Proceedings of the 5th International Symposium on Communications, Control and Signal Processing,
ISCCSP 2012, Rome, Italy, 2-4 May 2012

Perceptual watermarking robust to JPEG compression attack

Phi-Bang NGUYEN, Azeddine BEGHDADI, Marie LUONG

L2TI Laboratory, Institute of Galilee, University Paris XIII, France
Lastname@univ-paris13.fr

ABSTRACT

A new transparent and robust watermarking method against JPEG compression attack based on a perceptual approach is proposed. The main idea is to embed the watermark into regions where the blocking artifact has no effect on the watermark. These regions are identified by using a spatial prediction model of blocking effect. A JND (Just-Noticeable-Difference) model is then used to control the embedding strength in order to ensure robustness and transparency of the watermarking. The performance of the proposed method is evaluated on a set of images and some common "signal processing" attacks.

Index Terms— Blocking Effect, Contrast Masking, Human Visual System, JND, JPEG, Perceptual Watermarking, Pyramidal Decomposition

1. INTRODUCTION

A plethora of watermarking methods has been proposed during the two last decades. One of the most challenging issues in watermarking is to realize a compromise between transparency, robustness and capacity. This could be expressed as a multi-criteria optimization problem. Unfortunately, these criteria are conflicting and solving this problem is rather a difficult task. One approach would be to relax one or two of these criteria. For example, we could limit the study to transparency and robustness. One of the most promising approaches is to incorporate some perceptual models in the design of the watermarking scheme due to the fact that the transparency is highly related to human visual perception. Indeed, many perceptual approaches have been proposed recently to solve the trade-off between perceptual transparency and robustness [1-3]. With the increase of digital image acquisition devices, such as a smart phone and internet, many images are delivered and stored in compressed form. Some common lossy compression such as JPEG and MPEG standards are considered as unintentional attacks. For still images JPEG is one of the most used compression format. Some methods for preventing watermarking from JPEG effect have been proposed in the literature. Most of the proposed methods are based on the selection of the frequency bands for embedding the watermark. The idea is to embed the watermark in low or middle frequencies to make the signal more resistant to quantization artifacts. The method proposed by Koch and Zhao [4] is based on this idea. It operates in the DCT domain by using a JPEG compression model and pulse position modulation technique. The watermark is randomly embedded into the quantized DCT coefficients in the middle-frequency bands. Another similar method has been proposed by Cox et al. [5] where the watermark is embedded in the low frequency. This provides a more robustness to JPEG compression attack but at the expense of weak transparency. To circumvent this problem, some HVS (Human Visual System) based approaches have been proposed in the literature [6]. In [7], a simple perceptual watermarking approach is proposed. The idea is control the visibility of the watermark by adjusting the level of quality by analyzing the residual signal between the original image and its compressed version. Another interesting DCT-based method has been proposed by Seo et al. [8]. However, it suffers from some weaknesses, especially on the transparency side. To overcome this drawback, a JND model is used to strengthen the transparency.

In this paper, we follow the same reasoning by using a perceptual approach that ensures high level of transparency and robustness against block-based compression methods and especially JPEG compression. Two models, namely, Blocking Effect Prediction (BEP) and Pyramidal JND (PJND) are introduced in the design of the proposed watermarking scheme. The BEP model, introduced in [9], is used here to predict the appearance of blocking effect so as to select the regions, less affected by JPEG compression, where the watermark could be embedded. The strength of the watermark is controlled through the use of the proposed pyramidal JND model. Finally, the combination of the two models offers more gain in robustness and transparency.

The paper is organized as follows. Section 2 introduces the perceptual models and the watermarking process. Section 3 is devoted to the performance evaluation of the method. Finally, concluding remarks and some perspectives are given in section 4.

2. THE PROPOSED METHOD

978-1-4673-0276-0/12/$31.00 ©2012 IEEE

In the following, we give a brief description of our proposed method. Firstly, we introduce the pyramidal representation of the image based on which the JND and BEP models have been designed. Then, we explain how to use these models for embedding the watermark.

