Scale-space theory-based multi-scale features for aircraft classification using HRRP

Jia Liu✉, Ning Fang, Yong Jun Xie and Bao Fa Wang

High-resolution range profile is the significant characteristic of radar targets in automatic target recognition. Traditional feature extractions of range profiles in target classification are constrained to the original scale. This Letter proposes a multi-scale target classification method based on the scale-space theory. Target range profile feature is extended from single scale to multiple scales. The minimum Kullback–Leibler mean divergence (MKMD) algorithm is developed to achieve the automatic optimal scale factor selection. Classification evaluations on aircraft models using support vector machine and 3-nearest neighbour classifiers demonstrate that the application of scale-space theory in multi-scale feature extraction could effectively enhance the classification performance. The feasibility of the proposed MKMD algorithm is also validated by an enumeration method.

Introduction: Automatic target recognition [1, 2] is a critical area in radar research and application. High-resolution range profile (HRRP) is the significant radar characteristics to achieve target's real-time classification. Relative works in this area could be categorised into three trends. One focuses on the research of classification methods, which evolves from early template matching [3] to contemporary pattern recognition techniques [4]. Another one improves range profile imaging algorithm [5] to enhance HRRP's quality and robustness. The third type is interested in feature extraction of HRRP, which optimises the feature vector distribution pattern in feature spaces [6].

This Letter emphasises on HRRP's feature extraction. Current research achievements are constrained to the original scale feature extraction, which has more reliance on HRRP's local and detail characteristics. Global features are not well presented. This is more prominent for low-range resolution case and results in the information loss. Relative works are proved to be feasible in areas such as image processing and remote sensing [7]. We choose the scale-space theory [8] for HRRP's multi-scale feature extraction. Voting strategy is taken to synthesise the multiple scales' classification results. The minimum Kullback–Leibler mean divergence (MKMD) algorithm is proposed to achieve the scale factor selection. Classification results on aircraft models prove that multi-scale features could optimise classifier's performance. The MKMD algorithm is also validated by an enumeration experiment.

Multi-scale feature extraction of HRRP: Scale-space theory is a typical multi-scale technique. Signals' multi-scale presentation is realised by the unified signal sampling under the continuous scale factors. For a 2D image $I(u, v)$, its multi-scale extraction is

$$L(u,\ v,\ \sigma) = G(u,\ v,\ \sigma) \otimes I(u,\ v) \tag{1}$$

$\otimes$ is the convolution operator, u and v are spatial coordinates of image cells, and σ is the scale factor. $G(u,\ v,\ \sigma)$ is the scaling kernel

$$G(u,\ v,\ \sigma) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{u^2+v^2}{2\sigma^2}\right) \tag{2}$$

Formula (1) indicates that the scale feature extraction on $I(u, v)$ is accomplished by Gaussian kernel's smooth filtering. Gaussian kernel's low-frequency characteristic leads to the loss of high-frequency features, which could be detected by Laplace operator. Since Laplace operator is sensitive to the noise, second derivative operation is taken after the smooth filtering, which is usually realised by Gaussian kernel. This is the derivation of the Laplace of Gaussian (LOG) kernel

$$\Delta[G(u,\ v,\ \sigma) \otimes I(u,\ v)] = \mathrm{LOG} \otimes I(u,\ v) \tag{3}$$

Δ is the Laplace operator. The definition of LOG kernel is

$$\mathrm{LOG} = \Delta G(u,\ v,\ \sigma) = \frac{u^2+v^2-2\sigma^2}{\sigma^4}\exp\left(-\frac{u^2+v^2}{2\sigma^2}\right) \tag{4}$$

Peak and tough characteristics are predominant in classification, using Gaussian kernel in multi-scale feature extraction is improper. We choose 1D LOG kernel to achieve multi-scale feature extraction. Figs. 1*a* and 1*b* compare multi-scale features between Gaussian and LOG kernels. Gaussian kernel has smooth effect on HRRP. In contrast, LOG kernel well preserves the peak and tough features of the dominant scatters.

