

Research Article

Aircraft Detection from VHR Images Based on Circle-Frequency Filter and Multilevel Features

Feng Gao,^{1,2} Qizhi Xu,^{1,2} and Bo Li^{1,2}

¹ Beijing Key Laboratory of Digital Media, School of Computer Science and Engineering, Beihang University, Beijing 100191, China

² State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China

Correspondence should be addressed to Bo Li; boli@buaa.edu.cn

Received 16 July 2013; Accepted 7 August 2013

Academic Editors: S. Bourennane, C. Fossati, and J. Marot

Copyright © 2013 Feng Gao et al. 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.

Aircraft automatic detection from very high-resolution (VHR) images plays an important role in a wide variety of applications. This paper proposes a novel detector for aircraft detection from very high-resolution (VHR) remote sensing images. To accurately distinguish aircrafts from background, a circle-frequency filter (CF-filter) is used to extract the candidate locations of aircrafts from a large size image. A multi-level feature model is then employed to represent both local appearance and spatial layout of aircrafts by means of Robust Hue Descriptor and Histogram of Oriented Gradients. The experimental results demonstrate the superior performance of the proposed method.

1. Introduction

With the advent of the very high-resolution (VHR) earth observation satellite programs, the spatial resolution of remote sensing images dramatically increased from tens of meters to tens of decimeters, such as QuickBird, GeoEye-1, and WorldView-2. High-resolution images of a small number of locations are publicly available via Google Earth at an astonishing ground sampling distance (GSD) of 0.15 meter [1]. The high-resolution satellite sensors can acquire high definition images of ground objects with abundant spatial details and contextual information, which make it feasible to detect and recognize artificial targets, such as aircraft [2, 3], vehicle [1, 4, 5], ship [6], and building [7, 8]. In this paper, we concentrate on the problem of detecting aircrafts from such high-resolution aerial and satellite imagery. Aircraft automatic detection from VHR images plays an important role in a wide variety of applications. Detecting and tracking aircrafts in aerial videos is an important component in visual surveillance systems. Images of military airport, along with the distribution of aircrafts, can provide valuable information for military monitoring. However, while image resolution upgrades to decimeter level, background of ground targets

becomes complex and disturbing. It is therefore a challenging task to detect ground targets in a cluttered background. The presence of complex structures, such as buildings and airport terminals, can cause many false alarms.

Various methods have been developed for object detection from remote sensing images. For example, Lei et al. [9] proposed a color-enhanced rotation-invariant Hough forest method for detecting aircrafts and buildings. Liu et al. [2] presented a coarse-to-fine method by integrating shape prior; the pose of an aircraft is roughly estimated by template matching, and a shape prior is used to segment the target. Li et al. [10] proposed detecting aircrafts using a contour-based spatial model. Inglada [11] proposed a supervised learning method which used geometric image features to characterize classes of objects. Akçay and Aksoy [12] presented a geospatial object detection method by using hierarchical segmentation which combined spectral and structural information. Recently, there are also some methods concentrating on detecting objects in synthetic aperture radar (SAR) image [13, 14] and hyperspectral image [15, 16].

Another group of methods utilizes a sliding window approach. In this approach, a classifier is first trained to recognize an object; a detection window is scanned over

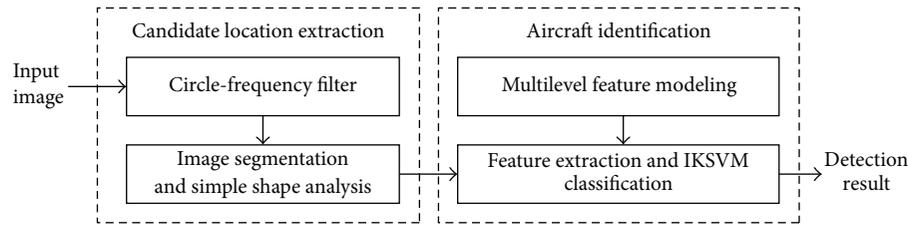


FIGURE 1: Flow diagram of the proposed method.

