# A ROBUST IMAGE WATERMARKING SCHEME BASED ON THE ALPHA-BETA SPACE

*P. Martins and P. Carvalho*

Centre for Informatics and Systems, University of Coimbra
Pólo II, Pinhal de Marrocos, 3030-290 Coimbra, Portugal
{pjmm,carvalho}@dei.uc.pt

## ABSTRACT

A robust image watermarking scheme relying on an affine invariant embedding domain is presented. The invariant space is obtained by triangulating the image using affine invariant interest points as vertices and performing an invariant triangle representation with respect to affine transformations based on the barycentric coordinates system. The watermark is encoded via quantization index modulation with an adaptive quantization step.

## 1. INTRODUCTION

The robustness with respect to geometric transformations has been a widely studied topic in digital image watermarking, leading to the introduction of several solutions trying to overcome the effects carried out by these distortions, specially the de-synchronization effect, since most of the watermarking schemes perform a detection based on a correlation measure; thus, if a marked image is geometrically distorted, the detector will correlate wrong image components [1]. Robust schemes can be be grouped into four categories: invariant domain based methods, template based methods, self-synchronization based methods and content based methods. The first category includes methods which exploit invariant or partially invariant domains for watermark insertion, *e.g.*, applying the Fourier-Mellin transform [2] or the Radon transform [3]. The main drawback of invariant based methods is that they usually require interpolation in order to obtain the invariant domain; at least two interpolation stages are present in an invariant-based scheme (one to obtain the invariant space and embed the watermark into it and another one to perform the inverse mapping). Because of the inaccuracy of interpolation methods, the performance of these solutions is highly affected. Methods from the second category provide robustness to geometric distortions by retrieving artificially embedded references which are used as a mean of identification of geometric transformations [4]. Template-based methods tend to affect severely the image fidelity due to the addition of the reference signal. Moreover, templates can be easily removed. Self-synchronization-based methods are similar to template-based methods in the sense that they achieve robustness by means of the identification of geometric distortions, but in this case, the watermark itself can be used to identify the transformation [5]. These approaches are quite sensitive to filtering. Feature based methods make use of perceptually significant portions of data to embed the information. For example, Bas *et al.* [6] exploited strategies based on feature points. The features were utilized to construct a Delaunay tesselation that later was applied to embed the watermark. In [7], Hang and Tang proposed a scheme based on the detection of feature points retrieved by the Mexican Hat wavelet scale interaction method. These feature points were used as references to embed the watermark into a normalized representation of the points neighborhood. The main disadvantages exhibited by feature-based schemes are the fact that the effectiveness of the watermarking depends on the effectiveness, *i.e.* robustness, of the feature detector/extractor, and they are usually computationally expensive.

In this paper, we present a content-based image watermarking scheme providing resilience to affine transformations, relying on the $(\alpha, \beta)$ space, an affine invariant embedding domain. The mapping is achieved without any interpolation, avoiding the inaccuracy carried out by this operation. The watermark embedding is performed via quantization index modulation [8], adjusting the quantization step in order to have the best trade-off between the robustness and imperceptibility of the watermark.

The remainder of this paper is organized as follows: Section 2 provides a description of the steps required to achieve the $(\alpha, \beta)$ domain; in Section 3, we describe a quantization-based technique which adopts the suggested space as the embedding domain; simulation results concerning the resilience of the watermarking scheme are given in Section 4; finally, in Section 5, concluding remarks are presented.

## 2. THE $(\alpha, \beta)$ SPACE

The main steps required to obtain the $(\alpha,\beta)$ space, an affine invariant domain, are outlined in Fig. 1. The first step includes an affine invariant interest point detection by means of a modified version of the detector introduced by Mikolajczyk and Schmid [9], a method based on an iterative process that converges to affine invariant key points by modify-

