Image recognition of laser radar using linear SVM correlation filter

Jianfeng Sun (孙剑峰), Qi Li (李 琦), Wei Lu (陆 威), and Qi Wang (王 骐)

National Key Laboratory of Tunable Laser Technology, Institute of Optoelectronics, Harbin Institute of Technology, Harbin 150001

Received February 2, 2007

Through deducing the relationship between support vector machine (SVM) and correlation principle, the optimal hyperplane is proved as a correlation filter when the kernel function is the linear kernel. So a new correlation filter, named linear SVM correlation filter (LSCF), is proposed. The filter has not only shift-invariance, but also SVM properties. The real images of laser radar are used as experiment data, and LSCF is used to solve the in-plane rotation invariance. The results show that the filter can recognize the different rotated objects, and the correlation output is stable. The filter is insensitive to the noise and gray change, and has good discrimination ability. In the same design way, LSCF is also suitable to solve other problems of correlation distortion.

OCIS codes: 100.5760, 070.6110, 280.3640.

In recent years, correlation recognition has attracted considerable attention again[1−5], because 1) the development of the high-speed digital signal processor (DSP) makes the correlation recognition easily meet the real-time requirement, and the algorithm complexity of correlation recognition is mainly decided by the field of view (FOV) resolution, but not the quantities of objects in FOV; 2) the advanced correlation filters are constantly presented, and the performance of algorithms is improved with respect to the distortion tolerance and discrimination. In addition, correlation filter itself has the shift-invariance, and therefore it has practical and valuable application. Normally, correlation filter is designed according to the optimal criteria[6,7], such as correlation peak criterion, energy criterion, and so on. However, the images that are collected by sensor, such as SAR or laser radar, usually have some changes like gray change and target distortion, which can influence on the correlation output performance and consequently reduce the correct recognition rate. That is to say, the generalization of correlation filter or its discrimination for untrained target is further improved.

The theory of support vector machine (SVM) is founded in the mid-1990's, and it is developed based on the statistical learning theory (SLT)[8,9]. It has behaved some special advantages in solving the problems of small sample, non-linear and high-dimension pattern recognition, and has become research hotspot in the field of machine learning.

In this paper, through deducing the relationship between SVM and correlation principle, the optimal hyperplane is proved as an advanced correlation filter when the kernel function of SVM is the linear kernel. So a new correlation filter, named linear SVM correlation filter (LSCF), is presented. The filter has not only shift-invariance, but SVM properties, such as good generalization. LSCF simultaneously has the distortion tolerance and good discrimination for different objects.

Supposing in space domain, the input image is denoted by $f(x,y)$, and the filter is denoted by $h(x,y)$. Correlation principle between them in space domain can be given as

$$f(x,y)\circ h(x,y)=\frac{1}{MN}\times\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}f(x+m,y+n)h^{*}(m,n), \quad (1)$$

where the superscript $*$ denotes the conjugate of a vector, and if h is the real function, then $h^{*}=h$. The size of h is $M\times N$. The correlation values corresponding to the locations are equal to the dot product between the filter and the sub-images. Among all the values, the maximum value denotes the sub-image corresponding to the location, which is optimally matched.

Supposing a set of labeled training patterns

$$(z_1,t_1),\cdots,(z_l,t_l),\quad z_i\in\{-1,+1\}.$$

The set is linearly separable if the optimal hyperplane $w\cdot t+b=0$ can be found. The optimal hyperplane means that it can correctly classify the training set, and the distance between the projections of the training vectors of two classes is maximal. The optimal hyperplane is a unique one. In order to find the optimal hyperplane, a quadratic programming problem is solved, which can be formulated as[9]

$$\min_{w,b}\text{imize}\,\psi(w)=\frac{1}{2}\|w\|^{2}, \quad (2)$$

$$\text{subject to } z_i\,(w\cdot t_i+b)\geq 1,\quad i=1,\cdots,l. \quad (3)$$

At last, the classify function based on the optimal hyperplane is denoted as

$$f(t)=\text{sgn}(\sum_{\text{SV}}z_i\alpha_i\Phi(t)\cdot\Phi(t_i)-b), \quad (4)$$

where the function Φ maps the input vector t into feature space vector, SV is abbreviation of support vector, sgn( ) is sign function. If $u>0$, $\text{sgn}(u)=1$; and if $u\leq 0$,

1671-7694/2007/090549-03

$\mathrm{sgn}(u) = -1$. We can subtract sgn( ) and b from Eq. (4), then

$$f'(t) = \sum_{\mathrm{SV}} z_i \alpha_i \Phi(t_i) \cdot \Phi(t). \quad (5)$$

Vapnik has proved that SVM need not compute the form of $\Phi(t_i) \cdot \Phi(t_j)$, and the kernel function can be denoted as dot product in high-dimensional Hilbert space:

$$\Phi(t) \cdot \Phi(t_i) = K(t, t_i). \quad (6)$$

At last, Eq. (5) can be rewritten based on Eq. (6) as

$$f'(t) = \sum_{\mathrm{SV}} z_i \alpha_i K(t_i, t). \quad (7)$$

At present, the types of kernel function have linear, polynomial, radial basis function (RBF), and multi layer perception. Specially, when the kernel is linear, the bridge between SVM and correlation is built, and Eq. (7) is written as

$$f'(t) = \sum_{\mathrm{SV}} z_i \alpha_i t_i \cdot t = w \cdot t. \quad (8)$$

Comparing Eq. (8) with Eq. (1), both of them carry out dot product. So, we can regard the optimal hyperplane w as the correlation filter,

$$w = h = \sum_{\mathrm{SV}} z_i \alpha_i t_i, \quad (9)$$

consequently, the filter h is named LSCF.

