Evaluation of standard watermarking techniques

Enrico Marini, Florent Autrusseau, Patrick Le Callet, Patrizio Campisi

HAL Id: hal-00250682
https://hal.science/hal-00250682

Submitted on 12 Feb 2008

Evaluation of standard watermarking techniques

Enrico Marini*♦, Florent Autrusseau*, Patrick Le Callet*, Patrizio Campisi♦

*Ecole Polytechnique de l'Université de Nantes, Nantes, France
E-mail: (florent.autrusseau, patrick.lecallet)@polytech.univ-nantes.fr

♦Dipartimento di Elettronica Applicata,
Università degli Studi di Roma "Roma Tre", Roma, Italy
E-mail: campisi@uniroma3.it

ABSTRACT

In the last decade digital watermarking techniques have been devised to answer the ever-growing need to protect the intellectual property of digital still images, video sequences or audio from piracy attacks. Because of the proliferation of watermarking algorithms and their applications some benchmarks have been created in order to help watermarkers comparing their algorithms in terms of robustness against various attacks (i.e. Stirmark, Checkmark). However, no equal attention has been devoted to the proposition of benchmarks tailored to assess the watermark perceptual transparency. In this work, we study several watermarking techniques in terms of the mark invisibility through subjective experiments. Moreover, we test the ability of several objective metrics, used in the literature mainly to evaluate distortions due to the coding process, to be correlated with subjective scores. The conclusions drawn in the paper are supported by extensive experimentations using both several watermarking techniques and objective metrics.

Keywords: Digital watermarking, image quality, objective quality metrics, subjective quality assessment.

1. INTRODUCTION

The spreading of digital formats for multimedia data, we have witnessed in the last decades, has introduced the possibility of making an infinite number of copies without any quality loss, thus leading to the necessity to devise copyright ownership protection methods in order to prevent fraudulent appropriation of the digital content. Although copyright protection was the very first application of watermarking, different uses have been proposed in the literature such as data authentication, fingerprinting, broadcast monitoring, multimedia indexing, content based retrieval applications, medical imaging, quality of service (QoS) assessment, to cite only a few.

When considering a watermarking scheme, depending on its specific application, different requirements need to be taken into account. One of them is the perceptual invisibility of the superimposed "mark" onto the host data. This implies that the alterations caused by the watermark embedding into the data should not degrade their perceptual quality. Moreover, when these techniques are used to preserve the copyright ownership with the purpose of avoiding unauthorized data duplications, the embedded watermark should be detectable both when malicious attacks and non-deliberate modifications, like filtering, compression, etc., affect the watermarked data. This requirement is known as watermark robustness. For some applications, when the robustness requirement is severely required, each attempt of removing the mark should result in irreversible data quality degradation. As a consequence the quality of the image must noticeably decrease before the removal attempt succeeds. Capacity is another watermarking requirement, referring to the number of bits of information that can be embedded in the original data, which needs to be fulfilled, depending on the specific application. These requirements hinder each other. Therefore the optimal trade-off is strictly tied to the target application.

Because of the proliferation of watermarking algorithms and their applications some benchmarks (i.e. Stirmark [1], Checkmark [2]) have been created in order to allow a fair comparison among watermarking algorithms in terms of robustness against various attacks. However, no equal attention has been devoted to the proposition of benchmarks

tailored to assess the watermark perceptual transparency or, equivalently to perform the watermarked image quality assessment. Thus, the Mean Square Error (MSE), or the Peak Signal to Noise Ratio (PSNR) are commonly used by the watermarking community to assess the fidelity of embedding algorithms. However, the PSNR is not well suited to take into account the visual impairments present in an image, since PSNR gives a measure of the statistical differences between an original image and a distorted version (marked one). However, statistical differences do not represent visual impairments. PSNR can be considered as a good indicator to provide qualitative rank order scores as long as the same content and the same algorithm are used while varying for instance only the watermarking embedding strength. Nevertheless, it is clear that PSNR fails miserably when comparing different algorithms.

In this paper, we evaluate, in terms of perceived visual quality, some watermarking algorithms. Specifically we have defined a subjective protocol suited to watermarking quality assessment purposes and we have run tests with human observers in a controlled environment defined accordingly to the ITU recommendation 500-11.

Moreover, in order to make the quality assessment an automatic process, some objective metrics, already validated in the literature to evaluate the effect of the compression distortions, have been used in this paper to assess the quality of images impaired by a watermark embedding.

Also the trade off between the robustness of the proposed algorithms and their perceptual transparency is discussed.

