Blind image watermarking in Wavelet-domain robust to printing and smart-phone acquisition.

Andrea Leopardi, Davide Soresina, Davide Marcantonio, Alain Malacarne, Nicola Conci, Giulia Boato
DISI - University of Trento

Abstract

In these days and age, printed and digital images are the principal means of communication chosen by companies to convey information about their products, since visual contents produce a direct and effective influence on people. At the same time, imagery can be digitally enriched with additional information, imperceptible to the human eye, yet still retrievable using specific software or hardware: this is the case of digital watermarking. In this work, we propose a digital watermarking pipeline that performs information embedding robust to printing operations and enables blind detection from digital acquisitions. We select a watermark from a set of orthogonal antipodal matrices and adaptively insert repeated copies in the horizontal, vertical and diagonal sub-bands of the Wavelet domain. Blind detection is achieved denoising the digitally acquired marked image, retrieving the watermark information from the Wavelet domain, restoring the original scaling with an optimization algorithm and computing a similarity score with each of the possible orthogonal marks. Our system is able to reconstruct the embedded information both in the cases of acquisition from digital and printed watermarked images. It is able to recognize the correct mark among the set of possible messages, even when considering poor-quality printing systems.

Introduction and Related Works

The field of digital watermarking is rich of solutions that are capable of embedding hidden information in digital images without degrading image quality, yet achieving reasonable robustness. [1][2][3][4]: attacks such as lossy compression or standard image processing operations are well countered in most of the proposed pipelines. At the same time, there is a growing interest towards marking for tagging purposes, especially on physical products. Common examples are bar- or QR-codes, which efficiently convey information when scanned, yet are visually bulky and may degrade product perception, covering parts of it. Embedding a hidden message into the visual product itself helps overcoming the waste of space of traditional solutions, even though technical challenges due to printing procedure and smart-phone acquisitions are still many and, to date, not much has been accomplished in this direction.

As far as printing is concerned, few works have been presented in the framework of digital watermarking. In [5], it is proposed a watermarking procedure able to resist the combination of printing and scanning operations, embedding the spectrum hologram of a given watermark, generated by conjugate-symmetric extension, in the Discrete Cosine Transform (DCT) domain. Instead, [6] presents a blind architecture both in the Discrete Wavelet Transform and DCT domains: in addition, authors propose a thorough modeling of the print-scan attack combination, including noise addition, non-linear effects and geometrical transformations. A more recent approach, presented in [7], performs the information embedding in image halftones, forcing a given correlation to pairs of pixels, while employing a Direct Binary Search algorithm for watermarking detection. However, all these works do not take into consideration digital acquisition of the watermarked content by means of smart-phone devices.

Digimarc Corporation [8] introduced the concept of Smart Images, physical imagery that is digitally enriched with more information, and proposed a blind watermarking pipeline that exploits an additive Spread Spectrum approach in the DCT domain combined with error-correction codes. We propose an alternative embedding procedure, while still maintaining a blind detection approach: the DWT domain is employed in combination with orthogonal watermarking dictionaries that help reducing misdetections in message retrieval. Furthermore, to better counter the effects of smart-phone acquisitions, we propose to use an optimization algorithm for scaling estimation, resorting to the correct image proportion and improving the information extraction process.

Proposed Method

In this section, a blind watermarking approach robust to printing and scanning operations is presented in all its steps. The embedding phase follows a traditional approach that aims at spreading watermarking information in the Wavelet domain, considering also the local variance to reduce visual degradation. The blind detection procedure relies on an optimization algorithm for scale restoration, that aims at countering the effects of digital acquisition through smart-phone devices, combined with a median filtering of the obtained image.

Watermark Generation

The watermarking information that will be considered in this approach is represented with square patches of antipodal values and side L. We opted to choose orthogonal square patches in order to maximize diversity between each watermark signal, obtaining a number of possible messages equal to L^2:

$$W_i \in \{-1, 1\}^{LxL}, \quad i = 1, ..., L^2 \tag{1}$$

$$\sum_{x=1}^{L} \sum_{y=1}^{L} W_i(x,y) * W_j(x,y) = 0, \quad i \neq j \tag{2}$$

In Figure 1 two examples of watermark patches are represented, considering side length $L = 16$.
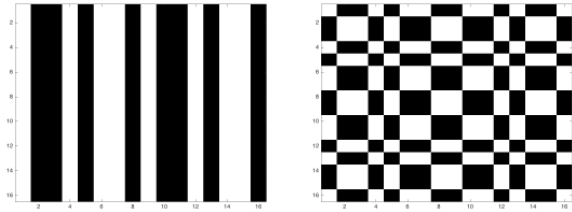
Figure 1: Examples of watermark patches (#15, #24) with $L = 16$

