MULTI-LEVEL WATERMARKING WITH INDEPENDENT DECODING

M. Butman

Bar-Ilan University
Department of Math and Computer Science
Ramat-Gan 52905, Israel

H.Z. Hel-Or

Department of Computer Science
Haifa University
Haifa 31905, Israel

ABSTRACT

A drawback of most watermarking techniques is the need for some additional information in order to retrieve the watermark. Additionally, the robustness of the watermark decreases as the number of information bits stored in the image increases. We present a Watermarking technique which requires no information for decoding in addition to the watermarked image. The watermark is multi-level with few bits embedded robustly at low levels and longer watermark sequences embedded less robustly at higher levels. This allows detection of ”tampered” and ”attacked” images by detection of existence of the watermark at low levels and deterioration of the watermark at high levels. In order to prevent interference of the watermarks at different levels various image representation spaces are used. The watermark is shown to be non visible and robust.

1. INTRODUCTION

Digital Watermarking is a technique for marking and labeling digital images. Methods have been developed to withstand attacks on watermarked images including image processing, image compression and geometric transformations attacks. The various methods previously developed can be categorized according to the following characteristics:

- Representation Space - Watermarking embeds a code by changing the image. The embedding can be performed in the image represented in various domains. The domain chosen for image representation affects the robustness and capabilities of the watermark to withstand attacks. Watermarking in the spatial domain [1, 2]has the advantage of maintaining locational information, thus geometrical attacks such as translation and especially cropping are more likely to be overcome. Watermarking in the frequency domain [3, 4] does not preserve locational information but attends to the frequency content of the image enabling robustness under filtering such as blurring, high pass image enhancement, etc. Local frequency transforms of the image (such as block DCT and Wavelet Transforms) [5, 6] have the advantage of both locality and frequency extraction, but are highly sensitive to attacks involving translation or cropping of the image.
- Code Length - The watermark itself is a code varying in length from 1 bit (interpreted as: a watermark exists or does not exist in the image), to a sequence of bits. The 1-bit watermarks (e.g. [1, 3]) are typically based on statistically evaluating the probability of the image containing a watermark. The watermark code containing a number of bits typically vary the values of the image representation in a sequence of changes that depend on a random generator and the bit values of the code (e.g. [2]. It is shown experimentally that 1-bit watermarks are far more robust under attacks than the watermarking techniques associated with a sequence of bits.
- Information Required for Decoding - A drawback of most watermarking techniques is the need for some additional information in order to retrieve the watermark. In many cases, typically for the 1-bit watermarks, only the RNGS (random number generator seed) is required, and possibly the image size (e.g. [1]). For decoding of longer bit sequences, the watermarking techniques typically require, in addition to the RNGS, also the original image and/or the watermark code (e.g. [3] requires the original image the RGNS and the watermark code). The external information required for decoding puts a heavy restriction on watermarking techniques, in that a list or database must be maintained to associate a given watermarked image with it’s watermark decoding parameters (which might include the original image).

Combining several watermarking techniques has already been suggested [7], however in this work, we exploit the advantages of various image representation domains to embed a multi-level watermark which is robust under cropping as well as filtering and compression attacks and requires no additional information for decoding. The watermarking technique suggested in this paper maintains very high robustness for the 1-bit content of the watermark, and lower robustness for the multi-bit sequence portion of the watermark. This method allows detection of ”tampered” and ”attacked” images by detection of existence of the watermark (1-bit) and the deterioration of the watermark (multi-bit).

2. MULTI-LEVEL WATERMARKING

The watermarking technique suggested is designed hierarchically. Every level of watermarking takes advantage of a different image representation space. Every level of the watermark supplies enough information to decode the following watermark level so that no external information is needed for decoding. The watermarking levels are such that few watermark bits are embedded in the first levels and a larger number of bits in higher levels. Accordingly, robustness under attacks is high in the first levels and decreases in following levels. A schematic diagram of the method is shown in Figure 1. The watermarking technique we developed consists of three Levels briefly described in the following sections. The encoding and the decoding process progress according to these levels. During decoding, the output of each level serves as input to the next level.