The paper is organized as follows. In Section 2 the watermarking algorithms chosen for quality assessment are briefly described. The modalities how the subjective tests are run as well as the obtained results are described in Section 3. In Section 4 the employed objective metrics are briefly detailed and their relations with the results obtained in the subjective experiments are described. An analysis of the trade off between robustness and perceptual transparency is provided in Section 5. In Section 6 the experimental results are presented and discussed. Eventually, conclusions are drawn in Section 7.

2. TESTED WATERMARKING ALGORITHMS

In this paper we perform both subjective and objective quality assessment for the set of watermarking algorithms [3-10] whose implementation is freely available online [11]. Specifically, the following algorithms, whose basic properties are briefly summarized below, are taken into account:

- "Cox" [3] approach is a spread spectrum based watermarking algorithm operating in the DCT domain,
- "Corvi" [4] method is a spread spectrum based watermarking algorithm operating in the wavelet domain,
- "Kim" [5] approach is a wavelet based algorithm. Perceptually significant wavelet coefficient, selected using a level-adaptive thresholding approach, are watermarked using different scale factors according to the level of decomposition,
- "Koch" [6] method is based on the random generation of the locations where to insert a watermark and in the subsequent embedding in the DCT coefficients of the blocks, corresponding to the selected locations, of a binary generated watermark,
- "Wang" [7] algorithm is a perceptual watermarking scheme, operating in the wavelet domain, where the significant wavelet coefficients search is motivated by the principle for the design of the multi-threshold wavelet codec,
- "Xia" [8] approach is based on the watermark embedding in the large coefficients at the high and middle frequency bands of the discrete wavelet transform of the host image,
- "Xie" [9] method is jointly designed together with a compression algorithm in order to prevent the watermark degradation after compression with the underlying assumption that coding is inevitably used during transmission. Specifically compression in the wavelet domain is taken into account,
- "Zhu" [10] algorithm embeds the watermark in high-pass bands in the wavelet domain.

It is worth pointing out that in this paper we want to provide a way to assess the perceptual quality of the watermarked images using both objective and subjective metrics, and that the approach we present is independent of the specific tested watermarking algorithms.

The five test images used in our experiments are shown in Figure 1.

Fig. 1 Selected contents for subjective tests. From left to right and from top to bottom: images "watch", "parrots", "kodie", "train", and "New york".

3. SUBJECTIVE QUALITY ASSESSMENT

As already pointed out in the Introduction, the PSNR is one of the most used metrics employed to test the perceptual transparency of a watermarking algorithm. However, the PSNR is highly inappropriate to fulfill this task since it gives a measure of the statistical differences between an original image and the watermarked one which are not directly related to represent visual impairments. As an example in Figure 2 (a) – (b) two watermarked images, obtained by using the approaches described in [4] and [5] respectively, are shown. Although they have the same perceptual appearance they are characterized by PSNR values significantly different (PSNR = 25.8 dB and PSNR = 45.25 dB respectively). In Figure 2 (c) – (d) two images watermarked using the approaches described in [8] and [4] respectively are shown. Although they have approximately the same PSNR = 31 dB their perceptual appearance is quite different.

The best way to assess images' quality is surely to run a subjective test. There is no standard for watermarking purpose. In this paper we propose to use a Double Stimulus Impairment Scale (DSIS) protocol. Effectively, one could consider the watermark process as a source of distortions and therefore evaluate the corresponding annoyance. In this protocol, the original and the watermarked images are presented explicitly to the observer. This latter is asked to rate the impairments on a scale of five categories: imperceptible corresponding to score 5, perceptible but not annoying/score 4, slightly annoying/score 3, annoying/score 2 and very annoying/score 1. Obviously, for watermarking applications, we should expect a quality score higher than 4.

We carried out this experiment in our subjective room, which is compliant with the ITU recommendation 500-11, we also use to perform subjective quality assessment as a member of the ILG (Independent Lab Group) of the VQEG (Video Quality Expert Group). We enrolled ten observers and checked their visual acuity. We watermarked five different contents by using eight algorithms and two embedding strengths per algorithm (the default strengths in [11] and 1.5 times the default value). Finally the observers rated 90 images in a session of 20 minutes at an observation distance of 6H the display size of the images. Both the original and the watermarked images were displayed side by side on the viewing monitor and the observers were asked to judge the quality of the marked image compared to the quality of the original. For each image an average of the scores given by the observers were evaluated thus obtaining the mean opinion score (MOS) for the image under analysis. Then, in order to compare the algorithms quality in terms of the perceptual transparency of the embedded mark, we averaged the MOS values over the content. It appears that only three algorithms, whose embedding strengths are set to the default values in their implementation in [11], fulfill the requirement of a