Moreover, in order to improve detection performance, the following condition is applied to discard similar messages, reducing confusion in the detection procedure:

$$sim(W_i, W_j) = \frac{\sum_x \sum_y \mathscr{F}^{-1}\{|\mathscr{F}\{W_i\} * \mathscr{F}\{W_j^*\}|\}}{L^2} \quad (3)$$

$$max(sim(W_i, W_j)) < 0.1 * sim(W_i, W_i), \quad i \neq j \quad (4)$$

In (3), W_j^* corresponds to a 180 rotation of watermark W_j, while $\mathscr{F}$ and $\mathscr{F}^{-1}$ refer to the Fourier and inverse Fourier transform, respectively. Such similarity metric relies on a matched filtering approach [9], commonly used to detect the presence of a specific template in an unknown signal.

Applying (3) and (4) to the initial set of L^2 watermarks results in a reduced dictionary M composed of messages with limited cross-similarity values. The number of messages included in M, defined as N_M, changes according to the chosen side L. Figure 2 represents the cross-similarity values calculated over the watermarks belonging to the reduced dictionary M with $L = 16$. As expected, when considering W_i and W_j with $i \neq j$, the similarity function results in reduced values, hence indicating great diversity among the watermarks that respect condition (4).
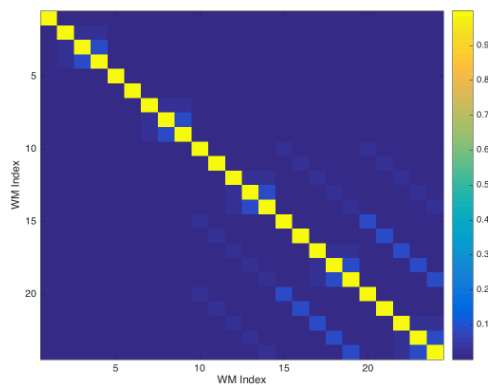

Figure 2: Cross-similarity among the watermarks in the reduced dictionary M.

Watermark Emebdding

The proposed embedding approach is presented in the block diagram of Figure 3.
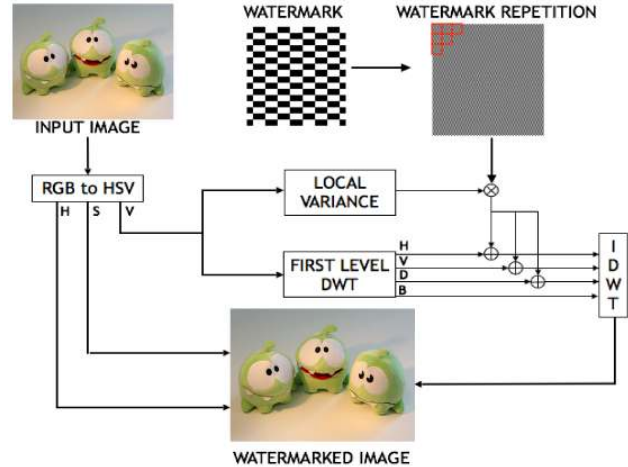

Figure 3: Proposed embedding scheme

The procedure starts with the conversion of the input image I from the RGB to the HSV color space, considering only the Value component for further processing. The first level Wavelet sub-bands are estimated using the image Value component and the Symlets wavelet family of order 4: all three directions are considered for watermarking embedding. Figure 4 shows examples of vertical V, horizontal H and diagonal D sub-bands as well as the approximation coefficients of the first Wavelet level applied to a given image. In order to introduce redundancy, we propose generate a matrix R composed by the spatial repetition of the following base structure:

$$M = \begin{bmatrix} W_s & W_m \\ W_m & W_s \end{bmatrix}$$

where W_m is the watermark message and W_s is a fixed synchronization signal exploited for scaling reconstruction in the detection phase. Structure M is spatially repeated until the dimensions of R match those of each wavelet subband.
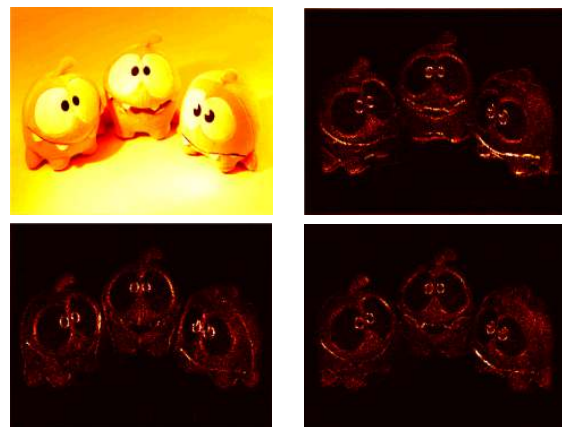

