DIGITAL IMAGE WATERMARKING ALGORITHM BASED ON SIFT AND SVR

[1]JIE ZHAO, [2]YUXIA ZHAO

[1]Department of Physics and Electronic Information Engineering, Shangluo University, Shangluo 726000, Shaanxi, China

[2]Department of Computer Science, Shangluo University, Shangluo 726000, Shaanxi, China

ABSTRACT

Digital color image watermarking algorithm based on scale-invariant feature transform(SIFT) and support vector regression(SVR) is proposed in this paper. The input feature vectors are selected in the wavelet domain and then the train model is obtained by applying the support vector regression theory. The watermark information can be embedded or extracted by utilizing the above trained SVR model. The proposed scheme can extract the digital watermark without the help of the original digital image. The scale-invariant feature transform(SIFT) is employed to against the scaling and rotation attacks. To improve the security and robustness, the original watermark is scrambled at first. Experimental results show that the proposed scheme is invisible and robust to common signal processing attacks such as adding noise, JPEG compression, sharpening, smoothing, filtering, contrast enhancement, cropping, rotation, scaling and so on.

Keywords: *Digital Image Watermarking, Scale-invariant Feature Transform, Support Vector Regression*

1. INTRODUCTION

The development of network and multimedia technology makes it convenient to access the multimedia information. These technologies bring people many conveniences but they also bring us some large side effects. The problem of multimedia copyright protection becomes increasingly serious. Digital image watermark technology is commonly achieved by certain modifications to the host signal. According to the embedding domain, the most traditional image watermark technology can be divided into two kinds: Spatial domain and transform domain method. Spatial watermarking method embeds watermark information into original digital image[1-2], while transform domain method performs watermark embedding in transform domain[3-6].

Generally, the spatial domain methods are of the advantage of simplicity and high capacity and the transform domain methods perform more robustness. In order to balance the watermarking imperceptibility and robustness, some scholars use the machine learning methods to optimize the embedding position, capacity or strength. Lou et al. used the neural networks to identify the optimum embedding strength and capacity of different image regions[7]. Davis and Najarian proposed a watermarking scheme in wavelet domain. The neural networks are used to implement an automated system of creating maximum-strength watermarks[8]. Shieh et al. proposed a watermarking scheme based on genetic algorithms(GA) in the transform domain[9]. They employed GA for optimizing the fitness function which includes both factors related to robustness and invisibility. Particle search optimization(PSO) is also used for watermark embedding[10]. Although these methods achieved some results, but have some limitations, such as “over learning” and the lack of a unified mathematical theory. Recently, support vector machine(SVM) has been applied in watermarking system. Fu et al. embedded the template and watermark into the original image in the same way, then the SVM training model is obtained by using the template samples, and the output of SVM model is obtained and the watermark is extracted[11]. They improved the algorithm in [12]. The support vector regression can be trained at the embedding procedure using the information provided by the reference positions. Then the watermark is adaptively embedded into the blue channel of the host image, and can be extracted by virtue of the good learning ability of support vector machine. Tsai et al. proposed a novel watermarking scheme based on SVM for image authentication[13]. It utilizes the set of training patterns to train the SVM and then applies the trained SVM to classify a set of test patterns. Following the results produced by the classifier, this method retrieves the hidden watermark without the original image during watermark extraction. Most of the algorithms do not resist the geometric attacks. A novel method based on support vector regression(SVR) is proposed in [14]. The SVR and Krawtchouk moments are used to correct the geometric attacks. The watermark is embedded by Pseudo-Zernike moments. The drawback of this method is the large calculation and high complexity.

This paper proposed a method using SVR and scale-invariant feature transform(SIFT). By the trained SVR model, watermark information is embedded into the wavelet domain of the carrier image. This method could against scaling and rotation attacks by virtue of the SIFT which is already implemented by hardware[15]. The rest of this paper is as follows: The description of support vector regression is overviewed in Section 2. Section 3 describes the geometric correction by SIFT. Section 4 covers the watermark embedding and detection procedure. The experimental results are showed in Section 5. Section 6 concludes this paper.