Substituting Eq. (9) into Eq. (1), then

$$f(x,y) \circ h(x,y) = \frac{1}{MN} \times \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{\mathrm{SV}} z_i \alpha_i t_i(m,n) \cdot f(x+m, y+n). \quad (10)$$

In addition, the subject condition can be described as

$$\begin{array}{ll} w \cdot t_i - b \geq 1 & z_i \in \{+1\}, \\ w \cdot t_i - b \leq -1 & z_i \in \{-1\}. \end{array} \quad (11)$$

Notice that if t_i belongs to the positive class, the correlation output ≥ 1, and if it belongs to the negative class, the correlation output ≤ -1. That is to say, the minimum difference of the decision output between the two classes is 2, which makes the filter have stable discrimination ability.

Laser radar can simultaneously produce intensity image and range image, and its resolution is high, which makes it be the sensor of automatic target recognition. Our team has developed the research work on the aspect of the image processing and target recognition for laser radar[10−14]. The images in Fig. 1 are collected by the real laser radar which is developed by us. Because of the limitation of experiment environment, the collected images are not rich on the different FOVs. Through rotating the image in Fig. 1 around z-axis, LSCF can be used to solve the in-plane rotation invariance.

Figure 1(a) is the picture of QinLin building which is far away from the real laser radar, Figs.1(b) and (c) are the intensity and range images of QinLin body, and Figs.1(d) and (e) are the images of QinLin top. They are regarded as two class targets, named "body" which is positive class and "top" which is negative class, respectively.

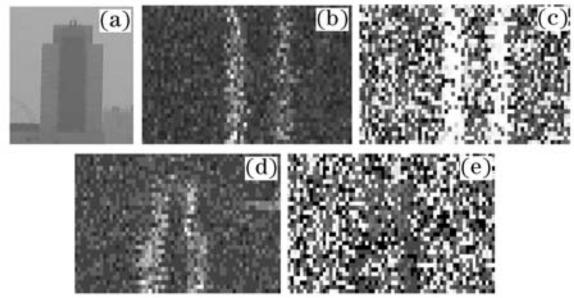

Fig. 1. Images of imaging laser radar. (a) Picture of QinLin building; (b) original intensity image; (c) original range image; (d) original intensity image; (e) original range image.

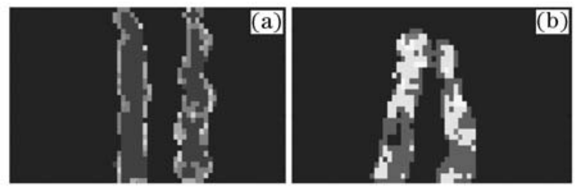

Fig. 2. Processed range images. (a) "Body" range image; (b) "top" range image.

Using the correlation relationship between intensity and range information, the range image is processed associating with intensity image. The pre-processed range images for Figs. 1(c) and (e) are shown in Fig. 2.

Artificially rotating the two objects in Fig. 2, the training images are taken at 5° intervals in aspect. After rotating 360°, each object can produce 72 training samples. Taking odd-number samples as training set, all samples are as testing set. Supposing the scene of laser radar is 150×150, it can include multiple samples. Taking two samples from each testing set, they are put on the scene at arbitrary position, and added background noise in the scene[15], as shown in Fig. 3. LSCF correlates with the input image, and the correlation output is shown in Fig. 4.

In Fig. 4, LSCF restrains the background noise, the correlation peaks of the positive class are maximal while the peaks of the negative class are minimal, and all sidelobes

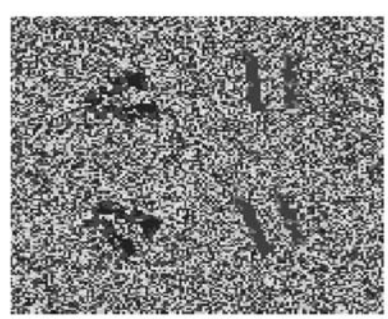

Fig. 3. Input image of laser radar.

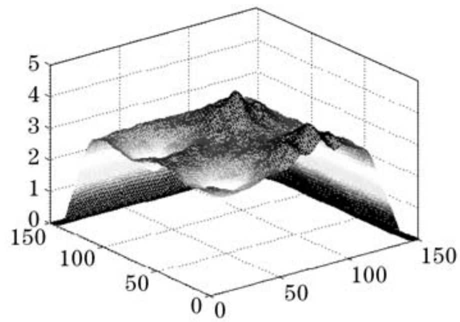

Fig. 4. Result of correlation output.

are not higher than the "body" peaks. LSCF detects or recognizes the testing set, the correlation output is similar to Fig. 4. That is to say, it can correctly recognize all the test samples of the two classes, consequently the recognition rate of LSCF is 100%.