quality score greater than 4. By using an embedding strength equal to 1.5 times the default value, only the "Koch" scheme satisfies this constraint. This is shown in Figure 3.
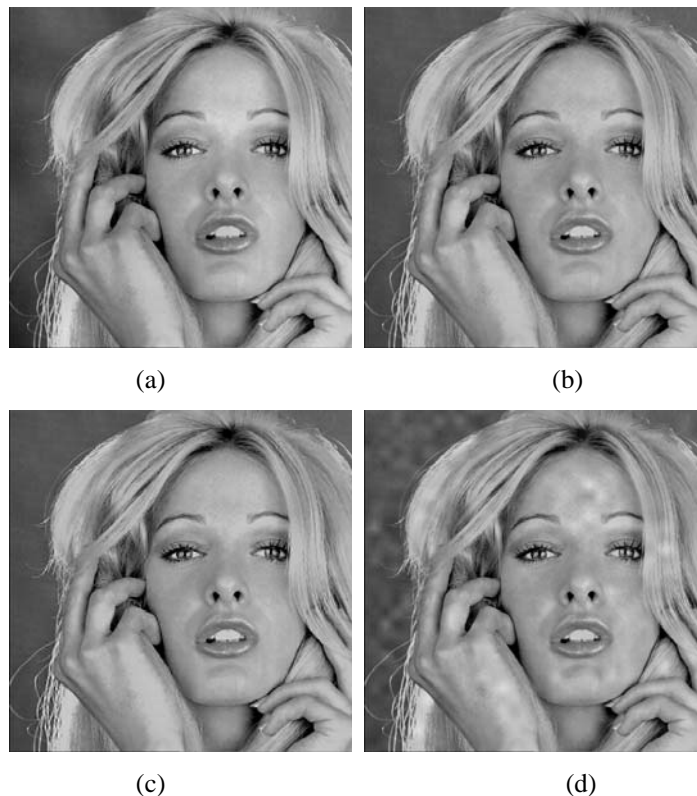

(a)　　　　　　　　　　　　　　　(b)

(c)　　　　　　　　　　　　　　　(d)

Fig. 2 (a): Image watermarked using the "Koch" algorithm [6] (PSNR = 25.8 dB). (b): image watermarked using the "Kim" algorithm [5] (PSNR = 45.25 dB). (c): image watermarked using the "Xia" algorithm [8] (PSNR = 31.05 dB). (d): image watermarked using the "Corvi" algorithm [4] (PSNR = 31.3 dB).
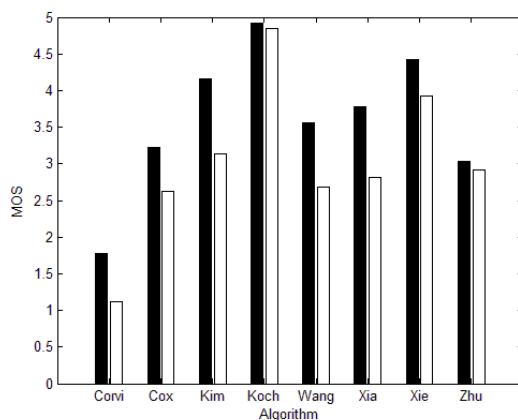

Fig. 3 MOS averaged for all tested contents using the default embedding strength (black bar) and 1.5 times the default embedding strength (white bar) for the tested algorithms.

Whereas Figure 3 depicts the MOS averaged for all tested contents, Figure 4 shows, only for the test image "train", the average MOS along with the confidence intervals for all tested watermarking algorithms. This figure clearly shows the invisibility performances of each embedding algorithm for a specific tested image.
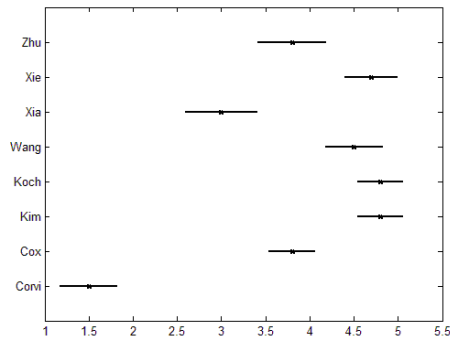

Fig. 4 "Train" image: MOS mean values with the confidence interval for all the tested algorithms.

4. OBJECTIVE QUALITY METRICS

Nowadays several objective metrics based on different approaches and with different accuracy levels can be used to perform quality assessment. These metrics can be classified into three different groups: **Full Reference (FR), Reduced Reference (RF),** and **No Reference (NR)**. The first one is surely the most accurate because it uses the original image to assess the quality of the distorted one. Although this approach generally gives very good results, the need of the original image is a very strong constraint limiting its appeal. Reduced Reference metrics make use of both the distorted image and a limited set of parameters of the original one in order to assess the quality of the distorted data. The set of parameters is usually representing the most important visual features of the image. No Reference methods, as it can be easily understood, only need the distorted version of the image. While this approach seems to be the most appealing, its complexity makes it suitable to evaluate only alterations introduced by some particular distortions like JPEG or JPEG 2000 compression.