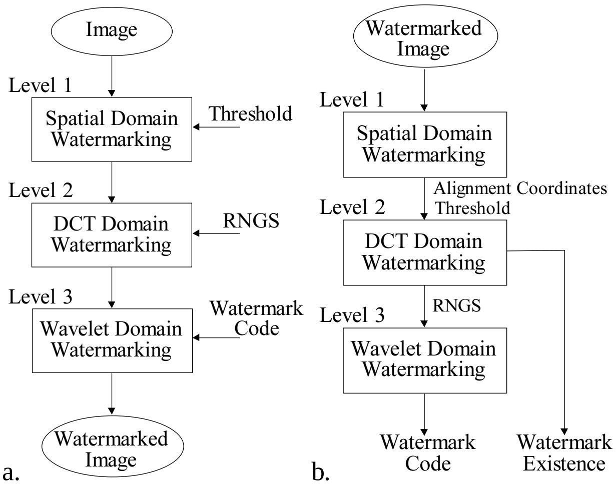

Fig. 1. Multi-Level Watermarking
a) the encoding process b) the decoding process

2.1. Level 1 - Spatial Domain

To overcome geometric attacks such as cropping, code embedding should be performed in the spatial domain [8, 9]. Since all other levels of watermarking assume that positioning of the cropped image is known (absolute pixel coordinates are required), the first level of watermarking must deal with positioning the possibly cropped image. Thus the first level performs embedding in the spatial domain. In a method similar to [1], embedded patches serve as markers to position the image. We extended the method, to embed two or more classes of patches to decrease run time significantly and to eliminate the need of knowing the image size. The method was tested and shown to be invisble and able to detect the watermark upto cropping of 75% of the image. Details of the method can be found in [9, 10]. In addition to the cropping resistant watermark, an additional watermark is embedded encoding a number. The outcome of the decoding is the positioning of the image and a single additional number which serves as input to the next level.

2.2. Level 2 - DCT Domain

At this level the goal is to embed a small number of code bits, however embedding must be robust. Embedding at this level must not interfere with the embedding of Level 1. Thus a local frequency domain is chosen for embedding, specifically the block DCT transform is used. A novel algorithm is suggested in which randomly chosen 8×8 blocks are altered in their mid frequency bands by nullifying coefficients below a given threshold. Decoding is based on statistical estimation of the appropriate coefficients. The threshold is image dependent and thus must be supplied to the decoder. In the hierarchical watermarking scheme, the threshold required for the DCT domain decoding is embedded as the additional number in Level 1 (see above) and is supplied to Level 2 during decoding.

Since the RNGS used to select the random DCT blocks during encoding, is unknown to the decoder, decoding in the DCT domain is performed by evaluating the nullified coefficients in randomly chosen blocks according to a range of RNGS values. The RNGS which produces a statistically significant value for the nullified coefficients (i.e. a value statistically close to zero compared to the values obtained for incorrect RNGSs) is determined as the correct RNGS. Details of the method can be found in [10].

The output of this watermarking stage is the RNGS that was found. This RNGS value is used as input to the next level. The watermarking method used in Level 2 is robust under attacks and under multiple watermark encodings. Thus, in practice we embed a few watermarks in the manner

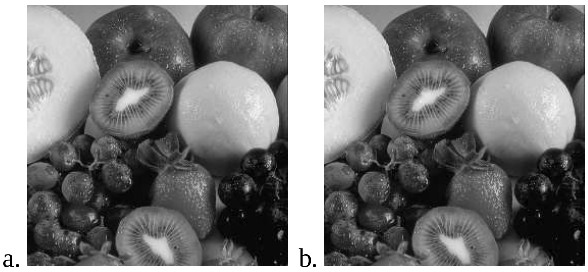

Fig. 2. a) Original image b) Watermarked image (Level 2)

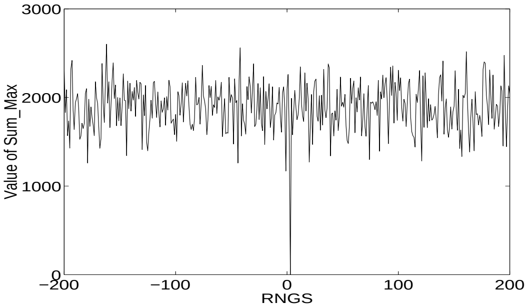

Fig. 3. Decoding results in the DCT domain

suggested, and supply a few RNGS values to Level 3.

Figure 2 shows that the watermarking technique in the DCT domain is invisible. Figure 3 shows the results of computing the value for the nullified DCT coefficients using 400 different RNGSs. Only for the correct RNGS (=3) the value is close to zero, and this deviation is statistically significant compared to the values obtain for all other RNGSs. The DCT domain watermarking was tested and found robust under various attacks including JPEG compression (upto 10% quality factor), Lowpass filtering (upto 50% filtering) and cropping (upto 75% cropping).

