EURASIP Journal on Applied Signal Processing 2002:2, 197–208

Robust Watermarking of Cartographic Images

Mauro Barni
Department of Information Engineering, University of Siena, Via Roma 56, 53100-Siena, Italy
Email: barni@dii.unisi.it

Franco Bartolini
Department of Electronics and Telecommunications, University of Florence, Via S. Marta 3, 50139-Firenze, Italy
Email: barto@lci.det.unifi.it

Alessandro Piva
Department of Electronics and Telecommunications, University of Florence, Via S. Marta 3, 50139-Firenze, Italy
Email: piva@lci.det.unifi.it

Filippo Salucco
Department of Electronics and Telecommunications, University of Florence, Via S. Marta 3, 50139-Firenze, Italy
Email: salucco@lci.det.unifi.it

Received 30 April 2001 and in revised form 15 October 2001

We present a method (Text-Based Geometric Normalization—TBGN) which, by exploiting the particular content of cartographic images, namely text content, permits to cope with global geometric transformations. First, text is extracted from the to-be-marked map, then text orientation and size are exploited to normalize the image geometry prior to watermark insertion. Watermarking is performed by means of any of the existing algorithms ensuring good robustness against image processing tools. At the decoder side, text is extracted again from the map and used to normalize image geometry. Owing to the robustness of text features with respect to common image manipulations, and to the likely spreading of text all across the digital map, the proposed system exhibits an excellent robustness.

Keywords and phrases: digital watermarking, map watermarking, robust watermarking, geometric normalization.

1. INTRODUCTION

Image watermarking has been widely studied in the last years for its importance in copyright protection applications involving the exchange of visual data in digital format. At the beginning, the exchange of general purpose data was considered without taking into account the peculiarities of the image data to be protected. Such an approach has lead to the development of a wide variety of general image watermarking algorithms which can be used virtually in any application environment [1]. As research has gone on, the necessity of applying digital watermarking in more restricted scenarios has come out, thus calling for the development of ad hoc watermarking solutions capable of coping with and exploiting the peculiarities of the application at hand. Possible examples include the watermarking of video surveillance data [2, 3], medical images [4], cartoons [5], and remote sensing imagery [6].

From a general point of view, the necessity of dealing with a class of images having peculiar characteristics has two main consequences. From one side, the invisibility requirement must be carefully revised, since it may assume a completely different meaning with respect to conventional multimedia applications. From the other side, image peculiarities may be used to develop ad hoc solutions to problems which are hard to solve when looked at from a general perspective. This is the case, for example, of video surveillance data authentication, where the invisibility, or, to better say, the unobtrusiveness requirement is less stringent than in multimedia applications, thus making it possible to improve watermark robustness [2]. At the opposite extreme, in biomedical applications, watermark unobtrusiveness is a very demanding constraint, thus making the design of a robust watermarking scheme even more difficult than usual.

This paper deals with robust watermarking of cartographic images such as geographical or road maps (Figure 1).

More specifically, the paper focuses on two main issues: watermark invisibility and robustness against global geometric manipulations. With regard to watermark invisibility, the

(a) (b)

Figure 1: Examples of digital maps: (a) geographical map (EU600) and (b) road map (ITA600). The peculiarities of this kind of images, such as the contemporary presence of crisp contours, flat areas, and text, are clearly evident.

particular nature of the to-be-marked images makes it is natural asking whether conventional hiding methods are still valid in this case. To answer such a question, we tested two different watermarking strategies developed for general purpose applications: one operating in the full-frame Discrete Fourier Transform (DFT) domain [7, 8], and one in the wavelet domain [9, 10]. The results we obtained demonstrate that, as far as images obtained by digitizing a hard copy of a cartographic map are concerned, no significant differences exist between digital maps and general purpose images (different conclusions may be drawn when synthetic, computer-generated, maps are considered). As a matter of fact, due to the richness of details contained in this kind of images, the watermark strength ensuring invisibility is equal to, or even higher than, that used for general purpose imagery.

As to robustness against geometric manipulations, we present a novel method which, by exploiting the particular content of cartographic images, permits to cope with global geometric transformations.

The proposed method relies on Feature-Based Geometric Normalization (FBGN) to correct geometric distortions applied to the image after watermark insertion [11, 12, 13]. According to such an approach, a set of image features is used to geometrically normalize the image prior to watermark insertion. Normalization is performed at the decoder side as well, in such a way to correct possible geometric distortions. In FBGN techniques, robustness ultimately relies on the stability of features used to normalize the image. As a matter of fact, it is very difficult to find a set of features which is robust against the wide variety of manipulations images may undergo. Possible solutions include the use of edges, corners, or image regions; however, the robustness of such image features is still an open issue which needs to be carefully analyzed. FBGN algorithms tend to be very sensitive to image cropping as well, since when cropping occurs some of the reference features are likely to be lost, thus compromising the effectiveness of geometric normalization.

The method we propose here permits to circumvent the above problems, since it relies on a very stable set of features: the text contained in cartographic images. First the text is extracted from the to-be-marked image, then it is used to rotate and resize the image so that its orientation and scale assume a given reference value. The watermark is inserted at this normalized orientation and scale, then the image is transformed back to its original format. The same sequence of operations is performed to detect the watermark. As it is witnessed by experimental results, the method is very robust, such a robustness deriving from the stability of text features.

With respect to existing algorithms adopting FBGN (e.g., the one described in [11]), our technique produces much more stable results, since text features are inherently more stable than features such as edges or image regions (one of the first *lessons* learned by any image processing researcher or practitioner is the extreme instability of edge extraction and image segmentation algorithms).

The rest of this work is organized as follows. In Section 2, we show the results we obtained with regard to the invisibility requirement. In Section 3, a general description of the new TBGN strategy is described. The watermarking algorithm is detailed in Section 4, whereas experimental results are given in Section 5. The work ends with some conclusions, drawn in Section 6.

2. MEETING THE INVISIBILITY REQUIREMENT

To investigate whether the particular nature of cartographic images has a significant impact on the invisibility constraint, we marked a set of images with two different watermarking algorithms, one operating in the DFT domain [7, 8] and one in the Discrete Wavelet Transform (DWT) domain [9, 10]. Both the algorithms insert the watermark in the luminance component of the host image.

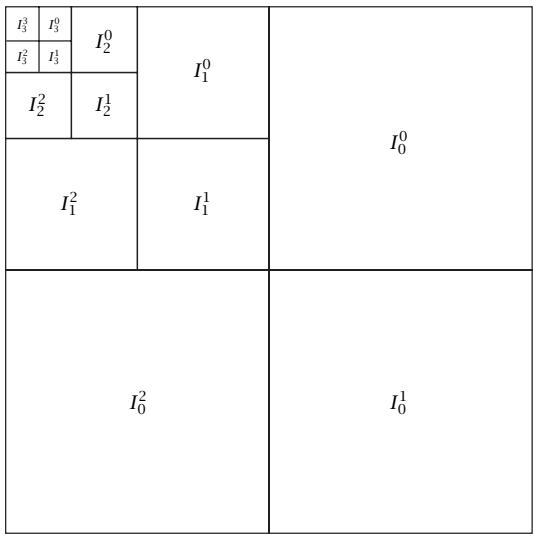


Figure 2: Wavelet decomposition of image I.