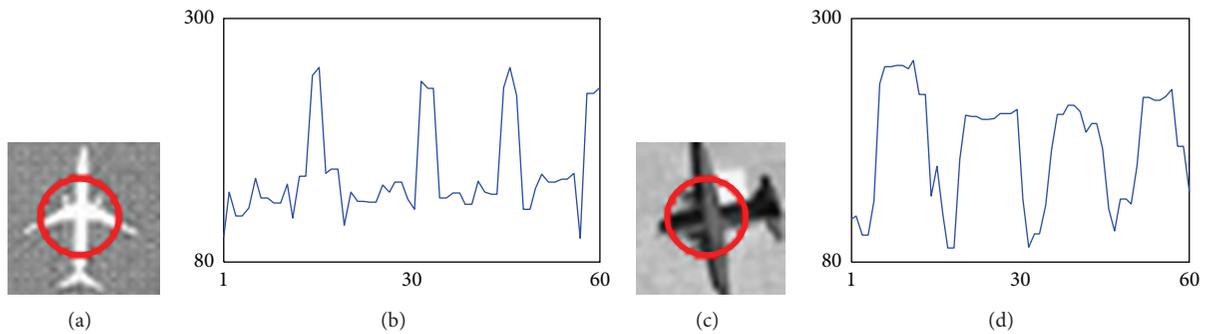


FIGURE 2: Illustration of the aircraft regular patterns. (a) Bright aircraft (32×32 pixels); (b) regular pattern of the bright aircraft; (c) dark aircraft (32×32 pixels); (d) regular pattern of the dark aircraft. The intensities along the red circle will change regularly from darkness to brightness and will repeat four times.

the image; then the classifier is used to determine whether the detection window contains an object or not. Kembhavi et al. [1] proposed a vehicle detector by using partial least square analysis to project high-dimensional features to a much lower dimensional subspace, and then the computational burden can be reduced. Sun et al. [17] presented an aircraft detector which applied a mapping strategy to encode the geometric information in the detection window. Zhang et al. [18] proposed a method to detect aircrafts by encoding the features of the rotated parts and objects. These methods rely on intensity-based features to capture the appearance of the object and obtain good results. However, all the pixels contained in the image are generally searched in these methods, thus resulting in a number of false positive detections. Moreover, due to the variability of color in a cluttered background, color cues have been ignored by most of these detection methods.

The proposed method improves the existing methods by exploiting circle-frequency filter (CF-filter) [3] to extract potential location of aircraft targets from the entire image. Consequently, most of the cluttered background regions are eliminated, and false alarms are therefore greatly reduced. Moreover, a multilevel feature model that incorporates both color and shape information is designed to identify aircrafts by checking potential locations. Our method includes a coarse-to-fine search process. The flow diagram of our method is shown in Figure 1.

Stage 1 (candidate location extraction by circle-frequency filter). Inspired by [3], CF-filter is applied to extract the candidate locations of aircraft. The goal of candidate extraction is

to reduce the false positives, for example, man-made objects that may be detected as aircrafts, and false negatives.

Stage 2 (aircraft identification by multilevel features). Robust Hue Descriptor [19] and Histogram of Oriented Gradients [20] are used to capture color and shape information of aircrafts, respectively. Meanwhile, a multilevel feature model is constructed to represent both local appearance and spatial layout of aircrafts.

We implement the proposed method on remote sensing images collected from aerial and WorldView-2 satellite images. Our methods are also compared with two previously proposed object detection methods, and experimental results demonstrate the superior performance of the proposed method.

2. Extraction of Aircraft Candidate Locations

The candidate location extraction is comprised of two steps: circle-frequency filter (CF-filter) and candidate locations extraction. To facilitate discussion, we begin this section with a general introduction to CF-filter. We follow that with an introduction on how candidate locations can be extracted.

