Dictionary Learning-Based Hough Transform for Road Detection in Multispectral Image

Weifeng Liu, *Senior Member, IEEE*, Zhenqing Zhang, Xinghua Chen, Shuying Li, and Yicong Zhou, *Senior Member, IEEE*

***Abstract*— It is of great importance to determine the location and orientation of a straight road in multispectral images for remote sensing. One of the classical methods for straight line detection is the Hough transform that is widely used in binary images. Although there are many previous works for straight road detection, it is still in its infancy to extract a straight road in multispectral images for remote sensing. In this letter, we propose a multiview dictionary learning formulation to approximate the Hough transform for straight road detection in multispectral images. Our formulation can exploit the complementary among the multiple spectral channels. Furthermore, it is natural to incorporate regularizations of prior information to significantly leverage the performance. We consider L_1-norm regularization as a case study and conduct extensive experiments on RSSCN7 data set to verify the proposed algorithm. The experimental results demonstrate the superiority of our method in comparison with traditional methods.**

***Index Terms*— Dictionary learning, Hough transform, multispectral image, Radon transform (RT), road detection.**

I. Introduction

DETERMINING the location and orientation of straight road in remotely sensed images is a challenging and important research topic [1], which offers many supports for human civilization [2], such as visual tracking [3], robot autonomous navigation, city planning, traffic management, GPS navigation, and so on. There is a lot of noise in the road detection, so a robust road detection algorithm is very important [4]. One classical method for straight road detection is the Hough transform [5]. Despite that many variants of the Hough transform have been reported, including randomized Hough transform [6], probabilistic Hough transform [7], and fuzzy Hough transform [8], the Radon transform (RT) [9] is a prominent version of the Hough transform, which extends the concept of the Hough transform from binary images to grayscale images. Aggarwal and Karl [10] reformulated the detecting straight lines in grayscale images as a regularization framework. And then, Liu *et al.* [11] proposed a dictionary learning-based RT for straight road detection in remote sensing images.

On the other hand, most of the remotely sensed images are multispectral images [12] that capture image data within multiple specific wavelength ranges. Multispectral images represent more information with multiple wavelengths compared with traditional images. Since different spectral bands may reflect different characters, combining multiple spectral information can obtain abundant information and utilize the complementary between the multiple spectral bands [13], [14]. Luo *et al.* [15] proposed tensor canonical correlation analysis (TCCA), which handles the data of multiple views by analyzing the covariance tensor of the different views. It will significantly boost the straight road detection performance by properly exploiting the multiple information. Unfortunately, most of the existed methods can only be applied for single channel images (i.e., gray images). Hence, it is still in its infancy to extract a straight road in remote sensing images especially for multispectral images that capture image data within different wavelength bands.

In this letter, we reformulate the RT as a multiview dictionary learning framework to tackle the straight road detection in multispectral images. The proposed method treats the multispectral image as multiview representation and simultaneously learns the straight road mapping in parameter space and the related multispectral dictionary. Hence, it can exploit the complementary among the multiple spectral channels and then boost the straight road detection. Furthermore, it is natural to incorporate specific regularizations into the proposed framework, which can significantly leverage the performance with prior information. We also consider L_1-norm regularization as a case study. To evaluate the proposed solution, we conduct extensive experiments on the RSSCN7 data set [16] for straight road detection. The experimental results demonstrate that the proposed method can achieve better performance in comparison with traditional methods.

In summary, the contribution of the proposed framework is listed in the following.

1) It effectively utilizes the complementary of multiple spectral information. The complementary information in multiple spectral is fully coordinated by the multiview

Manuscript received July 14, 2017; revised September 19, 2017; accepted October 12, 2017. Date of publication November 7, 2017; date of current version December 4, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61671480 and in part by the Fundamental Research Funds for the Central Universities, China University of Petroleum, China, under Grant 14CX02203A. *(Corresponding author: Shuying Li.)*

W. Liu and Z. Zhang are with the College of Information and Control Engineering, China University of Petroleum, Qingdao 266580, China (e-mail: liuwf@upc.edu.cn).

X. Chen is with the College of Information and Control Engineering, China University of Petroleum, Qingdao 266580, China, and also with the Nanjing University of Science and Technology, Nanjing 210094, China.

