

Research Article

Edge Detector Design Based on LS-SVR

Zhongdang Yu¹ and Hamid Reza Karimi²

¹ College of Engineering, Bohai University, Jinzhou 121013, China

² Department of Engineering, Faculty of Engineering and Science, University of Agder, 4898 Grimstad, Norway

Correspondence should be addressed to Zhongdang Yu; yuzhongdang@sina.com

Received 24 December 2013; Revised 12 January 2014; Accepted 13 January 2014; Published 6 March 2014

Academic Editor: Shen Yin

Copyright © 2014 Z. Yu and H. R. Karimi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For locating inaccurate problem of the discrete localization criterion proposed by Demigny, a new criterion expression of “good localization” is proposed. Firstly, a discrete expression of good detection and good localization criterion of two dimension edge detection operator is employed, and then an experiment to measure optimal parameters of two dimension Canny’s edge detection operator is introduced after. Moreover, a detailed performance comparison and analysis of two dimension optimal filter obtained via utilizing tensor product for one dimension optimal filter are provided which can prove that least square support vector regression (LS-SVR) is a smoothness filter and give the construct method of the derivate operator. This paper uses LS-SVR as the object function constructor and then realizes the approximation of two dimension optimal edge detection operator. This paper proposes the utility method of using singleness operator to realize multiscale edge detection by referencing the multiscale analysis technology of the wavelets theory. Experiment shows that the method has utility and efficiency.

1. Introduction

Edge detection is an important foundation of the image analysis such as image segmentation, texture feature extraction, and shape feature extraction, which is one of the most active research subjects in computer vision field. For the step edge of gray image, due to the edge position corresponding to the gradient extreme of gray function, edge detection may be resolved by solving numerical derivative. As we all know, the Roberts operator (Roberts, 1963), the Prewitt operator (Prewitt, 1970), and the Sobel operator (Sobel, 1970) are all proposed based on the above idea. Torre and Poggio [1] pointed that the image numerical derivative is an ill problem, and thus it is necessary for edge detection to use a low pass filter (LPF) to smooth image and turn it into well conditioned and then solve the derivative. Assuming that $G(X)$ is originality image, $H(X)$ is smooth operator, and $S(x)$ is the output after derivative, according to the convolution characteristic, we can have

$$S(x) = \left[\sum G(x-y)H(y) \right]' = \sum G(x-y)H'(y), \quad (1)$$

$x, y \in R^2.$

We can infer from (1) that edge detection may be realized by using the derivative operator of smooth filter to filter for image; clearly, the derivative operator of smooth filter, namely, the edge detection operator, determines the effect of edge detection.

To judge the performance of different edge detection operators, Canny [2] proposed the distinguished 3-criterion edge detection, took the one dimension step-like edge as an example, gave each criterion mathematics expression in continuous domain, and pointed that the first derivative of the Gaussian function may be the approximation of the optimal edge detection operator. Due to the rotation invariance and two dimension partible characteristic of the Gaussian function, Canny extended this conclusion to two dimension image edge detection. On the basis of Canny’s work, Demigny [3] proposed one dimension discrete format of Canny’s 3-criterion and gave the calculation method of optimal edge detection operator under this criterion. However, the image is two dimensional and thus the edge detection should be two dimensional. Therefore some academician suggested using tensor product format [4] to generate two dimension edge detection operator through one dimension smooth operator and one dimension edge detection operator, which idealize

image into separability signal of rows and columns. In fact, there is some relativity between the rows and columns of image, and in that respect the superiority of impartibility wavelet [5] in the image processing has proved that the tensor product format has bigger error. For the edge detection, taking one dimension signal as the research object could not clearly reflect the relativity of different rows and columns pixels, and thus it is necessary to go on researching the design problem of two dimension edge detection operator deeply. In fact, there are lots of other methods such as data driven [6–10] and support vector machine (SVM) which can deal with the issue.

SVM is one machine study method based on statistical learning theory [11], which is especially suitable for small training set, via realizing structural risk minimization to make study machine possess stronger generalization capability. SVM gained broad application for its special superiority on pattern recognition, function estimation, and so forth. Different from standard SVM, least squares support vector machine [12] (LS-SVM) turns convex quadratic function optimization problem into linear equation solution, which increases support vector number but avoids the complicated optimizing process. Therefore, SVM has been applied widely on function fitting [13–15]. The SVM function fitting is the process of looking for the target function from the function space, and the function format is decided by the setting of SVM parameters. Therefore, parameters selecting is always the puzzle problem in the SVM tradition application. But from another point of view, this characteristic of SVM just provides a new solving project for function construction problem; that is, utilizing the same input vector, we may obtain different format function via adjusting the SVM parameters. Due to edge detection operator design being essentially function construction problem, this paper tries to utilize least squares support vector regression (LS-SVR) to construct edge detection operator of two dimension image.

The paper is organized as follows. In Section 2, first, the Canny continuous criteria and Demigny discrete criterion are introduced, respectively. Next, localization criterion is improved in Section 2.3, and the good detection results criterion and localization criterion are extended to two dimension in Section 2.4. In Section 3, LS-SVR is proved to be smooth filter and the derivative operator constructor method is given. In Section 4, first, the approximation ability of LS-SVR is analyzed via experiment. The approximation of the optimal two dimension edge detection operator proposed in this paper is realized via adjusting LS-SVR parameters in Section 4.2; at the same time, the optimal parameters of Canny's edge operator are given and the LS-SVR operator in this paper is compared with Canny's operator. Section 4.3 uses tensor product format to build two dimension edge detection operator through one dimension optimal Canny's operator, compares the performance with two dimension optimal Canny's operator, and validates the conclusion that one dimension optimal filter tensor product is just approximate of two dimension optimal filter. Section 5 proposes the utility method using singleness operator to realize multiscale edge detection based on multiscale analysis technology of wavelet theory. A conclusion is made in Section 6.

2. Edge Detection Criterion

Noise is the main factor of influence edge detection effect; in the various types of noise, especially the additive white Gaussian noise is most familiar, and thus the evaluation criterion of edge detection operator is mainly to decrease the Gaussian noise effect as the main aim.