2.1. Circle-Frequency Filter. The candidate extraction is based on two common features of aircraft as follows. First, airplanes and their background have difference in brightness. Second, airplanes are comprised by four main bulges [3], that is, head, tail, and two wings. If a circle is located at the center of an aircraft, the intensities along the circle will change

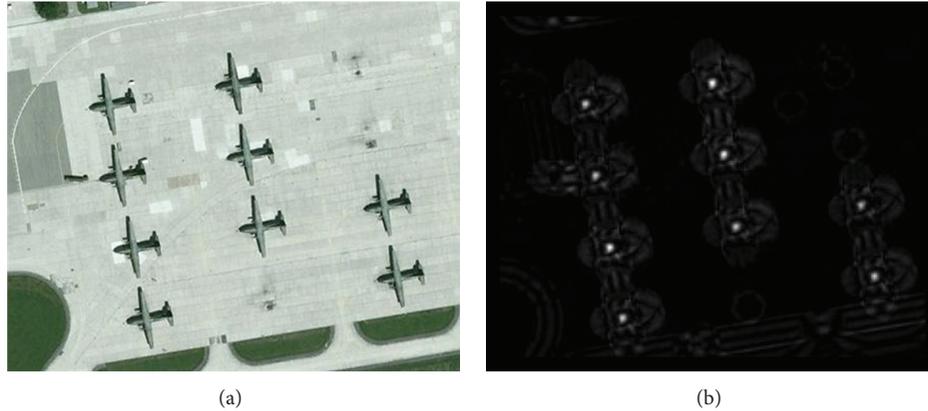


FIGURE 3: Illustration of CF-filter. (a) Original image; (b) CF-filtered result.

regularly from darkness to brightness and will repeat four times. Figures 2(a) and 2(c) are two examples of typical airplanes, that is, civil and military aircraft. In general, civil aircraft is brighter than background, while military aircraft is darker than background. Figures 2(b) and 2(d) illustrate the regular patterns of the two aircrafts.

Assume that f_i ($i = 0, \dots, N - 1$) are pixel values along the circle centered at (x, y) with radius r , the CF-filter can be expressed as

$$f(x, y) = \left(\sum_{k=0}^{N-1} f_k \cos\left(8\pi \frac{k}{N}\right) \right)^2 + \left(\sum_{k=0}^{N-1} f_k \sin\left(8\pi \frac{k}{N}\right) \right)^2. \quad (1)$$

If the whole image is filtered by CF-filter, a bright spot will be generated at the center of an airplane. Figure 3 gives an example of a CF-filtered image, where the values of CF-filtered image are normalized to $(0, 255)$. It can be seen from Figure 3 that only the center of airplane has a strong response. The CF-filter is rotation invariant; thus, airplanes with different layouts can be detected by CF-filter.

2.2. Aircraft Candidate Locations Extraction. As mentioned before, only the center of aircraft has a strong response in the CF-filtered image. Therefore, aircraft candidates can be extracted by removing low values of CF-filtered image. Here, Otsu's method [21] is used to calculate a global adaptive threshold T from CF-filtered image to remove low values, most of which are produced by clutters within background.

According to the specific size of aircraft candidate (bright spot), a simple shape analysis process is applied to eliminate obvious false candidates. The pixels remained are considered as true candidates; then these regions are forwarded to the second stage.

3. Identifying Aircrafts by Multilevel Features

The appearance of an aircraft can be better captured when color and shape information is combined together. In the second stage, two classes of feature are involved in the proposed solution: Robust Hue Descriptor and Histogram

of Oriented Gradients. We build a multilevel feature model, which is akin to the image pyramid representation proposed by Lazebnik et al. [22], to represent both local appearance and spatial layout of an aircraft.

3.1. Robust Hue Descriptor. Color is an indispensable aspect in describing the world around us. However, in a cluttered background, the color representation of an object is greatly interfered by photometric variance; thus, most of existing methods avoid exploiting color information. To obtain a good color descriptor, Van De Weijer et al. [19] proposed robust hue (HUE) descriptor. Experiments show that this descriptor is reliable under photometric and geometrical changes in image retrieval and classification. The HUE descriptor is represented by a histogram over hue computed from the corresponding RGB values of each pixel according to

$$\text{hue} = \arctan\left(\frac{\sqrt{3}(R - G)}{R + G - 2B}\right). \quad (2)$$

The HUE descriptor is invariant with respect to lighting geometry. In our implementation, the HUE descriptor has 6 dimensions.

