International Journal of Engineering Science Invention
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org ||Volume 4 Issue 5 || May 2015 || PP.01-06

Performance Analysis of Rotation Invariant parts Based Object Detection in High-Resolution Images

[1.] S.Prabhakher, [2].Mr.P.Sukumar,
[1]*PG Scholar, Nandha Engineering College, Erode.*
[2]*ASP of ECE, Nandha Engineering College, Erode.*

Abstract : Object detection is one of the essential applications and crucial tasks in remote sensing images analysis. With the enlargement of remote sensing knowledge, high-resolution(HR) optical remote sensing images are presented and the geospatial detail information of substances is more and more abundant.In this paper, rotation invariant parts based model to sense objects with difficult shape in high resolution remote sensing images. Explicitly, the geospatial matters with difficult shape are initially alienated into several main parts, and the assembly information among parts is labeled and controlled in polar coordinates to achieve the rotation invariance on configuration. In encoding the features of the rotated parts and objects, a new rotation invariant feature is projected by spreading histogram oriented gradients. Throughout the final recognition step, a clustering technique is introduced to discover the parts in objects, and that process can also be used to fuse the detection consequences.

Index Terms:-Geometric information, object detection, rotation invariant, random forest

I. INTRODUCTION

With the development of sensor technologies and the initiation of internet-based image databases, high-resolution remotely sensed (RS) images become widely reachable and increasingly used by researchers in object recognition as well as urban and social-economic studies. High-resolution RS images have different characteristics when compared to ground-shot images traditionally used in object detection. While they bring more detailed description of geospatial objects, the thorough detail could lead to complex inner structures and disturbing background. These raise new challenges in object detection. A lot of works in the area have been done with different focuses.

Recently, the part-based model has been very prevalent in the field of object detection due to their effectiveness in representing the conformation of objects. Each part delivers the local visual properties of the object and the geometrical preparation is characterized by certain connection between pairs of parts. There is a significant body of work on part based models [1]. In the constellation model [2], parts for the object are constrained to search in a sparse set of locations that are detected by an interest point detector [3], and their configuration is represented using a Gaussian distribution. The constellation model is flexible in application for it could describe the appearance, shape, and occlusion information of object simultaneously. But the detection results by the constellation model are not stable enough because the model is heavily dependent on the feature detector.

Another line of investigation in part-based models for object detection is the discriminatively deformable part-based models [4], particularly the discriminatively trained-part based models (DTPBM). The DTPBM have been widely used recently due to their strong presentation on difficult standards such as the PASCAL datasets. The object model in DTPBM consists of a global "root" filter and several part models. Each part model involves a spatial model and apart filter. In the spatial model, a deformation cost is defined to describe the displacement of each part relative to a detection window. The filters for root and parts stipulate weights fortheir histogram of oriented gradient (HOG) features. By discovering the dormant structure information, the arrangement of DTPBM is capable of demonstrating highly variable object classes. However, the DTPBM are rotation variant. First, the HOG feature that has been healthy used to detect human is can only make little alteration when the rotation revolution of an object is smaller than the alignment bin size. Moreover , the deformation cost in DTPBM can only be right when the object is upright and aligned. The shortcomings mentioned above make DTPBM difficult to apply in some scenes directly, such as object detection in the HR remote sensing images.

It is well known that there is a tradeoff between the discriminative power and invariance properties of detectors [5]. Since a rotation-invariant detector has to identify a prototypical object in many rotated directions; this added complexity requires a greater discriminative power as compared to a non-rotation invariant counterpart to deal with the increased number of possibilities. Then take two measures to enhance the discriminative power of our rotation-invariant detector: 1) using TF, to generate the codebook and 2) using color-invariant-gradient based local descriptors [6], to effectively capture color information.

In this paper, a new model to detect objects with complex shape in HR remote sensing images is proposed by spreading DTPBM. First, adapt the structure descriptor of DTPBM by regularizing the geometric information of portions in polar coordinates, which can attain the geospatial rotation invariance amongst parts. Meanwhile, our model defines a rotation cost for each part to specify their rotation deformation. Secondly, propose a rotation invariant HOG (RIHOG) feature extended from HOG by appraising and regulating the leading orientation of a region to crack the rotation variance of parts and objects. Finally, during the discovery phase introduce a clustering technique to reduce the large hypothesis spaces of parts that can fuse manifold overlapping detections. In the experiments, the presentation of aircraft detection in the HR remote sensing images from Quick Bird shows the efficiency of our model.

II. RELATED WORK

There is an important body of work on deformable models of many classes for object detection, with numerous types of deformable pattern models, and a variety of part-based models.

