Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

PERGAMON

Pattern Recognition 37 (2004) 1365–1375

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Localized image watermarking based on feature points of scale-space representation

Jin S. Seo*, Chang D. Yoo

Division of EE, Department of EECS, Korea Advanced Institute of Science and Technology, 373-1 Guseong Dong, Yuseong Gu, Daejeon 305-701, Republic of Korea

Received 16 June 2003; accepted 9 December 2003

Abstract

This paper proposes a novel method for content-based watermarking based on feature points of an image. At each feature point, the watermark is embedded after scale normalization according to the local characteristic scale. Characteristic scale is the maximum scale of the scale-space representation of an image at the feature point. By binding watermarking with the local characteristics of an image, resilience against affine transformations can be obtained easily. Experimental results show that the proposed method is robust against various image processing steps including affine transformations, cropping, filtering and JPEG compression.

Keywords: Watermarking; Feature points; Scale space; Geometric distortion; Content-based synchronization

1. Introduction

With the advent of Internet, there has been an explosion in growth of the use of digital media through electronic commerce and on-line services. Since digital media is easily reproduced and manipulated, anyone is potentially capable of incurring considerable financial loss to the media producers and content providers. Digital watermarking is introduced to safeguard against such loss. While the most prominent application of watermarking is copyright protection [1], others, including fingerprinting [2], image authentication [3] and copy protection [4] are important research areas.

With the development of watermarking technologies, the attacks against watermarking systems have become more sophisticated. In general, the attacks on watermarking systems can be categorized into noise-like signal processing and geometric distortions. While the noise-like signal processing, such as lossy compression, denoising, noise addition, low-pass filtering, reduces watermark energy, geometric distortions induce synchronization errors between the original and the embedded watermark patterns and therefore can mislead the watermark detector. Most methods in the past have addressed the robustness problem against noise-like signal processing attacks and only a few specialized watermarking methods have been proposed to address the geometric distortions. Those few can be classified into *non-blind scheme*, *invariant transform*, *embedding-based synchronization* and *content-based synchronization*.

- *Non-blind scheme*: Non-blind schemes use the original image to synchronize watermark. In Refs. [5,6], the meshes of the original image and the watermarked image are compared to recover synchronization errors. Non-blind scheme is effective for local distortions, but an extension to global affine transformations is computationally demanding [7].
- *Invariant transform*: The most obvious way to achieve resilience on geometric distortions is to use invariant transform. In Refs. [8,9], a watermark is embedded in affine-invariant domain such as the Fourier–Mellin

* Corresponding author. Tel.: +82-42-869-5470; fax: +82-42-869-8590.
E-mail addresses: jsseo@kaist.ac.kr (J.S. Seo), cdyoo@ee.kaist.ac.kr (C.D. Yoo).

0031-3203/$30.00
doi:10.1016/j.patcog.2003.12.013

transform. However, those techniques involving invariant domain suffer from implementation issues (transformation errors in Fourier–Mellin transform) and are vulnerable to cropping [9].

- *Embedding-based synchronization*: Another way to cope with geometric distortions is to use template [10,11] or periodic insertion [12] of watermark pattern. In Ref. [10], the template is embedded in DFT domain as local peaks in predefined positions. The embedded local peaks are searched during watermark detection in order to yield information about the affine transformations undergone by the image [11]. In Ref. [12], it is shown that the periodic insertion of watermark pattern can give synchronization information. The periodically embedded watermark pattern appears as a lattice of peaks in the autocorrelation domain and careful analysis of both the orientation of the lattice and the distance between the peaks can reveal the vital information about the affine transformations. However, this kind of approaches can be tampered by the malicious attack [13] since anyone without security information can access to the peaks in DFT or autocorrelation domain easily to eliminate them.
- *Content-based synchronization*: By binding the watermark synchronization with the image characteristics, watermark detection can be done without synchronization error. In Refs. [14,15], moment-based normalization approaches are proposed. In spite of the robustness against affine transformations, they are highly vulnerable against cropping attacks. In Ref. [16], region-based watermarking method is proposed after segmenting an image into regions. Two largest regions are watermarked with the affine normalized watermark. In Ref. [7], the tessellation of the feature points is used as a reference for watermarking to retain synchronization after geometric distortions. The method showed high robustness to most of geometric attacks. However, its robustness against cropping is not guaranteed, since the tessellation may be changed significantly after cropping. The weak point of the content-based approaches is the added computational complexity in calculating moments or features of the image before watermark detection.

The proposed method can be classified as content-based synchronization method. Watermark detection is analogous to pattern recognition in many respects. Their goals are to find a predefined pattern from an image. To effectively address the watermark detection problem, we turn our interest to the pattern recognition techniques. More specifically, the feature points of an image are used for referencing the watermark embedding and detection. The feature points are sufficiently invariant against various image processing steps to be considered as one of the genuine characteristics of an image. There have been approaches in the past that use feature points of an image as reference points of watermark [7,17]. However, these do not explicitly consider the scaling effect of an image and show limited robustness against scaling. In Ref. [17], scaling issue was not addressed. Feature points in one scale can be considerably different from those in another scale. This fact was not considered in Ref. [7]. By using scale-space theory in extracting feature points of an image, we will show that the watermark can be adapted to the scale changes. The details of scale space and scale invariant feature points are described in Section 2. Moreover, with the obtained synchronization the circularly symmetric spatial watermarking is applied at each selected feature point to improve robustness against rotation, cropping and local geometric attacks (random bending attack [18]). By using feature points of an image, the proposed method achieves both locality and affine resilience.

This paper is organized as follows. Section 2 describes the scale-space concepts and feature point detection. Section 3 describes the proposed watermarking method. Section 4 evaluates the performance of the proposed method. Section 5 summarizes the performance and the limitations of the proposed method.

2. Scale-space representation and feature detection

In content-based method, image characteristics appropriate for watermarking should be carefully selected. Feature points can be a suitable candidate for content-based watermarking since there exist feature point detection methods that are proven to be robust against many image processing steps such as sharpening, blurring, compression and geometric transformations [19–21]. For this reason, we have chosen feature points for watermark synchronization.