The DFT-domain algorithm operates by inserting the watermark in the medium portion of the frequency spectrum according to the following additive/multiplicative embedding rule:

$$y_i = x_i + \gamma m_i x_i, \tag{1}$$

where x_i indicates the magnitude of the ith DFT coefficient, y_i the magnitude of the marked coefficient, m_i is the ith coefficient of the watermark sequence, and γ is a parameter controlling the watermark insertion strength. The watermark sequence $\{m_i\}_{i=1,19200}$ consists of a set of independent random variables uniformly distributed in $[-1, 1]$. The invisibility of the watermark is improved by modulating the watermark in the spatial domain according to a masking image M, indicating, pixel by pixel, the sensibility of the human visual system to a disturb affecting that particular image area [14]. More specifically, a simple variance-based mask is used here as described in [15].

The DWT algorithm embeds the watermark, which consists of a binary (± 1) pseudorandom sequence, by modifying the wavelet coefficients belonging to the three detail bands at level 0 (i.e., I_0^0, I_0^1, and I_0^2 in Figure 2). Even in this case, watermark invisibility is ensured by modulating the watermark strength according to the local image content, thus leading to the following insertion rule:

$$\tilde{I}_0^\theta(i,j) = I_0^\theta(i,j) + \alpha w^\theta(i,j) x^\theta(i,j), \tag{2}$$

where α is a global parameter accounting for watermark strength, and $w^\theta(i,j)$ is a band-dependent weighting function considering the local sensitivity of the image to noise. We also investigated the visibility of the watermark when two DWT levels are marked (levels 0 and 1 in Figure 2).

In both cases (DFT and DWT watermarking), we did not adapt to the case of cartographic images the way the masking image M and the weighting function $w^\theta(i,j)$ are built, capitalizing on the fact that the main phenomena at the basis of human vision are the same in multimedia and cartographic applications.

Table 1: Maximum watermark strength, $\gamma_{\max}$, subject to watermark invisibility and corresponding PSNR (dB) obtained for 4 test images with the DFT-domain watermarking algorithm described in [7, 8].

	Lena	*Mandrill*	EU600	ITA600
$\gamma_{\max}$	0.12	0.27	0.15	0.2
PSNR	51.9	40.6	42.38	39.15

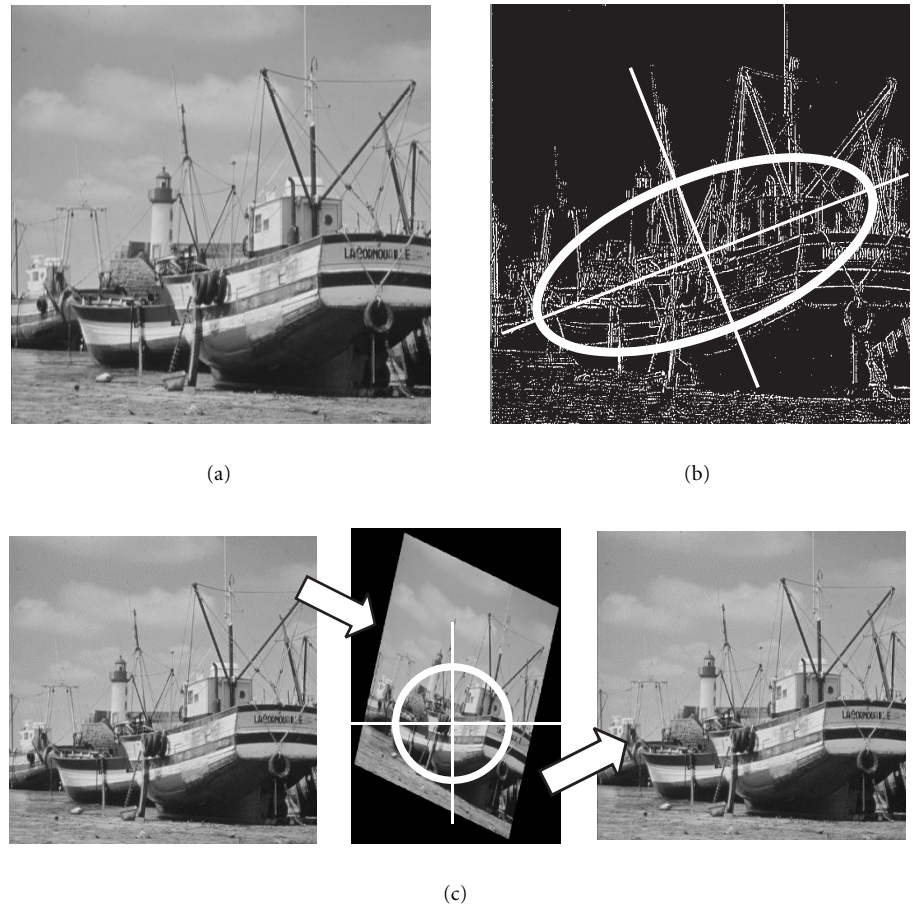
Table 2: Maximum watermark strength, $\alpha_{\max}$, subject to watermark invisibility and corresponding PSNR (dB) obtained for 4 test with the DWT-domain watermarking algorithm described in [8, 9]. Only the highest DWT subband is watermarked (I_0^θ, $\theta = 0, 1, 2$).

	Lena	*Mandrill*	EU600	ITA600
$\alpha_{\max}$	0.26	0.30	0.24	0.20
PSNR	34.7	29.5	31.0	30.0

Table 1 reports the results we obtained by using the DFT-domain algorithm. Though we used a large set of test images, the results in the table refer to the watermarking of two standard images, namely the *Lena* and *Mandrill* images, and the two cartographic images shown in Figure 1. The table was built by watermarking the images with different strengths γ. For each image, we report the maximum admissible γ, and the corresponding Peak Signal-to-Noise Ratio (PSNR). Upon inspection of the results it comes out that the behavior of cartographic images is much like that of general-purpose images. The maximum watermark energy is higher than that admissible in very flat images, like the *Lena* image, but lower than that of complex images like *Mandrill*. Note that a higher γ does not always correspond to a lower PSNR, due to the additive/multiplicative nature of the watermark.