2.1. Canny's Continuous Criterion

2.1.1. Good Detection. Good detection is to be a low probability of failing to mark real edge position and low probability of falsely marking nonedge position. Since both these probabilities will decrease monotonically with the increase of signal-to-noise ratio (SNR), this criterion is equivalent to maximizing the SNR of edge points after filtering by solving the smooth filter derivative operator $f(X)$. Assuming $G(X)$ is one dimension step edge signal and $N(X)$ is the Gaussian noise signal with zero mean and variance σ^2 , then after $f(X)$ filtering edge point SNR is

$$\text{SNR}(f) = \frac{H_G}{H_N} = \frac{\left| \int_{-W}^W G(-x) f(x) dx \right|}{\sigma \sqrt{\int_{-W}^W f^2(x) dx}}. \quad (2)$$

2.1.2. Localization Criterion. Good localization is that the edge points marked by the edge detection operator should be as close as possible to the real edge. Assuming that the real edge is centered at 0 and the edge point marked is x_0 , then minimizing the expectation $E(x_0^2)$ is to maximize the following function:

$$\text{Localization}(f) = \frac{\left[\int_{-W}^W G'(-x) f'(x) dx \right]^2}{\sigma^2 \int_{-\infty}^{\infty} f'^2(x) dx}. \quad (3)$$

2.1.3. Only One Response to a Single Edge. The response peak value of edge detection operator to noise is the main reason of repeat response to a single edge; thus, this criterion can be scaled by the average distance (x_{ave}) between the two peak values of noise response. Assuming that the width of filter windows is $2W$, we have

$$x_{\text{ave}}(f) = 2\pi \left(\frac{\int_{-\infty}^{+\infty} f'^2 dx}{\int_{-\infty}^{+\infty} f''^2 dx} \right)^{1/2} = kW. \quad (4)$$

Once k is fixed, the peak value number of the noise response is confirmed in filter windows. It can be proved that the result of this criterion is irrelative to the filter space scale; thus, we may take the 3rd criterion as the constraint, optimize the decision function which is composed of the first two criteria, and then solve the optimal edge detection operator.

2.2. Canny's Discrete Criterion. Canny obtained the results mentioned above under assuming that image and filter are all continuous function. However, images are discrete; accordingly, filters should be discrete. Therefore, in the application, only the sampling of continuous function is used to act on

image, but the sampling only satisfying sampling theorem can be taken as discrete expression of continuous function; otherwise, it will produce frequency overlap. So, the optimal result of continuous domain after figure image sampling is hardly optimal; moreover, according to statistic theory, the result of formula (3) also has a certain degree of deviation. Thus, by studying deeply [16, 17], Demigny proposed the one dimension edge detection discrete criterion [3] based on Canny's continuous criterion. Demigny found that the third criterion used by Canny may be replaced with threshold operation in the discrete domain; thus, this paper only introduces the first two criteria.

Assuming that h is the discrete edge detection operator based on gradient, scale of h is $2d + 1$, $\forall d > 0$, which should satisfy $h_0 = 0$, $\forall k > 0$, $h_{-k} > 0$, and $h_{-k} = -h_k$, and input signal is $e_n = A(U_n - U_{n-d})$, where U_n is the discrete Heaviside sequence, then, without noise, the filter output is

$$s_n = A \sum_{-d}^{-1} h_k, \quad (5)$$

where d is the scale of the filter size.

Only in the presence of noise, the variance of the filter output is

$$\sigma_s^2 = \sigma^2 \sum_{-\infty}^{+\infty} h_k^2. \quad (6)$$

2.2.1. Good Detection: Criterion Σ . Good detection should maximize the SNR of filter at the edge point, that is, maximize the following function:

$$\Sigma = \frac{\sum_{-d}^{-1} h_k}{\left(\sum_{-\infty}^{+\infty} h_k^2\right)^{1/2}}. \quad (7)$$

It is equivalent to maximize the following function:

$$F = \frac{\left(\sum_{-d}^{-1} h_k\right)^2}{\sum_{-\infty}^{+\infty} h_k^2}. \quad (8)$$

For the formulas above, solve the partial derivative and let it be 0; thus

$$\begin{aligned} -d \leq k \leq -1, \quad h_k &= \frac{\sum_{-d}^{-1} h_i}{F}, \\ k < -d, \quad h_k &= 0. \end{aligned} \quad (9)$$

We can see that the optimal edge detection filter is the box filter of size $2d + 1$, so the optimal value of formula (7) is

$$\Sigma_{\max} = \sqrt{\frac{d}{2}}. \quad (10)$$

Formula (9) gives a conclusion as follows: for one step edge of width d , the optimal edge detection operator should not be longer than $2d + 1$; that is to say, for the edge filter with $2d + 1$ coefficient, the edge would be detected reliably only when the positive and negative width are larger than d , in fact, which offers the basis for scale selection of edge detection.

2.2.2. Good Localization: Criterion Λ . Since noise may cause the offset of edge position, to confirm the good localization, we should maximize the $s'_{-1} > 0$ probability. If the white Gaussian noise N with variance σ^2 is added to the signal, then the total output at -1 point is

$$s_{-1} = A \sum_{-d}^{-1} h_k + \sum_{-\infty}^{+\infty} h_{-1-k} N_k. \quad (11)$$

Demigny use discrete difference to take place of derivation operation; thus

$$s'_{-1} = A (h_{-1} - h_{-1-d}) + \sum_{-\infty}^{+\infty} h'_{-1-k} N_k. \quad (12)$$

The probability to obtain $s'_{-1} > 0$ is then

$$P(s'_{-1} > 0) = \frac{1}{2} + \operatorname{erf}\left(\frac{\Lambda A}{\sigma}\right), \quad (13)$$

where $\operatorname{erf}(x) = 1/\sqrt{2\pi} \int_0^x e^{-t^2/2} dt$ and Λ is the localization criterion:

$$\Lambda = \frac{h_{-1} - h_{-1-d}}{\left(\sum_{-\infty}^{+\infty} h_k'^2\right)^{1/2}} = \frac{\sum_{-d}^{-1} h_k'}{\left(\sum_{-\infty}^{+\infty} h_k'^2\right)^{1/2}}. \quad (14)$$

We have

$$\Lambda^2 = \left(2 + 2 \frac{\sum_{-d}^{-1} h_k'^2}{\left(\sum_{-d}^{-1} h_k'\right)^2}\right)^{-1}. \quad (15)$$

When $h'_k = h'_{k+1}$, $-d \leq k \leq -1$, formula (15) obtains the maximization, so the optimal value of the localization criterion is

$$\Lambda_{\max} = \sqrt{\frac{d}{2(d+1)}}. \quad (16)$$

2.3. Improved Discrete Localization Criterion. For one dimension step edge, the reason of maximization the probability to obtain $s'_{-1} > 0$ is that the real edge position is between -1 and 0 . Demigny used discrete difference to calculate derivativeness, in fact, which gave the maximization probability to obtain $s_{-1} > s_{-2}$; that is to say, all the filters that satisfy $s_{-1} > s_{-2}$ are good localization operator, which permits the maximization of edge detection to be between -1 and -2 and lean to -1 , which has one pixel deviation with the real edge. Moreover, for the less scale filter, the operator derivativeness solved by difference has been influenced largely by the borderline. Therefore, only using the derivative operator of the edge detection filter that strictly satisfies the derivative definition can solve the existing problem.