3.2. Histogram of Oriented Gradients. Histograms of Oriented Gradients (HOG) [20] are used to capture the spatial distribution of gradients. Since the HOG operator has been proven to be robust and reliable for representing the shape of an object, it is therefore involved into our feature model to describe the shape information of aircrafts. In Dalal's work, the detection window of the HOG operator is split into small cells, in which a 9-bin histogram of gradient orientations is calculated. Every 2×2 cells are grouped together to form a block; thus, each block has a 36-bin histogram feature. Four blocks of features are then concatenated to form the HOG descriptor. Maji et al. [23] proposed a Spatial Histogram of Oriented Gradients (SPHOG) descriptor. Compared to Dalal's HOG descriptor, Maji's descriptor computes the oriented edge energy response by the magnitude of odd elongated oriented filters. Experiments show that, while combining with intersection kernel SVM (IKSVM),

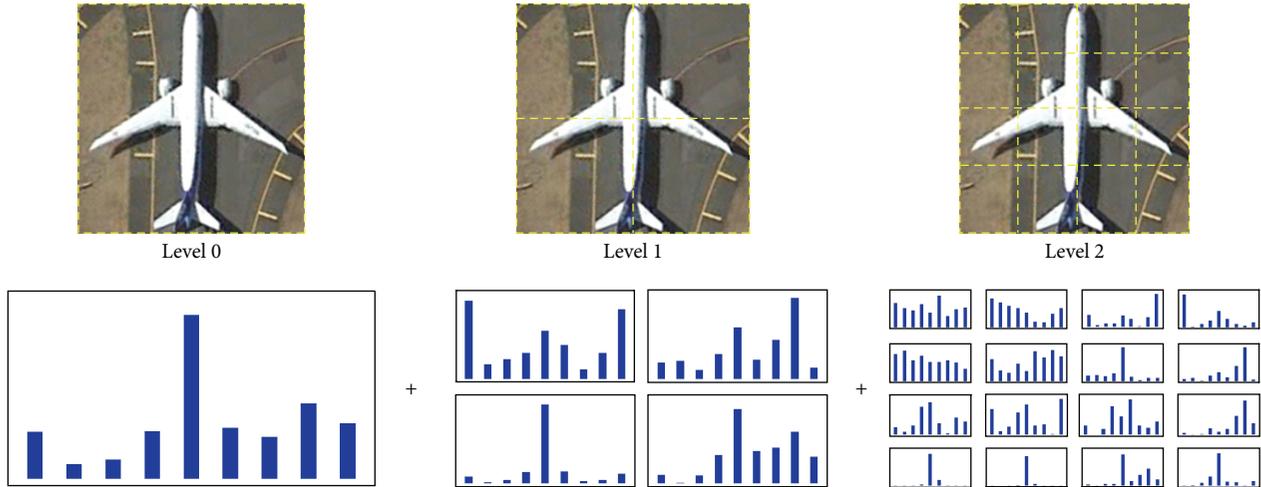


FIGURE 4: Multilevel feature representation.

the SPHOG descriptor outperforms the original single scale HOG. Hence, we follow Maji's method and use the odd elongated oriented filters to compute oriented edge energy and SPHOG descriptor.

