A method for the estimation and recovering from general affine transforms in digital watermarking applications

Frédéric Deguillaume, Sviatoslav Voloshynovskiy, and Thierry Pun

University of Geneva – CUI, 24 rue du Général Dufour, CH 1211, Geneva 4, Switzerland

ABSTRACT

An important problem constraining the practical exploitation of robust watermarking technologies is the low robustness of the existing algorithms against geometrical distortions such as rotation, scaling, cropping, translation, change of aspect ratio and shearing. All these attacks can be uniquely described by general affine transforms. In this work, we propose a robust estimation method using apriori known regularity of a set of points. These points can be typically local maxima, or peaks, resulting either from the autocorrelation function (ACF) or from the magnitude spectrum (MS) generated by periodic patterns, which result in regularly aligned and equally spaced points. This structure is kept under any affine transform. The estimation of affine transform parameters is formulated as a robust penalized Maximum Likelihood (ML) problem. We propose an efficient approximation of this problem based on Hough transform (HT) or Radon transform (RT), which are known to be very robust in detecting alignments, even when noise is introduced by misalignments of points, missing points, or extra points. The high efficiency of the method is demonstrated even when severe degradations have occurred, including JPEG compression with a quality factor of 50%, where other known algorithms fail. Results with the Stirmark benchmark confirm the high robustness of the proposed method.

Keywords: digital watermarking, auto-correlation, magnitude spectrum, affine transform, penalized Maximum Likelihood, Hough transform, Radon transform.

1. INTRODUCTION

Digital watermarking emerged as an efficient tool for document copyright protection, authentication and tamper proofing [6]. An important problem constraining the practical exploitation of watermarking technology is the low robustness of existing watermarking algorithms against global geometrical distortions such as translation, cropping, rotation, scaling, cropping, translation, change of aspect ratio and shearing, projective transforms, and any other kind of geometrical transformation applied on the image. Such distortions, known as *geometrical attacks*, desynchronizing the watermark detection and decoding.

Most of geometrical attacks can be uniquely described using the paradigm of general affine transforms, that can be represented by the 4 coefficients a, b, c, d forming a matrix A for the linear component, plus the two coefficients t_x, t_y for the translation part $\vec{t}$:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; \quad \vec{t} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} \tag{1}$$

Therefore, the affine transform maps each point of cartesian coordinates from (x, y) to (x', y'), according to the expression:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \vec{t} \tag{2}$$

where "." is the matrix product. The $\vec{t}$ component corresponds to the cropping and the translation which are easily estimated usually based on a cross-correlation [3]. Consequently, in the following we will only consider A which represents any rotation, shearing, scaling, change of aspect ratio or any combination of them. Furthermore, a succession of n arbitrary linear transforms A_i, $i = 1...n$ yields another linear transform, which can be expressed as

Correspondence (F. Deguillaume) : Email : frederic.deguillaume@cui.unige.ch; WWW : http://cui.unige.ch/~deguilla

$A = A_n \cdot A_{n-1} \cdot ... \cdot A_1$. Moreover, the Stirmark attack [9] or random bending attack (RBA) can be also expressed in term of local affine transforms [10].

The state-of-art methods capable to estimate and recover the undergone global affine transformations can be divided into several groups, depending on the reference structure used, and on the method applied to estimate the parameters of the affine transform. We will concentrate our analysis mostly on still image watermarking applications of the proposed approach. Depending on the reference features used the existing methods may be divided into 3 main groups: methods using a fully transform invariant domain [1], methods based on an additional template [2], and methods exploiting the self-reference principle based on an auto-correlation function (ACF) [3] or the Fourier magnitude spectrum (MS) of a periodical watermark [5]. The transform invariant domain approach mostly consists in the application of the Fourier-Mellin transform to the magnitude of the original (or cover) image spectrum, associated with a log-polar or a log-log coordinate mapping. However, the resulting stego image quality is poor due to interpolation errors. A more recent proposal [4] aims at overcoming the above problem using the general affine transform paradigm, based on an extra synchronization template. However, the necessity to spend part of the limited available energy for this template, as well as the relative ease with which attackers can remove template peaks, makes this approach not so robust as expected.

The algorithms for the estimation of affine transforms can be divided into 2 categories: first, algorithms based on log-polar and log-log mapping [1,2], and secondly, algorithms performing some constrained exhaustive search aiming at the best fitting of the reference pattern with the analyzed one [3,4]. These approaches have several drawbacks from both robustness, uniqueness and computational complexity points of view. Log-polar or log-log mapping are unable to estimate a rotation and a change of proportions simultaneously. Exhaustive search of templates is computationally expensive, and the accuracy of template points is strongly sensitive to any distortion such as lossy compression. These drawbacks caused the necessity to use therefore self-reference methods which uses the same affine paradigm. We will use here the self-reference watermark embedding approach presented in our previous paper [5] since it has a number of advantages in comparison with the other ones.