S. Li is with The 16th Institute, China Aerospace Science and Technology Corporation, Xi’an 710100, China (e-mail: angle_lisy@163.com).

Y. Zhou is with the Faculty of Science and Technology, University of Macau, Macau 999078, China (e-mail: yicongzhou@umac.mo).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2017.2764042

sparse coding. And then, the latent representation in parameter space can be obtained for exact straight road detection.

2) It can handle a batch of images at one time. The conventional RT can process only one image each time. The proposed method makes it easy to compute the parameter space image for multiple images. Hence, it will greatly save the computing complexity for the processing of a tremendous amount of multispectral images.
3) It allows to integrate specific regularization naturally. There is much prior information in the parameter space, which can be utilized. For example, the peak point corresponding to the straight line is usually rare. Then, adding a sparsity regularization will further suppress the influence of noises and enhance the detection accuracy.
4) We conduct extensive experiments for straight road detection to verify the proposed method including the L_1 regularized multispectral dictionary learning-based RT (L1_mDLRT), the multispectral dictionary learning-based RT (mDLRT), the L_1 regularized single spectral dictionary learning-based RT (L1_sDLRT), the single spectral dictionary learning-based RT (sDLRT), and the conventional RTRT.

The rest of this letter is assigned as follows. Section II introduces some related works on RT. Then, Section III describes the proposed multiview dictionary learning-based RT in detail. Section IV conducts the straight road detection experiments with discussion. Section V gives the conclusion.

II. Related Works

In this section, we briefly introduce some related works including RT and its variants.

The Hough transform [5] transforms shape detection in a binary image to a voting procedure in a parameter space.

Lu and Tan [6] proposed a iterative randomized Hough transform to overcome the limitations of the traditional Hough transform in detecting incomplete shapes of noisy images.

Matas *et al.* [7] proposed a progressive probabilistic Hough transform to exploit the difference in the fraction of votes needed to reliably detect lines with different numbers of supporting points and then minimized the amount of computation.

Han *et al.* [8] proposed a fuzzy Hough transform to approximately fit the data points for shape detection.

The RT projects a grayscale image in the Euclidean plane to the parameter space (i.e., Radon space) where each point in Radon space corresponds to a line in image space [9]. Considering an image $y(s_1, s_2)$, we can write its corresponding Radon projection $x(\rho, \theta) = \mathcal{R}y$ as

$$\begin{aligned} x &= \mathcal{R}y(s_1, s_2) \\ &= \int_{R^2} y(s_1, s_2)\delta(\rho - s_1\cos(\theta) - s_2\sin(\theta))ds_1 ds_2 \end{aligned} \quad (1)$$

where δ is the Dirac delta function.

Defining the inverse transform of $\mathcal{R}$ as $\mathcal{C} = \mathcal{R}^{-1}$, then we have

$$\begin{aligned} y(s_1, s_2) &= \mathcal{C}x(\rho, \theta) \\ &= \int_0^{\pi} z(\rho, \theta)d\theta \end{aligned} \quad (2)$$

where $z(\rho, \theta) \doteq \int_{-\infty}^{\infty} |\omega| X(\omega, \theta)e^{j2\pi\omega\rho}d\omega$ and $X(\omega, \theta) = \int_{-\infty}^{\infty} x(\rho, \theta)e^{-j2\pi\omega\rho}d\rho$.

Then, Aggarwal and Karl [10] employed the discretized form (2) to approximate the relationship between image y and parameter space image x as

$$y = Cx \quad (3)$$

which allows naturally to incorporate regularizations with RT.

Recently, Liu *et al.* [11] proposed to learn the discretized matrix C by using dictionary learning [17] algorithm and achieved promising performance. This method treats the RT as a linear transform, which can obtain the result of the RT by matrix multiplication $x = \mathcal{C}^{-1}y$.