Assume that width of the input signal is larger than $2d + 1$ and h'' is the second derivative operator of size $2d + 1$ that is deduced via formula (1), which should satisfy

$$\begin{aligned} \sum_{-d}^d h_k'' &= 0, \\ h_k'' &= h_{-k}'', \quad 0 < k \leq d, \\ h_k'' &= 0, \quad |k| > d. \end{aligned} \quad (17)$$

Thus,

$$\Lambda = \frac{\sum_{-d}^{-1} h_k''}{\left(\sum_{-d}^d h_k''\right)^{1/2}}, \quad \Lambda^2 = \frac{\left(\sum_{-d}^{-1} h_k''\right)^2}{\sum_{-d}^d h_k''^2}. \quad (18)$$

For $h_0'' = -2 \sum_{-d}^{-1} h_k''$, we have

$$\Lambda^2 = \frac{\left(\sum_{-d}^{-1} h_k''\right)^2}{4\left(\sum_{-d}^{-1} h_k''\right)^2 + 2 \sum_{-d}^{-1} h_k''^2} = \left(4 + 2 \frac{\sum_{-d}^{-1} h_k''^2}{\left(\sum_{-d}^{-1} h_k''\right)^2}\right)^{-1}. \quad (19)$$

Obviously, the optimal h'' obtained by solving the maximum of formula (19) should satisfy $h_k'' = h_{-k}' = h_{k+1}'', 1 \leq k \leq d - 1$, and $h_0'' = -2dh_k'', 1 \leq k \leq d$; thus, the optimal value of the localization criterion is

$$\Lambda_{\max} = \sqrt{\frac{d}{4d + 2}}. \quad (20)$$

2.4. Two Dimension Optimal Filter. One two dimension smooth filter of the good space domain should have centrosymmetric characteristic; that is, the filter coefficients are symmetric according to row, column, and diagonal. This paper discusses smooth filter with the above characteristic. According to the conclusion of formula (1), edge detection operator should be the derivative of a smooth filter operator and the gradient of smoothness operator for the two dimension operator. Supposing $x = (r, c)$ of formula (1), then, the response of two dimension edge detection operator should be

$$\begin{aligned} S'(x) &= \left[\left(\sum G(x-y) \frac{\partial}{\partial r} H(y) \right)^2 + \left(\sum G(x-y) \frac{\partial}{\partial c} H(y) \right)^2 \right]^{1/2}. \end{aligned} \quad (21)$$

We can see that two dimension edge detection operator is composed of row and column directional partial derivative of smooth filter operator. Assuming that input signal G is a two dimension column directional discrete Heaviside sequence, filter window is $N \times N$ neighborhood, where $N = 2d + 1$, and windows center coordinate is $(0, 0)$.

2.4.1. Optimal Filter for Criterion Σ . Assuming that dcH is column directional first partial derivative of formula (1) and drH is row directional first partial derivative, we have

$$\begin{aligned} dcH_{i,0} &= 0, \quad -d \leq i \leq d, \\ dcH_{i,j} &= dcH_{-i,j} = -dcH_{i,-j}, \\ -d \leq i \leq d, \quad -d \leq j \leq -1, \quad -d \leq i \leq d, \\ dcH &= drH^T. \end{aligned} \quad (22)$$

Since drH and dcH are transposing each other, to assure the column directional optimal can obtain the row directional optimal. Therefore, here only considering dcH and according to one dimension instance, "good detection" criterion should be

$$\Sigma = \frac{\sum_{-d}^d \sum_{-d}^{-1} dcH_{i,j}}{\left(\sum_{-d}^d \sum_{-d}^d dcH_{i,j}^2\right)^{1/2}}. \quad (23)$$

Thus, the optimal filter may be solved, which satisfies

$$\begin{aligned} dcH_{i,j} &= -dcH_{-i,-j} = k, \\ k > 0, \quad -d \leq i \leq d, \quad -d \leq j \leq -1, \\ dcH_{i,0} &= 0, \quad -d \leq i \leq d. \end{aligned} \quad (24)$$

So, the maximum of Σ is

$$\Sigma_{\max} = \sqrt{\frac{d(2d+1)}{2}}. \quad (25)$$

2.4.2. Optimal Filter for Criterion Λ . For formula (1), to introduce the Laplacian operator, we have

$$S''(x) = \sum G(x-y) \frac{\partial^2}{\partial r^2} H(y) + \sum G(x-y) \frac{\partial^2}{\partial c^2} H(y). \quad (26)$$

Localization criterion demands that there is minimal deviation between the marked edge position and the real edge position. Since the input signal is a column directional Heaviside sequence, simon-pure edge position should be between $(0, -1)$ and $(0, 0)$, and here $S''(0, -1) > 0$ and $S''(0, 0) < 0$. Good edge detection operator should make $S''(0, -1) > 0$ and $S''(0, 0) < 0$ have the maximal probability; therefore we discuss the maximization problem of the probability to obtain $S''(0, -1) > 0$ firstly. Assuming that $dccH$ is column directional second derivative operator of H and $drrH$ is row directional second derivative operator, then $dccH$ should have the symmetric characteristic according to row and column and satisfy $dccH = drrH^T$. Assuming that L is the Laplacian operator, where $L_{i,j}$ is the coefficient, $L = dccH + drrH$ holds. Thus, L is symmetric according to row and column and the elements of the zero row and the zero column are symmetric according to $(0, 0)$, and $\sum_{-d}^d \sum_{-d}^d L_{i,j} = 0$; we have

$$L_{0,0} = - \left(4 \sum_{-d}^{-1} L_{0,j} + 4 \sum_{-d}^{-1} \sum_{-d}^{-1} L_{i,j} \right). \quad (27)$$

According to one dimension format, localization criterion should be

$$\Lambda = \frac{\sum_{-d}^d \sum_{-d}^{-1} L_{i,j}}{\left(\sum_{-d}^d \sum_{-d}^{-1} L_{i,j}^2\right)^{1/2}}. \quad (28)$$

Maximization Λ may be equivalent to the minimization of the following formula:

$$F = \frac{\sum_{-d}^d \sum_{-d}^{-1} L_{i,j}^2}{\left(\sum_{-d}^d \sum_{-d}^{-1} L_{i,j}\right)^2} = \frac{L_{0,0}^2 + 4 \sum_{-d}^{-1} L_{0,j}^2 + 4 \sum_{-d}^{-1} \sum_{-d}^{-1} L_{i,j}^2}{\left(\sum_{-d}^{-1} L_{0,j} + 2 \sum_{-d}^{-1} \sum_{-d}^{-1} L_{i,j}\right)^2}. \quad (29)$$