In this work, we propose to overcome the above mentioned problems by exploiting apriory knowledge about regularity of ACF or MS of the periodically repeated watermark [5], which consists of a set of local maxima, or peaks, with periodical structure. These peaks can be considered as a regular grid of points aligned along 2 main axes within 2 periods. Therefore, keeping in mind this discrete approximation of the grid of lines, one can first extract point and then exploit a Hough transform (HT) [7], or apply a Radon transform (RT) [8] directly to the ACF or MS, in order to receive a robust estimate of the general affine transform parameters. Actually the method can be applied to any template with a known regular structure. This approach has a number of advantages in comparison with the previous methods. First, it is very general and makes possible to determine any affine transform or even a combination of sequentially applied affine transforms. Moreover, the false peaks or outliers on the grid due to lossy compression or any other attack do not decrease the robustness of the approach due to the fact that it exploits the inherent redundancy of the used set of peaks. Therefore, the proposed approach is tolerant even with very strong lossy compression and low quality printing, which is not a case for the previously proposed methods. Finally, the strict mathematical apparatus of the HT or RT avoids the necessity for an exhaustive search. This is demonstrated on a number of examples and results of performed tests are shown in the final part of the paper. Here we consider affine distortions at the global level, however an extension of our method to the recovering from non-linear or local distortions, such as the random bending attack (RBA), has been presented in [10].

2. PROBLEM FORMULATION

The general setup of this work is illustrated in Figure 1 where the underlying grid structure of input points for the proposed algorithm represents the peaks in the ACF or MS of a periodical watermark, before and after the application of an affine transform. The peaks or points are placed at intersections between two classes of parallel lines along 2 principal directions, or *main axes*; points are further equidistantly placed along each direction, with respect to 2 *periods*, one along each main axis. Affine transforms only change the main axes directions and the periods, keeping the grid's regularity. Therefore such a grid can be represented by two vectors $\vec{u}, \vec{v}$ of which directions indicate the main axes, and of which norms $\|\vec{u}\|, \|\vec{v}\|$ correspond to the respective periods. The affine transform can then be estimated in reference with $\vec{u}_0, \vec{v}_0$ associated with the known originally embedded structure.

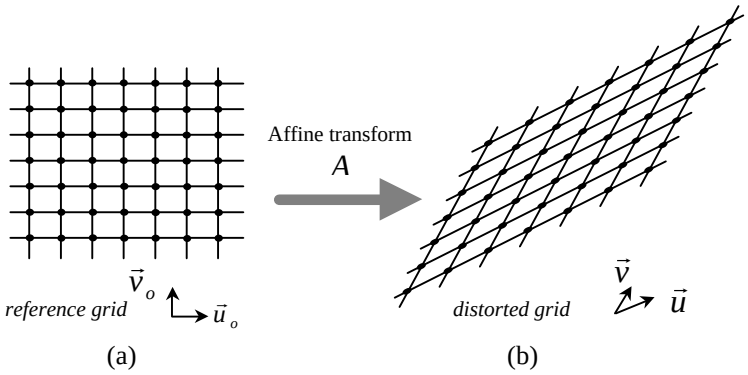


Figure 1. Grid distortion after affine transform: (a) grid and vectors $\vec{u}_o, \vec{v}_o$ corresponding to the known originally embedded information; (b) distorted grid and vectors $\vec{u}, \vec{v}$ after the application of the affine transform *A*.

3. PROPOSED APPROACH

The determination of the general affine transform applied to an image relies on the determination of the regular grid of points for estimating the affine matrix coefficients, and is based on the computation of the HT or RT. The entire approach lies on the regularity of the underlying grid, as extracted from an input and possibly distorted image, and by comparing it with a known reference grid, as explained below. If it is a colour image, in RGB or any other colour representation, the grid can still be extracted and the described method applied, either from the luminance component, or from each colour plane separately. The same approach can be applied to videos considering a regular grid along the time axis.

The first important aspect of the estimation problem is that the approach described here is not directly suitable for random templates, since it requires the grid of points to be regular, with points placed at intersections between two classes of equidistant spaced parallel lines as shown in Figure 1. Our method is typically applicable to the local maxima, or peaks, resulting either from the ACF, or from the MS generated by the periodic pattern, which have the property to keep this regularity under any affine transform. For this purpose a periodic pattern w (the watermark) should first be embedded into the original image. The ACF approach consists in computing the auto-correlation $\hat{w} * \hat{w}^T$ of the periodic pattern $\hat{w}$ which has been estimated from the possibly distorted image, where "*" is the convolution operator. The ACF can be computed as $\hat{w} * \hat{w}^T = F^{-1}(|F(\hat{w})|^2)$, where F is the discrete Fourier transform (DFT), F^{-1} is the inverse discrete Fourier transform (IDFT), and $|.|^2$ denotes the complex point-by-point square. In the MS approach peaks are extracted directly from the squared magnitude of the Fourier spectrum $|F(\hat{w})|^2$, where $|.|$ is the magnitude. This reduces the complexity of the approach, since only one DFT is required, and this also have a positive impact on the robustness.