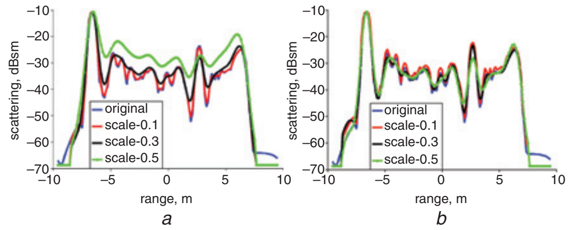

Fig. 1 *Multi-scale range profile features comparison*

a Multi-scale features using Gaussian kernel
b Multi-scale features using LOG kernel

Scale factor selection algorithm – MKMD: A scale factor selection algorithm named MKMD is proposed in this paper. The Kullback–Leibler divergence (KL divergence) describes range profile sample sets' statistical similarity. The algorithm is composed of three steps:

Step 1: Define the scale factor set $\Sigma = \{\sigma_1,\ \sigma_2,\ \ldots,\ \sigma_M\}$. This paper sets the scale factor range between 0.05 and 0.5 with interval of 0.05. Scale factor sub-spans $\Sigma_i (i = \{1,\ 2,\ \ldots,\ N\})$ evenly distribute within Σ.
Step 2: Compute the KL divergence of range profiles between original and other scales. Denote $\boldsymbol{Z}_0$ as range profiles under original scale, $\boldsymbol{Z}_s$ as the ones under scale s. The KL divergence is defined as

$$D_s(\boldsymbol{Z}_s,\ \boldsymbol{Z}_0) = \sum_a z_s(a)\ln\frac{z_s(a)}{z_0(a)} \quad z_s(a)\cdot z_0(a) \neq 0 \tag{5}$$

a is the sub-span of range profiles' amplitude information. $z_s(a)$ and $z_0(a)$ represent the distribution probability of $\boldsymbol{Z}_s$ and $\boldsymbol{Z}_0$ within a. The mean value of KL divergence is defined as

$$\overline{D} = \frac{1}{M}\sum_{s=1}^{M} D_s(\boldsymbol{Z}_s,\ \boldsymbol{Z}_0) \tag{6}$$

Step 3: Calculate the difference between $D_s(\boldsymbol{Z}_s,\ \boldsymbol{Z}_0)$ and $\overline{D}$ within each scale factor sub-span Σ_i. The scale factor with minimum difference is the optimum scale factor in Σ_i

$$\sigma_i = \arg\min_s \left|D_s(\boldsymbol{Z}_s,\ \boldsymbol{Z}_0) - \overline{D}\right| \quad s \in \Sigma_i \tag{7}$$

σ_i is the optimum scale factor of Σ_i. When KL divergences within Σ_i have apparent difference with $\overline{D}$, this criteria leads to the performance degradation. We define the deviation of 20% from $\overline{D}$ as the apparent difference. In this case, optimum scale factor selection jumps out of Σ_i, the searching extends to the neighbouring sub-span Σ_{i+1} (Σ_{i-1}). The scale factor σ'_{i+1} (σ'_{i-1}) is the optimum scale factor of Σ_i

$$\sigma_i = \sigma'_{i+1} = \arg\min_s \left|D_s(\boldsymbol{Z}_s,\ \boldsymbol{Z}_0) - \overline{D}\right| \quad s \in \Sigma_{i+1} \cap s \neq \sigma_{i+1} \tag{8}$$

The first reason of MKMD derivation is that larger-scale factor preserves more abundant low-frequency information. Selecting scale factors from evenly distributed Σ_i covers information from low-frequency to high-frequency bands. Secondly, this criterion guarantees the statistical similarity and relevance among multi-scale feature. It reduces the range profile distortion and reflects the difference of classification results among scale factors. These lead to the more comprehensive classification conclusions.