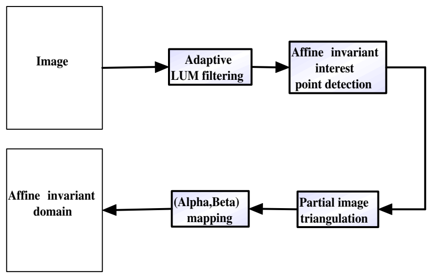


**Fig. 1**. Achieving the $(\alpha, \beta)$ space, an affine invariant domain.

ing the location, scale and shape of the initial points given by the Harris method [10], which are already translation and rotation invariant. Completely invariant points to any kind of affine transformation are obtained by introducing a scale-space representation for the Harris operator with pre-selected scales. Locations at which the Laplacian attains a maximum over scales are chosen (this procedure provides scale invariance). Invariance to other affine transformations is provided by estimating the affine shape of a pixel derived from the second moment matrix. In the affine scale-space, at pixel $\mathbf{x}$, the second moment matrix is defined by

$$M(\mathbf{x}, \Sigma_I, \Sigma_D) = \det(\Sigma_D) g(\Sigma_I) *$$
$$\begin{bmatrix} L_x^2(\mathbf{x}, \Sigma_D) & L_y L_x(\mathbf{x}, \Sigma_D) \\ L_x L_y(\mathbf{x}, \Sigma_D) & L_y^2(\mathbf{x}, \Sigma_D) \end{bmatrix}, \quad (1)$$

where $\Sigma_I$ and $\Sigma_D$ are the covariance matrices which determine the size of the integration and differentiation Gaussian kernels, respectively; $L_x$ and $L_y$ are the image derivatives in the $x$ and $y$ direction, respectively, computed with a Gaussian kernel whose size was determined by $\Sigma_D$; and $g$ is a Gaussian kernel determined by $\Sigma_I$. The aforementioned method is modified by performing an adaptive LUM filtering [11] over the image before the detection in order to improve the repeatability of the detector across a wide range of transformations over an image by setting adaptively the filter parameters which define the levels of sharpening and smoothing carried out by the filter (Fig. 2 illustrates the effects of this pre-filtering on the interest point detection). These parameters settings take mainly into account the inaccuracy of interpolation operations and the smoothing performed by Gaussian kernels in order to reduce noise. A more detailed description of the method can be found in [11]. The next step comprises a partial image triangulation using the interest points as vertices. The regions of interest identified by the triangles define potential areas for watermark embedding which are converted into an $(\alpha, \beta)$ coordinates representation. The conversion into the new coordinates system is merely a normalization process applied to the triangles relying on properties concerning triangular barycentric coordinates.

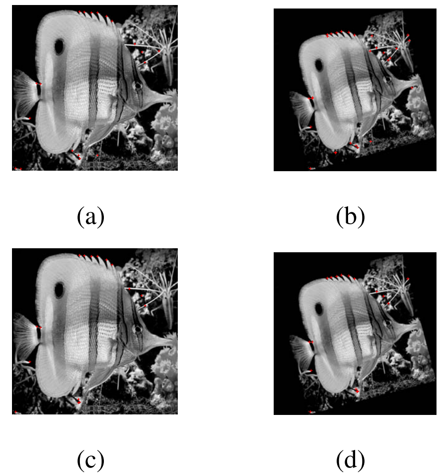


**Fig. 2**. "Fish" image and its affine invariant Harris interest points: (a) original; (b) rotated 15 degrees; (c) original (pre-filtered); (d) rotated 15 degrees (pre-filtered).

Given a triangle $T$ identified by the vertices $\mathbf{x_1}$, $\mathbf{x_2}$ and $\mathbf{x_3}$, and a point $\mathbf{x}$ inside $T$, the barycentric coordinates of $\mathbf{x}$ with respect to $T$ consist of the triplet $(\alpha, \beta, \gamma)$ such that

$$\mathbf{x} = \alpha \mathbf{x_1} + \beta \mathbf{x_2} + \gamma \mathbf{x_3}, \quad (2)$$

where $\alpha$, $\beta$, $\gamma \geq 0$ and $\alpha + \beta + \gamma = 1$.
The $(\alpha, \beta)$ coordinates system is derived from the barycentric coordinates system by discarding $\gamma$. Two properties regarding affine invariance are exhibited by this coordinates system:

**Property 1** *The coordinates* $(\alpha, \beta)$ *are invariant to affine transformations over* $\mathbf{x}$.

*Proof*: Let us assume that $\mathbf{x}$ goes under an affine transformation $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$. Thus,

$$f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b} = \mathbf{A}[\alpha(\mathbf{x_1} - \mathbf{x_3}) + \beta(\mathbf{x_2} - \mathbf{x_3}) + \mathbf{x_3}] + \mathbf{b}.$$