The second important aspect concerns the robustness of the proposed approach with respect to one or more distortions, including affine transformation combined with signal degradation or fading such as lossy compression like JPEG. Generally, the estimated grid is affected by noise, consisting of missing points, extra points, and slight errors in the location of the remaining correct points. Even if only 3 points are needed in theory to get the matrix A coefficients, this would mostly lead either to a wrong estimation (due to the selection of false points, unless an exhaustive search is performed), or to a lack in the precision of the estimate of the parameters. This is even more important keeping in mind the robustness against cropping that put certain constraints on the period; for the realistic assumption of small part cropping, we choose a small repetition period.

The idea of our approach is then: 1) to embed a periodical structure with many repetitions as presented in [5] in order to get a high number of peaks, meaning a high redundancy, at the opposite of M. Kutter's approach [3] where the watermark was repeated only four times; consequently in our case a large number of peaks survive, even under severe signal fading. 2) to exploit this redundancy using the fact that random points are less likely to present significant alignments than correct ones. Then the HT or RT, combined with a robust extraction of periods from aligned points,

gives us a robust estimator of the correct underlying grid from which the affine parameters can be computed and inverted prior to the decoding.

We would like also to emphasize the security aspect of our approach. Actually, to avoid the estimation removal attack, the watermark can be locally predistorted prior to the embedding: this will not change considerably the regular pattern in the ACF or the MS, but it will provide completely random patterns after block averaging watermark estimation.

4. ESTIMATION OF THE AFFINE TRANSFORM

The robust estimation of the applied affine transform can be expressed as a penalized Maximum Likelihood estimation as below:

$$\hat{A} = \arg\min_{A \in \Phi} \left\{ r\left(\begin{pmatrix} x' \\ y' \end{pmatrix} - A \begin{pmatrix} x \\ y \end{pmatrix} \right) + m\Omega(A) \right\} \tag{3}$$

where $\hat{A}$ is an estimate of the affine transform within the set of possible solutions $A \in \Phi$, m is a regularization parameter that controls the trade-off between the feasibility of the estimate with respect to the observed data (first term) and the prior $\Omega(A)$, $r(.)$ denotes the cost function that in the case of Gaussian misalignments is a quadratic norm, and for Laplacian is a ℓ_1 norm. More general cases can be expressed for a Generalized Gaussian distribution of misalignments of which the two previous cases are particular cases. The prior $\Omega(A)$ represents the possible variations of the four parameters a, b, c, d corresponding to the affine transform matrix A (equation 1). The idea of restricting possible combinations of shearing and flipping of the radii between template points in the magnitude spectrum proposed by S. Pereira [4] can be considered as a particular case of this prior.

Instead of performing the overall minimization of equation 3, we propose a simple practical solution using a two stage algorithm. First, the rotation angle or principal directions of the main axes are estimated based on the projections of the aligned points in the HT or RT domain. Secondly, the periods along the principal directions are robustly estimated along each main axis using a correlation-like method between a candidate period-based grid and aligned points. As a result the 2 main axes orientations and 2 periods estimates are obtained, from which the undergone affine transform is estimated.

4.1. Computing the affine parameters

Within the ACF or the MS domain, an affine transform maps the reference grid represented by vectors $\vec{u}_o$ and $\vec{v}_o$ to another one represented by vectors $\vec{u}$ and $\vec{v}$ (Figure 1). The four parameters a, b, c, d of the affine transform matrix A have then to be estimated. Assuming that we know the reference grid, specified by $\vec{u}_o, \vec{v}_o$, and that we got the correct estimations of $\vec{u}$ and $\vec{v}$ from the extracted points, the matrix can be expressed as:

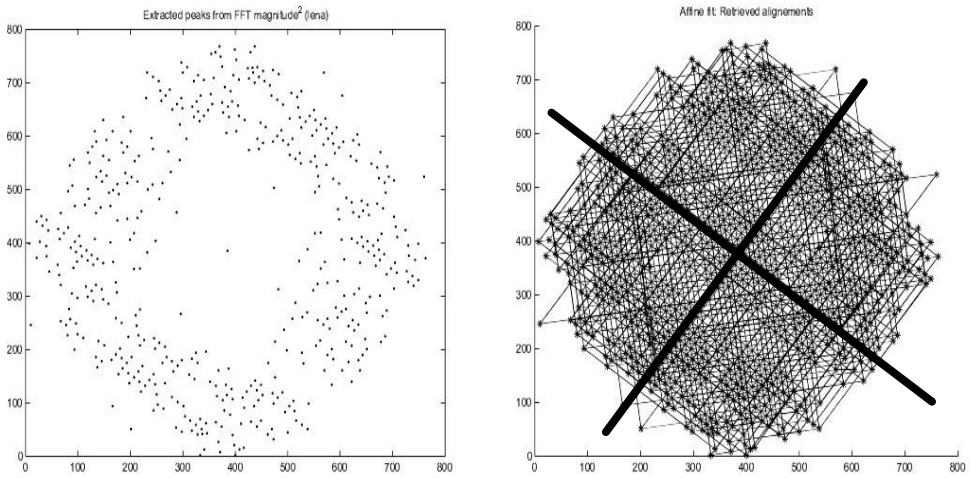
$$A = T \cdot T_o^{-1} \tag{4}$$

where the 2×2 matrices T and T_0 come from the pairs $(\vec{u}, \vec{v})$ and $(\vec{u}_o, \vec{v}_o)$ respectively. If the points came from the ACF, directions and norms of these vectors are directly related to the periodicity along the two main axes of the embeded pattern. If the MS was used, vectors of inverse norms should be considered to form T and T_0 since they come from the discrete Fourier transform (DFT) domain. In the case when a square pattern was embedded, repeated horizontally and vertically along the x- and y- axes within the same period t, T_0 becomes (for both ACF and MS cases):

$$T_o = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} \tag{5}$$

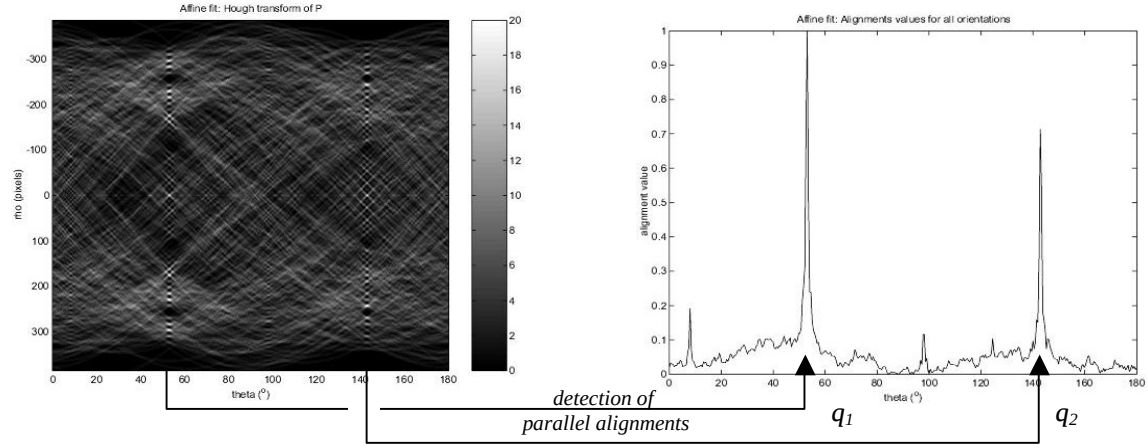
A final aspect to mention is the possibility of ambiguities in the estimate of the affine transform, since the extracted grid of point contains no information about the correct orientations of the vectors along the main axes. Therefore the estimated vectors $\vec{u}$ and $\vec{v}$ may have wrong orientations, resulting in 8 ambiguities namely horizontal, vertical flips,

and/or 90°-rotations. Then all of them need to be additionally checked. One means to overcome this inconvenient is to use patterns with central symmetry, presenting in this case only two 90°-rotations ambiguities according to equation 5.



(a) (b)

Figure 2. Estimating the affine parameters from a rotated and compressed image: (a) extracted peaks from the magnitude spectrum of the estimated watermark; (b) line fitting to these points, along the two main axes (indicated by the 2 thick lines) as well as two diagonals.



(a) (b)

Figure 3. Estimation of the principal directions: (a) the HT, showing vertically aligned peaks which correspond to the 2 main axes; (b) distribution function $h(q)$ of parallel alignments with respect to the angle of projection, with $q \in [0\ldots180°]$; the 2 highest peaks reflect the projection angles of the main axes q_1, q_2.

4.2. Estimating main axes based on Hough transform

One can either compute the HT of the discrete set of points, or compute the RT directly from the real ACF or the MS. Figure 2.a shows a noisy set of points due to some signal degradation, after the application of a rotation. The goal is to fit lines along significant alignments as shown in Figure 2.b, in order to estimate main axes directions and periods. The HT or RT is a parametric transform which converts the *x,y*-representation into a q,r-representation, r being a distance

of projection from the origin in function of the angle q of projection [7,8]. Figure 3.a shows the resulting HT from points in Figure 2, the projection r-axis being vertical and the angle q–axis being horizontal. The strong vertically aligned spots are clearly visible, corresponding to parallel alignments. Then we add these vertically aligned points to get the distribution of parallel alignments in function of the projection angle $h(q)$, of which the 2 highest peaks correspond to the 2 main axes with angles q_1, q_2. We then classify points belonging to each alignemnents, and groups of parallel alignements, while discarding remaining noisy points. Finally, any robust line fitting algorithm such as linear regression, least-squares or median least squares can be applied directly to the classified points (Figure 2.b) in order to get a precise estimate of the direction of each main axis. For higher robustness, one can consider more than one axis candidates (taking diagonals for example), and constrain the angle between two of them inside in a reasonable interval, e.g. between $p/8$ and $7p/8$.