A similar table was built for the DWT-based algorithm, see Table 2. Once again we inserted the watermark by using different strengths α. The table refers to the cases in which the watermark is inserted in the highest DWT level (I_0^θ, Figure 2). As for the DFT case, map images do not differ very much from standard images. The lower PSNR values obtained in the DWT case can be explained by observing that a much more efficient masking algorithm was used in this case, with respect to the variance-based mask adopted for the DFT algorithm. We also carried out some experiments in which we marked the two highest DWT levels (I_0^θ and I_1^θ): in such a case the watermark visibility increases significantly, thus making the watermarking of two DWT bands unadvisable.

The above results permit to conclude that no particular differences exist between cartographic and general purpose imagery. Such a result somewhat depends on the fact that the test images we used were scanned from the corresponding hard-copy maps. It can be argued that somewhat different results would have been obtained by dealing with synthetic, computer-generated, images.



(a) (b)

(c)

Figure 3: Example of feature-based geometric normalization. Image edges are first extracted (b). Then the image is rotated and scaled in such a way that the inertial axis of image edges assume a reference orientation and scale (c). After watermark insertion the image is brought back to its original geometry.

3. TEXT-BASED GEOMETRIC NORMALIZATION

In this section, we describe the various steps of the new text-based geometric normalization (TBGN) algorithm. Before delving into the details of the TBGN algorithm, we recall the main ideas behind feature-based geometric normalization.

3.1. The FBGN approach

Feature-based geometric normalization for improved robustness against geometric distortions relies on a very simple idea: always insert and detect the watermark when the image assumes a reference geometric configuration (by geometric configuration we mean here orientation and scale factor). To be meaningful, reference image geometry must be given with respect to a coordinate system which is known both to the encoder and the decoder. To achieve this, FBGN techniques define the reference geometric configuration with respect to a set of image features, for example, edges or corners. Hopefully, such reference features are chosen so that they are stable with respect to all the image manipulations the watermark must survive.

To be more specific, we consider an example in which geometric normalization is performed by relying on image edges (Figure 3). As a first step, image edges are extracted, then geometric normalization is performed by calculating the central inertial axis of edge pixels and the corresponding inertial moments. Before inserting the watermark, the image is rotated and scaled so that the central inertial axis and the corresponding moments assume given reference values. After watermark insertion the image is transformed back to its original format. The same operations are performed before attempting to detect the watermark, so that geometric transformations possibly applied after watermark insertion are automatically corrected.

Note that in this simple example, FBGN effectiveness mainly depends on edge stability, thus laying itself open to attacks which modify the edge content of the image, as in the case of image cropping. Actually, the lack of robustness against image cropping is one of the main weaknesses of FBGN techniques. Possible solutions consist in repeatedly applying geometric normalization to a set of image subparts, or choosing the reference features in such a way to minimize

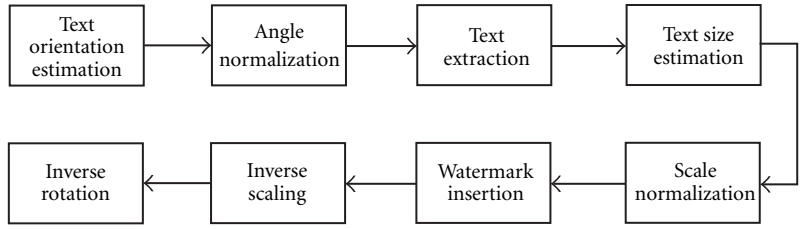


Figure 4: Overall scheme of TBGN-based watermark encoding.

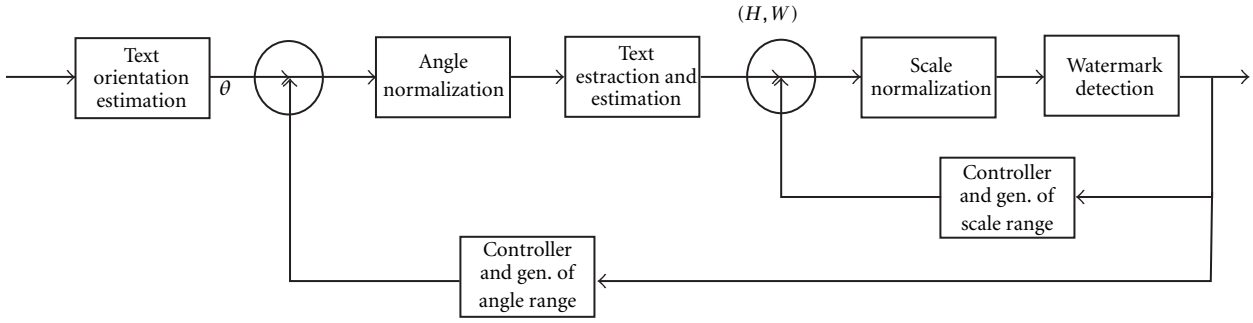


Figure 5: Overall scheme of TBGN-based watermark detection.

the impact of cropping on the establishment of a geometrical reference.

3.2. Overall system layout (TBGN)

The overall scheme of the watermark encoding/decoding system is reported in Figures 4 and 5. As a first step the orientation of the text contained in the map is estimated. Of course, we assume that a dominant text direction exists, which is the case in most cartographic images. Then the image is rotated so that text assumes a given reference orientation. At this point, a text recognition module is run to extract as much text as possible. This module also splits the text into single characters. The next module estimates text size by paying particular attention to group characters of the same size together. Then the image is resized so that the text assumes a reference dimension. Finally, the watermark is inserted within the geometrically normalized image. Of course, prior to distribution or storage, the image is brought back to its original orientation and size. Particular care must be taken here not to introduce visible degradation as a consequence of resampling. This can be achieved by choosing a suitable image resampling algorithm. It should be remembered, in fact, that no particular time constraint applies at the encoder, thus leaving room for the application of even the more sophisticated rotation/scaling techniques.

At the detection side, the same sequence of operations is performed, so that the watermark is always searched for in the same geometric configuration (Figure 5). To account for possible, minor, inaccuracies of geometric normalization, an exhaustive search is performed in a neighborhood of the recovered geometrical configuration. The exhaustive search is performed through two cycles: the first cycle considers a small range of angles in the neighborhood of the text orientation estimate, the latter takes into account several resizing factors to account for possible inaccuracy of text size estimation. For each step of the external cycle (different angle estimate), size normalization is performed from scratch to avoid that image misalignment corrupts size estimate.

It is worth noting that TBGN can be used in conjunction with virtually any watermarking schemes. The only constraint deriving from the algorithm used to insert and retrieve the watermark concerns the accuracy with which the reference geometric configuration must be reproduced to allow a correct detection. The knowledge of the precision required by the watermarking algorithm to detect the watermark is exploited to define the quantization steps used during the local exhaustive search.

The three most important/critical steps of the watermark encoding chain are now described in more details.