2.3. Level 3 - Wavelet Domain

At this level the goal is to embed a sequence of bits as the watermark code. As mentioned above, such watermarks are not very robust under attacks, and typically require either the original unmarked image or the watermark code itself. We propose a watermarking method which is less robust but requires no additional information for decoding except for an RNGS which is supplied by Level 2. To minimize interference with the watermarks embedded in previous levels, a different local frequency domain was chosen - the Wavelet domain. The method systematically sets randomly chosen Wavelet coefficients in the diagonally oriented, mid-frequency bands to predefined values. Thus at decoding, the watermark bits can be detected as well as tampered code bits. The output of the watermark decoder at this level is the sequence of watermark code bits.

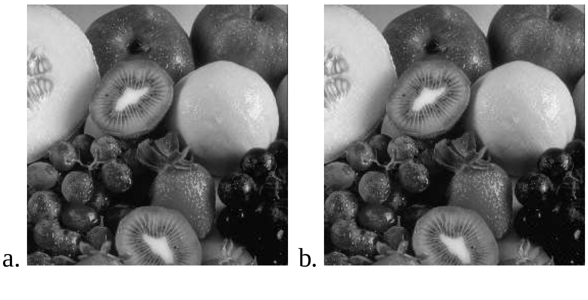

Fig. 4. a) Original image b) Watermarked image (Level 3)

Figure 4 shows that the watermarking technique in the wavelet domain is invisible. To evaluate the robustness of the wavelet domain watermarking, the percentage of correctly detected watermark bits was determined under various attacks including JPEG compression, Lowpass filtering, cropping and addition of salt&pepper noise. Figure 5 shows the results for JPEG and Lowpass attacks in terms of the percentage of successful watermark bit detection as a function of the attack strength.

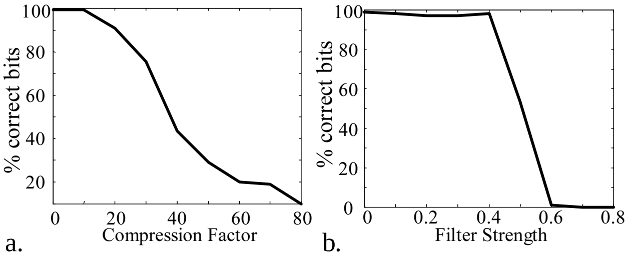

Fig. 5. Decoding results at Level 3 following:
a) JPEG compression. b) LowPass Filtering.

3. MULTI-LEVEL WATERMARKING - RESULTS

Combining the three levels of watermarking, a multi-level watermarking scheme is obtained. The watermark is not visable (Figure 6) and is shown to be robust under attacks including cropping,JPEG compression, Lowpass filtering and addition of salt&pepper noise. The watermark was also tested under StirMark attacks [11] giving satisfactory results. The robustness of the multi-level watermarking is dependent on the robustness of each level of watermarking. The inter-level interference is minimal due to the choice of different embedding domains. The existence of the watermark (1-bit) is highly robust and withstands attacks such as JPEG compression up to 80% whereas detection of the multi-bit code deteriorates at lower compression rates. Figure 7 shows decoding results under various attacks.

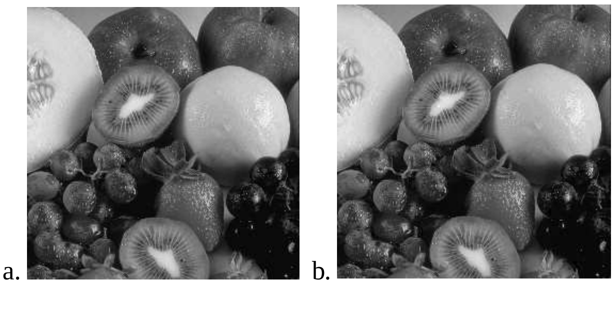

Fig. 6. Multi-level Watermarking
a) Original image. b) Watermarked image.