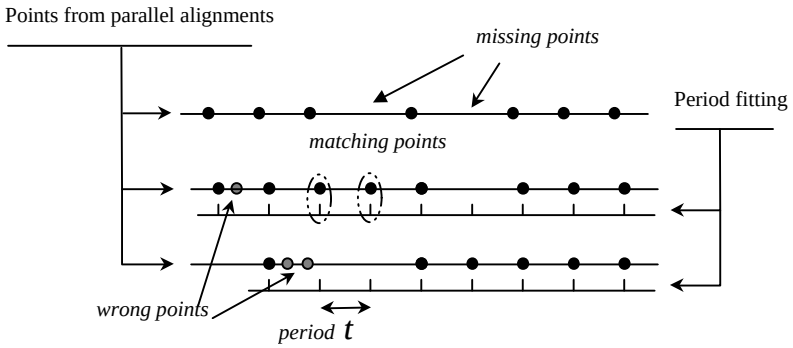


Figure 4. For each main axis, the period is estimated based on a ML-estimator, using the noisy set of points coming from parallel lines belonging to the same main axis.

4.3. Estimating the periods

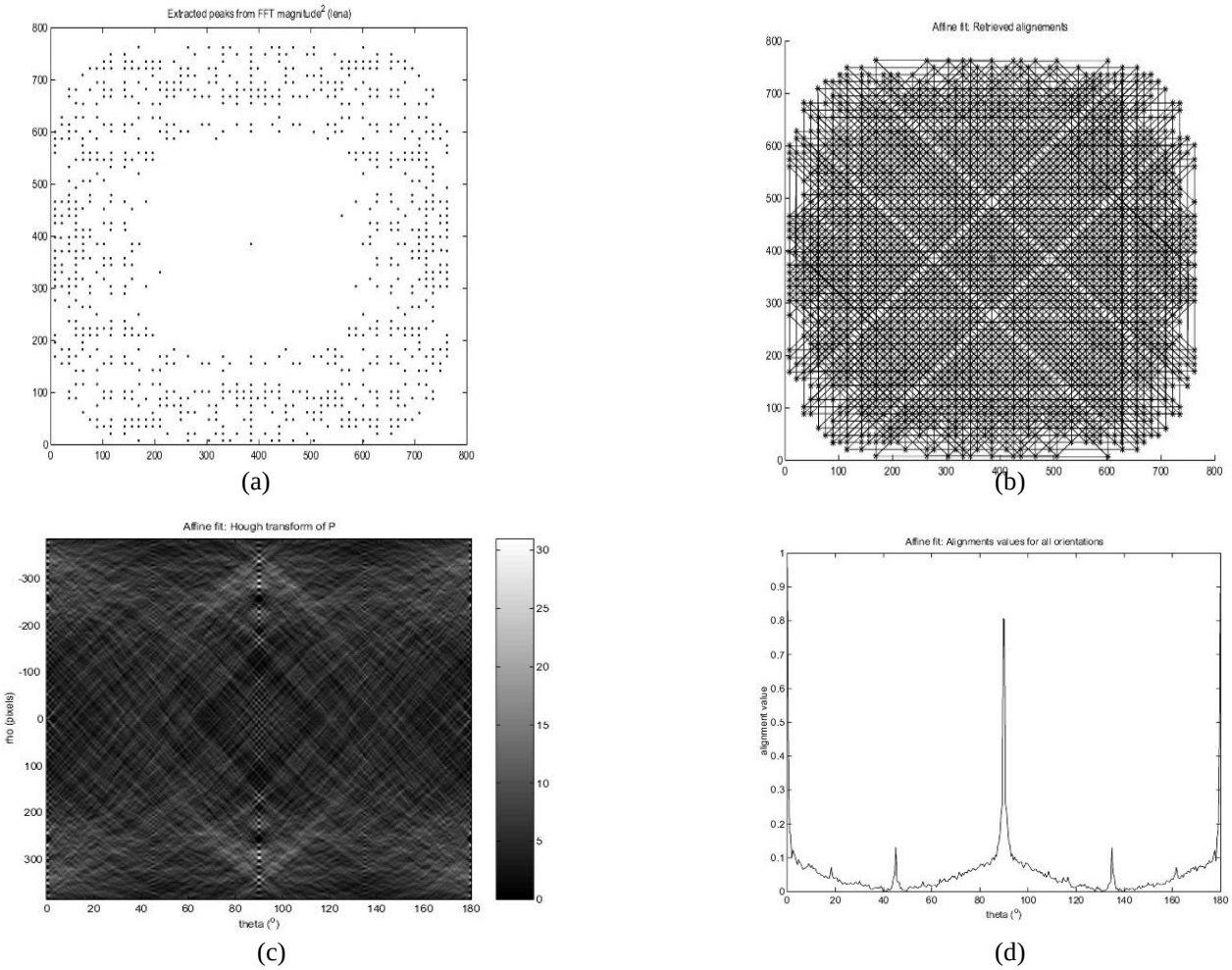
The periods can be estimated either from the distances between parallel lines or from the distances between points along each principal direction. Figure 4 demonstrates the problem, using all points from parallel alignments for better accuracy. We model the problem as finding the most likely period along a 1D-axis x between discrete peaks $y(x)$ according to equation 6:

$$y(x) = \sum_{k=0}^{n-1} a \cdot d(x - k \cdot t + e_t) + \sum_{i=0}^{m-1} g \cdot d(x - t_i) \tag{6}$$

where d is the Dirac function, t is the correct period, n is the number of repetitions, $a = \begin{cases} 1 & with\ prob.\ \ p \\ 0 & with\ prob.\ \ 1-p \end{cases}$ with p the probability that right points disappear, e_t is a zero mean Gaussian i.i.d. perturbation in the position of right peaks, m is the maximum number of outlier peaks, $g = \begin{cases} 1 & with\ prob.\ \ q \\ 0 & with\ prob.\ \ 1-q \end{cases}$ with q the probability of outliers, and t_i is an uniform i.i.d. random position for outlier peaks; also, $r \approx (p \cdot n)/(q \cdot m)$ gives a signal to noise ratio estimate of existing right peaks relatively to outliers. Therefore, we use a maximum likelihood (ML) estimator with one constrained period, for example $1/2 \le t \le 2$, in order to get a robust estimate of the correct periods t_1 and t_2 for both axes.

(a) (b) (c)

Figure 5. Example illustrating the performed experiments: (a) cover “Lena”; (b) stego “Lena”; (c) rotated stego “Lena” by 31°.



(a) (b)

(c) (d)

Figure 6. No distortion have been applied: (a) extracted peaks from the MS; (b) fitted lines, clearly showing vertical and horizontal alignments; (c) corresponding HT; (d) angle distribution function, with 2 strong peaks in 0°/180° and in 90°.