Figure 4: Lowpass component, vertical, horizontal and diagonal Wavelet sub-bands

In addition to the Wavelet sub-bands computation, image I is downsized, halving its dimensions and generating I_d. Local

variance Ω is calculated on I_d, employing a $[7 \times 7]$ sliding window according to (5).

$$\Omega(x,y) = \frac{1}{49} * \sum_{u=-3}^{3} \sum_{v=-3}^{3} |I_d(x+u,y+v) - \mu|^2 \qquad (5)$$

where μ is the mean value over the same window:

$$\mu = \frac{1}{49} * \sum_{u=-3}^{3} \sum_{v=-3}^{3} I_d(x+u,y+v) \qquad (6)$$

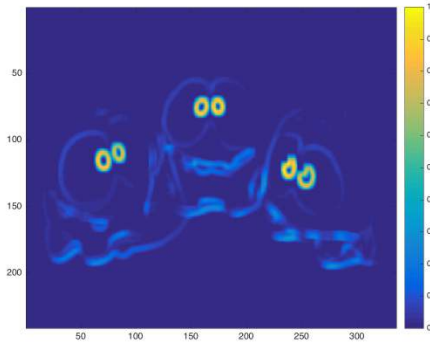For convenience, values of Ω are normalized to $[0,1]$.

Figure 5: Normalized local variance Ω.

Watermark embedding is performed using an additive approach in each of the Wavelet directions, according to the following:

$$E_i = O_i + \alpha \cdot \Omega \cdot R, \quad i = V, H, D \qquad (7)$$

where O_i refers to the original wavelet sub-bands and E_i to the corresponding embedded counterparts. Ω is used as weight to add more information in textured areas while reducing embedding strength in flat regions, exploiting the human visual system characteristics. Watermark intensity is also modulated by the scalar multiplier α, which is manually increased to introduce robustness or decreased to achieve imperceptibility. The watermarked image I_{wm} is finally obtained applying the inverse Wavelet transform using the modified sub-bands.

To monitor visual performance and evaluate the content degradation imposed by the inserted information, image quality assessment metrics ($PSNR$ [10], $WPSNR$ and $SSIM$ [11]) are applied to I and I_{wm}.

Printing rules

In order to properly print an embedded image and hence not to lose the watermark content, a guideline for printing is provided. Supposing that the bottleneck in the *print/acquire* process is the acquisition, then the dimension of the printed version of the image must be adjusted accordingly. Given a $M \times N$ resolution image, in order to find the *width* in centimeters of the image after the print phase the following equation is used:

Figure 6: Watermarked image I_{wm}

$$width = \frac{N}{ppi_{det}^{min}} \cdot 2.54[cm] \qquad (8)$$

where 2.54 accounts only for Inch to *cm* conversion, *ppi* (points-per-Inch) is the minimum number of points per Inch (of the watermarked image) expected to be read by the acquiring device (scanner or smart-phone) at the operational distance from this image. In the experiments $ppi_{det}^{min} = 100$ was set to decide for the printing size.

Watermark Detection

Figure 7 shows the proposed detection scheme, presenting the blind approach applied on a watermarked image I_{wm}^*.
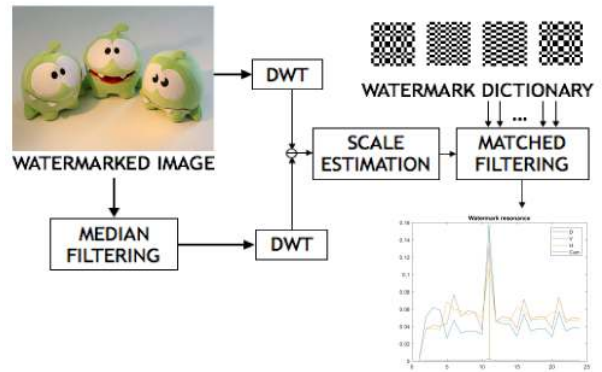

Figure 7: Proposed detection scheme

Since the embedding procedure introduces a noise-like layer on the original image, we propose in the detection phase to exploit a denoised version of the watermarked image, obtained applying a $[17 \times 17]$ median filter to the watermarked image. Figure 8 shows an image area where the noise pattern caused by the embedding operation is visible, and the same patch processed with the median filtering.

Consequently, the first level Wavelet sub-bands are extracted both from the filtered content and the watermarked image, and the corresponding directions are subtracted:

$$S_i = E_i - F_i, \quad i = V, H, D \qquad (9)$$

Figure 8: Noise-layer due to embedded information and result of the median filtering.