Solving partial derivative of F to coefficients and letting it be equal to 0

$$\begin{aligned} \frac{\partial F}{\partial L_{0,j}} = 0 &\implies L_{0,j} = \frac{1}{4} F \sum_{-d}^{-1} \sum_{-d}^{-1} L_{i,j} + L_{0,0}, \\ \frac{\partial F}{\partial L_{i,j}} = 0 &\implies L_{i,j} = \frac{1}{2} F \sum_{-d}^{-1} \sum_{-d}^{-1} L_{i,j} + L_{0,0}. \end{aligned} \quad (30)$$

Thus optimal filter for criterion Λ should satisfy

$$1 \leq |j| \leq d, \quad L_{0,j} = L_{j,0} = \frac{1-2d}{2} k, \quad L_{0,0} = -2dk. \quad (31)$$

Here, we can compute the maximal value for Λ :

$$\Lambda_{\max} = \frac{1}{2} \sqrt{d}. \quad (32)$$

It can be proved that the optimal filter for criterion Λ may assure the probability to obtain $S''(0,0) < 0$ to be maximization at the same time.

To sum up the result, the design process of two dimension edge detection operator is as follows: first, to construct one smooth operator; next, to solve the first derivative operator and take it as edge detection filter and, at the same time, to solve the second derivative of smooth filter; lastly, to utilize criterion Σ and criterion Λ to construct judge function and then be on seeking the optimization.

3. LS-SVR Filter

3.1. LS-SVR. For the training database, LS-SVR can be expressed by:

$$\begin{bmatrix} 0 & 1_N^T \\ 1_N & \Omega + \gamma^{-1} I_N \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix}. \quad (33)$$

Here, $Y = [y_1, \dots, y_N]^T$ is input vector; $1_N = [1, \dots, 1]^T$ and $\alpha = [\alpha_1, \dots, \alpha_N]^T$ are the Lagrange multipliers corresponding to each input sample; I_N is $N \times N$ identity matrix; $\Omega \in \mathbb{R}^{N \times N}$, $\Omega_{ij} = K(x_i, x_j)$, and $K(x, y)$ is kernel function; γ is penalty factorization.

After solving the value of parameters α and b , least squares support vector regression (SVR) function is

$$y = f(x) = \sum_{i=1}^N \alpha_i K(x_i, x) + b. \quad (34)$$

The key to assure formula (34) with generalization capability is to select kernel function $K(x, y)$. For the problem of including little prior information, the Gaussian kernel function is a good selection [18]; the Gaussian kernel function is selected as the kernel function in this paper:

$$K(x, y) = \exp\left(\frac{-|x - y|^2}{\sigma^2}\right). \quad (35)$$

Let $\bar{\Omega} = \Omega + \gamma^{-1} I_N$; thus $Y = 1_N b + \bar{\Omega} \alpha$; via transforming the augmented matrices of formula (32), the solving of formula (32) is

$$b = \frac{1_N^T \bar{\Omega}^{-1}}{1_N^T \bar{\Omega}^{-1} 1_N} Y = AY, \quad (36)$$

$$\alpha = \bar{\Omega}^{-1} (Y - 1_N b) = \bar{\Omega}^{-1} (I_N - 1_N A) Y = MY. \quad (37)$$

3.2. LS-SVR Filter Operator. Via analyzing LS-SVR, the following theorem can be obtained.

Theorem 1. Using LS-SVR to be fitting for the center of the equidistant discrete time sequence windows of size $2d + 1$, is equivalent to one discrete filter of size $2d + 1$.

Proof. Assuming that window center is $y(n)$ and the output by LS-SVR fitting is $f(n)$, then the input vector of LS-SVR is $Y = [y(n-d), \dots, y(n), \dots, y(n+d)]^T$. The fitting is only for the points which have fallen into the windows, therefore the points can be expressed by relative coordinate; that is, $Y_w = [y(-d), \dots, y(0), \dots, y(d)]^T$, and, obviously, $Y = Y_w$ holds. Thus, in the case of the kernel function and penalty factorization γ being determined, A is a constant coefficient row vector with $2d+1$ elements and M is a constant coefficient matrix of $[2d+1] \times [2d+1]$. Assuming b_i is the first element of A , m_i is the i th row of matrix M , and $m_{i,j}$ is the j th element of the i th row, then, $\alpha_i = m_i Y$ and $-d \leq i \leq d$ hold, and thus

$$\begin{aligned} f(n) &= \sum_{i=-d}^d \alpha_i K(i, 0) + b = \sum_{i=-d}^d K(i, 0) m_i Y + AY \\ &= \left[\sum_{i=-d}^d K(i, 0) m_i + A \right] Y. \end{aligned} \quad (38)$$

Let $H = \sum_{i=-d}^d K(i, 0) m_i + A$; thus, we have $f(n) = HY$, where H is the row vector with $2d+1$ constant coefficient and the j th element $H_j = \sum_{i=-d}^d m_{i,j} K(i, 0) + b_j$, and thus

$$f(n) = \sum_{-d}^d H_i y(n+i). \quad (39)$$

Therefore, the center point $y(n)$ of window with size $2d+1$ being fitted via LS-SVR is equivalent to one discrete filter of size $2d+1$. This completes the proof. \square

The meaning of the above conclusion is to use LS-SVR to be fitting discrete data points in the windows, to obtain a continuous domain function, whose value at the windows center points can be obtained via convolution of one discrete filter to each discrete point of windows, which avoid the probable arising problem of turning discrete data into continuous domain to be disposed and then sampled into discrete domain.