In this work, we assess the quality of the watermarking schemes in [11] using several objective metrics, namely SSIM [12], UQI [13], Komparator [14] (Full Reference metrics), as well as RRIQA [15], and C4 [16] (Reduced Reference metrics). The metrics outputs were compared with subjective scores provided by observers and a comparison with the PSNR performances is also given. A brief description of the aforementioned metrics is given hereafter.

- Komparator is an objective quality metric for color images based on Human Visual System (HVS) properties which doesn't need any a priori knowledge about the type of degradation introduced in the distorted image. As first step, a visual representation of the original and the distorted images is computed: in this stage a model based on results obtained by psychophysics experiments on color perception and on masking effects is used. The model adopted for early vision stages is inspired by the S. Daly's Visual Difference Predictor (VDP) extended to color. The visual representation stage is then followed by an error pooling process merging all the obtained distortion maps into a single value. This stage is made up by a frequency, a component, and a spatial pooling taking into account the density and the structure of the error.

- Structural SIMilarity (SSIM) is an objective metric for assessing perceptual image quality, working under the assumption that human visual perception is highly adapted for extracting structural information from a scene. Quality evaluation is thus based on the degradation of this structural information assuming that error visibility should not be equated with loss of quality as some distortions may be clearly visible but not so annoying. Finally SSIM does not attempt to predict image quality by accumulating the errors associated with psychophysically understood simple patterns, but proposes to directly evaluate the structural changes between two complex-structured signals.

- Universal Quality Index (UQI) is a universal objective image quality index, designed to model any image distortion as a combination of three factors: loss of correlation, luminance distortion, and contrast distortion. UQI is a mathematically defined metric such as the widely used Root Mean Squared Error (RMSE) or the PSNR.

- C4 is a metric based on the comparison between the structural information extracted from the distorted and the original images. What makes this metric interesting is that it uses reduced references containing perceptual structural information and exploiting an implementation of a rather elaborated model of the Human Visual System. The full process can be decomposed into two phases. In the first step, perceptual representation is built for the original and the distorted images, then, in the second stage, representations are compared in order to compute a quality score.

- Reduced Reference Image Quality Assessment (RRIQA) is a reduced reference metric, based on a natural image statistic model in the wavelet transform domain. The Kullback-Leibler distance between the wavelet coefficients marginal probability distributions of the reference and the distorted image is used as a measure of the image distortion. A generalized Gaussian model is employed to summarize the marginal distribution of the reference image wavelet coefficients, so that only a relatively small number of images' features are needed for quality evaluation. The basic assumption behind this approach is that most image distortions modify image statistics and make the distorted version "unnatural". The measured unnaturalness can then be used to quantify image quality distortion.

It is worth pointing out that all these metrics have not specifically designed to assess the watermark perceptual transparency. Therefore, although their performances for watermarking applications might not be very competitive, they might give a good clue about the watermark visibility, and most importantly, they might be very useful to compare visibility performances for several embedding algorithms. Ideally, the watermarking community would need an objective metric specifically designed for watermarking purpose, and furthermore, it would be very handy if such metric could be geometric distortions invariant. According to the specifications of the subjective experiment we ran in this study (DSIS protocol, as explained in Section 3), observers assess the quality of the distorted image on a range going from 1 to 5. For watermarking purpose, the optimum objective metric should be restrained to the scores ranking between 4 and 5.

In Figure 5 the subjective scores (Y-axis) are represented as a function of the predicted scores for all tested metrics (X-axis) thus allowing to determine if any of the objective metrics is able to provide a reliable perceptual quality assessment of the watermarking schemes under analysis as the human observers did. Specifically, for each embedding algorithm two values of the MOS, as a function of the predicted scores, for the default value and for 1.5 times the default value of the embedding strength are given. According to these plots, the Komparator metric supposedly provides the best correlation between the subjective and the objective scores.