3.3. Multilevel Feature Model. We build a multilevel feature model to represent the color and shape information of aircrafts. The detection window W_i is decomposed into a sequence of increasingly finer spatial cells by repeatedly doubling the number of decompositions. For cell C_i , the feature representation is obtained by concatenation:

$$C_i = [\text{HUE}_i, \text{HOG}_i]. \quad (3)$$

The feature vector for each cell is a 15-dimensional vector (HUE descriptor has 6 dimensions, and HOG descriptor has 9 dimensions). The number of pixels assigned to a specified color bin or gradient bin is the sum over those contained in the four cells it is divided into at a finer level. Hence, the multilevel feature model is a pyramid representation. The final features are the concatenations of all the 15-dimensional feature vectors. As shown in Figure 4, the cell at level l has 2^l cells along each dimension. Therefore, level 0 is represented by a 15-dimensional vector, level 1 by a 60-dimensional vector, and so forth. Total features of the detection window are a vector with dimensionality $15 \times \sum_{l \in L} 4^l$. Through a multilevel representation, the feature model captures both the local image appearance and its spatial layout. In our implementation, we use a feature model with three levels, yielding a total of 21 cells.

In this paper, SVM with histogram intersection kernel (IKSVM) [23] is applied as a classifier. For feature vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$, the intersection kernel $k(\mathbf{x}, \mathbf{y})$ can be expressed as

$$k(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \min(x(i), y(i)), \quad (4)$$

and the classification is based on evaluating

$$h(\mathbf{x}) = \sum_{i=1}^m \alpha_i y_i k(\mathbf{x}, \mathbf{x}_i) + b. \quad (5)$$

Maji et al. [23] noticed that the summation in (5) can be reformed as

$$\begin{aligned} h(\mathbf{x}) &= \sum_{i=1}^n \alpha_i y_i \left(\sum_{l=1}^m \min(x(i), x_l(i)) \right) + b \\ &= \sum_{i=1}^n \left(\sum_{l=1}^m \alpha_i y_l \min(x(i), x_l(i)) \right) + b \\ &= \sum_{i=1}^n h_i(x(i)) + b. \end{aligned} \quad (6)$$

Consider the functions $h_i(s)$ at fixed point i ; $\bar{x}_l(i)$ represents the sorted values of $x_l(i)$ in increasing order with corresponding values of α and labels given by $\bar{\alpha}_l$ and \bar{y}_l . Let r be the largest integer, such that $\bar{x}_r(i) \leq s$; then we can get

$$h_i(s) = \sum_{l=1}^m \alpha_l y_l \min(s, x_{l,i}) \quad (7)$$

$$= \sum_{1 \leq l \leq r} \bar{\alpha}_l \bar{y}_l \bar{x}_l(i) + s \sum_{r < l \leq m} \bar{\alpha}_l \bar{y}_l \quad (8)$$

$$= A_i(r) + s B_i(r), \quad (9)$$

where $A_i(r) = \sum_{1 \leq l \leq r} \bar{\alpha}_l \bar{y}_l \bar{x}_l(i)$, $B_i(r) = \sum_{r < l \leq m} \bar{\alpha}_l \bar{y}_l$. It is obvious that (8) is piecewise linear and the function A_i and B_i are independent of the input data. Therefore, s can be precomputed by finding the position of $s = x(i)$ in the sorted list. Maji et al. [23] noticed that the support distributions in each dimension tend to be smooth and concentrated. Therefore, $h(x)$ can be approximated by simpler functions, and the prediction can be accelerated. In this paper, $h_i(s)$ is computed by a look-up table with a piecewise constant approximation.

4. Experimental Results and Evaluation

4.1. Data Collection and Training. From the publicly available Google Earth service and WorldView-2 satellite images, we



FIGURE 5: Examples of positive and negative image patches of training dataset.

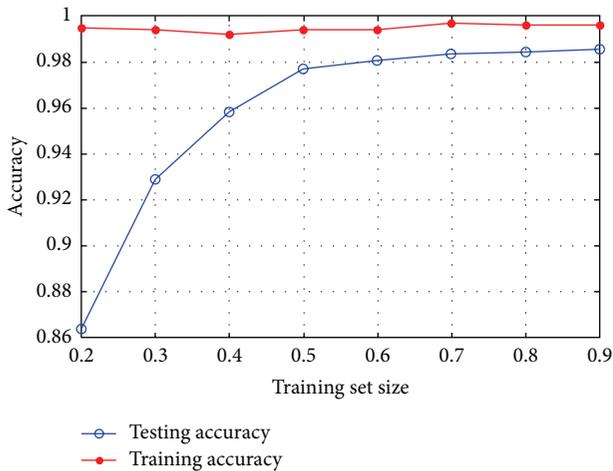


FIGURE 6: Cross-validation results.