2.1. Scale space and automatic scale selection

Detection of feature points, which are invariant to geometric transformations, has been one of the main issues in pattern recognition and computer vision. This leads to the study of scale space and automatic scale selection of an image. The scale-space representation is a set of images represented at different levels of resolutions [19]. More precisely linear scale-space representation $L:\mathbb{R}^2\times\mathbb{R}_+\rightarrow\mathbb{R}$ of an image $f:\mathbb{R}^2\rightarrow\mathbb{R}$ is defined as the solution to the diffusion equation

$$\partial_s L(\mathbf{x};s)=\tfrac{1}{2}\nabla^{\mathrm{T}}\nabla L(\mathbf{x};s)=\tfrac{1}{2}\operatorname{div}(\nabla L(\mathbf{x};s)) \tag{1}$$

with the initial condition $L(\mathbf{x};0)=f(\mathbf{x})$, where $\mathbf{x}=(x,y)$ refers to the image spatial coordinate and $\partial_s L(\mathbf{x};s)=(\partial/\partial_s)L(\mathbf{x};s)$. As a solution of the above equation, given a scale s, the *uniform Gaussian scale-space representation* [22] is defined by

$$L(\mathbf{x};s)=g(\mathbf{x};s)*f(\mathbf{x}), \tag{2}$$

where $g(\mathbf{x};s)$ is the associated uniform Gaussian kernel with standard deviation s and mean zero. A well-known property of the scale-space representation is that the amplitude

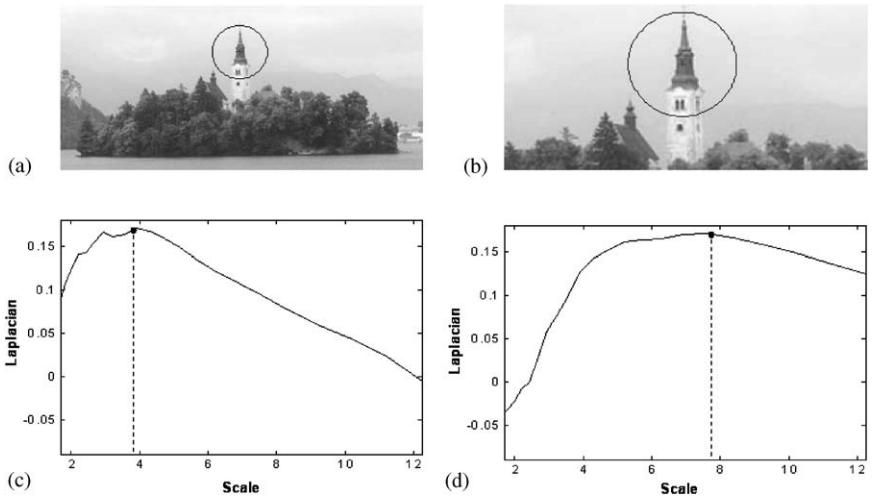

Fig. 1. (a) Original image, (b) 50% cropped and 200% scaled image, (c) scale-normalized Laplacian of the image (a) around the feature point at the center of disk, (d) scale-normalized Laplacian of the image (b) around the feature point at the center of disk; the radius of the circle is 10σ where σ is the characteristic scale at that point.

of spatial derivatives in general *decreases with scale* [19], i.e. if a signal is subjected to scale-space smoothing, then the numerical values of spatial derivatives computed from the smoothed data can be expected to decrease. Thus to obtain *characteristic scale* that reflects local characteristics of an image, normalized derivatives are introduced [19]. The scale-normalized derivative ∂^N is defined by

$$\partial^N_{x^{\alpha_1},y^{\alpha_2}} = s^{\alpha_1+\alpha_2}\partial_{x^{\alpha_1}}\partial_{y^{\alpha_2}}, \tag{3}$$

where α_1 and α_2 are the order of differentiation. The scale level, at which a combination of normalized derivatives assumes a local maximum over scales, can be treated as reflecting a characteristic length of corresponding structure. To give a formal characterization of this scaling property, consider two images f and $\tilde{f}$ related by $f(x) = \tilde{f}(tx)$. The scale-space representations L and $\tilde{L}$ of f and $\tilde{f}$ are defined by Eq. (2), respectively. Then the scale-space representations is related by

$$L(\mathbf{x}; s) = \tilde{L}(\tilde{\mathbf{x}}; \tilde{s}), \tag{4}$$

where $\tilde{\mathbf{x}} = t\mathbf{x}$ and $\tilde{s} = ts$. Differentiation of Eq. (4) by x and y with order α_1 and α_2 gives the following relationship:

$$\partial_{x^{\alpha_1}}\partial_{y^{\alpha_2}}L(\mathbf{x}; s) = t^{\alpha_1+\alpha_2}\partial_{\tilde{x}^{\alpha_1}}\partial_{\tilde{y}^{\alpha_2}}\tilde{L}(\tilde{\mathbf{x}}; \tilde{s}). \tag{5}$$

By multiplying $s^{\alpha_1+\alpha_2}$ to Eq. (5), we can obtain the *scale invariance* from Eq. (3) as follows:

$$\partial^N_{x^{\alpha_1},y^{\alpha_2}}L(\mathbf{x}; s) = \partial^N_{\tilde{x}^{\alpha_1},\tilde{y}^{\alpha_2}}\tilde{L}(\tilde{\mathbf{x}}; \tilde{s}). \tag{6}$$

The scale-invariance property of normalized derivatives ensures the invariance of scale-space maxima [19]. The normalized scale-space maximum is called *characteristic scale* of the image. If a normalized scale-space maximum is at $(\mathbf{x_0}; s_0)$ in the scale-space representation of an image f, then in the scaled image $\tilde{f}$, a corresponding scale-space maximum is assumed at $(t\mathbf{x_0}; ts_0)$ in the scale-space representation of $\tilde{f}$. Fig. 1 shows an example of scale selection. For the same point in the original and the cropped and scaled image (scale factor is 2.0), we compute the amplitude of normalized derivatives (Laplacian) [23] over scales. The figure shows that the characteristic scale is relatively invariant to scaling. The ratio of the scales at corresponding points in the two images, at which the maxima were found, is equal to the scale factor between the two images.