3.3. Estimation of text orientation and angle normalization

The problem of estimating text orientation in a digital map is similar to one of skew angle estimation of text documents, for which many solutions have been proposed in the relevant literature [16]. In order to demonstrate the feasibility of the TBGN approach, we implemented a simple skew estimation algorithm derived from the work described in [17]. The estimate is performed in the DFT-magnitude domain. Ideally, the sum of DFT coefficients magnitude is evaluated for all possible straight lines passing from the frequency spectrum origin and having different orientation angles θ. The sum is

normalized by dividing it by the length of the straight line. The orientation which maximizes the sum is taken as the direction orthogonal to text orientation.

To take into account the discrete nature of the DFT spectrum, the orientation angles are quantized and the line length is replaced by the number of DFT coefficients voting for each quantized angle. Of course, due to DFT symmetry, only angles ranging from $-\pi/2$ through $+\pi/2$ are taken into account. In addition, DFT coefficients belonging to a small neighborhood of the origin are not taken into account, to avoid that they bias the sum of DFT magnitudes.

In order to prevent the introduction of false horizontal and vertical frequencies, we calculate the DFT spectrum of the original image multiplied by a separable, raised-cosine, window. That is, given the original $N \times M$ image $I(i, j)$, angle estimation is performed on I' given by

$$I'(i,j) = \frac{1}{4} I(i,j) \left[1 + \cos\left(\frac{2\pi(i - N/2)}{N}\right)\right] \times \left[1 + \cos\left(\frac{2\pi(j - M/2)}{M}\right)\right]. \tag{3}$$

While the above skew angle estimation algorithm provides good results in the classical case of document images analysis, when cartographic images are dealt with, some problems may occur. Problems are mainly due to the sparse nature of text and to the presence of many other cartographic signs which contribute to augment the noisiness of the DFT spectrum. To enhance the robustness of the angle estimation module, we developed a simple modification of the algorithm described in [17]. Let $\lambda(\theta_i)$ be the normalized sum of DFT coefficients for a generic quantized angle θ_i (solid line in Figure 6b). Instead of searching for the maximum of $\lambda(\theta_i)$ over all possible θ_i, we look for the maximum of the following peakedness measure of $\lambda(\theta_i)$:

$$p(\theta_i) = \sum_{k=-l}^{+l} [\lambda(\theta_i) - \lambda(\theta_{i+k})]^2, \tag{4}$$

where l defines the width of the window taken into account to measure the peakedness of $\lambda(\theta_i)$ (during our experiments we set $l = 3$).

As an example of the improvement obtained by considering $p(\theta_i)$ instead of $\lambda(\theta_i)$, we consider the situations depicted in Figures 6 and 7. In both figures, $p(\theta_i)$ is represented by the dashed line and $\lambda(\theta_i)$ by the solid line. In Figure 6 the maximum of $\lambda(\theta_i)$ corresponds to true text orientation, however, this is not the case in Figure 7. On the contrary, the maximum of $p(\theta_i)$ permits to estimate text orientation correctly both in Figures 6 and 7.

One may wonder whether angle normalization still works for cartographic images distorted using a spherical projection. In such a case, in fact, maps are likely to contain letters with different orientation, for instance, following the longitudinal or latitudinal lines rather than one fixed direction. Of course, if the range of text directions is too wide and no preferred direction exists, angle normalization is likely to fail. Nevertheless, given that the direction spread is not too accentuated, angle normalization still performs satisfactorily. This is the case of maps that, though distorted due to spherical projection, represent a limited geographical area. An example of such an image is given in Figure 8. In Figures 9a and 9b, the corresponding plot of $\lambda(\theta_i)$ and $p(\theta_i)$ are given, respectively. As it can be seen, $\lambda(\theta_i)$ alone does not provide the right angle estimation (-53 degrees), since the presence of some vertical lines confuses the angle estimation module. However, the normalization angle is correctly recovered by resorting to $p(\theta_i)$. Note also that when strong linear features are present, angle normalization may choose a reference angle which is not based on text, however, this does not prevent the correct retrieval of the watermark, unless the map is cropped in such a way that linear features are removed.

Having estimated text orientation, angle normalization is a straightforward step: we simply rotate the image to make text orientation horizontal.

3.4. Text extraction

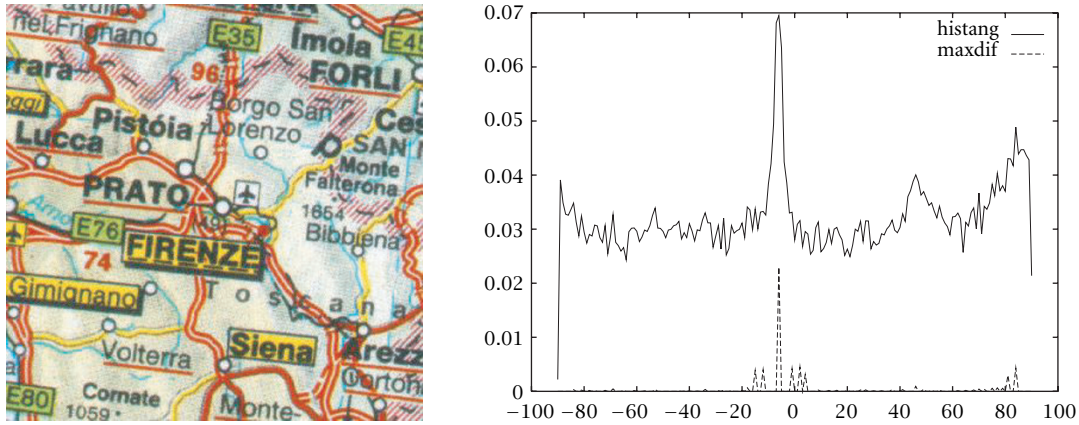
Text extraction from a generic image is quite a difficult problem for which an established and widely accepted solution does not exist [16, 18, 19]. For our system, we developed an ad hoc text extraction algorithm which exploits the peculiarity of our application. More specifically, the missed detection of some of the text is not a problem, since we only aim at estimating the dimension of text, not at reading it. The method we adopted originates from a work by Hasan and Karam [20]. We first segment the image by clustering pixels according to their gray level. More specifically, k-means clustering is applied to the luminance image histogram. As to the number of clusters k, we found that any value ranging from 3 through 5 may be used. Then, we try to identify the color of background pixels by identifying the most numerous cluster.