III. Multiview Dictionary Learning-Based Radon Transform

Suppose that we are given a multispectral image $y = \{y^{\upsilon}\}_{\upsilon=1}^{V}$ with V spectral channels. The RT-based straight road detection aims to find the related mapping in parameter space x and then to determine the location and orientation of the straight road. To effectively utilize the complementary of the multiple spectral, we present the multispectral RT as

$$y^{\upsilon} = C^{\upsilon}x, \quad \upsilon = 1, \dots, V$$

where C^{υ} is the dictionary corresponding to the υth spectral channel. Then, the multispectral RT can be reformulated as the following optimization problem:

$$x(\rho, \theta) = \arg\min_{X} \sum_{\upsilon=1}^{V} \|y^{\upsilon} - C^{\upsilon}x\|_2^2. \quad (4)$$

Formulation (4) can be easily extended by adding some regularizations with some specific prior information. And then, the regularized multispectral RT can be written as

$$x(\rho, \theta) = \arg\min_{X} \sum_{\upsilon=1}^{V} \|y^{\upsilon} - C^{\upsilon}x\|_2^2 + \lambda\varphi(x) \quad (5)$$

where λ is the balance parameter and $\varphi(x)$ is the regularization term to steer x under some specific purpose, e.g., sparsity constraint [18], smooth constraint, and so on.

To learn the dictionary of each spectral channel C^{υ}, we construct a training image set $Y^{\upsilon} = \{y_i^{\upsilon}\}_{i=1}^{N}$ with N samples and their related RT $X^{\upsilon} = \{x_i^{\upsilon}\}_{i=1}^{N}$. We have the following equation:

$$Y^{\upsilon} = C^{\upsilon}X^{\upsilon}. \quad (6)$$

Then, we can obtain the dictionary of the υth spectral channel C^{υ} as

$$C^{\upsilon} = Y^{\upsilon}X^{\upsilon T}(X^{\upsilon}X^{\upsilon T} + \alpha I)^{-1} \quad (7)$$

where α is a damping factor to ensure the stability and I is an identity matrix.

For a multispectral image $\hat{y} = \{y^{\upsilon}\}_{\upsilon=1}^{V}$, we now can compute its parameter image $\hat{x}$ by solving problem (4) or (5). Note that one advantage of the formulation lies in making it possible to detect the straight roads of multiple images at one time, which can significantly improve the

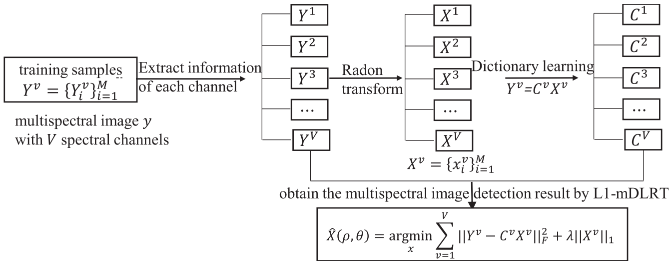

Fig. 1. Working framework of L1_mDLRT.

Fig. 2. Some example images with straight roads.

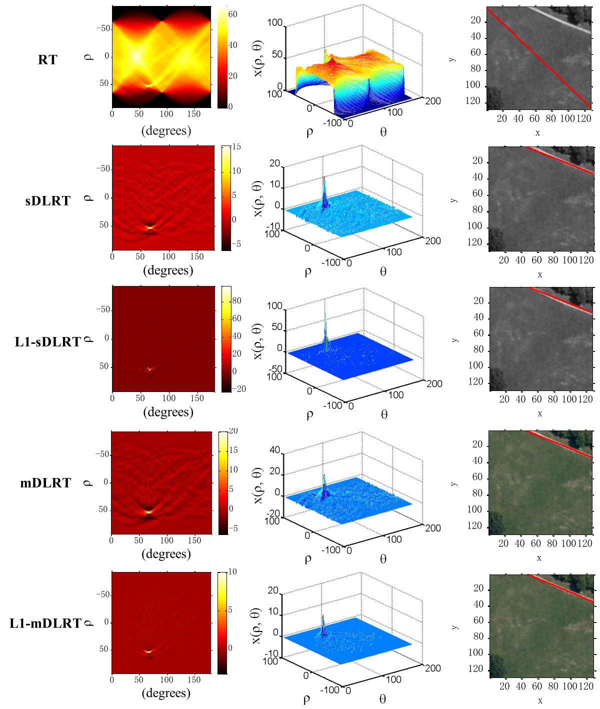

Fig. 3. Straight road detection on image a038 of different methods. (Left) RT of a038 in 2-D parameter space. (Middle) 3-D form of left. (Right) Detected line by different methods.