2.2. Detection of scale invariant feature points

Any suitable scale-space feature point detector can be used for the proposed watermarking method. In this paper, we use the *Harris–Laplacian* method in Ref. [20]. Scale invariant feature points are detected based on the scale selection at Harris corner points. The Harris–Laplacian method first builds a scale-space representation for the Harris corner strength measure. The scale-normalized second-moment matrix $U(\mathbf{x}, s)$ is given as follows:

$$U(\mathbf{x}, s) = s^2 g(\mathbf{x}; s) * \begin{bmatrix} L_x^2(\mathbf{x}; t) & L_xL_y(\mathbf{x}; t) \\ L_xL_y(\mathbf{x}; t) & L_y^2(\mathbf{x}; t) \end{bmatrix}, \tag{7}$$

where $L_x = \partial_x L(\mathbf{x}; t)$ and t is set to $s/2$ in this paper. The Harris corner strength measure $H(\mathbf{x}, s_n)$ at scale s_n and point $\mathbf{x}$ is given by a second-moment matrix $U(\mathbf{x}, s_n)$;

$$H(\mathbf{x}, s_n) = \det U(\mathbf{x}, s_n) - 0.04 * (\operatorname{trace} U(\mathbf{x}, s_n))^2. \tag{8}$$

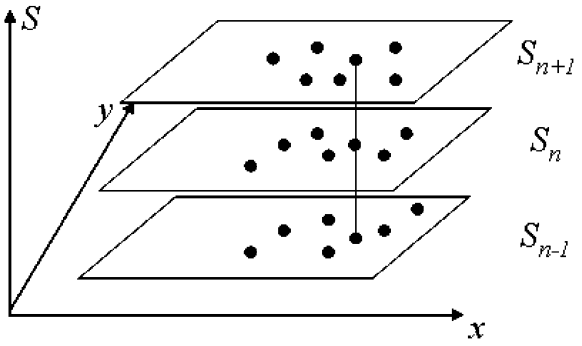

Fig. 2. Searching for maxima in scale space [20].

At each level of the scale space, corner points are detected as the local maxima in the image plane

$$H(\mathbf{x}, s_n) > H(\mathbf{x}_\mathbf{w}, s_n) \quad \forall\, \mathbf{x}_\mathbf{w} \in W,$$
$$H(\mathbf{x}, s_n) > t_h, \tag{9}$$

where W and t_h denotes the neighborhood of the point $\mathbf{x}$ and the detection threshold, respectively. Among the corner points we select the points at which a local Laplacian operator is maximal over scales as shown in Fig. 2. The scale-normalized Laplacian of an image $J(\mathbf{x}, s_n)$ at scale s_n and point $\mathbf{x}$ is given by

$$J(\mathbf{x}, s_n) = |s_n^2(L_{xx}(\mathbf{x}, s_n) + L_{yy}(\mathbf{x}, s_n))|. \tag{10}$$

For the candidate points (Harris maximum), we checked whether its Laplacian is a maximum at that scale or not by the following:

$$(J(\mathbf{x}, s_n) > J(\mathbf{x}, s_{n-1})) \cap (J(\mathbf{x}, s_n) > J(\mathbf{x}, s_{n+1}))$$
$$J(\mathbf{x}, s_n) > t_l, \tag{11}$$

where t_l denotes the detection threshold. Finally the scale invariant feature points are determined by satisfying both Eqs. (9) and (11). By using two measures, the feature points are characteristic to both the image plane and the scale dimension. Moreover, the characteristic scale information is automatically obtained during feature point extraction. In the comparative tests [20,21], the Harris corner points proved to be the most reliable under various image processing steps, and the Laplacian was determined to have the best repeatability of scale selection under large-scale changes. The scale invariant feature points selected by this method exhibit invariance to scale, rotation and translation as well as robustness to illumination changes [20].

3. Proposed watermarking scheme

Content-based watermarking schemes utilize the fact that the features of content can afford a reference invariant to geometric distortions. This idea has also been exploited for invariant pattern recognition and image retrieval [20,24–27]. In this paper, we apply the same idea to watermark synchronization. Watermark is shaped adaptively using the characteristic scale of the feature point. By using local characteristics of an image, the proposed method achieves both locality and affine resilience.

3.1. Watermark embedding

The proposed watermark embedding method is shown in Fig. 3. We extract feature points of the scale-space representation of an image appropriate for watermarking. In Eq. (8), the determinant and the trace of the scale-normalized second-moment matrix $U(\mathbf{x}, s)$ (as in Refs. [19,24]) are used in calculating the scale-normalized corner strength measure $H(\mathbf{x}, s)$. This means that corner strengths can be compared across different scales and the n strongest corners over all detected scales can be determined (typically $n=100$) [24]. Due to repeatability reason, we will only consider the n strongest corners for watermarking afterwards. Fig. 5(a) and (b) show the detected feature points and the n strongest feature points, respectively. Since the watermarked image can be scaled up or down, we first consider the middle scale bands in Fig. 4. Fig. 5(c) shows the feature points with middle scale. As shown in Fig. 5(c), we assume that each feature point represents a disk with radius $R=8\sigma_s$ where σ_s is the characteristic scale that is selected by Laplacian measure in Eq. (11). If two disks are overlapping, the scale-normalized corner strength measures of the two disks are compared, and the one with the strongest corner strength measure survives. By repeating the same process for all the feature points with middle scale, we can select feature points that has the strongest corner strength measure in their neighborhood. While retaining the selected feature points in the middle scale bands, the same process is performed for the feature points in the high- and the low-scale bands, respectively. In other words, if the disks of feature points with the low and high scales are overlapping with the disks of the selected feature points in the middle-scale band, the feature points are erased. Otherwise, the same selection process that is used in the feature points with the middle scale is applied to the feature points with the low and the high scales, respectively. Fig. 5(d) shows the feature points that survives after all. The watermark is embedded inside the disks in Fig. 5(d).