Since H is deduced by using LS-SVR, the first derivative operator H' and the second derivative operator H'' in strict accord with mathematics definition may be obtained as follows:

$$H'_j = \sum_{i=-d}^d m_{i,j} K'(i, 0), \quad (40)$$

$$H''_j = \sum_{i=-d}^d m_{i,j} K''(i, 0).$$

The conclusion above may be extended to two dimensions easily. Assuming that the fitting window is $[2d+1] \times [2d+1]$ neighborhood and windows center point is $y(m, n)$, then, M of formula (37) is the matrix of row $[2d+1] \times [2d+1]^2$ and A of formula (36) is the row vector of $[2d+1] \times [2d+1]$. Assuming that $m_{i,j}$ is the corresponding row of $\alpha_{i,j}$ in the matrix M , $m_{i,j}(k, l)$ is the corresponding element of input points $y(m-k, n-l)$ in $m_{i,j}$, and $b_{k,l}$ is the corresponding element of $y(m-k, n-l)$ in A , then

$$H_{i,j} = \sum_{k,l=-d}^d m_{i,j}(k, l) K((i, j), (0, 0)) + b_{i,j}. \quad (41)$$

Each coefficient value of filter H given by the theorem is related to enactment of LS-SVR kernel function format and penalty factorization γ ; therefore different kernel function and different penalty factorization will get the different smooth filter; in this sense, H is the function of kernel and penalty factorization; that is, $H = f(K, \gamma)$. For the Gaussian kernel function LS-SVR using the form of formula (35), we have

$$H = f(\sigma, \gamma). \quad (42)$$

It should be noted that smooth operator solved according to the theorem is a normalized operator, which may be used in application, but by which the first derivative operator is deduced and gives the first derivative value of windows center after images smoothing. When the smoothness function is stronger, the first derivative operator is less, and this is not for numerical processing. Therefore, the first derivative operator needs normalizing, and the maximal value of output is usually set as 0.5. For the first derivative operator of two dimension column direction, normalization method is

$$h_{i,j} = \frac{H'_{i,j}}{\left(\sum_{-d}^{-1} \sum_{-d}^{-1} H'_{k,l} \right)}. \quad (43)$$

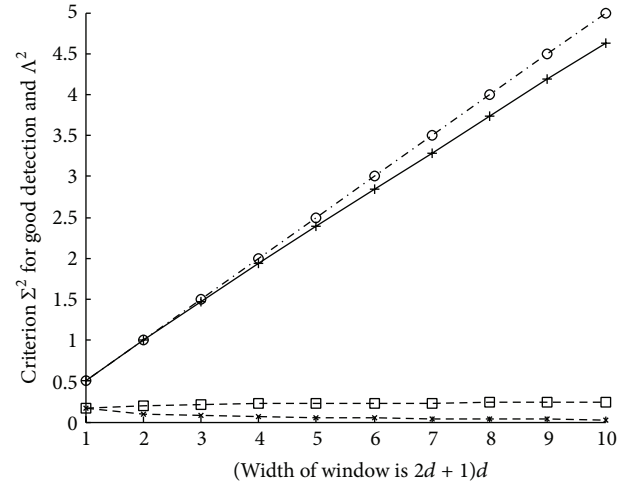


FIGURE 1: Curve of LS-SVR performance for good detection.

4. LS-SVR Approximate Ability of the Optimal Operator

4.1. LS-SVR Approximate Ability. Our aim is to construct the required derivative operator via utilizing LS-SVR. According to SVM theory, we can obtain object function that is expressed by the radix of the high dimension space, that is, kernel function, and is one point that belongs to the whole function space. Obviously, so far we have not found this kernel function, and its finite linearity combination may express the arbitrary function in the whole function space. Therefore, LS-SVR could only give the approximation of the required function. To review the approximate ability of LS-SVR to optimal edge detection operator, this paper provides experiment examples to illustrate.

Figure 1 presents the approximation curve of LS-SVR filter to the optimal edge detector for one dimension criterion Σ of size $2d+1$ and Table 1 presents the corresponding LS-SVR parameters value; we may see that LS-SVR could only construct the optimal filter for criterion Σ with $d=1$ and $d=2$ and the approximation filter for $d \geq 3$ and the optimization ability about the filter only reach 90% of optimal value for $d=10$. When LS-SVR obtains the maximum according to the criterion Σ , the corresponding Λ value will deviate with the optimal value largely.

Figure 2 presents the approximation curve of LS-SVR filter to the optimal filter for one dimension criterion Λ and Table 2 presents the corresponding parameters of LS-SVR. We can see from Figure 2 that the approximation curve of LS-SVR coincides with the theory curve; that is, LS-SVR can realize one dimension optimal edge detection filter for criterion Λ , but at this time the value of Σ^2 is around 0.5, which is widely discrepant from the optimal value.

Figure 3 is the approximation curve of LS-SVR filter to the two dimension optimal filter for criterion Σ ; we can see that, when $d=2$, LS-SVR can realize the optimal filter for criterion

TABLE 1: Optimal parameters of LS-SVR for good detection.

d	1	2	3	4	5	6	7	8	9	10
γ	9.78	4.69	0	0	0	0	0	0	0	0
σ^2	7.001	8.889	7.780	12.413	18.211	25.171	33.293	42.575	53.018	64.621

TABLE 2: Optimal parameters of LS-SVR for good localization.

d	1	2	3	4	5	6	7	8	9	10
γ	9.11	0.51	0.18	0.06	0.69	0.41	0.61	0.69	0.51	0.30
σ^2	9.602	0.030	0.030	0.032	0.023	0.044	0.037	0.017	0.036	0.027

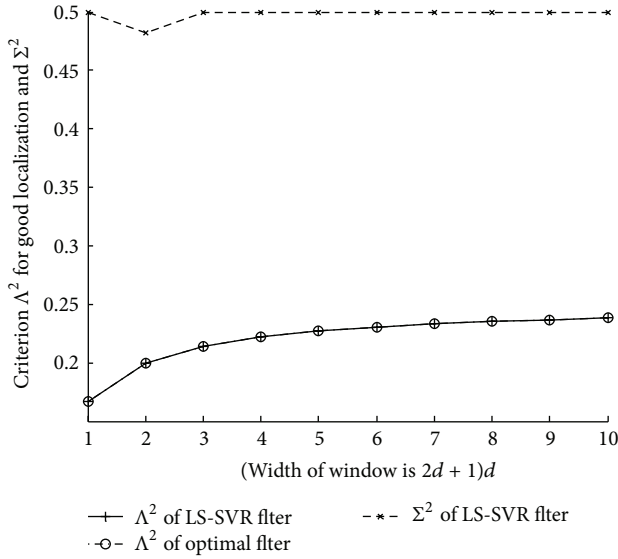


FIGURE 2: Curve of LS-SVR performance for good localization.

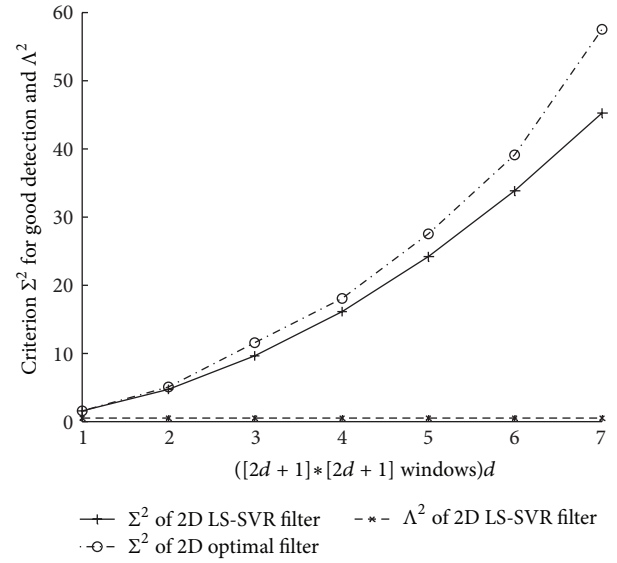


FIGURE 3: Curve of 2D LS-SVR performance for good detection.