Multi-scale feature-based target classification: The classification method based on multi-scale features includes the following steps:

Step 1: Construct the training set $\boldsymbol{X}_0$ and testing set $\boldsymbol{Y}_0$. This paper uses graphical electromagnetic computing [9] to simulate target's range profiles. Range bins' position and scattering amplitudes are chosen as the elements of feature vector.
Step 2: Combine $\boldsymbol{X}_0$ and $\boldsymbol{Y}_0$ into $\boldsymbol{Z}_0$. Use the LOG kernel to extract the multi-scale features of $\boldsymbol{Z}_0$. The multi-scale sample set $\boldsymbol{Z}_s$ is constructed using the extracted range profile's feature vector under scale s.
Step 3: Compute the KL divergence $D_s(Z_s,\ Z_0)$ between $\boldsymbol{Z}_s$ and $\boldsymbol{Z}_0$, collect the optimum scale factors using the MKMD algorithm. We

choose $N = 4$ as illustration. The optimum scale set is defined as $\Sigma_{optimum} = \{\sigma_{I}, \ \sigma_{II}, \ \sigma_{III}, \ \sigma_{IV}\}$. Decompose the $\boldsymbol{Z}_s$ ($s \in \Sigma_{optimum}$) into the training set $\boldsymbol{X}_s$ and testing set $\boldsymbol{Y}_s$.

Step 4: Choose the classifier to achieve the multi-scale classification by learning $\boldsymbol{X}_s$ ($s \in \Sigma_{optimum}$) and $\boldsymbol{X}_0$. For the test sample $y = \{y_j | y_j \in Y_0 \cup Y_s, \ s \in \Sigma_{optimum}\}$, the multi-scale classification result is defined as: $R_y = \{r_0, \ \ldots, \ r_j, \ \ldots, \ r_{IV} | r_j \in D\}$, r_0 represents the classification result under the original scale, and D denotes the target class set.

Step 5: Chooses the class with maximum polls as the conclusion

$$w_y = \arg\max_r V_r \quad r \in D \tag{9}$$

V_r is the poll for class r. Repeat the steps 2–5 until the class with most polls is found.

Results and discussion: Three aircraft models are chosen to validate the classification method. The radar frequency is 10 GHz, bandwidth ranges from 100 MHz to 1 GHz. Vertically transmitting and vertically receiving polarisation is taken. The sample repetition rate between training and testing sample sets is 0.3%. The correct classification rate R_c is taken as the descriptor of classifier's performance.

Classification performance evaluation: Support vector machine (SVM) and 3-nearest neighbour (3-NN) classifiers are taken to evaluate the method. Fig. 2 is the classification comparison of single scale and multiple scales. The multi-scale feature optimises both classifiers' performance. SVM usually possesses better performance than 3-NN in classification, noise tolerance, and the robustness. This makes the degree of R_c improvement on SVM is not that prominent as it reflects on 3-NN. However, the complexity of SVM leads to its limitation in time efficiency, the multi-scale feature extraction aggravates the computation burden. Since 3-NN does not need the machine learning process, its efficiency superiority is obvious. Results prove the adaptability of multi-scale features. The degree of improvements has close relevance with the classifier. Therefore, the multi-scale classification method is more suitable for classifiers with higher efficiency.

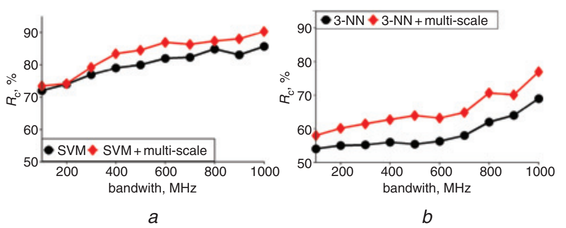

Fig. 2 *Classification performance comparison*
a SVM
b 3-NN

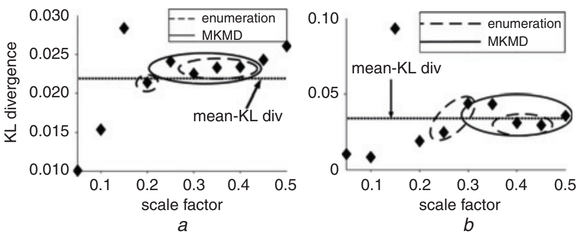

Fig. 3 *Scale factors comparison between enumeration and MKMD*
a 1 GHz bandwidth
b 500 MHz bandwidth

MKMD validation using enumeration algorithm: The enumeration algorithm is used to validate the MKMD algorithm. Enumeration algorithm randomly selects four scale factors from Σ and does the multi-scale classification. The process repeats until classification results from all scale factor combinations are collected. Scale factors with best classification results are compared with $\Sigma_{optimum}$ from MKMD. Figs. 3*a* and 3*b* are scale factors comparison between enumeration and the MKMD. Optimum factors of enumeration algorithm distribute around the mean value of KL divergence. This proves the reasonability of MKMD's scale factors selection criterion. Fig. 4 compares the results between enumeration and the MKMD algorithm. R_c values from enumeration have overall preponderance. However, the computation resource requirement of two methods are one and C_M^4. Considering the acceptable classification difference between two methods, the MKMD algorithm is feasible in multi-scale target classification method.