In spatial domain, watermark is embedded in circularly symmetric way centered at each selected feature point. The circularly symmetric watermarking is helpful in achieving resilience against rotation easily. We note that Solachidis and Pitas proposed a similar method in DFT domain [28]. The binary pseudo-random M by M pattern $O(m_1, m_2)$ is prepared as a watermark where $m_1, m_2 = 0, 1, \ldots, M-1$. Regarding a feature point as the center of a disk, the disk is separated in M homocentric circles of radius r_i ($i = 0, 1, \ldots, M-1, M$) and in M sectors as shown in Fig. 6. The radius r_i is set to make the area of each sector the same.

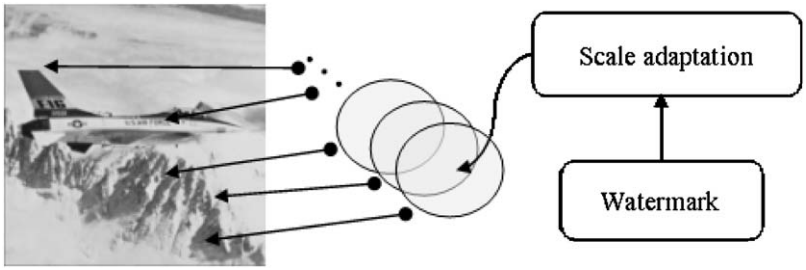

Fig. 3. Watermark embedding based on feature points.

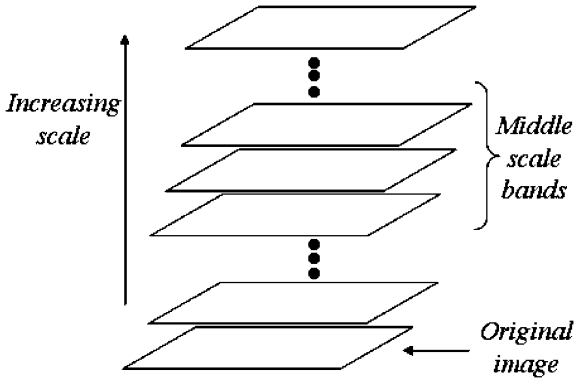

Fig. 4. Middle-scale bands of scale-space representation of an image.

For example, when we set $r_0 = R/4$ and $r_M = R$, r_i is given as $r_i = R\sqrt{1/16 + 15/16 * i/M}$. In each sector the value of the watermark is the same (1 or -1). For each disk, the embedded watermark $W(x, y)$ is obtained from the original watermark $O(m_1, m_2)$ as follows:

$$W(x, y) = O(m_1, m_2) \quad \text{if } (x, y) \in S_{m_1,\, m_2}, \tag{12}$$

where $S_{m_1,\, m_2} = \{(x, y), (x, -y), (-x, y), (y, x) | r_{m_1} \leqslant r < r_{m_1+1}$ and $\lfloor \theta M/0.5\pi \rfloor = m_2$ for $x, y \geqslant 0$, $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x)\}$, and $\lfloor u \rfloor$ is the biggest integer that is smaller than u. Thus the same watermark is embedded in the symmetric sectors as shaded in Fig. 6; i.e. $W(x, y) = W(x, -y) = W(-x, y) = W(y, x)$. This symmetry is conducive in making the watermark more robust. The watermark pattern $W(x, y)$ is embedded additively into the image centered at the feature point $\mathbf{x_p}$ as follows:

$$f'(\mathbf{x}) = f(\mathbf{x}) + \alpha(\mathbf{x}) W(\mathbf{x} - \mathbf{x_p}), \tag{13}$$

where $\alpha(\mathbf{x})$ is the local masking function calculated from the human visual system (HVS) [29] to make the embedded watermark imperceptible.

The watermark embedding procedure is summarized as follows:

1. Extract feature points of scale-space representation as shown in Fig. 5(a).
2. Select n feature points with the strongest scale-normalized corner strength measures as shown in Fig. 5(b).
3. Considering the spatial positions, corner strength measures and characteristic scales of the n feature points, final feature points for watermarking are selected as shown in Fig. 5(d).
4. The circularly symmetric watermark is embedded after adapting to the characteristic scale at each feature point.

3.2. Watermark detection

The proposed watermark detection method is shown in Fig. 7. We assume that the watermark exists at the feature points with the strong scale-normalized corner strength measure as the watermark was embedded. The first and second steps are the same to the embedding. As in the pattern recognition process using feature points of scale-space representation of an image [20,24,25], watermark detection is performed at each detected feature point. In our method, the watermark detection is performed at n selected feature points with the strongest scale-normalized corner strength measures. The detection mask D of each feature point is obtained in each sector as follows

$$D(m_1, m_2) = \begin{cases} \dfrac{\sum_{(x,y) \in S_{m_1,\, m_2}} I(x, y)}{\mathscr{CARD}(S_{m_1,\, m_2})} & \text{if } \mathscr{CARD}(S_{m_1,\, m_2}) \neq 0, \\ 0 & \text{if } \mathscr{CARD}(S_{m_1,\, m_2}) = 0, \end{cases} \tag{14}$$

where $\mathscr{CARD}(S_{m_1,\, m_2})$ is the number of elements in the set $S_{m_1,\, m_2}$. Watermark detection is based on correlation between the detection mask D and the original watermark pattern O. To achieve rotational invariance, we compute the cyclic convolution C between D and O along m_2 as follows:

$$C(k) = \frac{1}{M^2} \sum_{m_1, m_2} D(m_1, m_2) O(m_1, m_2 - k), \tag{15}$$

Fig. 5. (a) Detected feature points (b) 100 strongest feature points (c) feature points in the middle scale among the 100 strongest feature points (d) finally selected feature points; the radius of each circle is 8σ where σ is the characteristic scale at that point.