2.1. Pyramidal JND models

The pyramidal representation is one of the most widely used decompositions for image analysis. It provides an efficient and simple multi-resolution representation of the image [10]. This representation implicitly takes into account the way the HVS resolves details. One of the most studied characteristics of the HVS is its contrast sensitivity function (CSF). This function is exploited in many perceptual approaches for image processing and analysis. In the proposed JND model we express this function at different levels of the pyramidal representation as done in [1]. Following Burt and Adelson's scheme [10], two pyramids are defined for each image, namely Gaussian Pyramid (GP) and Laplacian Pyramid (LP). The contrast threshold (CT) is measured at each level of the pyramidal representation of the image. It could be noticed that LP contains details and hence the most relevant information on contrast. Here, we use the CSF proposed by Barten [11] thanks to its flexibility and relative simplicity. It is given by:

$$CSF(f,L) = a.f.\exp(-b.f).\sqrt{1+c.\exp(b.f)} \quad (1)$$

where *c* is a constant, *a* and *b* are functions of the global luminance *L* (in *cd/m2*) given by:

$$L(x,y) = L_0 + L_I(x,y) \quad (2)$$

where L_0 is the ambient luminance and L_I is the local luminance computed from the corresponding GP at the $(k+1)^{th}$ level, $G_{k+1}(x,y)$. Notice that the gray level should be transformed into luminance as follows:

$$L_I(x,y) = \max(L_{\max}(\frac{G_{k+1}(x,y)}{255})^{\gamma}, L_{\min}) \quad (3)$$

where L_{max} and L_{min} are respectively the maximum and minimum luminance of the display, whereas γ is the gamma correction factor. The local contrast threshold (CT), at each pixel, is then expressed as the contrast threshold at the peak frequency of the channel $CT(f_k^{peak})$ weighted by the contribution of the level. It is defined by:

$$CT_k(x,y) = CT(f_k^{peak})\frac{L_k(x,y)}{\sum_{k=1}^{K} L_k(x,y)} \quad (4)$$

where $L_k(x,y)$ is the LP coefficient of the pixel *(x,y)* at the kth level, $CT(f_k^{peak})$ is the contrast threshold at the peak frequency of that level.

The detection threshold $T_{0k}(x,y)$ which accounts for contrast sensitivity and luminance adaptation is then computed by*:*

$$T_{0k}(x,y) = CT_k(x,y).G_{k+1}(x,y) \quad (5)$$

The JND threshold is finally obtained by incorporating a contrast masking model inspired from [9] and given by:

$$JND_k(x,y) = \begin{cases} T_{0k}(x,y) & \text{if } |L_k(x,y)| \le T_{0k}(x,y) \\ T_{0k}(x,y).\left(\frac{|L_k(x,y)|}{T_{0k}(x,y)}\right)^{\varepsilon} & \text{otherwise} \end{cases} \quad (6)$$

where ε is a factor that describes the degree of masking, $0.6 \le \varepsilon \le 1$ [12].

2.2 The watermark embedding scheme

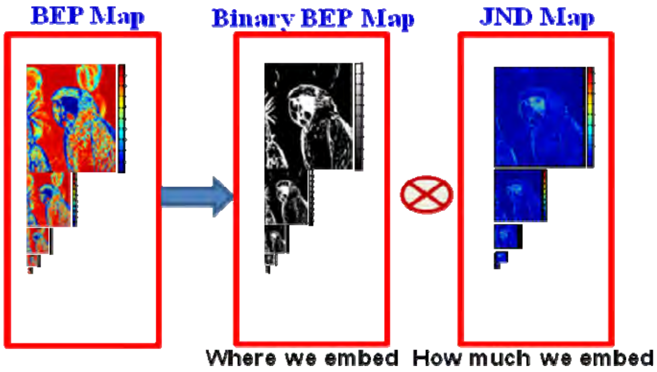
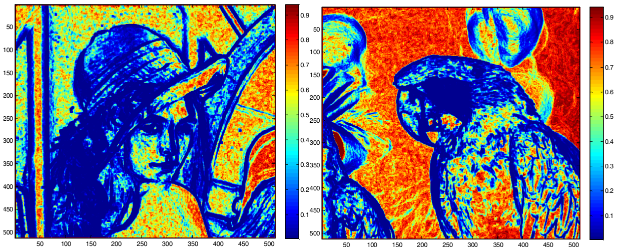


Figure 1. Illustration of the embedding process that combines JND and BEP models

The proposed watermark embedding scheme is illustrated in Fig. 1. The above JND is combined with a blocking effect prediction (BEP) map computed using a local analysis of the signal activity and a learning process as described in [9].

To have an idea about the BEP, Fig. 2 illustrates the BEP maps of four test images at a JPEG quality factor of 20. These BEP maps indicate at each pixel (*x,y*) and for a given quality factor, the weight of blocking effect appearance.



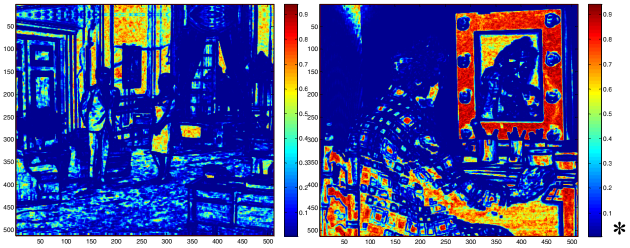


Figure 2. BEP heat maps of different images at the quality factor of 20

In this paper, we consider only the zero-bit additive scheme using a bipolar pseudo-random sequence $W_k \in \{-1,1\}$ embedded into each level of the LP.