From the linearity of $\mathbf{A}$, we get:

$$f(\mathbf{x}) = \mathbf{A}[\alpha(\mathbf{x_1} - \mathbf{x_3})] + \mathbf{A}[\beta(\mathbf{x_2} - \mathbf{x_3})] + \mathbf{A}\mathbf{x_3} + \mathbf{b} =$$
$$= \alpha \mathbf{A}(\mathbf{x_1} - \mathbf{x_3}) + \beta \mathbf{A}(\mathbf{x_2} - \mathbf{x_3}) + \mathbf{A}\mathbf{x_3} + \mathbf{b}.$$

**Property 2** *In the* $(\alpha, \beta)$ *space, a triangle in Cartesian coordinates is converted into a right-angled triangle identified by the vertices* $(0, 0)$, $(1, 0)$ *and* $(0, 1)$.

*Proof*:Let $T$ be a triangle defined by the vertices $\mathbf{x_1}$, $\mathbf{x_2}$ and $\mathbf{x_3}$ in Cartesian coordinates. (1,0), (0,1) and (0,0) are the corresponding $(\alpha, \beta)$ coordinates of $\mathbf{x_1}$, $\mathbf{x_2}$ and $\mathbf{x_3}$, and $T$ is mapped into the triangle $T'$, where

$$T' = \{(\alpha, \beta) \in \mathbb{R} | \alpha \geq 0, \beta \geq 0, \beta + \alpha \leq 1\}.$$

## 3. WATERMARKING SCHEME

The selection of which triangles are more suitable to be marked is a key stage in the watermark embedding process. This selection is performed by analyzing, for each triangle, the following features: (i) relative area; (ii) distance to the image centroid; (iii) texture level; (iv) strength of the vertices.

Suppose we have $p$ non-overlapped triangles, the relative area of a triangle $T_j$ is defined in Eq. (3), where $A_j$ is the number of pixels in $T_j$.

$$\bar{A}_j = \frac{A_j}{\max_{i\in\{1,\ldots,p\}}\{A_i\}}, \tag{3}$$

The image centroid is not an affine invariant measure, but it exhibits some stability with respect to geometric transformations. The image centroid of an image $m \times n$ $I$ has the form

$$\mathbf{c} = (\sum_{x=1}^{m}\sum_{y=1}^{n} xh(x,y), \sum_{x=1}^{m}\sum_{y=1}^{n} yh(x,y)), \tag{4}$$

where $h(x,y) = \frac{I(x,y)}{\sum_{u=1}^{m}\sum_{v=1}^{n} I(u,v)}$. The distance between $T_j$ and $\mathbf{c}$ is defined via the equation

$$\text{dist}(T_j,\mathbf{c}) = \frac{\left\|\mathbf{c} - \mathbf{c}_{T_j}\right\|}{\max_{i\in\{1,\ldots,p\}}\{\|\mathbf{c} - \mathbf{c}_{T_i}\|\}}, \tag{5}$$

where $\|\cdot\|$ is the Euclidean norm and $\mathbf{c}_{T_j}$ is the centroid of the sub-image defined by $T_j$.

Texture level is evaluated by computing the mean of the energy of the detail sub-bands of a single-level (Daubechies-8) wavelet transform decomposition of each triangle as defined in Eq. (6):

$$Text_j = \text{mean}(\text{e}(f_{j_h}), \text{e}(f_{j_v}), \text{e}(f_{j_d})), \tag{6}$$

where $f_{j_h}$, $f_{j_v}$ and $f_{j_d}$ are the horizontal, vertical and diagonal detail sub-bands, respectively and $\text{e}(\cdot)$ is the energy function defined according to Eq. (7), where $n$ and $m$ are, respectively, the number of rows and columns of a given matrix $I$.

$$\text{e}(I) = \frac{1}{m \times n}\sum_{x=1}^{m}\sum_{y=1}^{n} I^2(x,y) \tag{7}$$

The strength of $T_j$ is given by

$$str_j = \sum_{i=1}^{3}\frac{6}{\text{r}(v_{j_i})}, \tag{8}$$

where $\text{r}(v_{j_i})$ denotes the rank of vertex $i$ in triangle $T_j$ according to Harris interest point strength [10]. The measure of these features is then applied to establish a triangles hierarchy, by assigning a score $S_j$ to $T_j$:

$$S_j = \omega_1\bar{A}_j + \omega_2(1-\text{dist}(T_j,\mathbf{c})) + \omega_3\frac{\min\{t, Text_j\}}{t} + \omega_4 str_j, \tag{9}$$

where $\omega_i$, $i = 1,\ldots,4$, are user-defined positive weights and $t$ is a threshold energy value (empirically set to 40) used to normalize the texture level values.

