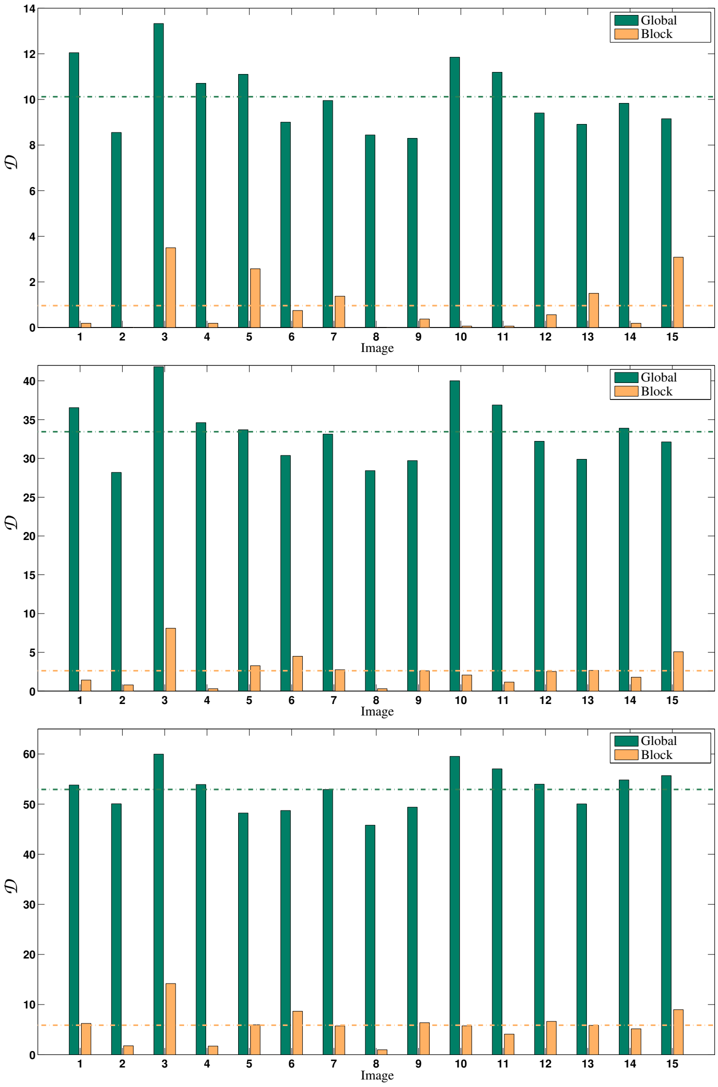

Fig. 5: JPEG Compression robustness at compression ratea 90 % (top), 80% (center), and 70% (bottom), based on the degradation of the correlation between extracted and inserted watermarks. Block SVD Watermarking Method (orange) and Global SVD Watermarking Method (green). Colored dashed lines indicate the mean value of $\mathcal{D}$ for the whole image set.

Figure 6 shows the values of the degradation coefficient when the watermarked images are shrunk to a 75% of the original size and then enlarged back to the original size (using bi-cubic interpolation).

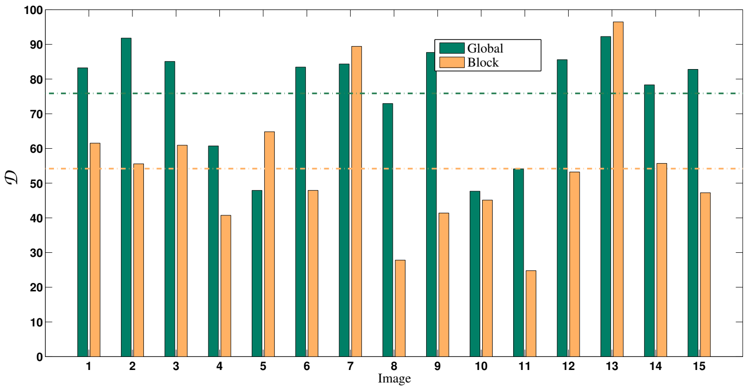

Fig. 6: 75% resizing robustness based on the degradation of the correlation between extracted and inserted watermarks. Block SVD Watermarking Method (orange) and Global SVD Watermarking Method (green). Colored dashed lines indicate the mean value of $\mathcal{D}$ for the whole image set.

As it can be observed from Figures 5 and 6 the Block SVD watermarking scheme outperforms the Global SVD one regarding robustness against JPEG compression and resizing.