For sake of readability, the gray level of text is usually very different from that of the background, then we make the hypothesis that the cluster which is furthest from the background cluster contains text pixels. We denote by μ_t the centroid of the text cluster, and by μ_{next} the centroid of the cluster closest to the text cluster. By considering a case in which the text is darker than the background, we classify a pixel as a text pixel if its grey level is lower that $\mu_t + \Delta$, where

$$\Delta = \frac{|\mu_t - \mu_{\text{next}}|}{10}. \tag{5}$$

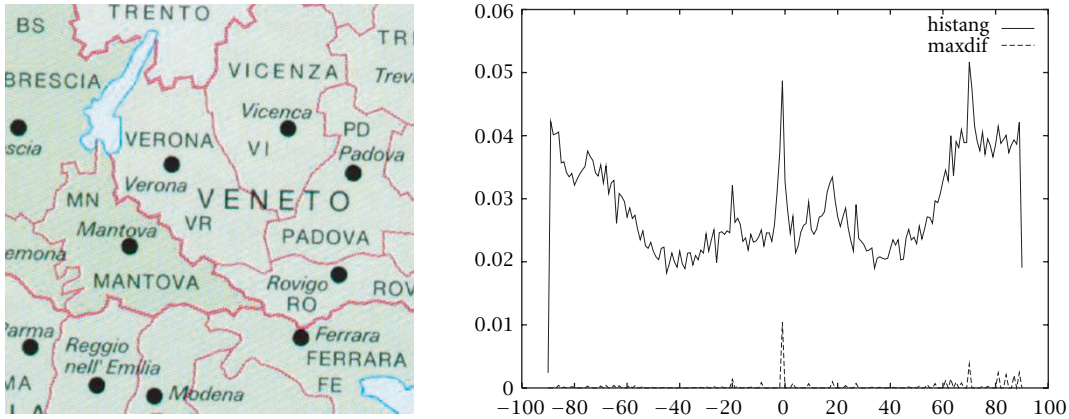
A similar control is adopted when text is brighter than the background.

After thresholding, text pixels are grouped into connected regions, each of which is a candidate for corresponding to a single text character. To decide whether a text-candidate region represents a true character or not, a set of heuristic rules is applied. The rules are very simple and tend to be rather restrictive, since we are more interested in avoiding to extract false text regions than in preventing missed detection events. Decision rules rely on the bounding box of candidate regions, which is easily extracted since text extraction is performed after angle normalization, that is, with text aligned horizontally (we constrain the bounding box to be parallel to text orientation). To be more specific, we indicate by H_i and W_i the height



(a) (b)

FIGURE 6: A rotated map image (a) and the corresponding plot of $\lambda(\theta_i)$, solid line, and $p(\theta_i)$, dashed line. The maximum of both $\lambda(\theta_i)$ and $p(\theta_i)$ corresponds to the true text orientation (-6 degrees).



(a) (b)

FIGURE 7: A rotated map image (a) and the corresponding plot of $\lambda(\theta_i)$, solid line, and $p(\theta_i)$, dashed line. Only the maximum of $p(\theta_i)$ permits to determine the true text orientation (-1 degree).

and width of the bounding box of the ith text-candidate region, respectively, and by M and N the height and width of the image at hand. For a text-candidate region to be classified as a text region we ask that

$$6 \le H_i \le 70, \qquad 6 \le W_i \le 70; \tag{6}$$

$$\frac{H_i}{M} \le 0.12, \qquad \frac{W_i}{N} \le 0.12, \tag{7}$$

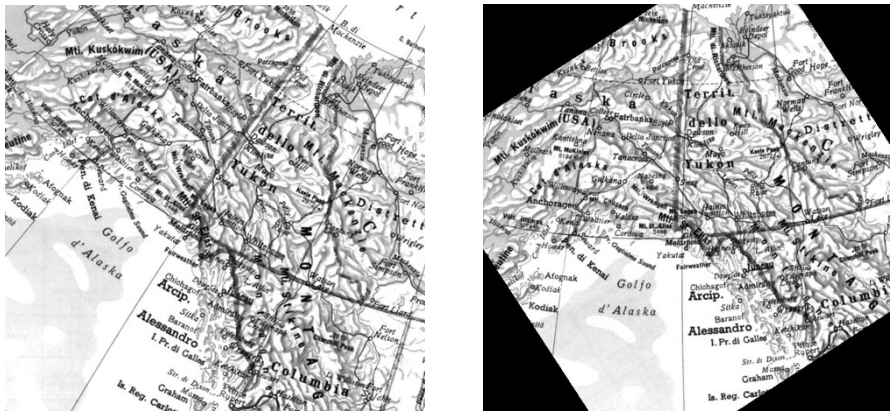
$$0.5 \le \frac{H_i}{W_i} \le 1.5. \tag{8}$$

In other words, we ask that characters are neither too small nor to large (equation (6)), that their dimensions with respect to the image size are not too large (equation (7)), and that they are characterized by a medium aspect ratio (equation (8)), the latter condition being essential to distinguish between characters and other cartographic signs. Of course, the condition expressed by equation (8) excludes from our analysis groups of characters merged in a single connected region. However, this is not a big problem, since the missed detection of some of the characters contained in the map does not prevent a correct estimation of text size. On the contrary, considering both single-character and multiple-character regions would make size estimation much more difficult.

An example of the results produced by the text extraction module is given in Figures 10c and 10g.

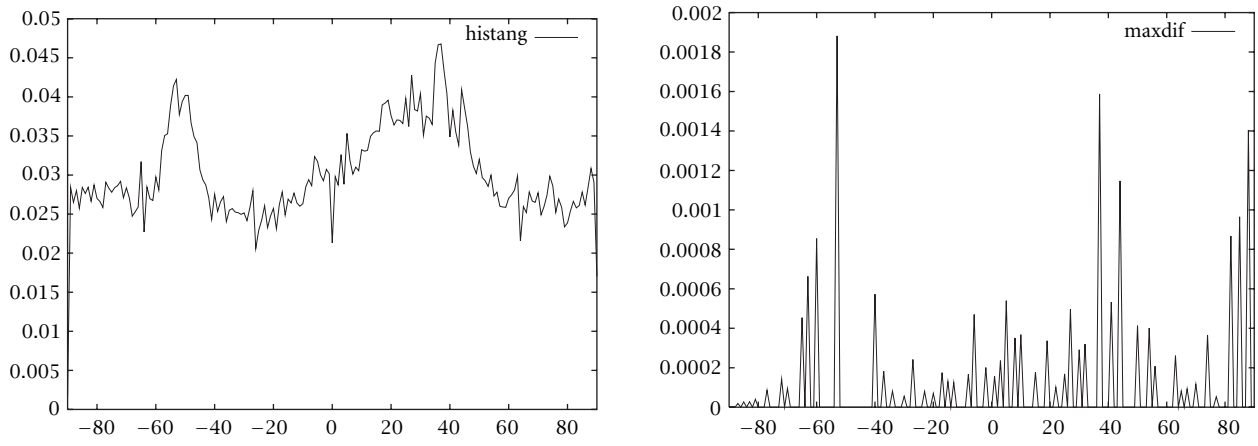
3.5. *Text size estimation and scale normalization*

Text size estimation is the most critical part of the whole TBGN procedure. As a matter of fact, it is likely that the



(a) (b)

FIGURE 8: Angle normalization of a map containing letters with different orientation. (a) Original image and (b) normalized image. The normalized image was clipped and scaled for sake of readability.