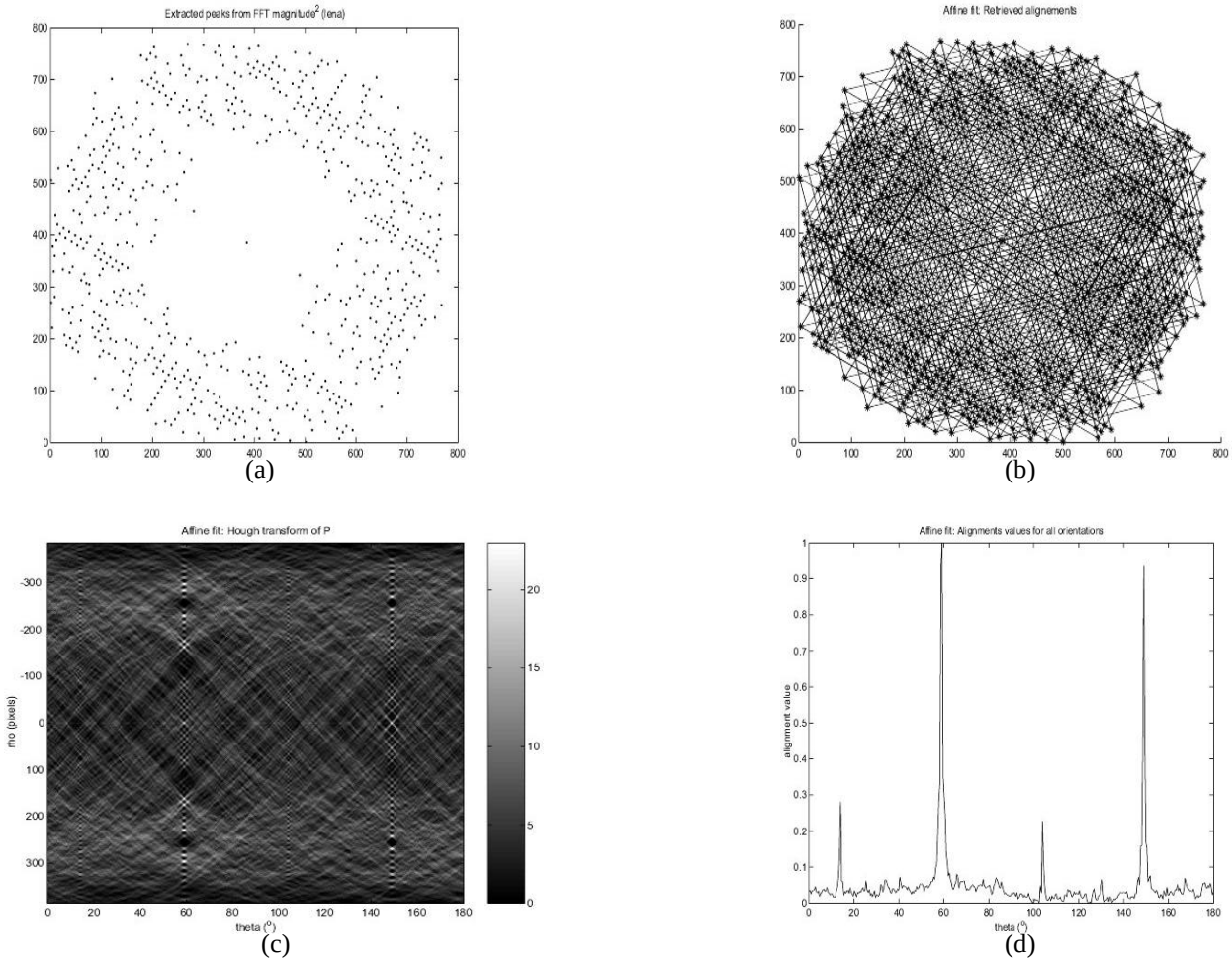


Figure 7. The rotation of 37° was applied: (a) extracted peak from the MS, rotated as well and with missing peaks; (b) fitted lines, clearly showing the rotation; (c) corresponding HT; (d) angle distribution function, with peaks shifted according to the angle of rotation.

5. EXPERIMENTAL RESULTS

The experiments were performed on several images. The robustness of the estimation of the HT or the RT based on the extracted grid of points from the MS of the estimated watermark is illustrated in Figures 5 to 8, without and with geometrical distortion, as well as with compression, using the "Lena" image as an example. Figure 5 shows cover image, stego image, and stego image rotated by 31°. Figures 6 to 8 show the extracted set of peaks (a), the fitted lines including two diagonals as well as the two main axes (b), the corresponding HT (c), and the angle $h(q)$ with angle varying from 0 to 180° (d). In Figure 6 no distortion occurred, and most of peaks survived; in Figure 7 the rotation of 31° was applied, resulting in a rotation of the grid as well and disappearing peaks; and in Figure 8 a JPEG compression with a quality factor (QF) of 50% was applied in addition to the rotation, leading to the removal of almost all peaks. However, in every case the angle of the main axes were robustly estimated, the undergone rotation correctly compensated and the watermark successfully decoded.

The high performance of the proposed approach with respect to the geometrical attacks has been confirmed by the Stirmark 3.1 benchmark [9]. Six proposed images were watermarked with the required PSNR of about 38 dB with a 64 bits watermark, the results of the attacks are reported in Table 1. The obtained results show the high robustness of the proposed approach, as compared to other existing methods.