In the assemblage models from, parts are embarrassed to be in a sparse set of locations resolute by an interest point operator, and their geometric preparation is captured by a Gaussian distribution. In divergence, pictorial structure models describe a matching problem where parts have a separate match cost in a dense set of locations, and their geometric preparation is captured by a set of “springs” joining pairs of parts. The patchwork of parts model from is similar, but it obviously considers how the presence model of overlying parts interrelates.

The overview of new local and semi local features has played a significant role in preceding the presentation of object recognition approaches. These features are characteristically invariant to illumination deviations and small deformations. Numerous current methods use wavelet-like features or locally normalized histograms of gradients. Other approaches, such as, learn dictionaries of local as assemblies from training images. In this work, use histogram of gradient (HOG) features from as a initial point, and familiarize a variation that decreases the feature size with no loss in presentation. As in, used PCA to discover low dimensional features, but note that the eigenvectors obtain have a clear structure that leads to a new set of “analytic” features. This eliminates the need to perform a costly projection step when calculating dense feature maps.

Significant variations in outline and appearance, such as those produced by extreme viewpoint variations, are not well captured by a 2D deformable model. Feature graphs area classical formalism for capturing important changes that are due to viewpoint variation. Mixture models deliver a simpler alternative method. For example, it is common to use manifold templates to encode frontal and side views of faces and cars. Combination models have been used to capture other features of appearance variation as well, such as when there are manifold natural subclasses in an object group.

Matching a deformable model to an image is a problematic optimization problem. Local search approaches necessitate initialization near the precise solution. To assurance globally optimal match, more violent search is needed. One prevalent approach for part-based models is to constrain part locations to a small set of conceivable locations reverted by an interest point detector. Tree (and star) designed pictorial structure models allow for the use of dynamic programming and widespread distance transforms to professionally search over all possible object shapes in an image, without restricting the possible locations for each part. Then use these techniques for matching our models to images.

III. Proposed Rotation Invariant Part Based Model

RIPBM treats an object as a assortment of numerous main parts extracted from local patches in a high resolution and a full object in a coarse resolution. The resolution of all parts is indistinguishable with the original image, twice the spatial resolution of the full object. In representing the formation of objects, propose the RIHOG feature to characterize the rotated objects and the parts, and the RIPBM to define a new deformation cost and a rotation cost. The deformation cost here labels the relative movements among parts, and the rotation cost defines the pose variance of each part relative to a discovery window.Fig. 1 shows the procedure of proposed model to train and test. The details of proposed model are described as follows.

Noise Reduction

Noise removal is the procedure of eliminating noise from an image. In the proposed scheme noise reduction is done by using Gaussian filter function. The noise standard deviation of the image is appraised using the Immerkaer’s fast method. The complete difference between the center pixel and nearby pixels in the filtering window is obtained by deducting each