where $k = 0, 1, \ldots, M-1$. Using the cyclic convolution C, the watermark detection problem could be formulated as the following hypothesis testing:

- H_0: An image I' is watermarked by O if $\max(C) \geqslant T$.
- H_1: An image I' is not watermarked by O if $\max(C) < T$,

where T is the watermark detection threshold and $\max(C)$ is the maximum value of vector C. We will discuss how to determine threshold T in Section 4.1. It is well known that the detection performance can be improved by preceding correlation by matched filtering that decorrelates (or whitens) the noise (image). The symmetrical phase only filtering (SPOMF) is known to be effective as a matched filter for pattern recognition problems [30] and watermarking [31]. By adopting SPOMF, the correlation vector C is given as follows:

$$C = \mathrm{IFFT} \times \left(\frac{1}{M} \sum_{m_1} \times [\mathrm{phase}(\mathrm{FFT}_{m_2}(D)) \times \mathrm{phase}(\mathrm{FFT}_{m_2}(O))^*] \right), \quad (16)$$

where $\mathrm{phase}(z) = z/|z|$ for complex signal z.

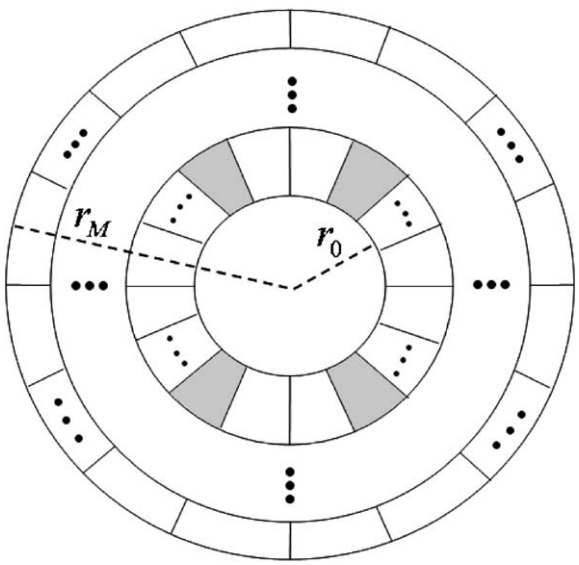

Fig. 6. Sector-based watermarking; the center of the circle is feature point.

The watermark detection procedure is summarized as follows:

1. The first two steps are the same to the watermark embedding.
2. SPOMF correlation detector is applied to n disks centered on the selected feature points. Fig. 8 shows the detector output.
3. According to the detection results in Fig. 8, we finally make a decision whether the image is watermarked or not by testing hypothesis with threshold T.

4. Experimental results

In the experiments, we have used scale-space representation with 39 scale levels. The scale-space representation starts with a initial scale of 2.1 and the scale factor between two levels of resolution is 1.05. Since watermark pattern is pseudo-randomly generated, watermark correlation detector is highly sensitive to the synchronization errors. The circularly symmetric watermark showed vulnerability to geometric distortions more than ± 0.3 pixel translation and $\pm 1.5\%$ scaling. However, the feature point detector can not have such a high accuracy since we deal with sampled spatial (digital images) and scale spaces. Most of the detected feature points typically show deviation up to the amount of ± 1.5 pixel of the location and $\pm 2.5\%$ of the characteristic scale after affine transformations. To resolve this mismatch, we need local search. In the experiment, we search ± 0.5 and ± 1 pixel around the detected feature points (± 0.25 pixel accuracy) and $\pm 2.5\%$ precision of scaling ($\pm 1.25\%$ scaling accuracy). Then local search points amount to 75 (five in x and y direction and three in scale) for each feature point.

4.1. False alarm analysis

For the selection of watermark detection threshold T, the false alarm rate P_{FA} and the false rejection rate P_{FR} should be considered. The false alarm rate P_{FA} is the probability to declare an unmarked image as *marked*. The false rejection rate P_{FR} is the probability to declare a marked image as *unmarked*. There is a tradeoff between the two probabilities in selecting threshold T. In practice P_{FR} is difficult to analyze since there are plenty of image processing steps of those we do not know the exact characteristics. Thus it is common to select a threshold T of minimizing P_{FR} subject to a fixed P_{FA}. First we examine the false alarm rate of each disk. By assuming that the watermark has zero mean and unit variance and is independent with the image, the mean (E) and standard deviation (Std) of correlation $C(k)$ is given as follows:

$$E[C(k)] = 0,$$

$$\mathrm{Std}[C(k)] = \frac{\sigma_D}{M} \quad \text{for } k = 0, 1, \ldots, M-1, \tag{17}$$

where σ_D is the standard deviation of the detection mask D. By assuming that $\sigma_D = 1$ from the SPOMF detection and the detection mask D has Gaussian distribution, the false alarm rate of the correlation output $P_{FA\text{-}cor}$ is given as follows for a certain value of the threshold T:

$$P_{FA\text{-}cor} = \int_T^\infty \frac{M}{\sqrt{2\pi}} \times \exp\left[\frac{-x^2M^2}{2}\right] \mathrm{d}x = \frac{1}{2}\,\mathrm{erfc}\left(\frac{MT}{\sqrt{2}}\right). \tag{18}$$