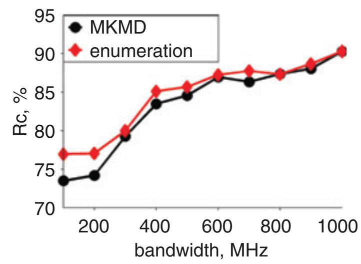

Fig. 4 *Classification results of MKMD and enumeration*

Conclusion: A multi-scale target classification method based on the scale-space theory is proposed. The MKMD algorithm is developed for scale factors selection. The voting strategy synthesises multi-scale classification results. Aircraft models are chosen to verify the proposed methods. Classification results indicate that multi-scale technology is helpful for classification performance improvement. The MKMD algorithm is also validated by the enumeration algorithm.

Submitted: *30 October 2015* E-first: *9 February 2016*
doi: 10.1049/el.2015.3583
One or more of the Figures in this Letter are available in colour online.

Jia Liu, Ning Fang, Yong Jun Xie and Bao Fa Wang (*Department of Electronic and Information Engineering, Beihang University, Beijing 100191, People's Republic of China*)

✉ E-mail: liujia1985219@gmail.com

References

1 Wu, R., Gao, Q., Liu, J., and Gu, H.: 'ATR scheme based on 1-D HRR profiles', *Electron. Lett.*, 2002, **38**, (24), pp. 1586–1588, doi: 10.1049/el:20021094
2 Peng, L., Wang, J., and Sun, J.P.: 'Classification of free rigid targets with micro-motions using inertial characteristic from radar signatures', *Electron. Lett.*, 2014, **50**, (13), pp. 950–952, doi: 10.1049/el.2014.1091
3 Li, H.J., and Yang, S.H.: 'Using range profiles as feature vectors to identify aerospace objects', *IEEE Trans. Antennas Propag.*, 1993, **41**, (3), pp. 261–268, doi: 10.1109/8.233138
4 Pilcher, C.M., and Khotanzad, A.: 'Maritime ATR using classifier combination and high resolution range profiles', *IEEE Trans. Aerosp. Electron. Syst.*, 2011, **47**, (4), pp. 2558–2573, doi: 10.1109/TAES.2011.6034651
5 Kim, D.H., Seo, D.K., and Kim, H.T.: 'Efficient radar target recognition using the MUSIC algorithm and invariant features', *IEEE Trans. Antennas Propag.*, 2002, **50**, (3), pp. 325–337, doi: 10.1109/8.999623
6 Du, L., Liu, H., Bao, Z., and Zhang, J.: 'Radar automatic target recognition using complex high-resolution range profiles', *IET Radar Sonar Navig.*, 2007, **1**, (1), pp. 18–26, doi: 10.1049/iet-rsn:20050119
7 Hong, T., Li, S., Qi, Y.F., Chen, Y.H., Yang, S., Li, J., and Clausi, D.A.: 'A multi-scale latent Dirichlet allocation model for object-oriented clustering of VHR panchromatic satellite images', *IEEE Trans. Geosci. Remote*, 2013, **51**, (3), pp. 1680–1692, doi: 10.1109/TGRS.2012.2205579
8 Lindeberg, T.: 'Scale-space theory: a basic tool for analysing structures at different scales', *J. Appl. Stat.*, 1994, **21**, (2), pp. 224–270, doi: 10.1080/757582976
9 Liu, J., Fang, N., Wang, B.F., and Zhang, L.Q.: 'An efficient ray-tracing method for RCS prediction in GRECO', *Microw. Opt. Technol. Lett.*, 2013, **55**, pp. 586–589, doi: 10.1002/mop.27349