In [1], a JND based embedding is proposed, using only the PJND model. The idea is to hide the watermark just beneath the detection threshold so as to achieve the trade-off between robustness and transparency. This could be done using the following scheme at different levels of the decomposition:

$$L_{wk}(x,y) = L_k(x,y) + JND_k(x,y).W_k(x,y) \tag{7}$$

Now, with the BEP map at hand, we can embed the watermark as follows to prevent it from being attacked by the Jpeg compression:

$$L_{wk} = L_k(x,y) + \alpha.BPBEP_k(x,y).W_k(x,y) \tag{8}$$

where α is a global factor for controlling the watermark strength, PBEP and BPBEP are the Pyramidal BEP map and its binary version, respectively.

The PBEP map is the Gaussian pyramid representation of the BEP map. Its binary version, noted BPBEP is obtained by thresholding the PBEP map. Here, we set the threshold to 0.3 to ensure that selected zones are the least affected.

So, with the BEP model, we can obtain gain in robustness against Jpeg compression. Nevertheless, the embedding strength is not optimally defined since it requires an additional global factor. On the other hand, the JND model offers an explicit and more sophisticated manner to determine the visual detection threshold for the watermark strength. Hence, it is necessary to take advantages of these two models. Such a solution can be obtained by combining the BEP map and the JND map. The final embedding rule is now redefined as follows:

$$L_{wk}(x,y) = L_k(x,y) + BPBEP_k(x,y).JND_k(x,y).W_k(i) \tag{9}$$

For watermark detection, a simple linear correlation based detector is used as follows:

$$Cor_{lc} = \frac{1}{N}\sum_{i=1}^{N} f(i).w(i) \tag{10}$$

where w is the watermark, f is the watermarked coefficient (possibly attacked) and N is the number of watermarked coefficients.

3. PERFORMANCE EVALUATION

A series of experiments has been performed in order to evaluate the efficiency of the proposed watermarking scheme. Here we limit the investigation of the method performance in terms of transparency and robustness. Although we mainly focus on gray-level image, the extension of the proposed method to color images is straightforward. The method is evaluated on a set of 8 test images.

In Table 1, the robustness of the proposed method is demonstrated in comparison with the method in [1] which uses only the JND model. It is clear that the proposed method outperforms the method in [1] for Jpeg compression attack. Further evaluation is given in Fig. 4. In this figure, the detector output of the proposed scheme and the JND based scheme against Jpeg compression (tested on 8 images) is plotted as a function of the quality factor. It can be seen that using both BEP and JND maps offers a significant gain in robustness to Jpeg compression.

Now, it is also necessary to know if the overall robustness (against other attacks) decreases or increases with the combined schemes. The results in Table I indicate that there are no significant difference in robustness between these two schemes to other attacks (exclude Jpeg compression). Furthermore, robustness against some attacks “like Jpeg” (Jpeg2000) is even slightly improved. This can be explained by interesting remark on common characteristic of the two maps. The JND map also presents very high values at regions “selected” by the BEP map so that most of watermark energy concentrates at these common regions. This leads to the fact that embedding at other regions does not help to increase much more in terms of watermark energy.

In terms of transparency, it is also observed that the proposed method is equal or even better than the method in [1] (see Fig. 3). This can be explained by the fact that the transparency is ensured by using the JND model and moreover, by using the BEP model. Furthermore, in contrast to the scheme proposed in [1], the watermark is embedded only in a small portion of the image.

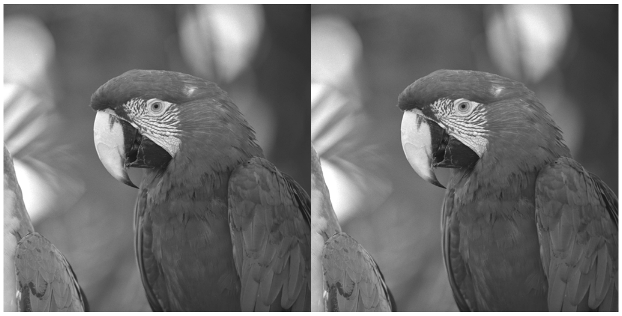


Figure 3. Watermarked images using the JND based scheme [1] (left) and the proposed scheme (right)

Table I: Robustness Evaluation with some “signal processing attacks” and Photoshop manipulations.