collected 914 positive image patches (aircrafts) and 1953 negative image patches. Negative image patches are randomly chosen from the airport backgrounds. Each image patch has a size of 32×32 pixels. Figure 5 shows parts of training image patches of the dataset. We build our testing set by collecting 214 images of other airports containing 597 aircrafts in China, Germany, and France. Each image contains multiple instances of aircrafts with different orientations and sizes.

The IKSVM [23] is applied as a classifier in our method. To find an optimal training set size, we carried out cross-validation. A portion of the training set is selected to train the IKSVM classifier, and then the performance of the classifier is tested by the remaining images in the training set. This process is repeated, and average accuracy is calculated. The average training accuracy and testing accuracy versus training set size are illustrated in Figure 6. The training accuracy is high for all training set sizes, indicating that the multilevel feature model is effective to distinguish the images in the training set. The testing accuracy tends to be stable at

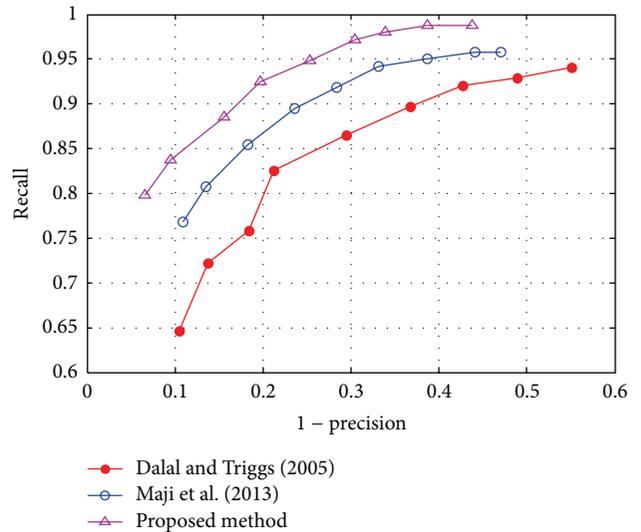


FIGURE 7: Performance of the three detectors.

a high rate when the number of training images is more than half of training set, which indicates that our method is capable of representing aircraft from a small training set.

4.2. Quantitative Evaluation. We manually label the aircrafts appearing in all testing images as a ground truth. When a detection to be marked is a true positive, more than 50% of it must be detected. M denotes the total number of aircrafts in testing images. The recall-precision curve (RPC) is chosen to exhibit the tradeoff between recall and precision. Recall and $1 - \text{precision}$ are defined as

$$\text{Recall} = \frac{TP}{(TP + FN)},$$

$$1 - \text{Precision} = \frac{FP}{(TP + FP)},$$
(10)