Σ basically, but, when $d \geq 3$, the approximation capability of LS-SVR filter will be low apparently.

In the approximation experiment for criterion Λ , we find that LS-SVR approximation capability is the worst; when $d = 2$, LS-SVR could only obtain 70% optimization ability about ideal state, moreover, when $d \geq 3$, the maximum of Λ could not be improved obviously via utilizing LS-SVR. In fact, the worse approximation capability of LS-SVR to the criterion Λ could not influence us to use the LS-SVR to construct edge detection filter. We can see from the experiment that Λ changes very slowly with the LS-SVR parameter modification but Σ has a larger rate of change. The final result of the decision function according to product criterion is decided by the larger entry of change rate. Therefore, we can use LS-SVR to design the edge detection operator with better performance.

4.2. Optimization of a Combination of Σ and Λ . Criteria Σ and Λ have the mutex characteristic; we can see from Figures 1 and 2 that the localization capability of the optimal filter for criterion Σ is poor and the signal SNR of the optimal filter for criterion Λ is poor. Therefore, we can only make a trade-off on the two criteria in the real application, and the concrete method is to utilize criteria Σ and Λ to construct one decision

function, so as to find the filter that makes the two criteria reach the larger value at the same time and thinks it is the optimal.

4.2.1. Summation Criterion. Use the summation of criteria Σ and Λ to construct decision function; that is,

$$J = \Sigma + \mu\Lambda. \quad (44)$$

Here $\mu > 0$ is called adjustment coefficient. In the case of different scale, there is a great difference between the value of Σ and Λ , so different scale μ should choose different value.

4.2.2. Product Criterion $\Sigma\Lambda$. In the method the decision function is

$$J = \Sigma \cdot \Lambda. \quad (45)$$

This method is very often used. Canny has used the decision function as (43) to compute and obtain a conclusion that the Gaussian function is the approximation of the one dimension optimal filter and extended it to two dimension edge detection. According to this method, this paper utilizes LS-SVR to approximate two dimension column directional

TABLE 3: The parameters of LS-SVR for 2D optimal filter.

d	1	2	3	4	5	6	7
γ	1	$0.1E-9$	$0.1E-9$	$0.1E-9$	$0.1E-9$	$0.1E-9$	$0.1E-9$
σ^2	85.00	4.86	8.62	13.80	20.30	28.12	37.25
Σ^2	1.5000	4.5284	9.1280	15.2770	22.9654	32.1946	42.9631
Λ^2	0.2500	0.2246	0.2261	0.2267	0.2270	0.2271	0.2272

TABLE 4: Optimal parameters of Canny's edge operator.

d	1	2	3	4	5	6	7
σ^2	0.41	0.80	1.32	2.00	2.82	3.76	4.84
Σ^2	1.0773	2.4563	4.5218	7.2343	10.5046	14.2572	18.5671
Λ^2	0.0924	0.1849	0.2061	0.2094	0.2109	0.2126	0.2133

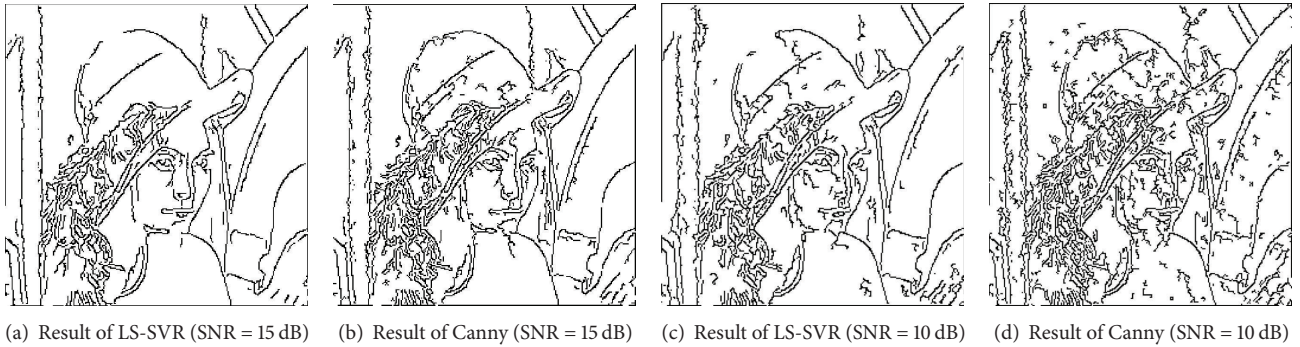


FIGURE 4: Result of Lena by LS-SVR operator and Canny operator.

optimal filter; Some approximation results and corresponding parameters under several common scales are shown in Table 3.

In this paper, we test Canny's edge detection operator according to the proposed two dimension discrete criterion, and in the testing the Gaussian smoothness filter function used is

$$H_g(i, j) = \exp\left(-\frac{(i^2 + j^2)}{2\sigma^2}\right). \quad (46)$$

The first derivative of the column directional and the Laplacian operator are, respectively,

$$H'_g(i, j) = -\frac{j}{\sigma^2} \exp\left(-\frac{(i^2 + j^2)}{2\sigma^2}\right), \quad (47)$$

$$\nabla_g^2(i, j) = \left(\frac{j^2 + i^2}{\sigma^4} - \frac{2}{\sigma^2}\right) \exp\left(-\frac{(i^2 + j^2)}{2\sigma^2}\right).$$

Since Canny's operator is a linear operator, formula (45) and the standard Gaussian function have the same calculation result. Table 4 is the optimal parameters of Canny's operator of generated scale.

From the data of Tables 3 and 4, we can see that the localization capability of Canny's operator is poorer than that of the operator obtained by LS-SVR, but on the noise

resistance ability this operator solved in this paper has the greater advantages.

In order to compare the two operators' performance, this paper utilizes the Lena image and the Cameraman image of the 256×256 pels to be on simulation experiment. In the experiment, the used filter window is 5×5 neighborhood; the edge detection algorithm is as follows.

Step 1. Solve the row and column directional derivative of image by using edge detection operator and then obtain the gradient image and angular image.