We compared the used objective metrics by calculating the correlation coefficients between the objective metrics scores and the MOS values obtained by subjective experiments, both averaged over all contents. Average values for both the subjective metrics ($MOS_{(algo,s)}$) and for the objective ones ($OM_{(algo,s)}$), were computed independently for each embedding algorithm over all the N tested contents as in (1) and (2):

$$MOS_{(algo,s)} = \frac{1}{N} \sum_{n=1}^{N} MOS_{(algo,s)}(n) \tag{1}$$

$$OM_{(algo,s)} = \frac{1}{N} \sum_{n=1}^{N} OM_{(algo,s)}(n) \tag{2}$$

where "algo" represent the embedding algorithm under analysis, "s" stands for the embedding strength, and n represents the current tested image. Specifically, the correlation between the subjective and predicted scores is 0.68 for PSNR and it appears that most of the objective metrics under consideration clearly outperform PSNR. Among these metrics, Komparator provides the best results in terms of correlation between the subjective and predicted scores with a value equal to 0.92. The other examined metrics give the following correlation values: C4 ~ 0.83, UQI ~ 0.82, RRIQA ~ 0.82, and SSIM ~ 0.58.

5. ROBUSTNESS EVALUATION

Although the first goal of this work was to highlight the best objective metric for watermarking purpose, we nevertheless evaluated the robustness of the watermark embedding techniques, in order to propose a full evaluation benchmarking tool.
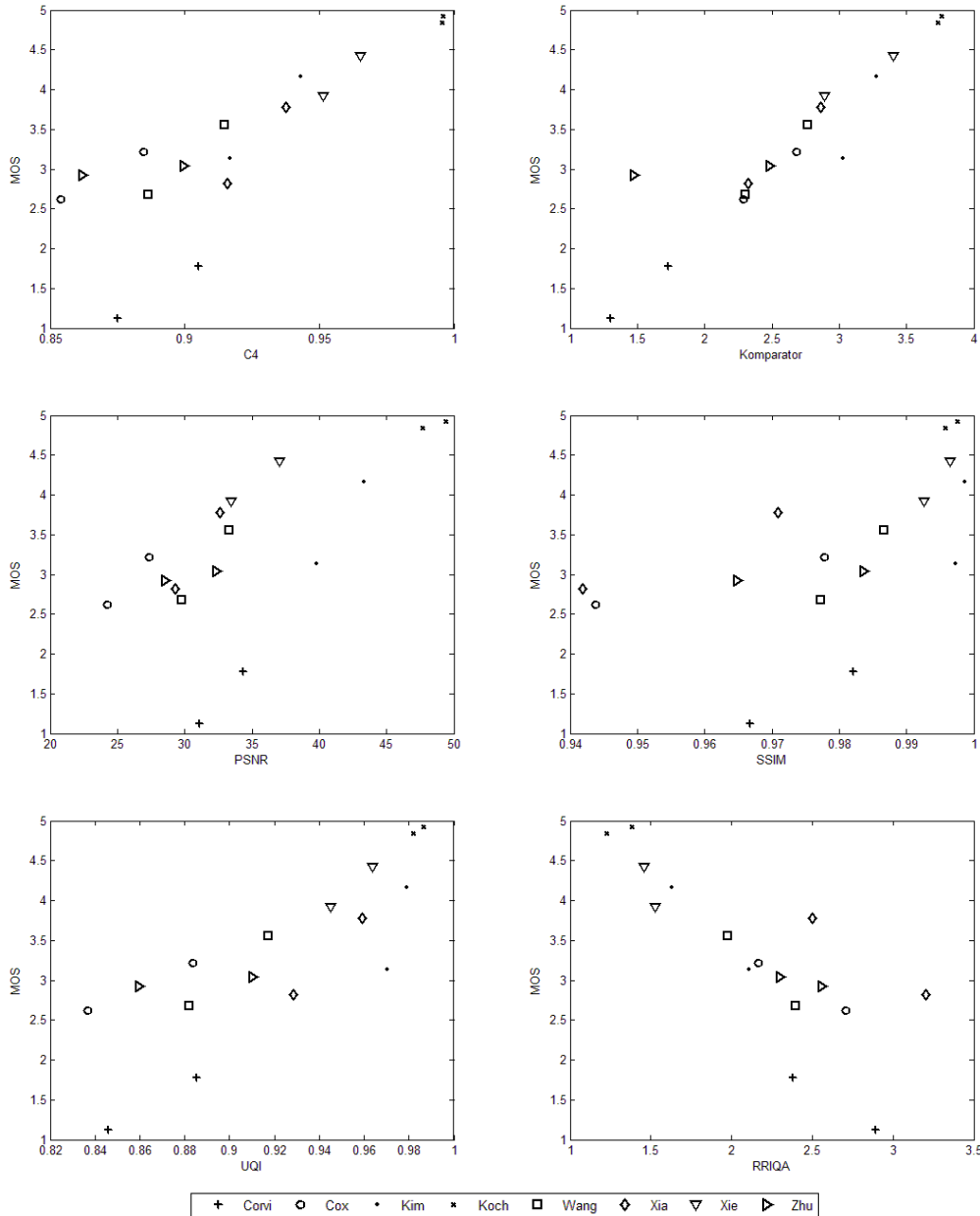
Fig. 5 MOS as a function of C4 (top left), Komparator (top right), SSIM (middle left), UQI (middle right), RRIQA (bottom left), and PSNR (bottom right) score.