Figure 9: Example of cost function

with E_i, F_i and S_i referring to the embedded, filtered content and resulting sub-bands, respectively. Antipodal values are obtained applying a threshold at 0 to each difference sub-band S_i:

- $S_i \geq 0 \rightarrow S_i = 1$
- $S_i < 0 \rightarrow S_i = -1$

As a consequence, each S_i is an approximation of the inserted content R, which hosts both the selected watermark and the synchronization information.

Scaling Factor Estimation

Given that, in a real case scenario, I^*_{wm} is the result of a smartphone acquisition, user imprecisions are bound to affect the acquired content, especially in terms of scale: the distance of the device from the printed watermark product may result in content whose proportions differ from those of the digitally embedded image. In order to counter this issue, we propose to employ a Differential Evolution (DE) approach to find the best scaling factor, according to the cost function (10) that takes into account the synchronizing information inserted during the embedding phase.

$$y = -\sum_i sim_{pool}(SC_i, W_s), \quad i = V, H, D \tag{10}$$

where SC_i indicates the i-th directional band of the scaled content and $sim_{pool}(\cdot)$ as the similarity metric applied after a max pooling passage as defined in:

$$sim_{pool}(W_i, W_j) = \frac{\sum_x \sum_y pool^{max}_{16 \times 16}\{\mathcal{F}^{-1}\{|\mathcal{F}\{W_i\}\} * \mathcal{F}\{W_j^*\}|\}}{L^2} \tag{11}$$

where $pool^{max}_{16 \times 16}(\cdot)$ is the *max-pooling* operation with a 16×16 window. It is applied before summing up the coefficients and it consists in a substitution of each 16×16 non-overlapping block with the maximum value contained in it. This passage reduces the noise generated by the raster convolution (since the peak is expected ideally only on perfect match) and does not affect detection since there is spatial redundancy and the filter window is sized to the dimension of the watermark blocks.

In Figure 9 an example of cost function behavior is presented, highlighting a global minimum in correspondence to the inverse of the scaling factor that maximizes similarity between the synchronization mark and the respective estimated counterpart.
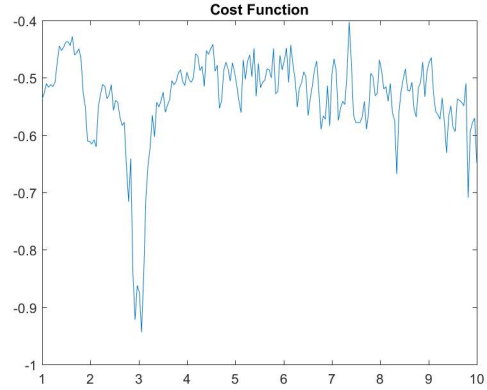
The choice of a global optimization algorithm, such as the DE, comes from the non-convexity of the cost function, that presents local minima at wrong scales. To successfully explore the cost function, the DE algorithm is initialized with a population of $P = 6$ scale factors and the initial positions are chosen using a 1D Latin Hypercube Sampling to avoid under or over sampled regions. For each scale factor, a resizing operation is applied to I^*_{wm} and the corresponding directional sub-bands are extracted and thresholded. Subsequently, the pooled similarity function (11) is evaluated between each thresholded sub-band and the synchronization mark W_s, and their result is used to evaluate the fitness function. The particular algorithm is the DE/best/2/bin and the optimized variable is the inverse of the scale factor mentioned. This choice was motivated by the speed-up of the evaluation that results from the non-linear compression of the search space. At each iteration of the DE each of the particles undergoes a crossover operation where the best solution so far is mixed with four other particles chosen at random in the current population in a differential fashion. If the resulting *child* particle has a better fitness than its parents one, then the parent is substituted with its child. This procedure is repeated for $N_{iter} = 3$ iterations, then the scale corresponding to the minimum fitness is selected as optimal for image I^*_{wm}.

The optimal scale is applied to I^*_{wm}, along with the very same detection procedure, until the thresholded sub-bands extraction. With the optimal scale application, the information contained in each S_i sub-band should result in reduced errors in the watermark evaluation.