Step 2. Obtain nonmaxima suppression.

Step 3 (two thresholds detection edge). Here, assume that high threshold (Th) is 0.8 of cumulative histogram and low threshold (Tl) is 0.4 Th.

Figure 4 is the output result of two edge detection after adding the Gaussian noise to the Lena image and Figure 5 is the detection result of the Cameraman image. We can see from the figures that the operator proposed in this paper has the better performance.

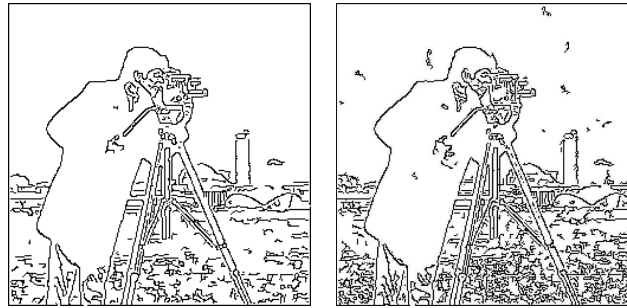
4.3. Comparison of One Dimension and Two Dimension Operator. For the separable signal, two dimension edge detection windows operator can be obtained via one dimension edge detection operator according to tensor product format [4].

TABLE 5: Optimal parameters of 1D Canny's edge operator.

d	1	2	3	4	5	6	7
σ^2	0.33	0.73	1.07	1.50	2.04	2.66	3.34
Σ	0.7071	0.8605	0.9900	1.1069	1.2157	1.3129	1.3994
Λ	0.3774	0.4278	0.4180	0.3888	0.3610	0.3383	0.3199

TABLE 6: 2D Σ and Λ criterion value of Canny's operator from 1D optimal filter.

d	1	2	3	4	5	6	7
σ^2	0.33	0.73	1.07	1.50	2.04	2.66	3.34
Σ^2	0.9449	2.2305	3.5903	5.3181	7.4819	9.9642	12.6860
Λ^2	0.0878	0.1977	0.2355	0.2434	0.2455	0.2467	0.2475



(a) Result of LS-SVR

(b) Result of Canny

FIGURE 5: Result of 15 dB Cameraman by LS-SVR operator and Canny operator.

Assuming that the scale of one dimension edge detection operator H' educed from formula (1) is $2d + 1$, the two dimension edge detection operator can be obtained via the following formula:

$$Hdx = HH'^T, \quad Hdy = Hdx^T, \quad (48)$$

where Hdx and Hdy are the column and row directional two dimension edge detection operator of size $[2d + 1] \times [2d + 1]$, respectively.

The Gaussian function has the two dimension separability; for example, the column directional Canny's edge detection operator may be decomposed into product of the column directional one dimension Gaussian function derivative and the row directional Gaussian smooth filter operator; that is, the two dimension optimal edge detection operator may be obtained via one dimension optimal operator according to formula (48). Just according to the characteristic of the Gaussian function, Canny extended the one dimension optimal operator to a two dimension operator. From the view of optimal filter, we think the two dimensions extend form of the one dimension optimal filter obtained via formula (48) which just only is the approximation of two dimension optimal filter. We take Canny's edge detection operator as an example to validate the conclusion as follows.

According to the improved one dimension criterion that is proposed in this paper, this paper carries out experiments for one dimension Canny's edge detection operator according to (45), where the Gaussian function is of one dimension form as (46). Table 5 gives the optimal parameters of one

dimension optimal Canny's edge detection operator via experiment measurement. Utilizing variance presented in Table 5 to build two dimension edge detection operator via formula (48), the values of Σ and Λ obtained according to two dimension criterion are shown in Table 6. Comparing Table 4 with Table 6, we can see that the SNR of the two dimension operator built from the one dimension optimal filter declines more largely than the two dimension optimal filter; just the localization capability rises a little. But from the view of formula (45), the performance difference is larger. Therefore, we can draw a conclusion that utilizing tensor product format to extend one dimension optimal operator to two dimension operator just can obtain the approximation of two dimension optimal filter.

5. Multiscale Fusion

Image edge has different local intensity characteristics for the different factors such as obstruction, shadow, highlight, peak, and veins, so Marr pointed out that, in order to reliably detect edge, multi-non-scale edge operator should be used [19], and then Witkin and others developed the idea into scale space filtering technology. For the variance σ^2 of the Gaussian filter controls the image smoothness degree, in multiscale filter technology, take σ^2 as scale gene, via altering the value of σ^2 to reach the aim of multiscale filter, and apply it into multiscale Canny's edge detection. But we can see from Table 4 that, for the definite edge detection windows, the Gaussian function that is only σ^2 is a particular value

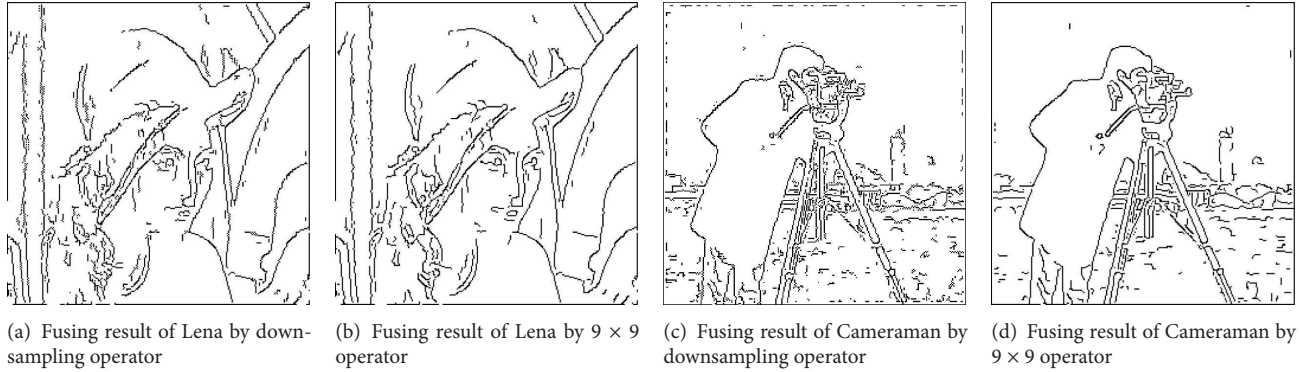


FIGURE 6: Fusing result of downsampling and scale-up operator.