We thus present here some preliminary results about the robustness issue and try to link the embedding strength to both its effects on robustness and visibility, since an optimal trade-off between robustness and quality must be achieved when watermarking an image. Only the knowledge about both the achieved invisibility and robustness allows choosing, according to the intended application and to individual needs, the best algorithm and the optimal embedding strength to use. The embedding strength factor plays a very important role in this task: evidently, by increasing its value the robustness increases and as a drawback the visual quality decreases.

To evaluate the robustness of a watermarking algorithm in terms of correct detection, for a fixed value of the embedding strength, it is necessary to set an optimal threshold value. This latter must be well tuned so that the lowest False Positive and False Negative rates are granted during watermark detection.
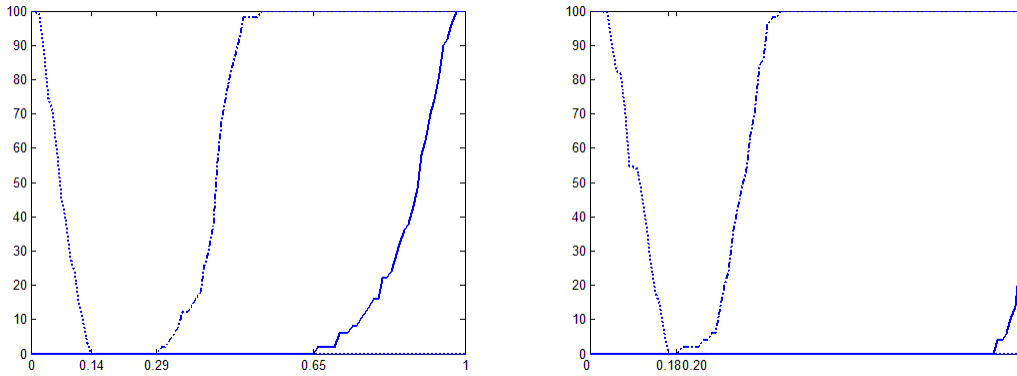
Fig. 6 False Positive (dotted line), False Negative (solid line) and False Positive for distorted images (dash-dotted line) curves *vs.* the threshold value for the "Corvi" (left) and "Wang" (right) embedding algorithms (default embedding strength).
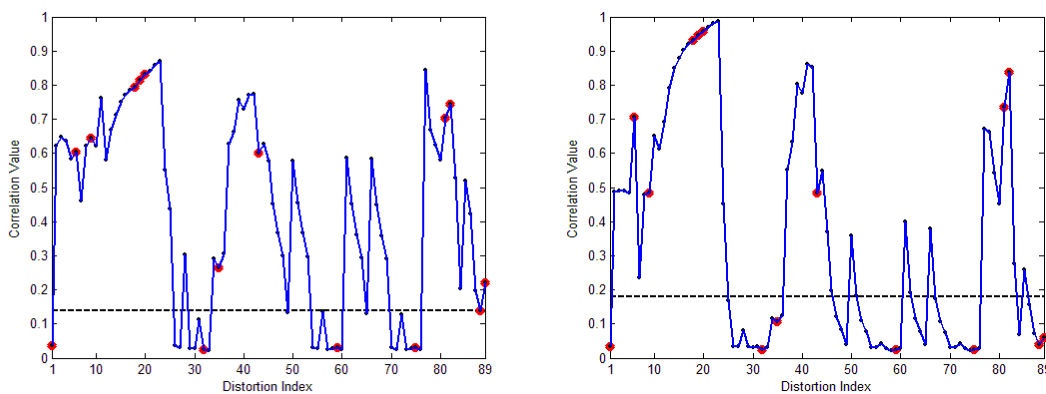

Fig. 7 Cross-correlation between the original watermark and the distorted one as a function of the Stirmark distortion index for the "Corvi" (left) and "Wang" (right) embedding algorithms (default embedding strength). Threshold value: dotted line.

Both False Positive and False Negative curves were evaluated for all tested algorithms, and a detection threshold was chosen for each embedding technique using the default embedding strength, in such a way that False Positives were strictly avoided. In Figure 6 the False Positive, False Negative, and False Positive for distorted images curves *vs.* the threshold value are given for two of the tested watermarking algorithms.