Watermark Selection

The similarity function (3) is evaluated on each sub-band S_i against each of the watermarks signals in dictionary M (excluding the synchronization mark W_s), resulting in N_M responses per sub-band:

$$C_i(j) = sim_{pool}(S_i, W_j), \quad i = V, H, D, \quad j = 1, ..., N_M \tag{12}$$

The cumulative response information is obtained combining in a multiplicative way the responses evaluated on each sub-band:

$$C_{cum} = \prod_i C_i, \quad i = V, H, D \tag{13}$$

In Figure 10, an example response analysis is presented, considering a set of 24 marks. The response of each watermark is evaluated in every directional sub-band, and the cumulative response is obtained combining such information. This example shows how combining information helps reducing secondary peaks that may arise in the sub-bands, while maintaining a high response for the correct detection.
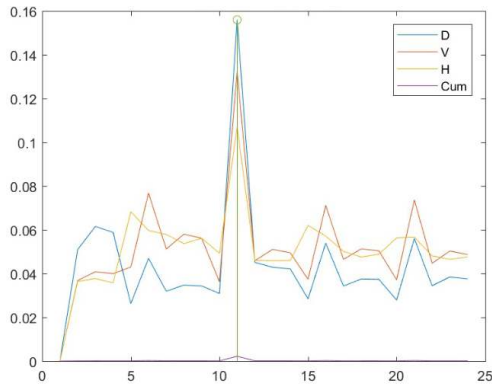

Figure 10: Similarity measure in each sub-band and cumulative response for each of the watermarks.

The watermark with the highest response in C_{cum} is selected. In the example of Figure 10, watermark #21 is selected as detected information. In addition, the confidence value for a given watermark selection is computed as:

$$CF = \frac{max(C_{cum}) - c}{max(C_{cum})} \tag{14}$$

where c corresponds to the second highest response in C_{cum}.

Experimental Setting

Dataset

The evaluation of our proposed technique is performed considering selected pictures that are likely to be employed in real scenarios, such as product leaflets or advertising imagery. We considered the Office-Home Dataset [12], originally created for object recognition purposes using deep learning approaches, as a reliable source of such imagery. The dataset consists of 65 categories of everyday objects, represented in various ways, including sketches, cliparts, real-environment pictures and real objects with removal of background content. In this context, we selected pictures of real objects with and without background information. In Figure 11, examples of such image configurations are presented, including a bed in a real unprocessed context and a set of real markers represented discarding any other object from the visualization.

Testing Phase

The proposed watermarking approach is tuned to be robust to printing operations and smart-phone acquisitions. For this reason, we selected specific image processing operations aimed at reproducing imperfections that arise from a real scenario, in order to perform digital tests. Subsequently, we performed analysis on

Figure 11: Example images extracted from the Office-Home dataset [12].

smart-phone acquisitions of printed watermarked content, to confirm the results of the first phase. The printer and scanner used for the experiments is the *Kyocera TASKalfa 4551ci*. The smart-phone test was done with a *One Plus 5T* (16 MP camera).

As to the digital reproduction of the combination of printing and smart-phone acquisition, we considered three sets of image processing operators:

- **contrast adjustments** and **histogram equalization** to take into consideration over- and under-exposure, typical consequences of reduced printing quality or illumination condition in the acquisition environment;
- **geometrical operations**, including rotation, cropping, and scaling, to simulate a non-optimal position of the smart-phone device with respect to the printed watermarked content;
- **blurring filter** to consider de-focused acquisitions;

Most processing operations require the definition of a specific parameter: Table 1 includes the various configurations that have been considered. In particular, contrast adjustments are applied to map intensity values to a new range $[G_{min}, G_{max}]$, saturating all the pixel values exceeding G_{min} and G_{max}.

Table 1: Parameter configuration for the first test phase

Operation	Parameter	Min	Max	Step
Over-exposition	G_{max}	0.387	0.9125	0.087
Under-exposition	G_{min}	0.387	0.9125	0.087
Rotation	ϕ [°]	0.5	4	0.58
Cropping	c [%]	0	99	5
Down-Scaling	s $\in$ [0,1]	0.9	0.99	0.015
Up-Scaling	s >1	2	7	0.66
Gaussian Blurring	σ	0.5	2.5	0.33

The embedding procedure was applied to a set of images (see Table 2) selected from [12], considering a dictionary of $N_M = 24$

marks: the same watermark W was inserted in all images with strength $\alpha = 10$. All the watermarks proved to be retrievable in general so we decided to keep the watermark fixed for all the images for a fair comparison of the performance. We randomly chose the watermark #6 as the hidden tag to retrieve (W_m) and the watermark #1 as the synchronization one (W_s), they are visible in Figure 15. Moreover, Table 2 reports for each embedding case the resulting *PSNR*, *WPSNR* and *SSIM*.