6 Concluding Remarks

In this paper, the perceptual fidelity metric based on wavelet domain SVD introduced in [22] is briefly reviewed and used to quantify the perceptual transparency of two state-of-the-art SVD-based watermarking schemes. The metric resorts to a widely used perceptual model of the Human Visual System in the DWT domain. It is important to note that other HVS models (such as the one in [14]) could be easily incorporated by the proposed metric without significant changes in the algorithm. As expected, the watermarking method based on a global processing of the image shows better results regarding fidelity in comparison to the method based on block processing. Additionally, the watermarking schemes were tested for robustness against JPEG compression and image resizing. In this case the method based on block processing outperforms the one based on global processing. In contrast to other works in the literature, the methods were tested on an image dataset of natural color images rather than on a single image (usually the "politically incorrect" Lenna image). The number of images in the dataset was limited to fifteen in order to have a reasonable duration of the subjective tests.

References

1. Barni, M., Bartolini, F.: Watermarking Systems Engineering - Enabling Digital Assets and Other Applications. Marcel Dekker, Inc., New York (2004)
2. Cox, I., Miller, M., J.Bloom: Digital Watermarking. Morgan Kaufmann, San Francisco (2002)
3. Langelaar, G., Setyawan, I., R.Lagendijk: Watermarking digital image and video data. IEEE Signal Processing Magazine **17**(5) (2000) 20–46
4. Petitcolas, F.: Watermarking schemes evaluation. IEEE Signal Processing Magazine **17**(5) (2000) 58–64

5. Podilchuk, C., Delp, E.: Digital watermarking: Algorithms and applications. IEEE Signal Processing Magazine **18**(4) (2001) 33–46
6. Barni, M., Bartolini, F., Piva, A.: Improved wavelet-based watermarking through pixel-wise masking. IEEE Transactions on Image Processing **10**(5) (2001) 783–791
7. Podilchuk, C., Zeng, W.: Image-adaptive watermarking using visual models. IEEE Journal on Selected Areas in Communications **16**(4) (1998) 525–539
8. Chang, C., Tsai, P., Lin, C.: SVD-based digital image watermarking scheme. Pattern Recognition Letters (26) (2005) 1577–1586
9. Changzhen, X., Fenhong, G., Zhengxi, L.: Weakness analysis of singular value based watermarking. In: Proceedings of the 2009 IEEE International Conference on Mechatronics and Automation, Changchun, China. (2009) 2596–2601
10. Liu, R., Tan, T.: An SVD-based watermarking scheme for protecting rightful ownership. IEEE Transactions on Multimedia **4**(1) (2002) 121–128
11. Mohammad, A., Alhaj, A., Shaltaf, S.: An improved SVD-based watermarking scheme for protecting rightful ownership. Signal Processing (88) (2008) 2158–2180
12. Bhartnagar, G., Raman, B.: A new robust reference watermarking scheme based on DWT-SVD. Computer Standards & Interfaces (31) (2009) 1002–1013
13. Li, Q., Yuan, C., Zhong, Y.: Adaptive DWT-SVD domain image watermarking using human visual model. In: Proceedings of the 9th International Conference of Advanced Communication Technology (ICACT 2007), Korea (2007) 1947–1951
14. Lewis, A., Knowles, G.: Image compression using the 2-D wavelet transform. IEEE Transactions on Image Processing **1**(2) (1992) 244–250
15. Wang, Z., Bovik, A.: A universal image quality index. IEEE Signal Processing Letters **9**(3) (2002) 81–84
16. Shnayderman, A., Gusev, A., Eskicioglu, A.: An SVD-based grayscale image quality measure for local and global assessment. IEEE Transactions on Image Processing **15**(2) (2006) 422–429
17. Wang, S.H., Lin, Y.P.: Wavelet tree quantization for copyright protection watermarking. IEEE Transactions On Image Processing **13**(2) (2004) 154–165
18. Le Callet, P., Barba, D.: A robust quality metric for color image quality assessment. In: Proceedings of the IEEE International Conference on Image Processing. Volume 1. (2003) 437–440
19. Eskicioglu, A.: Quality measurement for monochrome compressed images in the past 25 years. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Volume 4., Turkey (2000) 1907–1910
20. Winkler, S.: A perceptual distortion metric for digital color images. In: Proceedings of IEEE International Conference on Image Processing. Volume 3. (1998) 175–184
21. Marini, E., Autrusseau, F., Le Callet, P., Campisi, P.: Evaluation of standard watermarking techniques. In: Proc. of SPIE-IS& Electronic Imaging. Volume 6505., San Jose, CA, USA (2007) 1–10
22. Del Colle, F., Gómez, J.C.: A new DWT-SVD based perceptual fidelity metric for quality assessment of watermarking schemes. In: Proc. of the 18th European Signal Processing Conference, EUSIPCO 2010, Aalborg, Denmark (2010) 815–819
23. Watson, A., Yang, G., Solomon, J., Villasenor, J.: Visibility of wavelet quantization noise. IEEE Transactions on Image Processing, **6**(8) (1997) 1164–1175
24. Mallat, S.: A wavelet Tour of Signal Processing. Academic Press, San Diego (1998)