Through changing the carry-to-noise ratio (CNR) of target and background, it is found that when the target CNR is less than the background CNR, LSCF can not detect or recognize the objects, and the image signal-to-noise ratio (SNR) is equal to −2.7 dB or so.

The factors, such as laser power being not stable, always influence on the range gray. Choosing 15 images whose gray changes constitute test set A, and the test set B can be obtained through preprocessing set A. Choosing 15 training samples constitute test set C. The correlation outputs for the test sets are shown in Fig. 5. Groups I and II label the correlation results for the test set A, groups III and IV label the correlation results for the test set B, groups V and VI label the correlation results for the test set C. LSCF has the stable correlation output for the three test sets and is not sensitive to the noise and the gray change. The noise increases the whole value of correlation output, however, it does not effect on the discrimination ability of the filter for the two targets.

The statistical analysis for the correlation curves in Fig. 5 is done, and the detailed data is shown in Table 1. It shows that the standard deviation for the curves is small, the maximum is 0.181, and the minimum is 0.006. Then, the swing change of correlation peak height

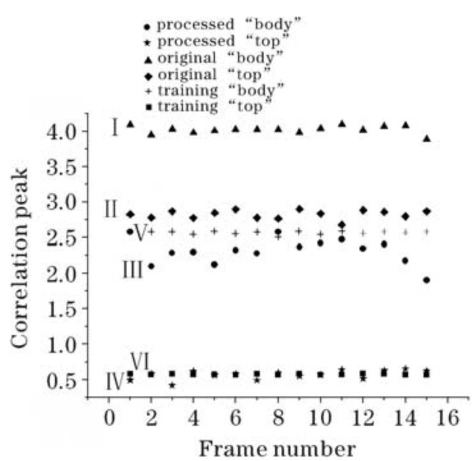

Fig. 5. Curve chart of correlation peak.

Table 1. Curve Analysis

Image		Mean	STD	Mean Difference
Original (Test Set A)	Body	4.01	0.055	1.19
	Top	2.82	0.061	
Processed (Test Set B)	Body	2.31	0.181	1.75
	Top	0.56	0.065	
Training (Test Set C)	Body	2.57	0.016	2.00
	Top	0.57	0.006	

is small, and correlation output is stable. The average peak height is 2 for the training samples, which tallies the expected theory. For the test set A and B, with gray change and noise, the average peak heights decrease 1.19 and 1.75, respectively, which states that the preprocessing can enhance the discrimination ability of the filter. If considering the real-time problems, the filter can directly correlate with the original image.

LSCF is a good correlation filter that it removes the restriction of a closed form solution, so it can solve the different types of distortion recognition, such as position, rotation, and scale invariance. The experiments verify the filter performance, and state that the filter has good discrimination ability, and is insensitive to the noise and gray change. LSCF can ensure the distance between the peaks of two classes, so its correlation output is stable. LSCF has the attractive attributes of correlation and SVM, which makes it have practical application value, such as embedding it into DSP.

J. Sun's e-mail address is hit_sunjianfeng@yahoo.com.cn.

References

1. B. V. K. V. Kumar, A. Mahalanobis, and A. Takessian, IEEE Trans. Image Processing **9,** 1025 (2000).
2. J. Yao, S.-M. Tan, and Y.-B. Liew, Opt. Eng. **41,** 81 (2002).
3. S. Roy, H. H. Arsenault, and D. Lefebvre, Opt. Eng. **42,** 813 (2003).
4. J. Sun, W. Lu, Q. Li, and Q. Wang, Proc. SPIE **6027,** 602731 (2006).
5. J. Sun, Q. Li, W. Lu, and Q. Wang, Chin. Opt. Lett. **5,** 118 (2007).
6. A. Mahalanobis, B. V. K. V. Kumar, S. Song, and S. R. F. Sims, Appl. Opt. **33,** 3751 (1994).
7. B. V. K. V. Kumar, Appl. Opt. **31,** 4773 (1992).
8. X. Zhang, Acta Automat. Sin. (in Chinese) **26,** 32 (2000).
9. V. N. Vapnik, IEEE Trans. Neural Networks **10,** 988 (1999).
10. Q. Li, Y. Wang, Q. Wang, and Z. Li, Acta Opt. Sin. (in Chinese) **25,** 581 (2005).
11. Z. Li, Q. Wang, Q. Li, and J. Sun, Chin. J. Lasers (in Chinese) **32,** 356 (2005).
12. Z. Li, Q. Li, Z. Tian, Y. Wang, J. Sun, W. Du, and Q. Wang, Chin. Opt. Lett. **2,** 210 (2004).
13. J. Sun, Q. Li, and Q. Wang, Proc. SPIE **5640,** 434 (2005).
14. J. Sun, Q. Li, W. Lu, and Q. Wang, Chin. J. Lasers (in Chinese) **33,** 1467 (2006).
15. T. J. Green, Proc. SPIE **3070,** 200 (1997).