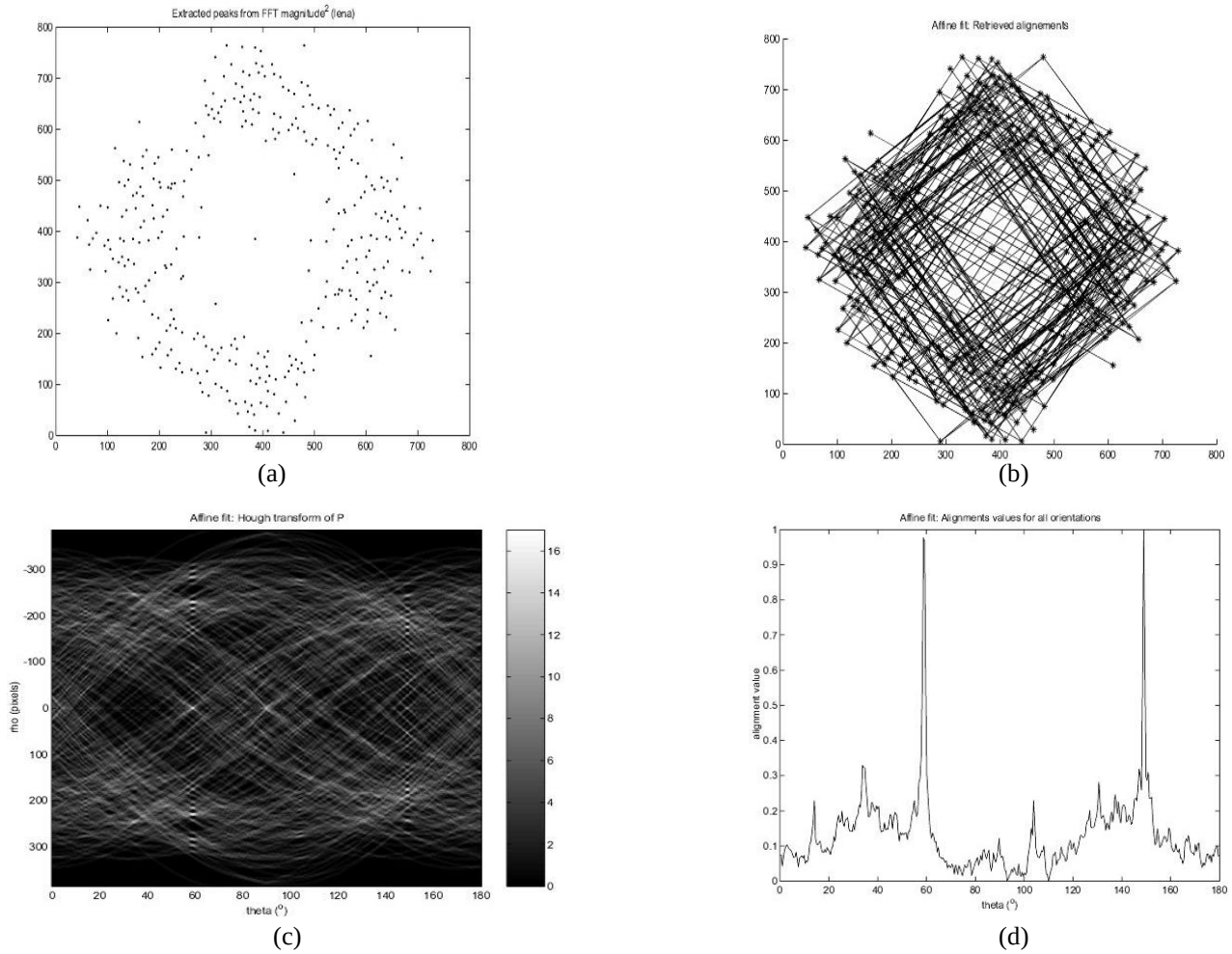


Figure 8. A rotation of 37° combined with a compression with classical JPEG, with QF=50%, was applied; the main axes detection and the estimation of the affine transform are still easily performed: (a) extracted peak from the MS, most of them are lost; (b) fitted lines, including lines along the correct main axes; (c): the corresponding HT; (d): angle distribution function, still presenting strong peaks for the correct angles of the main axes.

Table 1. Results of the Stirmark 3.1 geometrical attacks.

Geometrical attack	Stirmark score
Scaling	1.00
Cropping	0.99
Shearing	1.00
Rotation (auto-crop, auto-scale)	0.99
Column and line removal	1.00
Flip	1.00

6. CONCLUSIONS

In this paper, we presented a robust approach for the determination of general affine distortions which have been applied to an image, for watermarking applications. The problem of the robust determination of the geometrical transform, based on regular templates or on self reference watermarks associated with auto-correlation, relies on a robust estimation of a penalized Maximum Likelihood function relating the extracted points and a set of a-priory known reference points. We showed that this problem can be efficiently solved by Hough transform or Radon transform in the case of general affine transforms, avoiding the need for exhaustive search. This approach robustly filters points resulting from templates or from self reference information, exploiting most of them while discarding outliers, based the expected regularity of their underlying structure. As a result we obtain a high probability of estimation of the correct affine transform, even when severe degradations have occurred including JPEG compression with a quality factor of 50%. The robustness of the method is illustrated by the Stirmark tests. The proposed approach can also be used in any application where images need to be resynchronized, such as satellite imaging, or interactive digital maps, avoiding the need to embed visible marks.

REFERENCES

1. J. O'Ruanaidh, and T. Pun, "Rotation, Scale and Translation Invariant Spread Spectrum Digital Image Watermarking", *Signal Processing*, v. 66, No 3, pp. 303—317, 1998.
2. S. Pereira, J. J. K. O'Ruanaidh, F. Deguillaume, G. Csurka and T. Pun, "Template Based Recovery of Fourier-Based Watermarks Using Log-polar and Log-log Maps", in *Int. Conference on Multimedia Computing and Systems*, June 1999.
3. M. Kutter, “Watermarking resistant to translation, rotation and scaling”, in *Proc. SPIE Int. Symp. on Voice, Video, and Data Communication*, November 1998.
4. S. Pereira, T. Pun, “Fast Robust Template Matching for Affine Resistant Watermarks”, *Lecture Notes in Computer Science: Third International Workshop on Information Hiding*, Springer, vol. 1768, pp. 199-210, 1999.
5. S.Voloshynovskiy, F. Deguillaume and T. Pun, “Content adaptive watermarking based on a stochastic multiresolution image modeling”, *EUSIPCO2000, X European Signal Processing Conference*, September 2000.
6. F. Hartung, and M. Kutter, "Multimedia watermarking techniques", Proceedings of the IEEE, vol. 87, no. 7, pp. 1079-1107, July 1999.
7. P. Hough, "Method and means for recognizing complex patterns", U.S. Patent 3069654, 1962.
8. S.R. Deans, "Hough transform from Radon transform", *IEEE Transactions on PAMI*, vol. 3, no. 2, pp. 185-188, March 1981.
9. M. Kutter, and F.A.P. Petitcolas, "A fair benchmark for image watermarking systems", *SPIE, Electronic Imaging '99, Security and watermarking of multimedia contents*, vol. 3657, pp. 219-239, January 1999.
10. S. Voloshynovskiy, F. Deguillaume, and T. Pun, “Multibit Digital Watermarking Robust Against Local Nonlinear Geometrical Distortions”, In *IEEE International Conference on Image Processing*, ICIP2001, pp. 999-1002, Thessaloniki, Greece, 2001.