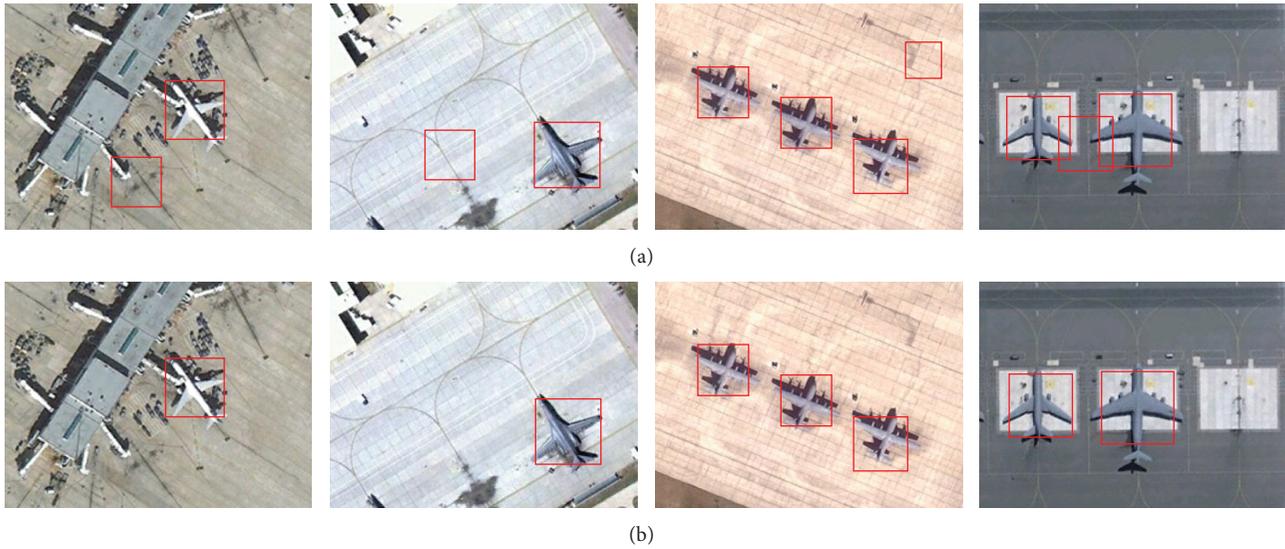


FIGURE 8: The effectiveness of CF-filter. (a) Detection results without CF-filter; (b) detection results with CF-filter.

where TP (true positive) denotes the number of true detected aircrafts, FN (false negative) is the number of missed detections, and FP (false positive) is the number of false detected aircrafts. Therefore, recall denotes the number of true detected aircrafts divided by the total number of aircrafts in testing images and $1 - \text{precision}$ denotes the number of false detected aircrafts divided by the elements detected. The ratio $\text{recall}/(1 - \text{precision})$ is used to represent the performance of the algorithm.

Since the size of aircraft in the testing set is unknown, the testing images are scanned at multiple scales. At each scale, the radius r and the number of pixels N in CF-filter are set to 8 and 60, respectively. After candidate extraction, we use the sliding window approach to detect aircrafts. The size of the sliding window is 32×32 pixels, and the step is set to be 8 pixels.

We compared our method with two classical methods. The first method is proposed by Dalal and Triggs [20], in which a linear Support Vector Machine is exploited to classify aircrafts and background by HOG features. The second method is Maji's approach [23], where the IKSVM is applied as a classifier and spatial histograms of oriented gradients (SPHOG) are calculated for the classifier. From Figure 7, we can observe that the proposed method outperforms the other two detectors. At the same recall level, the precision of the proposed method is the best, which means that the false alarms rate of the proposed method is the lowest. While at the same precision level, the proposed method detects more true positives.

In order to test the impact of CF-filter, our original method is compared to the modified method without CF-filter, see Figure 8. The first row is the experimental results without CF-filter and the second row is the results with CF-filter. It is obvious that false positives of our original method are much less than the modified method. This is because CF-filter can identify the shape of aircrafts in cluttered background.

5. Conclusion and Future Work

In this paper, we proposed a coarse-to-fine method for aircraft detection in VHR remote sensing images. CF-filter is applied to extract the candidate locations of aircrafts in the coarse stage. In the fine stage, a multilevel feature model is designed to capture both the local appearance and spatial layout of aircrafts. Meanwhile, Robust Hue Descriptor and Histogram of Oriented Gradients are utilized to describe the color and shape information of aircrafts. The experiment results show the good performance of the proposed method.

Many satellite images provide more than red, green, and blue channels; other channels such as infrared, coastal blue, and red edge are often acquired. In the future work, we will study how to utilize more spectral information to represent aircrafts.

Conflict of Interests

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

Acknowledgments

The authors would like to thank Dr. S. Maji for kindly providing IKSVM [23]. The authors would also like to thank Dr. Y. Yang from the Ocean University of China for her helpful discussion and the anonymous reviewers for their helpful comments and suggestions. This work was partially supported by the 973 Program (2010CB327900) and the National Science Fund for Distinguished Young Scholars (no. 61125206).