2. SUPPORT VECTOR REGRESSION

Support vector machine(SVM) is a universal classification algorithm proposed by Vapnik in the middle of 1990s, which is regarded as a new innovation of learning machine based on the statistical learning theory[11]. The basic theory of SVM can be depicted by a typical two-dimensional case shown in Figure 1, in which ● and ■ denote two categories of samples, H is the separating hyperplane, H_1 and H_2 are parallel to H and no training points fall between them. The optimal separating hyperplane in the case of structural risk minimization is the separating hyperplane which can separate the two categories with the maximum margin. Thus the problem of optimal separating hyperplane can be transformed into a constraint optimization problem.

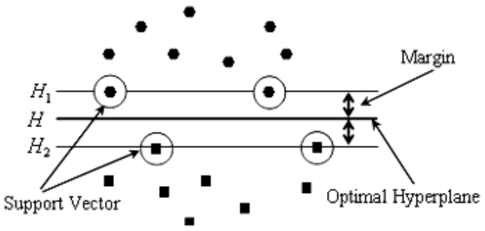


Figure 1: Support Vector Machine

Support vector regression(SVR) is an important application of support vector machine on the regression learning. Generally, for the training sets: $\{(x_1,y_1),(x_2,y_2),...,(x_n,y_n)\}$, where $x \in R^n$, and $y \in R$ to get the relation between the input x_i and output y_i, it can seek an optimal regression function $f(x)$ by SVR training. Then the difference between the output value and the corresponding objective value of every input sample is not more than error ε. For the linear situation, the form of function is: $f(x) = \omega \cdot x + b$, and $\omega \in x$, $b \in R$. In order to get an optimal regression function, it needs a minimum ω, then the above problem can be described as an optimization problem:

$$\min \frac{1}{2}\|\omega\|^2 \text{, s.t.}$$

$$\begin{cases} y_i - \omega \cdot x_i - b \le \varepsilon \\ \omega \cdot x_i + b - y_i \le \varepsilon \end{cases}, i = 1,2,3,...,n \quad (1)$$

Considering the existence of regression error is permitted, so the positive slack variables $\xi_i \ge 0$ and $\xi_i^* \ge 0$ are introduced, and the above formula can be described as follows:

$$\min \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \text{, s.t.}$$

$$\begin{cases} y_i - \omega \cdot x_i - b \le \varepsilon \\ \omega \cdot x_i + b - y_i \le \varepsilon \end{cases}, \ i = 1,2,3,...,n \quad (2)$$

where C is the penalty parameter which controls the trade-off between errors of the SVM on training data and margin maximization.

The optimization problem is converted to the following problem according to this:

$$\max \ \omega(\alpha,\alpha^*) = -\varepsilon\sum_{i=1}^{n}(\alpha_i^* + \alpha_i) + \sum_{i=1}^{n} y_i(\alpha_i^* - \alpha_i)$$
$$-\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(x_i - x_j)$$

$$\text{s.t.} \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0, \ i = 1,2,3,...,n \quad (3)$$

where α_i and α_i^* are Lagrange multiplying factors.

Then the regression function is:

$$f(x) = \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)(x_i - x) + b^* \quad (4)$$

where α_i and α_i^* are few non-zero factors. The corresponding samples are called support vectors. b^* is the parameter which determines the position of the separating hyperplane.

For the nonlinear situation, the data is mapped to a high dimensional feature space by the nonlinear map and then the linear regression could be implemented.

3. GEOMETRIC CORRECTION BY SIFT

The scale-invariant feature transform(SIFT) is a great powerful feature point diction method and proves the invariant to image rotation and scaling and translation[16]. It is used widely and also can correct the geometric distortion. The basic idea is to get features through a series of filtering operation to extract stable point in the image scale space.

(1) The scale space of an image is defined as a function $L(x,y,\sigma)$, that is produced from the convolution of a variable-scale Gaussian function $G(x,y,\sigma)$, with an input image $I(x,y)$:

$$L(x,y,\sigma)=G(x,y,\sigma)*I(x,y) \qquad (5)$$

where $*$ is the convolution operation, and

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2} \qquad (6)$$

$D(x,y,\sigma)$ is used to get stable key point locations in scale space, which can be calculated from the difference of two nearby scales separated by a constant multipartite factor k:

$$D(x,y,\sigma)=L(x,y,k\sigma)-L(x,y,\sigma) \qquad (7)$$

(2) Local extrema detection is implemented by comparing the pixel of difference-of-Gaussian images to its neighbors in the current and adjacent scales. It is selected only if it is larger than all of these neighbors or smaller than all of them.