Table 2: Perceptual quality of selected images after watermarking, the originals are visible in Figure 16

Image	PSNR	WPSNR	SSIM
bed12	30.32	53.37	0.67
couch9	30.31	54.70	0.62
knives13	30.43	54.88	0.79
marker12	31.48	46.21	0.73
mug27	30.52	46.77	0.64
printer25	30.10	54.32	0.63
soda12	31.46	51.44	0.87
toys16	30.06	54.58	0.56

The detection logic was instead carried out on a printed version of the embedded contents, subsequently digitalized by means of a scanner device, to avoid all issues related to smart-phone acquisitions. A preliminary test was carried out directly on the digitalized images, without further processing: confidence results of this scenario are reported in Table 4. Table 4 also reports confidence results for the histogram equalization case, which did not require any specific parameter configuration. When the *confidence* value is non-zero then the correct watermark has been found. In addition, the higher the confidence the higher the certainty of correct choice (assuming the correct scale was estimated).

Table 3: Confidence values in cases of no additional processing and histogram equalization, the originals are visible in Figure 16

Image	CF [%] - No Proc	CF [%] - HistEQ
bed12	73.68	72.57
couch9	69.14	70.11
knives13	29.30	25.06
marker12	24.67	32.99
mug27	99.14	99.15
printer25	93.99	93.70
soda12	50.10	50.02
toys16	94.73	94.64

The plots reported in Figure 12 refer to the confidence results employing all the operations and parameter configurations of Table 1. We can notice that the behavior is not stable across pictures: the watermark is almost always correctly detected but only in some cases with a high confidence level for almost all considered attacks. Further research should be carried out in order to better understand which are the image characteristics allowing good performance in different scenarios.

Finally, we consider smart-phone acquisition of printed embedded contents and analyze the performance of the proposed method in such very challenging scenario. A total of 51 acquisitions have been tested with our detection approach, where ac-
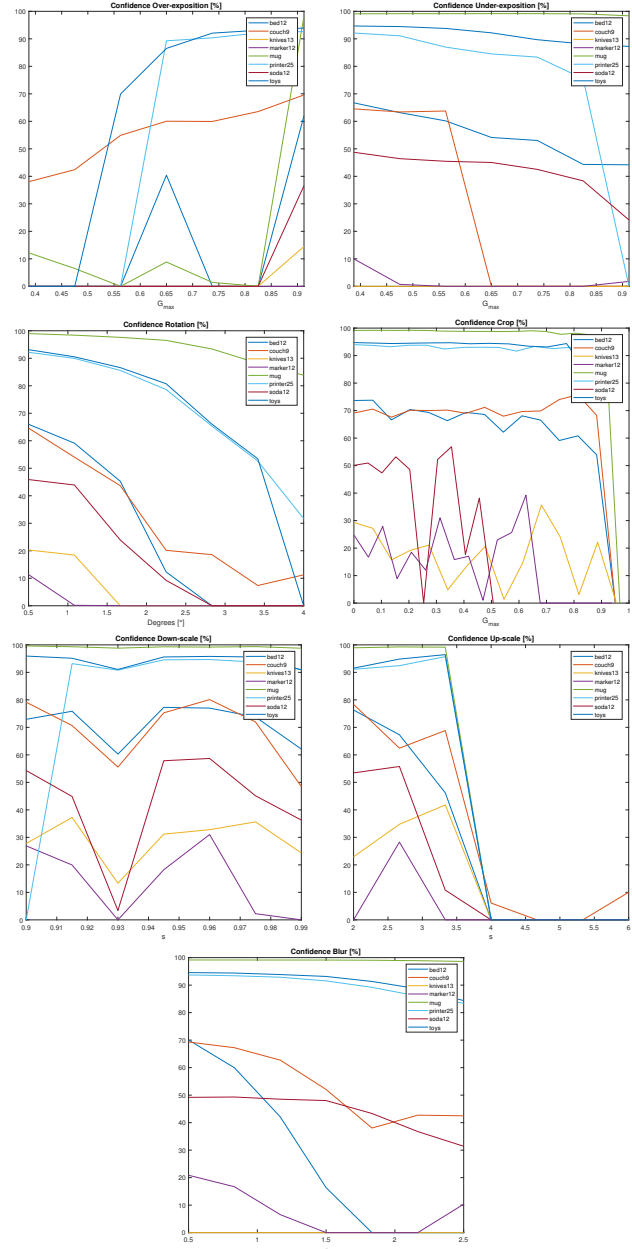

Figure 12: Confidence plots considering the processing operations and parameter configurations of Table 1.

quisition of each printed image is performed with realistic configurations of non ideal smart-phone position or lighting conditions, see examples in Figure 13.