### 3.1. Watermark encoder

For a triangle $T_i$ in $(\alpha,\beta)$ coordinates, the embedding strategy comprises the steps described below:

1. Store the pixel intensities of a pre-selected set of coordinates $(\alpha_j,\beta_k)$ ,$j = 1,\ldots,L_1$ and $k = 1,\ldots,L_2$ into a vector $x_i$. Pre-selected coordinates which do not have a corresponding intensity, are removed from the vector.

2. Generate pseudo-randomly (based on a secret key), the dither vector $d[\cdot,0]$ over $[\frac{\Delta_i}{2},\frac{\Delta_i}{2}]$, where $\Delta_i$ is one of the quantization steps $\{20{,}40{,}60\}$ chosen according to the triangle's texture level, and $d[\cdot,1]$ is given by

$$d[n,1] = \begin{cases} d[n,0] - \frac{\Delta_i}{2} & \text{if } d[n,0] > 0 \\ d[n,0] + \frac{\Delta_i}{2} & \text{if } d[n,0] \le 0 \end{cases}. \tag{10}$$

3. Encode each bit $m_n$ from message $m$ into the $n^{th}$ sample of $x_i$ according to Eq. (11), where $Q$ is the quantizer, *i.e.*, $Q(x,\Delta_i) = round(\frac{x}{\Delta_i})\Delta_i$.

$$\widetilde{x}_{i_n} = Q(x_{i_n} + d[n,m_n],\Delta_i) - d[n,m_n] \tag{11}$$

4. Compute $\widetilde{T}_i$, the modified version of $T_i$, based on $\widetilde{x}_i$.

### 3.2. Watermark decoder

The initial steps performed by the decoder are similar to the ones performed by the encoder: pre-filtering followed by interest point detection and image triangulation. However, the strength of the vertices tends to be modified after some image distortions, changing the hierarchy of the interest points (the adaptive LUM filtering stage tries to attenuate these changes), hence, if at the embedding stage, $\frac{n!}{(n-3)!3!}$ triangles are considered, at the decoder, $\frac{n!}{(n-3)!}$ triangles have to be considered, *i.e.*, each triangle $T$ is seen as six different triangles by rearranging the vertices sequence. Once again, the triangles hierarchy is set, and from the triangle with the highest rank to the one with lowest rank, the watermark extraction is performed. Decoding is performed on each $n^{th}$ sample of a vector $z$ containing the pixel-intensities of the pre-selected $(\alpha,\beta)$ coordinates, by estimating the message bit $\hat{m}_n$ in a way that it minimizes the following Euclidean distance:

$$\hat{m}_n = \arg\min \|z_n - Q(z_n + d[n,m_n],\Delta)\|^2, \tag{12}$$

where $m_n \in \{0,1\}$.

## 4. SIMULATION RESULTS

Table 1 lists the extraction results after geometric distortions, JPEG compression, median filtering or Gaussian noise addition over marked versions of images "Lena" and "Peppers" ($512 \times 512$, 256 graylevels) depicted in Fig. 3. The message

"ICME" was previously embedded into the images by encoding it into the triangle exhibiting the highest score defined in Eq. (9) (the weights were all set to 1). The watermark was successfully decoded in all the cases, except for JPEG compression with a quality factor of 25%, however after this attack, the perceptual quality of images was considerably degraded. Mainly due to a local affine invariant embedding domain, the algorithm has revealed to be effective dealing with geometric transformations such as affine transformations or cropping. However, geometric distortions tend to alter the hierarchy of interest points established by the Harris measure, particularly in the case of scale changes and shearing. Hence, not only one triangle was checked for watermark extraction, but all the triangles that can be identified by the 10 strongest affine invariant interest points. The word "ICME" was found among the extracted sequences.