Since the maximum of $C(k)$ is used in the hypothesis testing, the false alarm rate of the SPOMF convolution output $P_{FA\text{-}conv}$ is given as follows:

$$P_{FA\text{-}conv} = \sum_{i=1}^{M} \binom{M}{i} P_{FA\text{-}cor}^{i} (1 - P_{FA\text{-}cor})^{(M-i)} \tag{19}$$

$$= 1 - (1 - P_{FA\text{-}cor})^{M}. \tag{20}$$

Moreover, local search around the feature points is performed K times (in our case, $K = 75$) to compensate for the inaccuracy of the feature point. The final false alarm rate of the disk $P_{FA\text{-}disk}$ is given as follows:

$$P_{FA\text{-}disk} = \sum_{i=1}^{K} \binom{K}{i} P_{FA\text{-}conv}^{i} (1 - P_{FA\text{-}conv})^{(K-i)} \tag{21}$$

$$= 1 - (1 - P_{FA\text{-}conv})^{K}. \tag{22}$$

By viewing all the watermarked disks as independent communication channels, we claim the existence of watermark if the same watermark is detected from at least μ number of disks [17]. Then the false alarm probability for an image $P_{FA\text{-}image}$ is given as follows:

$$P_{FA\text{-}image} = \sum_{i=\mu}^{N} \binom{N}{i} P_{FA\text{-}disk}^{i} (1 - P_{FA\text{-}disk})^{(N-i)}, \tag{23}$$

where N is the number of disks in an image that are possible for watermarking. In our case, we try to detect watermark for 100 strongest feature points. Thus we set $N = 100$. Fig. 9 shows $P_{FA\text{-}image}$ versus $P_{FA\text{-}cor}$ for $\mu = 1, 2, 3$. As shown in Fig. 5(c) many of the feature points coincide with one another. If we detect watermarks at two points closer than their scales, that should be counted one. Thus we can expect that the actual $P_{FA\text{-}image}$ is lower than that in Fig. 9. Table 1 shows the watermark detection threshold with the settings $M = 32$, $K = 75$ and $N = 100$. All the tests in Section 4.2 are done with the threshold at $P_{FA\text{-}image} = 10^{-4}$.

4.2. Robustness test

The proposed watermarking scheme is tested on the four popular test images: 512×512 Airplane, Lena, Baboon and Peppers. Original watermark pattern O is prepared as a zero-mean pseudo-random 32×32 pattern. To test the robustness of the proposed method, all the watermarked images were subjected to various kinds of image processing steps (see Ref. [18] for a detailed description of the

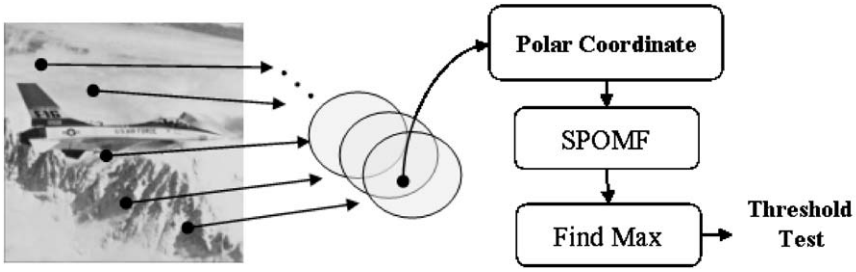

Fig. 7. Watermark detection based on feature points.

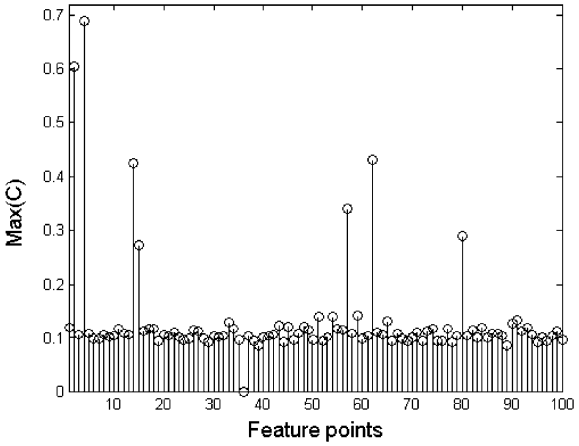

Fig. 8. Maximum of SPOMF correlation C for n selected feature points ($n = 100$ in this case).

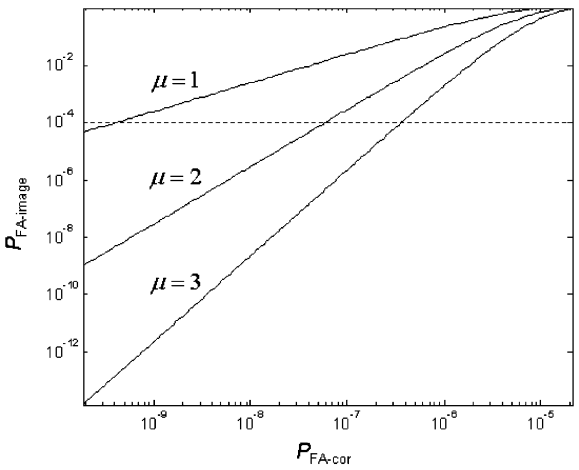

Fig. 9. $P_{FA\text{-}image}$ versus $P_{FA\text{-}cor}$ (in log scale).

processing steps). The PSNR of watermarked images was about 45 dB after masking. For the experiments, we choose $P_{FA\text{-}image} = 10^{-4}$. Watermark detection results for common signal processing steps and geometric distortions are shown in Tables 2 and 3, respectively. Three numbers in the tables indicate the number of detected disks for $\mu = 1, 2, 3$, respectively. The results denote that the proposed method is robust against affine transformations, which preserve aspect ratio, local geometric attacks (such as random bending attack), and other image processing steps including compression and various filtering. In fact the watermark detection under geometric attacks strongly depends on the location of watermarked disks. If a disk is located in the border of an image, it might be removed after geometric attacks. Moreover, the detection performance depends on the characteristics of the region around the feature point. The watermarked disks in the textured area were less robust since many fake feature points may show up, and the watermarked feature points may shift considerably (more than 1 pixel) after geometric attacks [7,17]. The feature points in the intersection of strong edges showed best repeatability. These facts are reflected in Tables 2 and 3, and the textured image Baboon showed less robustness than the other images.