(a) (b)

FIGURE 9: (a) Plot of $\lambda(\theta_i)$ and (b) $p(\theta_i)$ for the map in Figure 8. Though less evident that in Figures 6 and 7 the maximum of $p(\theta_i)$ permits to determine the true text orientation (-53 degree).

text extraction module picks out characters of different size, and even some wrong regions that do not correspond to true characters. Text estimation should be as immune as possible to the errors of the text extraction module.

To do so, the histogram of the height of extracted characters is built. Then histogram peaks are looked for, and the one corresponding to the largest characters selected. Of course, peaks must correspond at least to a minimum number of characters. The height $H_{\max}$, corresponding to the peak of the largest characters, is used as a reference size measure. Before inserting the watermark, in fact, the image is scaled so that $H_{\max}$ assumes a reference value H_{ref} (in our experiments we set $H_{\text{ref}} = 10$), a value that will be used by the decoder to re-scale the image so to search the watermark at the same scale used to insert it.

The choice to rely on the largest characters in the image is motivated by the robustness of large characters, which are likely to survive better image manipulations such as filtering or shrinking. Feature stability also motivated the choice of using characters height only, without taking width into account. As a matter of fact, the height of a character is much more stable than its width, since different characters belonging to the same font family may have a large number of different widths, but usually can have only two different heights.

4. THE WATERMARKING ALGORITHM

As noted previously, the TBGN scheme does not depend on the watermarking algorithm actually used to mark the image.

Nevertheless, the maximum rotation angle and scaling factor the watermarking system is immune to, must be known, to verify if they are larger or smaller than TBGN accuracy, and, if necessary, to set the extent and the quantization step used in the exhaustive search of the watermark.

In order to choose one watermarking scheme to carry out the experimental validation of the TBGN algorithm, we considered the two algorithms briefly outlined in Section 2. Then we compared the intrinsic sensibility of such algorithms with respect to small geometric distortions. We found that, from this point of view, the DFT algorithm is slightly superior to the DWT one. More specifically, the maximum rotation angle the DFT algorithm is immune to, is about 0.5 degrees, whereas for the DWT algorithm such an angle is only 0.3 degrees. As to resizing, the DFT algorithm survives a scaling factor up to seven pixels out of 512, whereas for the DWT scheme, the maximum admissible scaling factor is about three pixels out of 512.

Thereby, throughout our experiments we adopted the watermarking algorithm described in [7, 8]. It is worth noting that in [7, 8] robustness to cropping is achieved by always zero-padding the image up to a fixed size (here we set such a size to 512×512). Apart from zero padding, we applied the core algorithm without including the synchronization mechanism used in [7] to cope with geometric attacks, such as rotations and scaling.

5. EXPERIMENTAL RESULTS

We carried out a set of experiments to validate various aspects of the TBGN algorithm. We divided the experiments in two main categories: (i) evaluation of the effectiveness of geometric normalization, and (ii) evaluation of the overall robustness of the algorithm. Note that whereas point (i) takes into account only the TBGN algorithm, the results concerning point (ii) are influenced by the effectiveness of the watermarking algorithm used in conjunction with TBGN.

5.1. Geometric normalization

With reference to angle estimation, we evaluated the performance of the algorithm both in the absence and in the presence of attacks. Given an image, we first rotated it, then we tried to estimate the rotation angle through the algorithm described in Section 3.3. We repeated the test on a set of five images containing geographical, political and road maps. We obtained a mean absolute error of 0.8 degrees.

To evaluate the robustness of the angle estimation algorithm in the presence of attacks we repeated the above experiments by JPEG coding (45% quality factor) the images prior to estimation. The results did not change at all, confirming a mean absolute error of 0.8 degrees. The same results were obtained by coding the images with a quality factor of 25%.

As expected, more critical results were obtained for the scale estimation algorithm. Table 3 reports the results we obtained on three test images. In all cases, the images were resized with a scale factor of 1.5. As it can be seen, for no compression or moderate compression conditions, the error on scale factor is lower than 0.1. Significantly worse results are obtained when 25% quality JPEG coding is applied.

Table 3: Estimation of scale factor for three test images. For each image the initial size, the initial letter size, and the estimated scale factor under different coding conditions are given. The true scale factor is 1.5.

Image	Size	Letter size	No coding	JPEG 45%	JPEG 25%
Europe	300×300	9	1.55	1.55	1.22
England	290×290	12	1.53	1.53	1.58
Italy	300×300	9	1.5	1.5	1.33

5.2. Overall performance

As an example of the overall performance achievable through the TBGN scheme, we consider the map reported in Figure 10a (256×256). Prior to watermark encoding, rotation and scale normalization are performed (in Figure 10b the map after rotation normalization is given). The output of the text extraction module is shown in Figure 10c. As it can be seen, most of the text is recovered along with some wrong regions, that do not affect TBGN significantly. The text in Figure 10c is used to normalize the image scale, then the image is watermarked. The final watermarked image (after going back to the original geometry) is given in Figure 10d.

To test the watermark for robustness, the watermarked image was scaled up to a size of 315×315 (1.23 factor), slightly cropped to 308×308, rotated by 22 degrees and JPEG-coded with a quality factor of 70% (the attacked image is depicted in Figure 10e). To detect the watermark, the skew angle of the attacked map is estimated to re-align the text horizontally (Figure 10f). Then text is extracted, leading to the image in Figure 10g. The text extracted from the attacked image is not exactly the same as that extracted from the original image, nevertheless, it is sufficient to normalize the image scale. Finally, a local exhaustive search is carried out, resulting in the correct retrieval of the watermark (in Figures 10h and 10i the detector response for the true watermark and 1000 fake watermarks is given both in the absence and presence of attacks). To be more specific, the estimated skew angle and scale factor were 22.5 degrees and 1.24, respectively. The exhaustive search was performed for skew angles ranging from 21 through 24 (step 0.5 degrees) and scaling factors in the range [1.2,1.3] with step 0.01, for a total of 70 iterations.