The next procedure is to perform a detailed fit to the nearby data for location, scale, and ratio of principal curvatures after the key point candidate has been found by comparing a pixel to its neighbors. This information allows points to be rejected that have low contrast. The rejection to key points with low contrast is not enough to stability. The difference of Gaussian function will have a strong response along edges. These unstable responded points will also be rejected.

(3) The main orientation of the key points should be identified. The sample process is implemented by using a region around the key point. The statistical histogram of the gradient direction of neighborhood pixels is obtained. The peak of histogram represents the principal direction of the key point. The gradient magnitude m and orientation θ are calculated as:

$$m=\sqrt{(L(x+1,y)-L(x-1,y))^2+(L(x,y+1)-L(x,y-1))^2} \qquad (8)$$

$$\theta=\tan^{-1}(\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)})) \qquad (9)$$

(4) The SIFT feature vector is generated. If the image is scaled s times, the number of matching points is n, the scale factor of key point P_i in the original image is d_i, and the scale factor of the corresponding matching key point Q_i in the scaled image is q_i,then:

$$s=(\sum_{i=1}^{n} q_i/d_i)/n \qquad (10)$$

If rotation angle is α, the number of matching points is n, the center angle of key point P_i in the original image is ϕ_i, and the center angle of the corresponding matching key point Q_i in the rotated image is φ_i, then:

$$\alpha=(\sum_{i=1}^{n}(\varphi_i-\phi_i))/n \qquad (11)$$

4. PROPOSED METHOD

Traditional approaches, which are based on the principle of experiential risk minimization instead of expected risk minimization, achieve the best, when the number of training samples is infinite. The support vector machine is a new statistical learning method. It can solve small-sample, non-linear and high dimensional problems by using structural risk minimization instead of empirical risk minimization. Statistical learning theory points out that if there are only finite samples, the minimization of experiential risk cannot guarantee the minimization of real risk. The typical situation is the over learning of neural networks. Because supporting vector regression is based on the minimization of structural risk, under the circumstance of finite samples, it can not only guarantee the highest generalization ability of the pattern and the smoothness of the output function, but also posses good abilities of learning and generalization. Therefore, as long as certain points are chosen from the carrier images which contain masses of data, together with the partial information around, the mode of their relation can be well-built, and the watermark information can be embedded and extracted through the model.

Generally, let I represents an RGB color image with the size of $M \times N$, and $I = \{I_r, I_g, I_b\}$ where I_r, I_g, I_b are the red, green and blue components respectively. The watermark image w is a meaningful binary image with the size of $P \times Q$. The watermark is embedded into the blue channel since the human eye is relatively insensitive to the blue component.

4.1 Watermark Embedding

(1) The SIFT feature vectors should be extracted at first, and the binary watermark w is scrambled by secret key and scanned to a sequence $w0$.

(2) The blue component I_b of image I to be watermarked is decomposed through DWT in four levels, and the low-pass subband is denoted as f.

(3) Some adjacent pixels in f are selected randomly. These pixels and their neighborhood pixels in every 3×3 window are used to train the support vector machine.

$$\begin{aligned} S = \{ & f(i-1, j-1), f(i-1, j), f(i-1, j+1), f(i, j-1), f(i, j+1), \\ & f(i+1, j-1), f(i+1, j), f(i+1, j+1), f(i, j)\} \\ = \{ & O(i, j), f(i, j)\} \end{aligned} \quad (12)$$

The pixels set $O(i, j)$ are the feature vectors for support vector regression training. The pixels set $f(i, j)$ are the training objective of support vector regression. The support vector regression machine can be trained by appropriate kernel function and learning coefficients.

(4) For each embedding position (x, y), the eight pixels in 3×3 window are collected to form the dataset $U(x, y)$. The $P \times Q$ embedded positions are adjacent, not overlap and not equal to the pixels in step3. These embedded pixels are also not on the boundary.

$$\begin{aligned} U(x, y) = \{ & f(x-1, y-1), f(x-1, y), f(x-1, y+1), \\ & f(x, y-1), f(x, y+1), f(x+1, y-1), \\ & f(x+1, y), f(x+1, y+1)\} \end{aligned} \quad (13)$$

The dataset $U(x, y)$ is extracted as the input vectors of the trained support vector regression and the output of support vector regression $\delta(x, y)$ is obtained.