Table 1
Thresholds for correlation detector for $P_{FA\text{-}image}$ with settings $M = 32$, $K = 75$ and $N = 100$

$P_{FA\text{-}image}$	$\mu = 1$	$\mu = 2$	$\mu = 3$
10^{-3}	0.1801	0.1586	0.1498
10^{-4}	0.1918	0.1656	0.1547
10^{-5}	0.2029	0.1720	0.1593
10^{-6}	0.2096	0.1782	0.1639

5. Conclusion

This paper shows that watermark detection resilience against geometric distortions can be significantly improved by using the synchronization based on feature points. The robustness against geometric distortions is essential because it is quite easy to impose geometric distortions to images with modern computers. By using the scale invariant feature points, the search space for the affine transformations is much reduced. Translation, scaling and rotation invariance is obtained by the feature points, automatic scale selection

Table 2
Watermark detection results under non-geometric signal processing steps

Processing	Airplane	Lena	Baboon	Peppers
Gaussian filter	3, 4, 4	3, 4, 5	1, 1, 1	5, 5, 5
Sharpening filter	1, 1, 1	1, 1, 2	0, 0, 0	5, 5, 5
Median filter (4×4)	4, 5, 5	5, 6, 6	1, 2, 2	4, 4, 4
Additive uniform noise (scale=0.1)	5, 7, 7	5, 5, 5	3, 4, 4	6, 6, 6
Additive uniform noise (scale=0.2)	2, 2, 2	3, 4, 5	1, 2, 2	5, 5, 6
Histogram equalization	1, 1, 1	4, 4, 4	2, 2, 2	2, 2, 2
JPEG ($Q = 70\%$)	3, 4, 5	3, 3, 4	1, 2, 4	6, 7, 7
JPEG ($Q = 60\%$)	3, 3, 3	3, 4, 4	1, 2, 2	6, 6, 7
JPEG ($Q = 50\%$)	2, 3, 3	1, 2, 3	1, 2, 3	4, 6, 6
JPEG ($Q = 40\%$)	1, 2, 2	1, 2, 2	1, 1, 2	4, 5, 5
JPEG ($Q = 30\%$)	1, 1, 1	0, 0, 0	0, 1, 1	4, 4, 4
Gaussian filter + JPEG ($Q = 90\%$)	3, 3, 3	2, 4, 4	0, 1, 1	5, 5, 5
Sharpening filter + JPEG ($Q = 90\%$)	1, 1, 1	1, 1, 1	0, 0, 0	5, 5, 5
Median filter + JPEG ($Q = 90\%$)	2, 4, 4	5, 5, 5	0, 1, 2	2, 4, 4
FMLR	0, 0, 0	1, 3, 4	4, 6, 6	0, 1, 3

Three numbers in the tables indicate the number of detected disks for $\mu = 1, 2, 3$, respectively. The number of watermarked disks are 7 for Airplane, Lena and Baboon and 8 for Peppers.

Table 3
Watermark detection results under geometric signal processing steps

Processing	Airplane	Lena	Baboon	Peppers
17 column				
5 row removed	3, 3, 3	5, 6, 6	1, 2, 2	5, 5, 5
Cropping 15% off	5, 5, 5	6, 6, 6	4, 4, 4	2, 2, 2
Cropping 25% off	5, 5, 5	4, 4, 4	1, 2, 2	2, 2, 2
Rotation 20° + autocrop	3, 5, 5	5, 5, 5	1, 3, 3	3, 4, 4
Rotation 45° + autocrop	5, 5, 5	2, 2, 3	1, 1, 1	1, 1, 1
Scaling 90%	1, 2, 2	4, 5, 5	2, 3, 4	6, 6, 6
Scaling 75%	3, 3, 3	3, 4, 4	0, 2, 2	6, 6, 6
Scaling 50%	1, 1, 1	2, 2, 2	0, 0, 0	2, 3, 3
Shearing (1%)	4, 4, 4	5, 5, 5	2, 3, 4	4, 4, 4
Shearing (5%)	0, 0, 1	1, 1, 1	0, 0, 0	0, 1, 1
Linear geometric transform $(1.007, 0.010, 0.010, 1.012)$	3, 4, 4	6, 6, 6	3, 3, 3	5, 5, 5
Linear geometric transform $(1.010, 0.013, 0.009, 1.011)$	4, 4, 4	7, 7, 7	1, 5, 5	7, 7, 7
Linear geometric transform $(1.013, 0.008, 0.011, 1.008)$	5, 5, 5	7, 7, 7	0, 3, 3	5, 5, 5
Random bending attack	2, 2, 2	4, 5, 5	0, 2, 3	3, 3, 3

Three numbers in the tables indicate the number of detected disks for $\mu = 1, 2, 3$, respectively. The number of watermarked disks are 7 for Airplane, Lena and Baboon and 8 for Peppers. Note that autocrop refers to cropping to original size after rotation.

and SPOMF detection. The original image is not needed in the watermark detector. Two limitations are commonly quoted on the content-based watermarking methods: the inaccuracy of feature point detector [7,17] and the added computational complexity [16]. In this paper, the inaccuracy was overcome by using the local search. Obviously, the proposed method is computationally more demanding than watermarking methods that do not handle geometric distortions. However, it is certainly less burdensome than exhaustive search. The experimental results show that the robustness of content-based watermarking, especially to scaling and cropping, is improved by using scale-space feature points. This implies that there may be room for improvement of watermarking performance by using refined pattern recognition techniques. Future work includes the development of more robust feature point detector, especially on textured areas, and watermarking method that is adaptable to nonisometric geometric distortions.