The robustness of the watermarking algorithms is then tested by evaluating the cross-correlation between the original watermark and the distorted one for each attack that can be performed by using the Stirmark benchmarking tool. Then the output of the watermark detector is compared with the chosen threshold (see Figure 7). In detail along the X-axis of Figure 7 the employed distortions are indexed as follows: number 1 represents flipping, from 2 to 6 rows and/or column removal, from 7 to 9 median filtering, from 10 to 18 cropping, 19 FMLR, 20 Gaussian filtering, from 21 to 32 JPG compression, from 33 to 35 linear filtering, from 36 to 43 ratio modification, from 44 to 59 rotations, from 60 to 75 cropping plus scaling, from 76 to 81 simple scaling, 82 sharpening, from 83 to 88 shearing and 89 random Stirmark bend. Geometrical distortions are then represented by indexes from 2 to 6, from 10 to 18 and from 36 to 81. Points on lines in Figure 7 represent the end of each distortion group and help to identify if correlation, for a particular class of distortions, is higher than detection threshold. The horizontal dashed line represents the detection threshold determined from the false positive and false negative analysis.

Among the tested algorithms the "Cox" algorithm appears to resist to almost all soft geometrical distortions, whereas all other techniques presented much weaker robustness results. For instance, the "Kim" technique is totally unsuccessful against geometrical attacks, "Koch", "Wang", and "Zhu" algorithms are resistant against ratio modifications, line and/or columns removal, scaling and cropping, being fragile against all kind of rotations, while "Xie" can survive only to cropping and lines and/or columns removal. At last the Xia algorithms can resist cropping and ratio modification.

6. EXPERIMENTAL RESULTS

It is clear that robustness and quality are conflicting requirements for watermarking applications. This is clearly shown in Figure 8 where the visual quality of the marked images evaluated by means of subjective experiment as a function of the robustness is depicted. The X-axis represents the percentage of detected watermarks over all Stirmark distortions. Figure 8 shows, for each algorithm, the simultaneous drop of the visual quality along with the rise of the robustness when the embedding strength increases. In order to build these plots detection rates calculated over the subset of the five images used in the subjective test have to be considered.

Whereas Figure 8 represents the subjective scores as a function of the robustness, Figure 9 shows both the Komparator predicted scores along with the PSNR results. Like for Figure 8 the plots in Figure 9 are obtained averaging scores over contents after selecting the embedding algorithm and the embedding strength. Because of the different scales, only relative positions of the curves can be compared. Relative positions for most algorithms are very faithfully respected. For the Komparator metric only the "Zhu" algorithm presented important variation in terms of objective and subjective scores (especially for the overmarked version), nevertheless its general performances are very good. The goodness of each tested objective quality metric was assessed by taking into account this quality score versus robustness plot, and the weak correlations given at the end of Section 4 were confirmed. Komparator metric appeared to be the most accurate quality assessment metric. Furthermore, Figure 9 clearly shows the weak capability of the PSNR to judge the perceived quality of a watermarking algorithm. An analysis of the embedding algorithms based on the PSNR measurements, would suggest that the "Corvi" algorithm would induce much less noticeable distortions, whereas both subjective experiments and the Komparator metric show that the distortions induced by the "Cox" algorithm are much less noticeable.
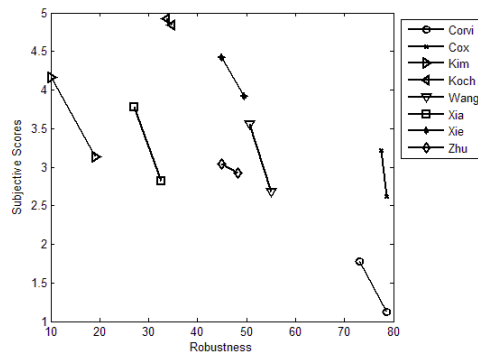

Fig. 8 Subjective scores as a function of robustness (percentage of detected watermarks over all Stirmark attacks) for all the tested algorithms. For each algorithm the "left" marked point corresponds to a default embedding strength and the "right" marked point corresponds to 1.5 times the default embedding strength.
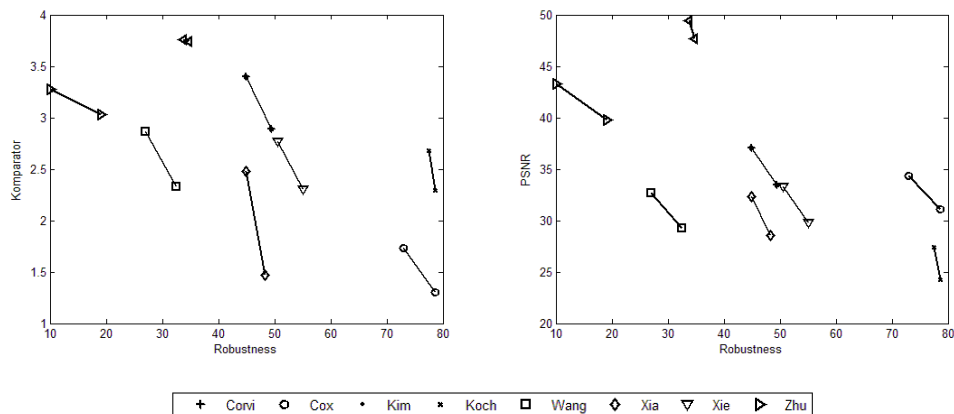