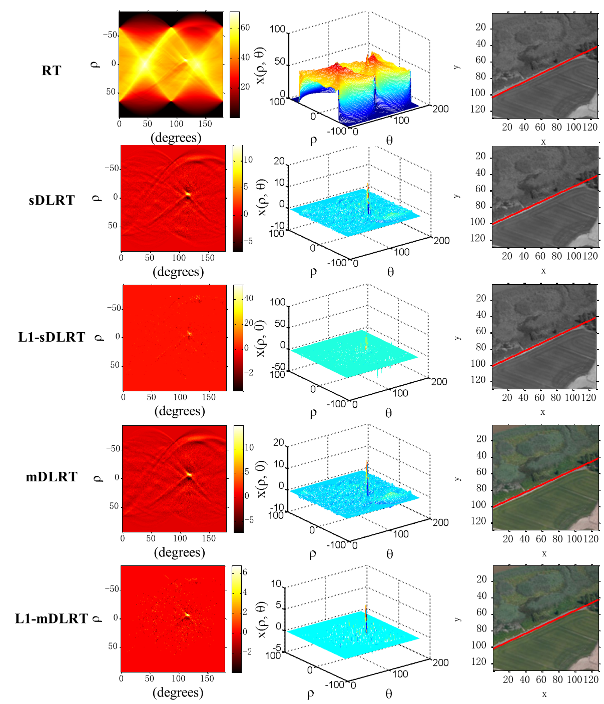

Fig. 4. Straight road detection on image a185 of different methods. (Left) RT of a185 in 2-D parameter space. (Middle) 3-D form of left. (Right) Detected line by different methods.

computing efficiency. When we are given a set of multispectral images $\hat{Y}^v = \{\hat{y}_i^v\}_{i=1}^M$, $v = 1, \ldots, V$ and their related RT X. We can achieve their related parameter images $\hat{X} = \{\hat{x}_i\}_{i=1}^M$ by optimizing the reformulated problems (4) and (5), respectively, as

$$\hat{X}(\rho,\theta) = \arg\min_X \sum_{v=1}^{V} \|\hat{Y}^v - C^v X\|_F^2 \tag{8}$$

and

$$\hat{X}(\rho,\theta) = \arg\min_X \sum_{v=1}^{V} \|\hat{Y}^v - C^v X\|_F^2 + \lambda\varphi(X). \tag{9}$$

For problem (8), we have a close solution

$$\hat{X} = (\boldsymbol{C}^T\boldsymbol{C} + \beta I)^{-1}\boldsymbol{C}^T\hat{\boldsymbol{Y}} \tag{10}$$

where

$$\boldsymbol{C} = \begin{pmatrix} C^1 \\ C^2 \\ C^V \end{pmatrix}, \quad \hat{\boldsymbol{Y}} = \begin{pmatrix} \hat{Y}^1 \\ \hat{Y}^2 \\ \vdots \\ \hat{Y}^V \end{pmatrix} \tag{11}$$

and β is the damping factor.

On the other hand, if we choose L_1 constraint for $\varphi(X)$, problem (9) can be rewritten as

$$\hat{X}(\rho,\theta) = \arg\min_X \sum_{v=1}^{V} \|\hat{Y}^v - C^v X\|_F^2 + \lambda\|X\|_1. \tag{12}$$

Then, we can use the fast iterative shrinkage-thresholding algorithm [19] to efficiently solve problem (12). Fig. 1 shows the working framework of L1_mDLRT. In Section IV, we will conduct the multiview dictionary learning-based RT and L_1 regularized case for straight road detection in multispectral images.

IV. Experiments

To evaluate the performance of the proposed method, we carry out extensive experiments for straight road detection on a select subset of the remote sensing database RSSCN7 [16]. The RSSCN7 database contains about 2800 remote sensing