(5) The watermark is embedded by modifying the pixel $f(x, y)$ at the embedding position as:

if $w0(t) = 1$ $(t = 1, 2, 3, ..., P \times Q)$ then

$$f^*(x, y) = \max(f(x, y), \delta(x, y) \times (1 + h(x, y))) \quad (14)$$

else

$$f^*(x, y) = \min(f(x, y), \delta(x, y) \times (1 - h(x, y))) \quad (15)$$

where $h(x, y)$ is the embedding strength function and $f^*(x, y)$ is the modified wavelet coefficient at position (x, y).

Given the perceptibility of the human visual system, the strength function $h(x, y)$ hasn't been set as numerical variable. Generally human eyes are not very sensitive to the additional noise in the high luminance region, which means that high luminance background region can embed stronger information. Human eyes are more sensitive to the smooth region than the highly textured area; therefore, the stronger information can be added to the highly textured area of the image. The strength function $h(x, y)$ is given by:

$$h(x, y) = \beta \cdot B(x, y) \cdot T(x, y) \quad (16)$$

where β is the intensity modulation factor, $B(x, y)$ is the weighing function according to the local sensitivity of the image luminance and $B(x, y)$ is given as:

$$B(x, y) = \sum_{l=-1}^{1} \sum_{k=-1}^{1} f(x + l, y + k) / 9 \quad (17)$$

$T(x, y)$ is the function according to the local sensitivity of the image texture and is given as:

$$T(x, y) = \sum_{l=-1}^{1} \sum_{k=-1}^{1} \left| f(x + l, y + k) - B(x, y) \right| \quad (18)$$

4.2 Watermark Detection

(1) If the test image is suffered scaling or rotation attack, the geometric correction by SIFT should be implemented at first. The blue component I'_b of the test image I' to be watermarked is decomposed through DWT in four levels, and the low-pass subband is denoted as f'.

(2) For each detecting position (x, y) which is equal to the embedding procedure, the eight pixels in 3×3 window are collected to form the dataset $U'(x, y)$:

$$\begin{aligned} U'(x, y) = \{ & f'(x-1, y-1), f'(x-1, y), f'(x, y-1), \\ & f'(x-1, y+1), f'(x, y+1), f'(x+1, y-1), \\ & f'(x+1, y), f'(x+1, y+1)\} \end{aligned} \quad (19)$$

(3) The dataset $U'(x,y)$ is extracted as the input vectors of the trained support vector regression and the output of support vector regression $\delta'(x,y)$ is obtained.

(4) The watermark information is extracted as:

if $f'(x,y) > \delta'(x,y)$ then $w0'(t) = 1$

else $w0'(t) = 0$.

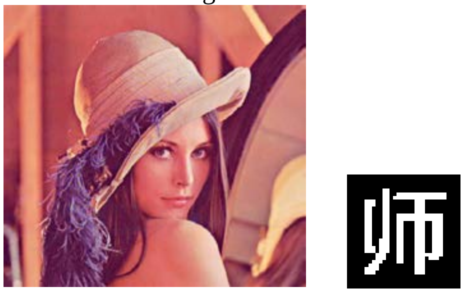
(5) The watermark w' is obtained from $w0'$ by inverse scrambled which used the key which is used in the embedding procedure.

5. EXPERIMENTAL RESULTS

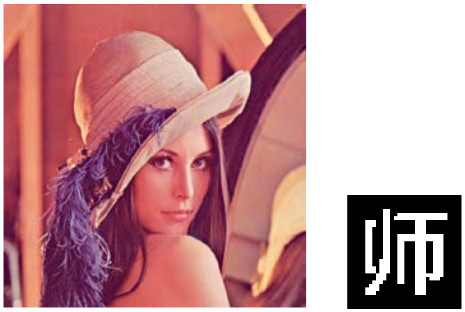
The color image Lena with the size of 512×512 is used as the original carrier image, and the meaningful binary image with the size of 32×32 is used as the original watermark image, which are shown in Figure 2. The watermarked image and the extracted watermark image are also shown in Fig.3. Some parameters of the training support vector regression are briefly determined through experiments. The radius-based function(RBF) is adopted as the kernel function of the support vector regression since it performed better than others[14]. In order to get more appropriate parameters, a lot of experiments are implemented to the training support vector regression. Then the parameters are determined as: $\sigma = 10$, $\varepsilon = 0.008$, $C = 1$. The support vector regression could not get enough training if the learning sample number is too small, and it also leads to over learning if the number is too large. The number is set as 200 by trial experiments. The bit error rate(BER) is used to show the watermark detection results. If the BER is less than the predetermined threshold, the watermark is considered to be existed, otherwise, the watermark does not exist. The BER is defined by:

$$BER = \frac{\sum_{m=1}^{P}\sum_{n=1}^{Q} w(m,n) \oplus w'(m,n)}{P \times Q} \qquad (20)$$

where w is the original watermark and w' is the extracted watermark image.



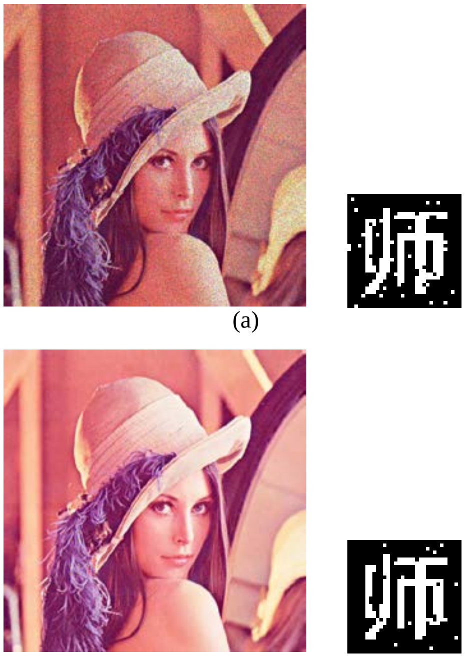
(a)



(b)

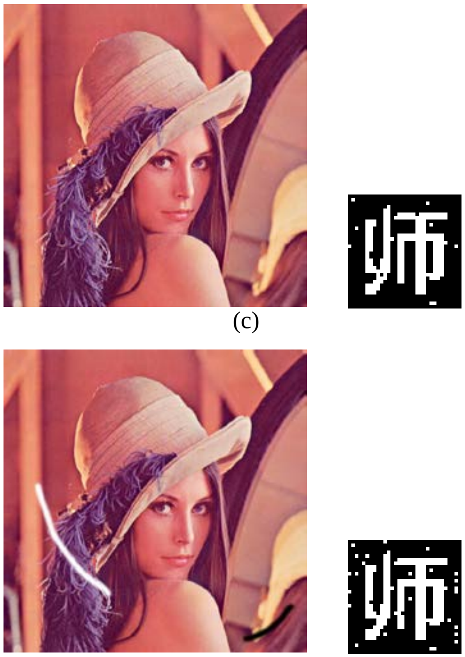
Figure 2: (A) Original Carrier Image And Watermark (B) Watermarked Image And Extracted Watermark

The speckle noise, image enhance, image sharpening and random printing attacks are tested, and are shown in Figure 3. In order to test the robustness of this watermark method, more common signal processing attacks are implemented to the watermarked image, such as add noise, JPEG compression, filtering, cropping, scaling, rotation and so on. Table 1 shows the test results compared with the method in literature[12] and[13]. The simulation results show that the proposed method performs better.



(a)

(b)



(c)

(d)

Figure 3: (a) Speckle noise attack(PSNR=19.39, BER=0.0471) (b) Image enhance attack(PSNR=19.01, BER=0.0209) (c) Image Sharpening attack(PSNR=38.75, BER=0.0166) (d) Random painting attack(PSNR=24.86, BER=0.0322)

Table 1 : The Results Of Common Processing Attacks

Attacks	This method	Method [12]	Method [13]
Gaussian noise(0,0.005)	pass	pass	pass
Salt &Pepper noise(0.04)	pass	pass	pass
JPEG compression (QF=90)	pass	pass	pass
JPEG compression (QF=40)	pass	fail	pass
Average filtering	pass	pass	pass
Median filtering	pass	pass	pass
Low pass filtering	pass	pass	pass
Cropping(25%)	pass	pass	pass
Scaling(200%)	pass	fail	fail
Rotation(120°)	pass	fail	fail
Luminance(+50%)+ Contrast(+50%)	pass	pass	pass

6. CONCLUSION

Image watermarking technology has been developed for decades but there are even many difficulties in the watermark system. The contradiction between the invisibility and robustness is still not easy to solve. In this paper, the watermark embedding strength is related to the image content based on the support vector regression. The SIFT is used to against scaling and rotation distortions. Experimental results show the robustness to common image processing attacks of this proposed algorithm.