Acknowledgements

This work was supported by grant No. R01-2003-000-10829-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

References

[1] F. Hartung, M. Kutter, Multimedia watermarking technique, Proc. IEEE 87 (7) (1999) 1079–1107.
[2] D. Boneh, J. Shaw, Collusion-secure fingerprinting for digital media, IEEE Trans. Inf. Theory 44 (5) (1998) 1897–1905.
[3] M. Schneider, S.-F. Chang, A robust content based digital signature for image authentication, Proceedings of IEEE International Conference on Image Processing, Lausanne, Switzerland, 1996
[4] J.A. Bloom, I.J. Cox, T. Kalker, J.-P.G. Linnartz, M.L. Miller, C.B.S. Traw, Copy protection for DVD video, Proc. IEEE 87 (7) (1999) 1267–1276.
[5] P. Dong, J.G. Brankov, N. Galatsanos, Y. Yang, Geometric robust watermarking based on a new mesh model correction approach, Proceedings of the IEEE International Conference on Image Processing, 2002.
[6] F. Davoine, Triangular meshes: a solution to resist to geometric distortions based watermark-removal softwares, Proceedings of European Signal Processing Conference, 2000.
[7] P. Bas, J.-M. Chassery, B. Macq, Image watermarking: an evolution to content based approaches, Pattern Recognition 35 (2002) 545–561.
[8] J. Ruanaidh, T. Pun, Rotation, scale and translation invariant spread spectrum digital image watermarking, Signal Process. 66 (3) (1998) 303–317.
[9] C-Y. Lin, M. Wu, J.A. Bloom, M.L. Miller, I.J. Cox, Y-M. Lui, Rotation, scale, and translation resilient public watermarking for images, IEEE Trans. Image Process. 10 (5) (2001) 767–782.
[10] S. Pereira, T. Pun, Robust template matching for affine resistant image watermarks, IEEE Trans. Image Process. 9 (6) (2000) 1123–1129.
[11] S. Pereira, T. Pun, An iterative template matching algorithm using the Chrip-Z transform for digital image watermarking, Pattern Recognition 33 (2000) 173–175.
[12] M. Kutter, Watermarking resisting to translation, rotation and scaling, Proceedings of the SPIE International Symposium on Voice, Video, and Data Communication, Boston, November 1998.
[13] A. Herrigel, S. Voloshynovskiy, Y. Rytsar, The watermark template attack, Proceedings of SPIE Security and Watermarking of Multimedia Contents III (vol. 4314), San Jose, January 2001.
[14] M. Alghoniemy, A.H. Tewfik, Geometric distortion correction through image normalization, Proceedings of IEEE International Conference on Multimedia and Expo, 2000.
[15] P. Dong, N.P. Galatsanos, Affine transformation resistant watermarking based on image normalizaton, Proceedings of IEEE International Conference on Image Processing, 2002.
[16] A. Nikolaidis, I. Pitas, Region-based image watermarking, IEEE Trans. Image Process. 10 (11) (2001).
[17] C.-W. Tang, H.-M. Hang, A feature-based robust digital image watermarking scheme, IEEE Trans. Signal Process. 51 (4) (2003).
[18] F.A.P. Fetitcolas, Watermarking Schemes Evaluation, IEEE Signal Process. Mag. 17 (5) (2000).
[19] T. Lindeberg, Feature detection with automatic scale selection, Int. J. Comput. Vision 30 (2) (1998) 77–116.
[20] K. Mikolajczyk, C. Schmid, Indexing based on scale invariant interest points, Proceedings of the International Conference on Computer Vision, 2001, pp. 525–531.
[21] C. Schmid, R. Mohr, C. Baukhage, Comparing and evaluating interest points, Proceedings of the International Conference on Computer Vision, 1998, pp. 230–235.
[22] J. Babaud, A.P. Witkin, M. Baudin, R.O. Duda, Uniqueness of the Gaussian kernel for scale-space filtering, IEEE Trans. Pattern Anal. Mach. Intell. 8 (1) (1986).
[23] A.K. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, Englewood Cliffs, NJ, 1989.
[24] A. Baumberg, Reliable feature matching across widely separated views, Proceedings of Computer Vision and Pattern Recognition, 2000, pp. 774–781.
[25] C. Schmid, R. Mohr, Local gray value invariants for image retrieval, IEEE Trans. Pattern Anal. Mach. Intell. 19 (5) (1997).
[26] K. Xin, K.B. Lim, G.S. Hong, A scale-space filtering approach for visual feature extraction, Pattern Recognition 28 (8) (1995) 1145–1158.
[27] M. Kreutz, B. Volpel, H. Janben, Scale-invariant image recognition based on higher-order autocorrelation features, Pattern Recognition 29 (1) (1996) 19–26.
[28] V. Solachidis, I. Pitas, Circularly symmetric watermark embedding in 2-D DFT domain, IEEE Trans. Image Process. 10 (11) (2001).
[29] B. Girod, The information theoretical significance of spatial and temporal masking in video signals, Proceedings of the SPIE/SPSE Conference on Human Vision, Visual Processing and Digital Display, Los Angeles, USA, 1989, pp. 178–187.
[30] L. Brown, An overview of image registration techniques, ACM Comput. Surv. 24 (1992) 325–376.
[31] M. Maes, T. Kalker, J.-P.M.G. Linnartz, J. Talstra, G.F.G. Depovere, J. Haitsma, Digital watermarking for DVD video copy protection, IEEE Signal Process. Mag. 17 (5) (2000).

About the Author—JIN S. SEO received B.S. and M.S. degrees from Korea Advanced Institute of Science and Technology in 1998 and 2000, respectively, all in electrical engineering. Now he is a Ph.D. student at the same institute, where he is a research assistant at the Multimedia processing lab. While working toward Ph.D. degree, he was a thesis trainee at Philips Research Eindhoven in 2002. His main interests are multimedia watermarking, image processing, pattern recognition and adaptive signal processing.

About the Author—CHANG D. YOO received the B.S. degree in Engineering and Applied Science from California Institute of Technology in 1986, the M.S. degree in Electrical Engineering from Cornell University in 1988 and the Ph.D degree in Electrical Engineering from Massachusetts Institute of Technology in 1996. From January 1997 to March 1999 he worked at Korea Telecom as a Senior Researcher. He joined the Department of Electrical Engineering at Korea Advanced Institute of Science and Technology in April 1999. His current research interests include the application of digital signal processing theory in speech and audio, speech recognition, digital communication, and adaptive signal processing. He is a member of Tau Beta Pi and Sigma Xi.