Attack Type	JND [1]	BEP+JND
Centered cropping	0.5%	1%
Jpeg compression	q=9	q=3
Jpeg 2000	0.1 bpp	0.08 bpp
Gaussian Noise	$\sigma = 70\%$	$\sigma = 70\%$
Wiener filtering	Ok	Ok
Median filtering	5x5	5x5
Sharpening	Ok	Ok
Blurring	Ok	Ok
Bit plan reduction	Ok	Ok
Histogram Equal.	Ok	Ok
Rescale (50%)	Ok	Ok

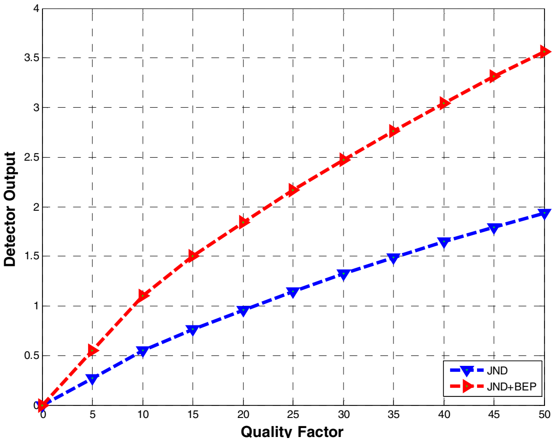


Figure 4. Robustness against Jpeg compression at different quality factors. The results are averaged on a set of 8 test images.

4. CONCLUSION AND PERSPECTIVES

An efficient perceptual watermarking has been proposed for still image watermarking. Through the obtained results, it is shown that by incorporating some characteristics of the HVS, it is possible to achieve the trade-off between transparency and robustness. Indeed, it is demonstrated that the introduction of the pyramidal representation of the JND and the BEP map offers a strong robustness to JPEG compression attack while maintaining a high level of transparency. The proposed method could be extended to deal with other block-based compression attacks. It could be noticed that only gray-level images are considered in this study but it is straightforward to extend this approach to color images. Another issue to be addressed is the extension of the proposed technique to video, but at the cost of an increased computational burden. Furthermore, more elaborated spatio-temporal models of the JND should be introduced in order to deal with video.

7. REFERENCES

[1] P. B. Nguyen, A. Beghdadi, and M. Luong, “Perceptual watermarking using pyramidal JND maps”, in Proc. of 10th IEEE ISM’08, pp. 418-423, Berkeley, USA, 2008.

[2] P. B. Nguyen, A. Beghdadi and M. Luong, “A robust-transparent watermarking scheme based on perceptual modelling”, 7th IEEE International Workshop on Systems, Signal Processing and their Applications, Corne d’Or, Tipaza, Algeria, 2011.

[3] Y. Niu, M. Kyan, L. Ma, A. Beghdadi and S. Krishnan, “A visual saliency modulated just noticeable distortion profile for image watermarking”, 19th European Signal Processing Conference, Barcelona, Spain, 2011.

[4] E. Koch and J. Zhao, “Toward robust and hidden image copyright labeling,” in IEEE Workshop on Nonlinear Signal and Image Processing, Greece, June 1995, pp. 452–455.

[5] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, “Secure spread spectrum watermarking for multimedia,” IEEE ICIP’97, vol. 6, no. 12, pp. 1673–1687, December 1997.

[6] M. Kutter and S.Winkler, “A vision-based masking model for spread spectrum image watermarking,” IEEE Transactions on Image Processing, vol. 11, no. 1, 2002.

[7] H. J. Lee, J. H. Park, and Y. Zheng, “Digital watermarking robust against jpeg compression,” in Information Security, LNCS. 1999, vol. 1729, pp. 167–177, Springer-Verlag.

[8] H-U. Seo, J-S. Sohn, B-I. Kim, T-G. Lee, S-I. Lee, and D-G. Kim, “Robust image watermarking method using discrete cosine decompostion and just noticeable distortion,” Proc. of the 23rd ITC-CSCC’ 2008, pp. 765–768.

[9] A. Chetouani, G. Mostafaoui and A. Beghdadi, “Predicting blocking effects in the spatial domain using a learning approach”, in Proc. of SIGMAP 2008, pp. 197-201, Porto, Portugal, 2008.

[10] P. J. Burt and E. H. Adelson, “The Laplacian Pyramid as a Compact Image Code”, in IEEE Transactions on Communications, April 1983, pp. 532-540.

[11] P. G. J. Barten, “Evaluation of Subjective Image Quality with the Square-Root Integral Method”, in Jour. of the Opt. Society of America A: Vol. 7, Issue 10, Oct. 1990, pp. 2024-2031.

[12] G. E. Legge and J. M. Foley, “Contrast Masking in Human Vision”, in Jour. of the Opt. Soc. of America, 1980, pp. 1458-1471.