(a) (b)

**Fig. 3**. Marked versions of test images: (a) "Lena"; (b) "Peppers".

| | "Lena" | "Peppers" |
|---|---|---|
| Rotation(5 deg.) | "ICME" | "ICME" |
| Rotation(15 deg.) | "ICME" | "ICME" |
| Rotation(25 deg.) | "ICME" | "ICME" |
| Scaling(1.2%) | "ICME" | "ICME" |
| Scaling(0.9%) | "ICME" | "ICME" |
| Scaling(0.8%) | "ICME" | "ICME" |
| Rotation(20 deg.) + scaling(0.8%) | "ICME" | "ICME" |
| Shearing($410 \times 460$) | "ICME" | "ICME" |
| Cropping | "ICME" | "ICME" |
| Flipping | "ICME" | "ICME" |
| JPEG(Q=75%) | "ICME" | "ICME" |
| JPEG(Q=50%) | "ICME" | "ICME" |
| JPEG(Q=25%) | – | – |
| Median filter ($3 \times 3$) | "ICME" | "ICME" |
| Gaussian noise (3%) | "ICME" | "ICME" |

**Table 1**. Extraction results.

## 5. CONCLUDING REMARKS

A robust image watermarking scheme based on the $(\alpha, \beta)$ space was introduced. The robustness of the scheme with respect to affine transformations is achieved by embedding the watermark in the $(\alpha, \beta)$ space, an affine invariant domain, which is obtained, firstly, by detecting affine invariant interest points and defining triangle-shaped image regions identified by the interest points. The partial image triangulation is followed by a triangle normalization, carried out by mapping into a coordinate system, based on barycentric coordinates. Since feature point based watermarking schemes effectiveness depends on the effectiveness of the feature detection, an adaptive pre-filtering stage was added to the interest point detector, yielding a better repeatability across several geometrically distorted versions of an image. Unlike most of the invariant-based methods, no interpolation is required to achieve the invariant domain and embed the watermark into it, or ,consequently, extract the hidden information from it. Furthermore, the invariant space does not limit the embedding strategy to be applied and can be used as the starting point of any other invariant domain.

## 6. REFERENCES


[1] B. Abdel-Aziz, J. Zhao, and J.-Y. Chouinard, "On the Limits of Second Generation Watermarks," in *Proc. of the First Canadian Conference on Computer and Robot Vision*, 2004.

[2] J. J. K. O'Ruanaidh and S. Pereira, "A secure robust digital image watermark," in *Proc. of SPIE – Electronic Imaging: Processing Printing and Publishing in Color*, 1998, pp. 150–163.

[3] D. Smitopoulos, D. E. Koutsonanos, and M. G. Strintzis, "Robust Image Watermarking Based on Generalized Radon Transformations," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 732–745, 2003.

[4] S. Pereira and T. Pun, "Fast robust template matching for affine resistant image watermarks," in *Proc. of the 3rd International Information Hiding Workshop*, 1999, pp. 207–218.

[5] F. Deguillaume, S. Voloshynovskiy, and T.Pun, "Method for the estimation and recovering from general affine transform in digital watermarking," in *Proc. of SPIE – Secutity and Watermarking of Multimedia Contents IV*, 2002, pp. 313–322.

[6] P. Bas, J. M. Chassery, and B. Macq, "Geometrically invariant watermarking using feature points," *IEEE Trans. on Image Processing*, vol. 11, no. 9, pp. 1014–1028, 2002.

[7] C.-W. Tang and H.-M. Hang, "A Feature-Based Robust Digital Watermarking Scheme," *IEEE Trans. on Signal Processing*, vol. 51, no. 4, pp. 950–959, 2003.

[8] B. Chen and G. W. Wornell, "Quantization Index Modulation Methods: A Class of Provably Good Methods for Digital Watermarking and Information Embedding," *IEEE Trans. Information Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[9] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

[10] C. G. Harris and M. J. Stephens, "Combined corner and edge detector," in *Proceedings of the Fourth Alvey Vision Conference*, 1988, pp. 147–151.

[11] P. Martins and P. Carvalho, "Adaptive Lum Pre-Filtering for Affine Invariant Interest Point Detection," in *Proc. of the IEEE Int. Workshop on Intelligent Signal Processing*, 2005, pp. 248–253.