ACKNOWLEDGEMENTS

This work was supported by the Science and Technology Foundation Project of Shangluo University(NO. 09SKY032, 09SKY033) and Scientific Research Program Funded by Shaanxi Provincial Education Department(Program No. 12JK0747).

REFERENCES:

[1] G. Voyatzis, I. Pitas, “Chaotic watermarks for embedding in the spatial digital image domain”, *Proceedings of International Conference on Image Processing*, IEEE Conference Publishing Services, October 4-7, 1998, pp. 432-436.

[2] Xianyong Wu, Zhihong Guan, “A novel digital watermark algorithm based on chaotic maps”, *Physics Letters A*, Vol. 365, No. 5, 2007, pp. 403-406.

[3] P. Bas, N. Le Bihan, J. M. Chassery, “Color image watermarking using quaternion Fourier transform”, *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, IEEE Conference Publishing Services, April 6-10, 2003, pp. 521-524.

[4] Haohao Song, Songyu Yu, Xiaokang Yang, Li Song, Chen Wang, “Contourlet-based image adaptive watermarking”, *Signal Processing: Image Communication*, Vol. 23, No. 3, 2008, pp. 162-178.

[5] Dazhi Zhang, Boying Wu, Jiebao Sun, Heyan Huang, “A new robust watermarking algorithm based on DWT”, *International Congress on Image and Signal Processing*, IEEE Conference Publishing Services, October 17-19, 2009, pp. 1-6.

[6] Ting Luo, Mei Yu, Gangyi Jiang, Aihong Wu, Feng Shao, Zongju Peng, “Novel DCT-based blind stereo image watermarking algorithm”, *Future Wireless Networks and Information Systems*, Vol. 143, 2012, pp. 297-304.

[7] DC. Lou, JL. Liu, MC. Hu, "Adaptive digital watermarking using neural network technique", *International Carnahan Conference on Security Technology*, IEEE Conference Publishing Services, October 14-16, 2003, pp. 325-332.

[8] KJ. Davis, K. Najarian, "Maximizing strength of digital watermarks using neural networks", *International Joint Conference on neural networks*, IEEE Conference Publishing Services, July 15-19, 2001, pp. 2893-2898.

[9] CS. Shieh, HC. Huang, FH. Wang, JS Pan, "Genetic watermarking based on transform domain techniques", *Pattern Recognition*, Vol. 37, No. 3, 2004, pp. 555-565.

[10] Xiaoxia Li, Jianjun Wang, "A steganographic method based upon JPEG and particle swarm optimization algorithm", *Information Sciences*, Vol. 177, No. 5, 2007, pp. 3099-3109.

[11] Y. Fu, R. Shen, H. Lu, "Watermarking scheme based on support vector machine for color images", *IEEE Electronics Letters*, Vol. 40, No. 16, 2004, pp. 986-987.

[12] R. Shen, Y. Fu, H. Lu, "A novel image watermarking scheme based on support vector regression", *The Journal of Systems & Software*, Vol. 78, No. 1, 2005, pp. 1-8.

[13] HH. Tsai, DW. Sun, "Color image watermark extraction based on support vector machines", *Information Sciences*, Vol. 177, No. 2, 2007, pp. 550-569.

[14] XY. Wang, ZH. Xu, HY. Yang, "A robust image watermarking algorithm using SVR detection", *Expert Systems with Applications*, Vol. 36, No. 5, 2009, pp. 9056-9064.

[15] Leonardo Chang, José Hernández-Palancar, "A hardware architecture for SIFT candidate keypoints detection", *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, Vol. 5826, 2009, pp. 95-102.

[16] David G. Lowe, "Distinctive image features from scale-invariant key points", *International Journal of Computer Vision*, Vol. 60, No. 2, 2004, pp. 91-110.