References

- [1] A. Kembhavi, D. Harwood, and L. S. Davis, "Vehicle detection using partial least squares," *IEEE Transactions on Pattern*

- Analysis and Machine Intelligence*, vol. 33, no. 6, pp. 1250–1265, 2011.
- [2] G. Liu, X. Sun, K. Fu, H. Wang, and Aircraft recognition in high-resolution satellite images using coarse-to-fine shape prior, *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 3, pp. 573–577, 2013.
 - [3] H. Cai and Y. Su, “Airplane detection in remote-sensing image with a circle-frequency filter,” in *International Conference on Space Information Technology*, vol. 5985 of *Proceedings of SPIE*, Wuhan, China, November 2005.
 - [4] B. Salehi, Y. Zhang, and M. Zhong, “Automatic moving vehicles information extraction from single-pass worldView-2 imagery,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 1, pp. 135–145, 2012.
 - [5] J. Leitloff, S. Hinz, and U. Stilla, “Vehicle detection in very high resolution satellite images of city areas,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 7, pp. 2795–2806, 2010.
 - [6] C. Zhu, H. Zhou, R. Wang, and J. Guo, “A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 9, pp. 3446–3456, 2010.
 - [7] L. Chen, S. Zhao, W. Han, and Y. Li, “Building detection in an urban area using lidar data and QuickBird imagery,” *International Journal of Remote Sensing*, vol. 33, no. 16, pp. 5135–5148, 2012.
 - [8] A. O. Ok, C. Senaras, and B. Yuksel, “Automated Detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 3, pp. 1701–1717, 2013.
 - [9] Z. Lei, T. Fang, H. Huo, and D. Li, “Rotation-invariant object detection of remotely sensed images based on texton forest and hough voting,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 4, pp. 1206–1217, 2012.
 - [10] Y. Li, X. Sun, H. Wang, H. Sun, and X. Li, “Automatic target detection in high-resolution remote sensing images using a contour-based spatial model,” *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 5, pp. 886–890, 2012.
 - [11] J. Inglada, “Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 3, pp. 236–248, 2007.
 - [12] H. G. Akçay and S. Aksoy, “Automatic detection of geospatial objects using multiple hierarchical segmentations,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 7, pp. 2097–2111, 2008.
 - [13] G. E. Atteia and M. J. Collins, “On the use of compact polarimetry SAR for ship detection,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 80, no. 6, pp. 1–9, 2013.
 - [14] M. T. Alonso, C. Lopez-Martinez, J. J. Mallorqui, and P. Salembier, “Edge enhancement algorithm based on the wavelet transform for automatic edge detection in SAR images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 222–235, 2011.
 - [15] N. Renard and S. Bourennane, “Improvement of target detection methods by multiway filtering,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2407–2417, 2008.
 - [16] S. Bourennane, C. Fossati, and A. Cailly, “Improvement of target-detection algorithms based on adaptive three-dimensional filtering,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 4, pp. 1383–1395, 2011.
 - [17] H. Sun, X. Sun, H. Wang, Y. Li, and X. Li, “Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model,” *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 1, pp. 109–113, 2012.
 - [18] W. Zhang, X. Sun, K. Fu, C. Wang, and H. Wang, “Object detection in high-resolution remote sensing images using rotation invariant parts based model,” *IEEE Geoscience and Remote Sensing Letters*, 2013.
 - [19] J. Van De Weijer, T. Gevers, and A. D. Bagdanov, “Boosting color saliency in image feature detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 150–156, 2006.
 - [20] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR ’05)*, pp. 886–893, June 2005.
 - [21] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
 - [22] S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: spatial pyramid matching for recognizing natural scene categories,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR ’06)*, pp. 2169–2178, June 2006.
 - [23] S. Maji, A. C. Berg, and J. Malik, “Efficient classification for Additive Kernel SVMs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 66–77, 2013.



Hindawi

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