The inserted watermark is correctly detected (green line) in various acquisition settings: standard light-bulb, no flash-light used, close-up picture; presence of surrounding environment, distant picture; flash-light, close-up. Out of the complete set of acquisitions, 10/51 (19.6%) tests result in mis-detections: many of such cases refer to the use of the smart-phone flash-light. If we do not consider the flash-light case the mis-detections decrease to 4/38 (10.5%). Figure 14 report the cumulative responses for a few cases out of the 51 tested (due to lack of space): the inserted watermark is correctly detected (green line) despite the poor quality

Figure 13: Example of printed and acquired watermarked images.

of analyzed pictures.

Table 4: Confidence values for smart-phone detection on the 8 images (Figure 16) used in the benchmark testing case.

Image	CF [%]
bed12	10.00
couch9	13.79
knives13	23.79
marker12	14.64
mug27	6.20
printer25	0.00
soda12	7.00
toys16	17.58

Conclusions

We have presented a novel watermarking system for inserting an invisible information robust to printing operations and for enabling blind detection from digital acquisitions. Our system is able to reconstruct the embedded mark both in the cases of acquisition from digital and printed watermarked images. It is able to recognize the correct mark among the set of possible messages, even when considering poor-quality printing systems. Future work will be devoted to further increase the robustness of the methodology, also extending the watermark testing with an increased number of image configurations to understand wether specific watermarks or images result in increased performance. In addition, an expansion of the dictionary to increase the number (and hence dimensions) of possible messages that can be hidden is a possible direction for future investigation.
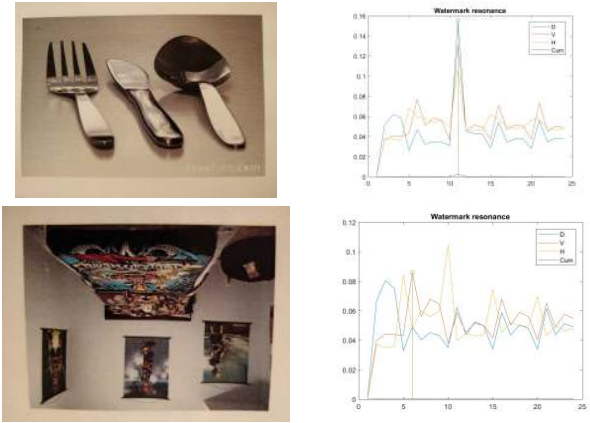

Figure 14: Photographs and corresponding similarity scores for two tested, printed and acquired images: the green vertical line highlights the correct watermark.
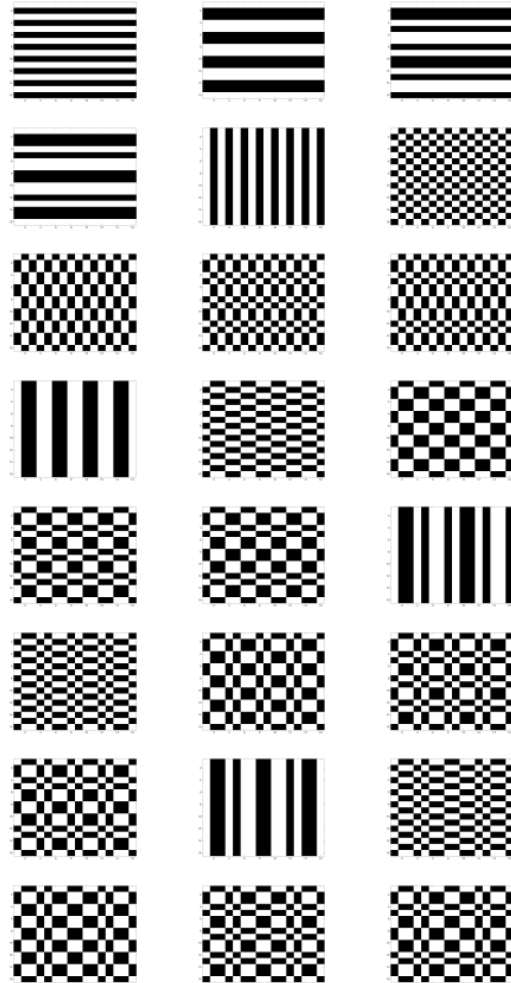

Figure 15: The complete set of the 24 generated watermarks in order of index from the top-left as #1 to the down-right as #24.

References

[1] F. Battisti, M. Cancellaro, G. Boato, M. Carli, and A. Neri, "Joint watermarking and encryption of color images in the fibonacci-

Figure 16: The set used in the benchmark, from top left: *bed12, couch9, knives13, marker12, mug27, printer25, soda12, toys16*

haar domain," *EURASIP Journal on Advances in signal Processing*, vol. 2009, no. 1, 2009.