satisfying the optimal filter for product criterion $\Sigma\Lambda$, so this method via altering σ^2 to carry out multiscale edge detection is improper. Therefore, from the view of the optimal edge detection operator, multiscale edge detection can be realized only via converting the size of filter windows, but for the different filter windows parameters of edge detection filter are different, which only needs using different filter operator to realize the aim of multiscale detection, and with the size of filter windows increase, operation magnitude will increase geminately, which will bring much inconvenience to us in real application. In 1992, Mallat proposed the wavelet multiscale edge detection technology [20] and realized the multiscale edge detection via using approximate 3D spline function with the Gaussian function to construct two dimension spline function and using modulus maximum method in different scales. The aim of Mallat constructing the two dimension spline function is to make the edge detection operator of wavelet domain satisfy the wavelet admitting condition, to realize the signal reconstruction after transformation. But for the edge detection, we only need to judge whether one point is edge or not, namely, whether the modulus is the maximum or not after transformation. As long as the transformation satisfies Canny's criteria, we may think it is one good transformation and need not to consider the reconstruction after the transformation. So, we can utilize one definite optimal edge detection operator to replace the filter of wavelet transformation in order to obtain the different scale edge and then realize multiscale edge detection via multiscale fusion technology. That is to say, we can construct big scale filter windows via downsampling method and obtain different scale edge via using single edge detection operator. In order to validate the presumption, experiments are carried out in this paper. The concrete method is as follows.

Step 1. Solve the derivative of row and column direction via utilizing LS-SVR operator whose size is 5×5 .

Step 2. Construct filtering windows via row and column directional downsampled by two for each point of the original image and utilize 5×5 LS-SVR operator to solve the row and column derivative of the image.

Step 3. Multiply row directional derivative by column directional derivative obtained from Steps 1 and 2, to obtain row directional and column directional derivative of two kinds of scale fusion and to solve the gradient value.

Step 4. Carry out nonmaxima suppression and two-threshold edge detection.

The filter windows data, which is the 4 extraction of the originality image, has been obtained by Step 2. Therefore, the comparative object employed the windows size which is defined 9×9 operator, and the same edge extraction method as the proposed 5×5 filter operator method. Figure 6 presents the comparsion results under the two methods about Lena and Cameraman image with 10 dB SNR. From the image we can see that the two methods have the close detection result, but the method proposed in this paper has a faster operation speed.

6. Conclusion

Starting from the problem that the solving numerical derivative of image is an ill problem, this paper studies the edge detection operator design criterion. For locating inaccurate problem of the discrete localization criterion proposed by Demigny, this paper puts forward a new criterion expression of "good localization," establishes the discrete expression of good detection and good localization criterion of two dimension edge detection operator, and then introduces the experiment to measure optimal parameters of two dimension Canny's edge detection operator. In this paper, compare the obtained two dimension optimal Canny's edge detection operators with the two dimension extended forms of one dimension optimal Canny's edge detection operator and validate the conclusion that one dimension optimal filter tensor product just only is the approximation of two dimension optimal filter. We prove that LS-SVR is a smoothness filter, give the construct method, and realize the approximation of two dimension optimal edge detection operator via adjusting LS-SVR parameters. This paper proposes the utility method of using singleness operator to realize multiscale edge detection by referencing the multiscale analysis technology of

the wavelets theory. This paper uses LS-SVR as the object function constructor and obtains LS-SVR parameter values of the optimal edge detection filter in the case of generated scale via the experiment. These LS-SVR parameters are irrelative from concrete application object, so we need not to consider LS-SVR learning problem in different application. Utilizing the given LS-SVR parameters, we can obtain not only the better edge detection operator but also the smoothness filter with the edge-preserving capability, which has the reference meaning for the smoothness filter design. This paper uses the Gaussian kernel function as kernel function of the least squares support vector machine, that is, takes the Gaussian function as the basis of function space to construct new function. For the limitation of the Gaussian function, this paper obtains merely the approximation that needs function, so construct the better kernel function is the father research task.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This present work was supported partially by the Polish-Norwegian Research Programme (Project no. Pol-Nor/200957/47/2013). The authors highly appreciate the above financial support.

References

- [1] V. Torre and T. Poggio, "On edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 2, pp. 147–163, 1986.
- [2] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
- [3] D. Demigny, "On optimal linear filtering for edge detection," *IEEE Transactions on Image Processing*, vol. 11, no. 7, pp. 728–737, 2002.
- [4] M. Peter and G. Bogdan, "Edge detection with embedded confidence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 12, pp. 1351–1365, 2001.
- [5] G. Uytterhoeven and A. Bultheel, "The red-black wavelet transform," in *Proceedings of the IEEE Benelux Signal Processing Symposium*, pp. 191–194, Leuven, Belgium, 1998.
- [6] S. Yin, H. Luo, and S. Ding, "Real-time implementation of fault-tolerant control systems with performance optimization," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2402–2411, 2013.
- [7] S. Yin, G. Wang, and H. R. Karimi, "Data-driven design of robust fault detection system for wind turbines," *Mechatronics*, 2013.
- [8] S. Yin, X. Yang, and H. R. Karimi, "Data-driven adaptive observer for fault diagnosis," *Mathematical Problems in Engineering*, vol. 2012, Article ID 832836, 21 pages, 2012.
- [9] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, 1996.
- [10] L. Wang, L.-F. Bo, F. Liu, and L.-C. Jiao, "Least squares hidden space support vector machines," *Chinese Journal of Computers*, vol. 28, no. 8, pp. 1302–1307, 2005.
- [11] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, 1995.
- [12] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.
- [13] S. Zheng, J. Liu, and J. W. Tian, "A new efficient SVM-based edge detection method," *Pattern Recognition Letter*, vol. 25, no. 10, pp. 1143–1154, 2004.
- [14] J. Zhu, S. T. Wang, X. S. Wu, and F. L. Chung, "A novel adaptive SVR based filter ASBF for image restoration," *Soft Computing*, vol. 10, no. 8, pp. 665–6672, 2006.
- [15] Z. D. Yu and L. S. Wang, "Research on image noise suppression algorithm based On LS-SVR," *ACTA Automatica Sinica*, vol. 35, no. 4, pp. 364–370, 2009.
- [16] D. Demigny and M. Karabernou, "An effective resolution definition or how to choose an edgedetector, its scale parameter and the threshold?" in *Proceedings of the International Conference Proceedings of the International Conference on Image Processing*, vol. 1, pp. 829–832, September 1996.
- [17] D. Demigny and T. Kaml, "A discrete expression of canny's criteria for step edge detector performances evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 11, pp. 1199–1211, 1997.
- [18] S. S. Keerthi and C. J. Lin, "Asymptotic behaviors of support vector machines with Gaussian kernel," *Neural Computation*, vol. 15, no. 7, pp. 1667–1689, 2003.
- [19] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 617–643, 1992.
- [20] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, 1992.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