As a further test, we evaluated the robustness of the watermark against a combination of rotation, cropping, scaling, and JPEG coding with decreasing quality factor. The results we obtained are summarized in Figure 11. The image reported in Figure 10a was marked with watermark strength $\gamma = 0.15$. Then we rotated the image by 67 degrees, and we scaled it with a scale factor equal to 1.5. Finally, we JPEG-coded the map image with a quality factor ranging from 100% through 5%. As it is shown in the figure, the detector was able to correctly recover the watermark down to a quality factor of

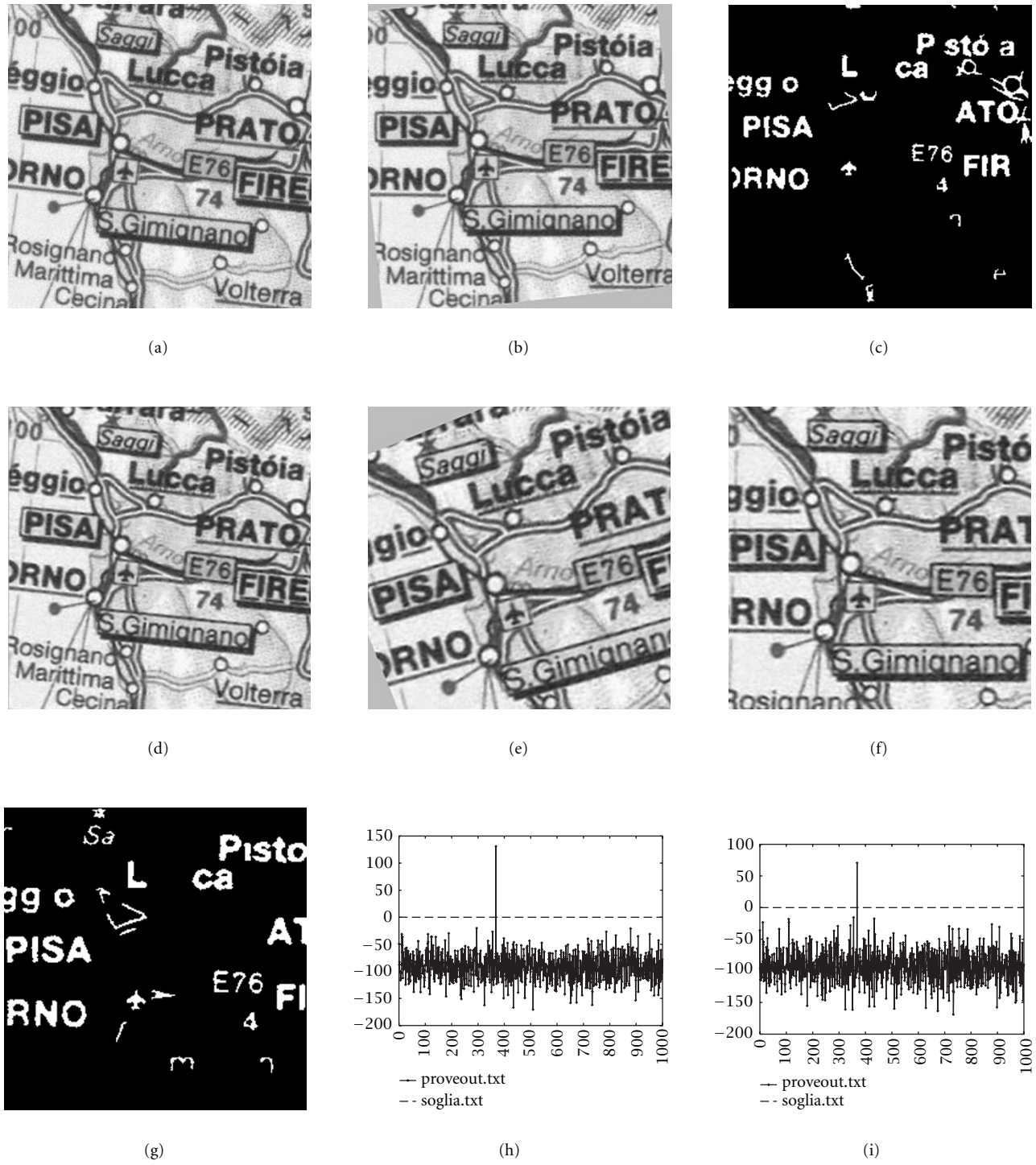


Figure 10: (a) Original map image (256×256); (b) original image after rotation normalization; (c) text extracted from the aligned original image; (d) final watermarked image; (e) particular of the watermarked image after resizing to 315×315, rotation of a 22 degrees angle and 70% quality JPEG coding; (f) attacked image after re-alignment; (g) text extracted from the attacked image; (h) detector response in the absence of attacks and (i) after attacks.

25%. It should be noted, though, that as it is suggested by the data reported in Table 3, the scale estimation error for quality factors as low as 25% is rather high, thus making it necessary a much longer exhaustive search.

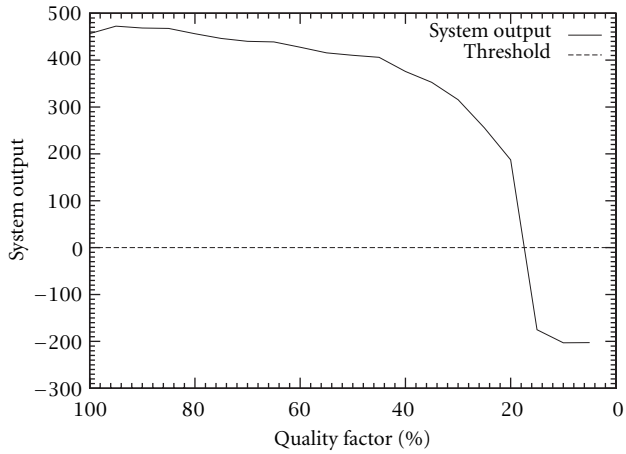


FIGURE 11: Robustness against JPEG coding with decreasing quality factor. The image in Figure 10a was rotated by 67 degrees, scaled with a scale factor equal to 1.5, and JPEG-coded with a decreasing quality factor ranging from 100% through 5%.

6. CONCLUSIONS

By exploiting the peculiarities of cartographic images, we have proposed a TBGN-based technique to achieve watermark robustness against geometrical transformations, as well as against the most common image processing manipulations. TBGN robustness ultimately relies on the robustness of the text contained in the digital map and its spreading all over the map: two assumptions that are satisfied by virtually all cartographic images. Actually, the algorithm does not need that all the text contained in the image is retrieved and identified as such, it only needs that a sufficient number of characters extracted.

The proposed methodology can be used in conjunction with any watermarking algorithm. It only has to be noted that by the comparison of the accuracy of TBGN and the robustness of the watermarking algorithm against small geometric transformations, it may arise the need of a locally exhaustive search of the watermark. As for all FBGN-like algorithms, the worst attack against the proposed TBGN technique is cropping. Nevertheless, for the cropping attack to be successful, the cropped part of the image must be chosen so that it contains virtually no text (or just a few letters), thus allowing only the extraction of areas of minor interest.