[2] G. Boato, V. Conotter, F. G. De Natale, and C. Fontanari, "Watermarking robustness evaluation based on perceptual quality via genetic algorithms," *IEEE transactions on information forensics and security*, vol. 4, no. 2, pp. 207–216, 2009.

[3] M. Cancellaro, F. Battisti, M. Carli, G. Boato, F. G. De Natale, and A. Neri, "A commutative digital image watermarking and encryption method in the tree structured haar transform domain," *Signal Processing: Image Communication*, vol. 26, no. 1, pp. 1–12, 2011.

[4] V. Conotter, G. Boato, M. Carli, and K. Egiazarian, "High capacity reversible data hiding based on histogram shifting and non-local means," in *Local and Non-Local Approximation in Image Processing, 2009. LNLA 2009. International Workshop on*, pp. 130–136, IEEE, 2009.

[5] S. Wang, S. Huang, X. Zhang, and W. Wu, "Hologram-based watermarking capable of surviving print-scan process," *Applied optics*, vol. 49, no. 7, pp. 1170–1178, 2010.

[6] S. H. Amiri and M. Jamzad, "Robust watermarking against print and scan attack through efficient modeling algorithm," *Signal Processing: Image Communication*, vol. 29, no. 10, pp. 1181–1196, 2014.

[7] F. Wang and J. P. Allebach, "Printed image watermarking using direct binary search halftoning," in *Image Processing (ICIP), 2016 IEEE International Conference on*, pp. 2727–2731, IEEE, 2016.

[8] A. M. Alattar, "Smart images' using digimarc's watermarking technology," in *Security and Watermarking of multimedia contents*, pp. 264–273, 2000.

[9] G. Turin, "An introduction to matched filters," *IRE transactions on Information theory*, vol. 6, no. 3, pp. 311–329, 1960.

[10] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of psnr in image/video quality assessment," *Electronics letters*, vol. 44, no. 13, pp. 800–801, 2008.

[11] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[12] H. Venkateswara, J. Eusebio, S. Chakraborty, and S. Panchanathan, "Deep hashing network for unsupervised domain adaptation," in *(IEEE) Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

Author Biography

Andrea Leopardi received his B.S. degree in Electronics and Telecommunications Engineering at the University of Trento in 2016. He is currently enrolled in the EIT Digital M.S. double degrees programme in Information and Communications Engineering at the University of Trento and KTH Royal Institute of Technology in Stockholm. His research interests are: media processing and communications, recognition systems and machine learning.

Davide Soresina received his B.S. degree in Electronics and Telecommunications Engineering at the Univeristy of Trento in 2016. He is currently enrolled in M.S. degree in Information and Communications Engineering at the University of Trento. His research interests are: image and signal processing, remote sensing and machine learning.

Davide Marcantonio received his B.S. degree in Electronics and Telecommunications Engineering at the University of Trento in 2016. He is currently enrolled in M.S. degree in Information and Communications Engineering at the University of Trento. His research interests are: signal processing, recognition systems and electromagnetics.

Alain Malacarne received his B.S degree in Electronics and Telecommunications Engineering in 2013 and his M.S. degree in Telecommunications Engineering in 2016 at the University of Trento. He has been involved with the Multimedia Signal Processing and Understanding Lab from 2016 to 2018. His main research interest include image processing with particular focus on digital forensics.

Nicola Conci received the B.S. and M. S. degrees in telecommunication engineering from the University of Trento, Trento, Italy, in 2002 and 2004 respectively, and the Ph.D degree from the Uni- versity of Trento in 2007. He was a Visiting Student with the Image Pro- cessing Laboratory, University of California at Santa Barbara CA, USA, and a Postdoctoral Researcher with the Multimedia and Vision Research Group, Queen Mary University of London, U.K. He has authored and coauthored more than 90 scientific papers in international journals and conferences. He is the cofounder of the startup Xtensa. His main research interests of image and video processing for human behavior analysis with particular application to environmental monitoring, surveillance, and assisted living.

Giulia Boato is Assistant Professor at the Department of Informa-tion Engineering and Computer Science of the University of Trento. From 2008 to 2011 she has been coordinating the Multimedia Signal Processing and Understanding Lab. Since 2018 she is associated editor for the SPIE Journal of Electronic Imaging and area editor for the EURASIP journal Signal Processing: Image Communication. Her research interests are focused on image and signal processing, with particular attention to multimedia data protection, data hiding and digital forensics, but also intelligent multidimensional data management and analysis. She is author of 85 papers in international conferences and journals. She is member of the IEEE Multimedia Signal Processing Technical Committee and in the EURASIP Special Area Team on Biometrics, Data Forensics, and Security.