Fig. 9 Komparator (left) and PSNR (right) scores as a function of robustness for all the tested algorithms. For each algorithm the "left" marked point corresponds to a default embedding strength and the "right" marked point corresponds to 1.5 times the default embedding strength.

This reversed ranking of the "Cox" and "Corvi" algorithms in terms of visual perception was indeed suggested by two of the objective metrics (C4, SSIM), whereas both UQI and RRIQA granted the exact same quality to the two algorithms.

7. CONCLUSIONS

The purpose of this work was to highlight the inefficiency of the PSNR as a fidelity measure for watermarked images. This fact was briefly demonstrated by means of few examples and several objective quality metrics were tested for some watermarking algorithms. Results were compared to the PSNR measures and clearly highlighted its unreliability. While all the employed objective metrics are known to be well suited to provide quality assessment for coding purposes, the results obtained in this paper confirm the design of metrics tailored to assess the quality of watermarked image is still an open issue. Nevertheless, the Komparator[1] metric seems to give quite reliable results, although it should be pointed out that its reliability has not been checked for all the possible classes of image watermarking algorithms.

ACKNOWLEDGMENTS

REFERENCES

[1] Fabien A. P. Petitcolas, Ross J. Anderson and Markus G. Kuhn, *Attacks on copyright marking systems, Information hiding*, Second International Workshop, Proceedings, LCNS 1525, Springer-Verlag, ISBN 3-540-65386-4, pages 219-239, April 1998.

[2] Checkmark *http://watermarking.unige.ch/Checkmark/*

[3] I.J. Cox, J. Kilian, T. Leighton and T.G. Shamoon, *Secure spread spectrum watermarking for multimedia*, Proceedings of the IEEE International Conference on Image Processing, 6:1673-1687, 1997.

[4] M. Corvi and G. Nicchiotti, *Wavelet based image watermarking for copyright protection*, Scandinavian Conference on Image Analysis, SCIA, 1997.

[5] J.R. Kim and Y.S. Moon, *A robust wavelet based digital watermark using label adaptative thresholding*, Proceedings of the 6th IEEE International Conference on Image Processing, pages 202-205, 1999.

[6] E. Koch and J. Zhao, *Towards robust and hidden image copyright labeling*, Proceedings of the IEEE International Workshop on Nonlinear Signal and Image Processing, pages 452-455, 1995.

[7] H.-J. Wang, P.-C. Su and C.C. JayKuo, *Wavelet based digital image watermarking*, Optic Express, 3:491-196, 1998.

[8] X.-G. Xia, C.G. Boncelet and G.R. Arce, *Wavelet transform based watermark for digital images*, Optic Express, 3:497-511, 1998.

[9] L. Xie and G. Arce, Joint wavelet compression and authentication watermarking, Proceedings of the IEEE International Conference on Image Processing, October 1998.

[10] W. Zhu, Z. Xiong, and Y.-Q. Zhang, *Multiresolution watermarking for images and video: a unified approach*, Proceedings of the IEEE International Conference on Image Processing, October 1998.

[11] Peter Meerwald, *Digital Image Watermarking in the Wavelet Transform Domain*, Master's Thesis, Department of Scientific Computing, University of Salzburg, Austria, January 2001, http://www.cosy.sbg.ac.at/~pmeerw

[12] Z. Wang, A.C. Bovik, H.R. Sheikh and P.E. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing, 13(4):600-612, 2004.

[13] Z. Wang and A.C. Bovik, *A universal image quality index*, IEEE Signal Processing Letters, 2001.

[14] D. Barba, P. Le Callet, *A robust quality metric for color image quality assessment*, Proceedings of the IEEE International Conference on Image Processing, pp.437-440, 2003.

[15] Z. Wang and E.P. Simoncelli, Reduced-reference image quality assessment using a wavelet-domain natural image statistic model, Human Vision and Electronic Imaging X, 5666, Jan. 2005.

[16] M. Carnec, P. Le Callet and D. Barba, *An image quality assessment method based on perception of structural information*, Proceedings of the IEEE International Conference on Image Processing, 3:185-188, 2003.

---

[1] The Komparator metric will soon be released; interested readers might enquire by e-mail FA for a preliminary version of the software.