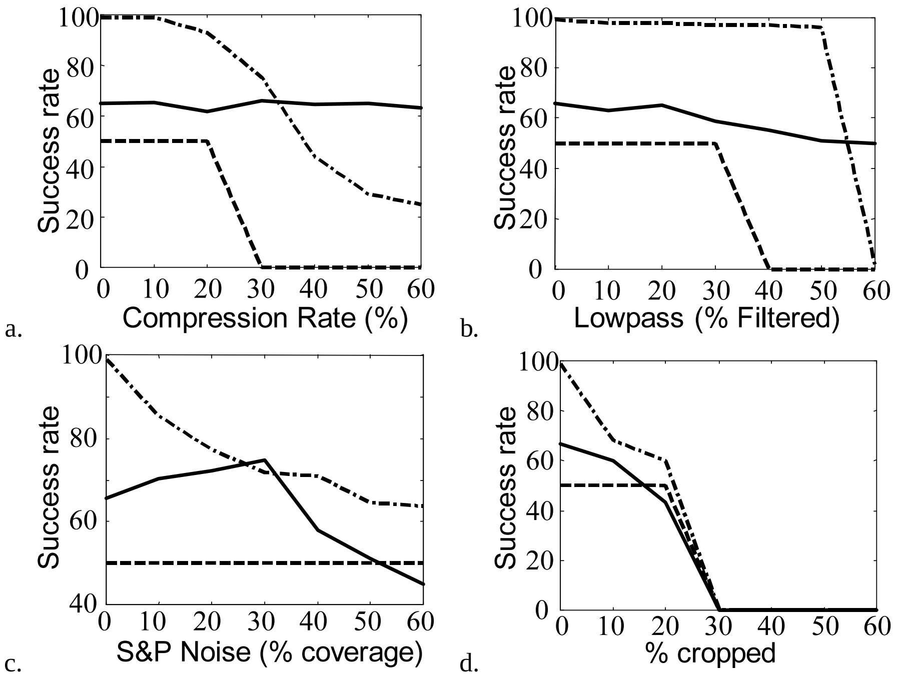

Fig. 7. Decoding results for the Stand-alone watermarking technique.
a) JPEG compression. b) Lowpass filtering. c) Salt&Pepper noise. d) Cropping.
Level 1: The dashed line represents successful detection of the alignment coordinates (value ’50’ represents success and ’0’ represents failure). Level 2: The continuous line represents the distance (in s.t.d. units) between the difference value for the correct RNGS and the average value over all RNGSs (scaled by 10).
Level 3: The Dashed dot line represents the percentage of successfully detected watermark bits.
The horizontal axis represents the attack strength.

4. REFERENCES

[1] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, “Techniques for data hiding,” *IBM Systems Journal*, vol. 35, no. 3/4, pp. 313–336, 1996.

[2] M. Kutter, F. Jordan, and F. Bossen, “Digital signature of color images using amplitude modulation,” in *Proc. SPIE - The Int. Soc. for Opt. Eng.*, San Jose CA, Feb 1997, vol. 3022, pp. 518–526.

[3] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, “A secure, robust watermarking for multimedia,” in *1st Int. Workshop on Information Hiding*, R. Anderson, Ed., Cambridge, UK, 1996, vol. LNCS 1174, pp. 185–206, Springer-Verlag.

[4] J. J. K. O Ruanaidh, W. J. Dowling, and F. M. Boland, “Phase watermarking of images,” in *IEEE Int. Conference on Image Processing*, Lausanne, Switzerland, Sept 1996, vol. 3, pp. 239–242.

[5] A. G. Bors and I. Pitas, “Image watermarking using dct domain constraints,” in *IEEE Int. Conference on Image Processing*, Lusanne, Switzerland, Sept 1996, vol. 3, pp. 231–234.

[6] D. Kundur and D. Hatzinakos, “Digital watermarking using multiresolution wavelet decompression,” in *IEEE Int. Conference on Image Processing*, Santa Barbara, CA, October 1997, vol. 1, pp. 544–547.

[7] J. Friedrich, “Combining low frequency and spread spectrum watermarking,” in *Proc. SPIE - The Int. Society for Optical Engineering*, San Jose, CA, 1998, vol. 3456, pp. 2–12.

[8] R. W. Hwang, *A Robust Algorithm for Information Hiding in Digital Pictures*, Ph.D. thesis, Massachusetts Institute of Technology, Media Laboratory, Boston, MA, 1999.

[9] M. Butman and H. Hel-Or, “Watermarks for overcomming cropping attacks,” *Submitted to ICIP01*, 2001.

[10] M. Butman, “Stand-alone multi-level watermarking,” M.S. thesis, Bar-Ilan University, Dept of Computer Science, Ramat-Gan, Israel, 2000.

[11] R.J. Anderson F.A.P. Petitcolas and M.G. Kuhn, “Attacks on copyright marking systems,” in *Information Hiding, 2nd Int. Workshop*, 1998, pp. 219–239.