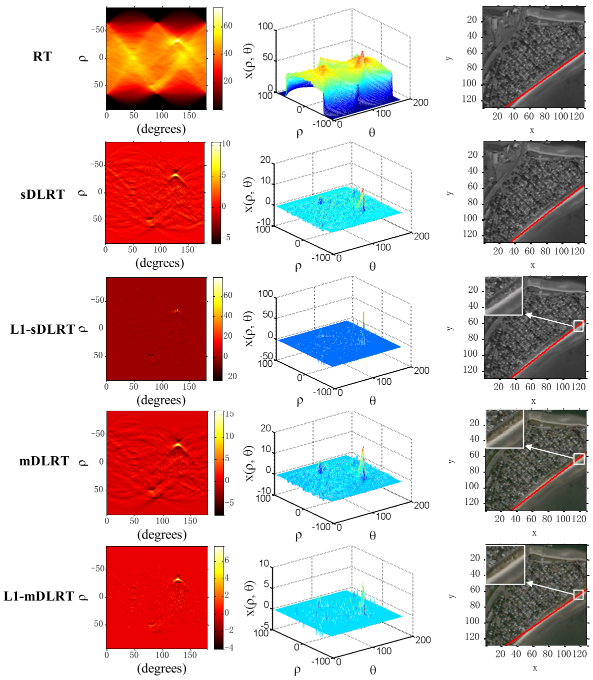

Fig. 5. Straight road detection on image f326 of different methods. (Left) RT of f326 in 2-D parameter space. (Middle) 3-D form of left. (Right) Detected line by different methods.

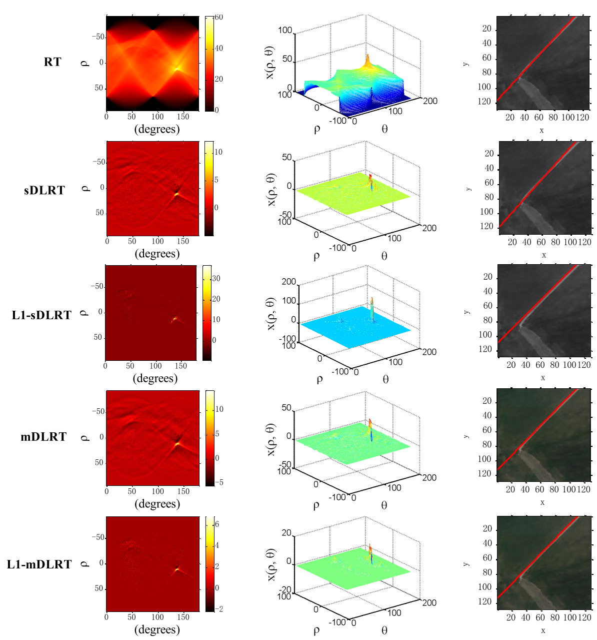

Fig. 6. Straight road detection on image b197 of different methods. (Left) RT of b197 in 2-D parameter space. (Middle) 3-D form of left. (Right) Detected line by different methods.

images that can be grouped into seven categories, including the grass land, forest, farm land, parking lot, residential region, industrial region, river, and lake. We select about 175 images that contain straight road for our experiments. Some image examples are shown in Fig. 2.

We choose 150 images for the multispectral dictionary learning and the rest for performance evaluation. Considering the small number of training set, we rotate the 150 images to get new examples and finally build a training set of 2850 images.

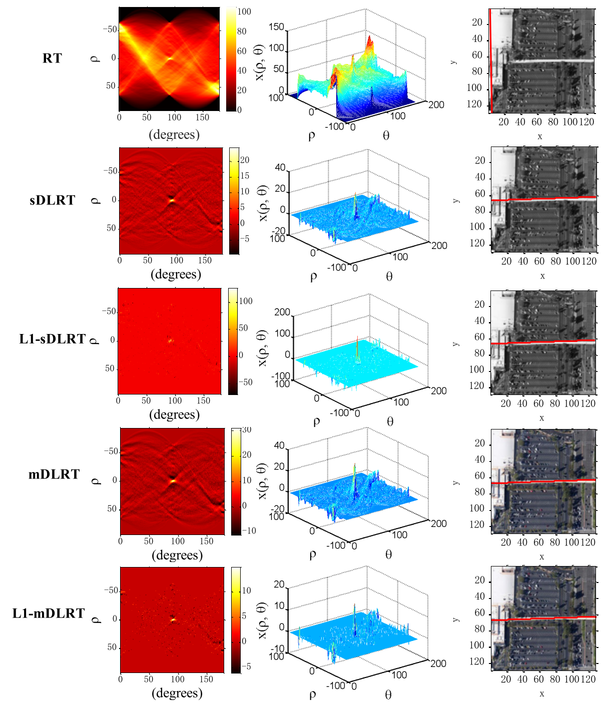

Fig. 7. Straight road detection on image g146 of different methods. (Left) RT of g146 in 2-D parameter space. (Middle) 3-D form of left. (Right) Detected line by different methods.