Possible improvements of the proposed algorithm can be obtained by adopting more sophisticated angle estimation and text extraction techniques. Another guideline for future research consists in the use of automatic character recognition techniques to distinguish between different characters, thus augmenting the robustness of the size estimation module.

As a minor contribution, we also investigated whether the invisibility constraint assumes a particular meaning in the context of cartographic image watermarking, demonstrating that only minor differences exist with respect to the watermarking of general purpose imagery.

REFERENCES

[1] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1079–1107, 1999.

[2] M. Barni, F. Bartolini, J. Fridrich, and A. Piva, "Digital watermarking for the authentication of AVS data," in *Proc. 10th European Signal Processing Conference*, Tampere, Finland, September 2000.

[3] F. Bartolini, A. Tefas, M. Barni, and I. Pitas, "Image authentication techniques for surveillance applications," *Proceedings of the IEEE*, pp. 1403–1418, October 2001.

[4] G. Coatrieux, B. Sankur, and H. Maitre, "Strict integrity control of biomedical images," in *Proc. SPIE on Security and Watermarking of Multimedia Contents III*, vol. 4314, San Jose, Calif, USA, January 2001.

[5] J. Fridrich and D. Rui, "Secure steganographic methods for palette images," in *Proc. 3rd Information Hiding Workshop*, pp. 47–60, Dresden, Germany, 29 September–1 October 1999.

[6] M. Barni, F. Bartolini, V. Cappellini, A. Garzelli, E. Magli, and G. Olmo, "Watermarking techniques for electronic delivery of remote sensing images," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium*, Sidney, Australia, July 2001.

[7] A. De Rosa, M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "Optimum decoding of non-additive full frame DFT watermarks," in *Proc. 3rd International Workshop on Information Hiding*, vol. 1768, pp. 160–172, Springer, Dresden, Germany, 29 September–1 October 1999.

[8] M. Barni, F. Bartolini, A. De Rosa, and A. Piva, "A new decoder for the optimum recovery of nonadditive watermarks," *IEEE Trans. Image Processing*, vol. 10, no. 5, pp. 755–766, 2001.

[9] M. Barni, F. Bartolini, V. Cappellini, A. Lippi, and A. Piva, "DWT-based technique for spatio-frequency masking of digital signatures," in *Proc. SPIE on Security and Watermarking of Multimedia Contents*, vol. 3657, pp. 31–39, San Jose, Calif, USA, January 1999.

[10] M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Trans. Image Processing*, vol. 10, no. 5, pp. 783–791, 2001.

[11] M. Alghoniemy and A. H. Tewfik, "Geometric distortion correction in image watermarking," in *Proc. SPIE on Security and Watermarking of Multimedia Contents II*, vol. 3971, pp. 82–89, San Jose, Calif, USA, January 2000.

[12] A. Nikolaidis and I. Pitas, "A region-based technique for chaotic image watermarking," in *Proc. 10th European Signal Processing Conference (EUSIPCO 2000)*, Tampere, Finland, September 2000.

[13] P. Bas, J.-M. Chassery, and B. Macq, "Robust watermarking based on the warping of pre-defined triangular patterns," in *Proc. SPIE on Security and Watermarking of Multimedia Contents II*, vol. 3971, pp. 99–110, San Jose, Calif, USA, January 2000.

[14] F. Bartolini, M. Barni, V. Cappellini, and A. Piva, "Mask building for perceptually hiding frequency embedded watermarks," in *Proc. IEEE International Conference on Image Processing (ICIP'98)*, vol. I, pp. 450–454, Chicago, Ill, October 1998.

[15] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Processing*, vol. 66, no. 3, pp. 357–372, 1998.

[16] G. Nagy, "Twenty years of document image analysis in PAMI," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 38–62, 2000.

[17] G. S. Peake and T. N. Tan, "A general algorithm for document skew angle estimation," in *Proc. IEEE International Conference on Image Processing (ICIP '97)*, vol. II, pp. 230–233, Santa Barbara, Calif, USA, October 1997.

[18] R. G. Casey and E. Lecolinet, "A survey of methods and strategies in character segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 7, pp. 690–706, 1996.
[19] A. K. Jain and B. Yu, "Document representation and its application to page decomposition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 294–308, 1998.
[20] Y. M. Y. Hasan and L. J. Karam, "Morphological text extraction from images," *IEEE Trans. Signal Processing*, vol. 9, no. 11, pp. 1978–1983, 2000.

Mauro Barni was born in Prato, Italy, in 1965. He received his B.S. in electronic engineering at the University of Florence in 1991. He received the Ph.D. in informatics and telecommunications in October 1995. From 1991 through 1998 he was with the Department of Electronic Engineering, University of Florence, Italy, where he worked as a postdoctoral researcher. Since September 1998, he has been with the Department of Information Engineering at the University of Siena, Italy, where he works as Assistant Professor. His main interests are in the field of digital image processing and computer vision. His research activity is focused on the application of image processing techniques to copyright protection and authentication of multimedia data (digital watermarking), and to the transmission of image and video signals in error-prone, wireless, environments. He is author/co-author of more than 100 papers published in international journals and conference proceedings. Mauro Barni is member of the IEEE, where he serves as member of the Multimedia Signal Processing Technical Committee (MMSP-TC).

Franco Bartolini was born in Rome, Italy, in 1965. In 1991, he received his B.S. (cum laude) in electronic engineering from the University of Florence, Florence, Italy. In November 1996, he received the Ph.D. degree in informatics and telecommunications from the University of Florence. Since November 2001 he is Assistant Professor at the University of Florence. His research interests include digital image sequence processing, still and moving image compression, nonlinear filtering techniques, image protection and authentication (watermarking), image processing applications for the cultural heritage field; signal compression by neural networks, and secure communication protocols. He has published more than 100 papers on these topics in international journals and conferences. He holds two Italian patents in the field of digital watermarking. Dr. Bartolini is a member of IEEE and IAPR.

Alessandro Piva was born in Florence, Italy, in 1968. In 1995, he received his B.S. (cum laude) in electronic engineering from the University of Florence, Florence, Italy. In February 1999, he received the Ph.D. degree in informatics and telecommunications from the University of Florence. He is now a Postdoctoral researcher with the University of Florence. His research activity is focused on multimedia systems, digital image sequence processing, image protection, and authentication (watermarking), image processing techniques for cultural heritage applications and secure communication protocols. He has published more than 40 papers on these topics in international journals and conferences. He holds two Italian patents in the field of digital watermarking.

Filippo Salucco was born in Florence, Italy in 1968. In 1999 he started a collaboration as teacher on "Operative Systems" and "Networking computers" for a private institute. In 2001, he received his B.S. in electronic engineering from the University of Florence, Florence, Italy. He is now working as system engineer in a software house. His activity is focused on cryptographical methods for protection of legal electronic documentation, electronic data interchange (EDI), and secure communication protocols.