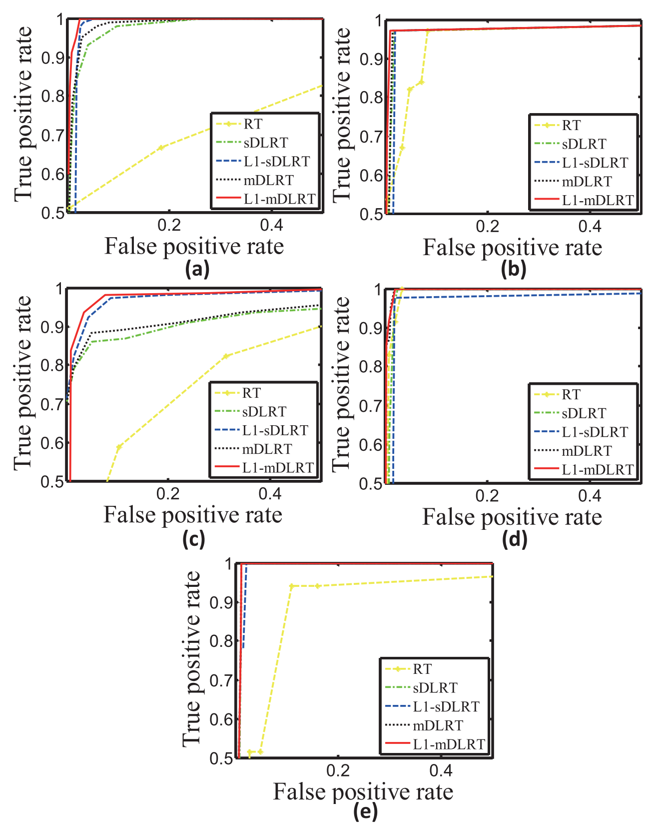

Fig. 8. (a) ROC curve on image a038. (b) ROC curve on image a185. (c) ROC curve on image f326. (d) ROC curve on image b197. (e) ROC curve on image g146.

TABLE I

LOCATION AND ORIENTATION (ρ, θ) OF DETECTED STRAIGHT ROAD OF DIFFERENT METHODS

Image	Groundtruth $(\rho, \theta/^\circ)$	RT	sDLRT	L1_sDLRT	mDLRT	L1_mDLRT
a038	(52.4, 67.115)	(52, 68)	(52,68)	(53, 68)	(52, 68)	**(52, 67)**
a185	(-7.94, 114.613)	(-8, 116)	(-7, 115)	(-7, 115)	(-7, 115)	**(-8, 115)**
b197	(11.34, 135.83)	(10, 137)	(10, 138)	(13, 135)	(11, 136)	**(11, 136)**
f326	(-32.45, 125.29)	(-32, 126)	(-33, 128)	(-33, 128)	(-33, 128)	**(-32, 125)**
g146	(-0.941, 91.6)	(-60, 1)	(0, 92)	(0, 92)	(-1, 92)	**(-1, 92)**

TABLE II

MEAN-ERROR AND VARIANCE OF FIVE METHODS

Parameter	RT	sDLRT	L1_sDLRT	mDLRT	L1_mDLRT
Mean error ρ	12.26	0.834	0.94	0.46	**0.262**
Mean error $\theta/^\circ$	18.95	1.31	1.04	0.91	**0.312**
Variance of error ρ	547.67	0.11	0.16	0.084	**0.14**
Variance of error $\theta/^\circ$	1283.47	4.56	3.69	0.86	**0.03**

For comparison, we conduct straight road detection by using L1_mDLRT, mDLRT, L1_sDLRT, sDLRT, and the conventional RT.

Fig. 3 shows the straight road detection results on image a038 of different methods, including RT, sDLRT, L1_sDLRT, mDLRT, and L1_mDLRT. Note that the first three algorithms, i.e., RT, sDLRT, and L1_sDLRT can only be applied on a single spectral image. The first and second columns demonstrate the mapped parameter space in 2-D and 3-D, respectively. The last column is the straight road detection. To demonstrate the robust of different methods, we draw the ROC curve of image a038 in Fig. 8(a). From Figs. 3 and 8(a), we can see that: 1) the dictionary leaning-based RT performs better than the conventional RT; 2) multiple spectral information leverages the accuracy; and 3) L_1 regularization can suppress the noise and enforce the detection. Some other exemplar results are showed in Figs. 4–8. The location and orientation of the detected straight road in exemplar images are listed in Table I for comparison. Table II illustrates the mean-error and variance of the five methods.

V. CONCLUSION

It plays a key role to determine the location and orientation of a straight road in multispectral images for remote sensing. Although there are many previous works for straight road detection, there are many unsolved problems to extract straight road in remotely sensed multispectral images. In this letter, we propose a multiview dictionary learning formulation to approximate RT for straight road detection in multispectral images. The proposed method can tackle the straight road detection in multispectral images by exploiting the complementary among the multiple spectral channels. Another advantage lies that it facilitates incorporating regularizations into RT. We further construct L_1-norm regularized multiview dictionary learning as case study and conduct extensive experiments on RSSCN7 data set to verify the proposed algorithm. The experimental results demonstrate that our algorithm is superior to the baseline methods including the traditional Hough transform.

REFERENCES

[1] Q. Wang, J. Fang, and Y. Yuan, "Adaptive road detection via context-aware label transfer," *Neurocomputing*, vol. 158, pp. 174–183, Jun. 2015.

[2] Y. Wei, Z. Wang, and M. Xu, "Road structure refined CNN for road extraction in aerial image," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 709–713, May 2017.

[3] Q. Wang, J. Fang, and Y. Yuan, "Multi-cue based tracking," *Neurocomputing*, vol. 131, pp. 227–236, May 2014.

[4] T. Liu and D. Tao, "Classification with noisy labels by importance reweighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 3, pp. 447–461, Mar. 2016.

[5] V. C. H. Paul, "Method and means for recognizing complex patterns," U.S. Patent 3 069 654, Dec. 18, 1962.

[6] W. Lu and J. Tan, "Detection of incomplete ellipse in images with strong noise by iterative randomized Hough transform (IRHT)," *Pattern Recognit.*, vol. 41, no. 4, pp. 1268–1279, Apr. 2008.

[7] J. Matas, C. Galambos, and J. Kittler, "Robust detection of lines using the progressive probabilistic Hough transform," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 119–137, Apr. 2000.

[8] J. H. Han, L. Kóczy, and P. Timothy, "Fuzzy Hough transform," *Pattern Recognit. Lett.*, vol. 15, no. 7, pp. 649–658, Jul. 1994.

[9] S. R. Deans, "Hough transform from the radon transform," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, no. 2, pp. 185–188, Mar. 1981.

[10] N. Aggarwal and W. C. Karl, "Line detection in images through regularized Hough transform," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 582–591, Mar. 2006.

[11] W. Liu, Z. Zhang, S. Li, and D. Tao, "Road detection by using a generalized Hough transform," *Remote Sens.*, vol. 9, no. 6, p. 590, Jun. 2017.

[12] Q. Wang, J. Lin, and Y. Yuan, "Salient band selection for hyperspectral image classification via manifold ranking," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 6, pp. 1279–1289, Jun. 2016.

[13] C. Xu, D. Tao, and C. Xu, "Multi-view intact space learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 12, pp. 2531–2544, Dec. 2015.

[14] Q. Wang, Z. Meng, and X. Li, "Locality adaptive discriminant analysis for spectral-spatial classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 2077–2081, Nov. 2017.

[15] Y. Luo, D. Tao, K. Ramamohanarao, C. Xu, and Y. Wen, "Tensor canonical correlation analysis for multi-view dimension reduction," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 11, pp. 3111–3124, Nov. 2015.

[16] Q. Zou, L. Ni, T. Zhang, and Q. Wang, "Deep learning based feature selection for remote sensing scene classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2321–2325, Nov. 2015.

[17] D. Tao, X. Li, X. Wu, and S. J. Maybank, "Geometric mean for subspace selection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 260–274, Feb. 2009.

[18] J. Yu, Y. Rui, and D. Tao, "Click prediction for Web image reranking using multimodal sparse coding," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 2019–2032, May 2014.

[19] W. Liu, Z.-J. Zha, Y. Wang, K. Lu, and D. Tao, "p-Laplacian regularized sparse coding for human activity recognition," *IEEE Trans. Ind. Electron.*, vol. 63, no. 8, pp. 5120–5129, Aug. 2016.