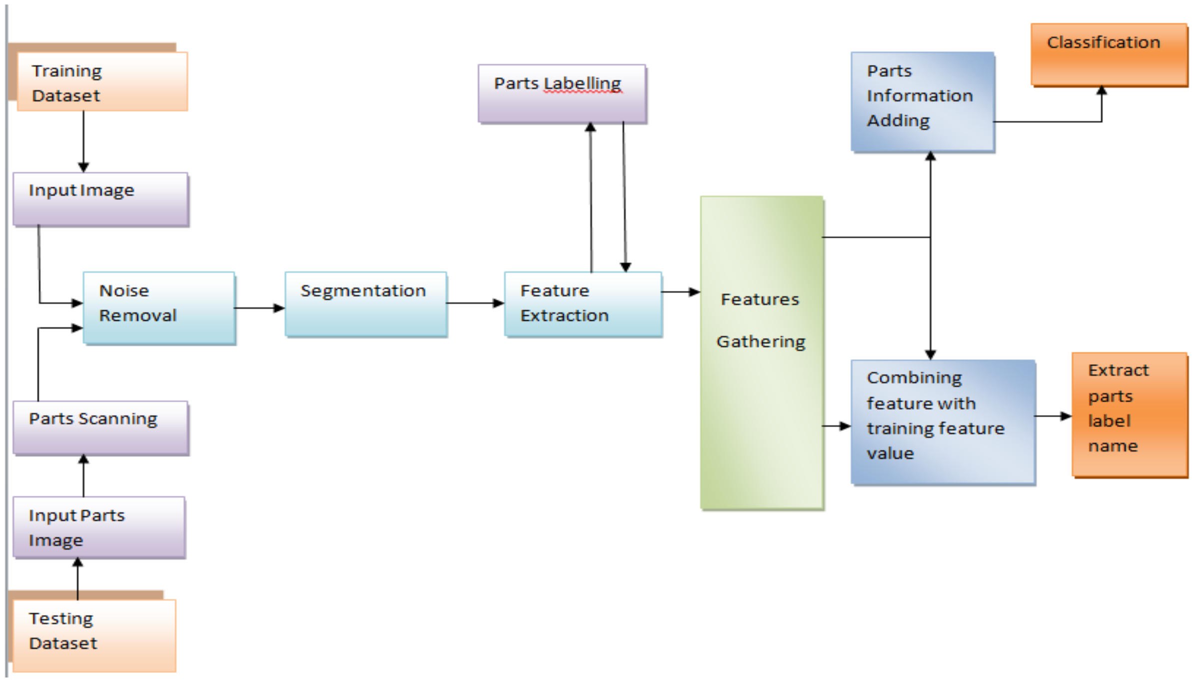

Fig.1 Architecture diagram of the proposed rotation invariant part based model

component in the filtering window with the center pixel. The difference will be large when the image is highly corrupted. This difference is compared with a threshold. The threshold can be demarcated as the invention of smoothing factor and noise standard deviation. The value of smoothing factor is chosen as two for optimal presentation. If the smoothing factor is designated to be a high value, the noise removal is healthier at the cost of loss of image details. If this complete difference is inside the threshold, the corresponding pixel values are alone taken for additional processing.

Part Based Segmentation

Afterward eliminating noise from an image segment the object by means of a part based segmentation. Image segmentation is the procedure of partitioning a digital image into manifold segments. Segment the part given wherever inside that object. This step is supported out using only color, texture and depth information. In order to choose the part inside dissimilar objects and pick the segments consistent to the objects, use border ownership information at boundary pixels.

Beforehand implement the segmentation; some prior information of the target object is needed. The knowledge is incorporated in a part- based shape model. To attain this model, the user only wants to di- vide an instance of the target object into numerous parts (components). The guideline is that the multifaceted object is decomposed to a few simple and controllable components after the partition. Each part has comparatively simple morphology. In addition, the user wants to stipulate “salient” parts in order to prepare the conforming in the next stage. The guideline for choosing the salient parts is that they typically have the largest size, the number of additional parts to which they are connected is large, they are the most illustrative parts of the object or they are the common parts or features of the object. For example, a shape of a person is able to disassemble into 10 parts, i.e., head, trunk, left and right arm parts (upper and lower), and left and right leg parts (upper and lower). Noticeable parts of the person are head and trunk.

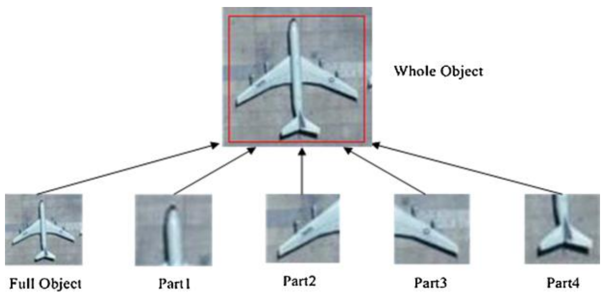

Fig.2 Sample object of part based segmentation

The salient parts of a chair can be the sitting surface and the chair back while the other parts are the arms and legs. It should be renowned that the major goal of the current work is an eloquent object or an object that can be easily decomposed into distinctive parts. From this partition, obtain the shape of each part and its estimated position. The shape and pose of each part in the part-based model are observed as the prior knowledge of the target object. Be- tween two adjacent parts in the model, there is a communal boundary. The two end points of this boundary are the anchor points (control points). The connection and interaction among adjacent parts are manipulated through these control points. Two non-adjacent parts can also interact by changing relative poses. The alterations of each part involve translation, rotation and scaling.

Feature Extraction

After the procedure of part based segmentation the features of the test image will be mined by using histogram based gradient function. An "image histogram" is a type of histogram that performances as a graphical protest of the tonal spreading in a digital image. It plots the quantity of pixels for each tonal value. By looking at the histogram for a detailed image a viewer will be able to judge the complete tonal distribution at a glance.

Image histograms are extant on many present digital cameras. Photographers can use them as an aid to display the delivery of tones captured, and whether image detail has been lost to blown-out highlights or blacked-out darks.

In measurements, a histogram is a graphical demonstration of the distribution of data. It is an estimate of the probability delivery of an incessant variable and was first introduced by Karl Pearson. A histogram is a representation of tabularized frequencies, shown as nearby rectangles or squares (in some situations), erected over discrete intervals (bins), with an area comparative to the frequency of the observations in the interval. The elevation of a rectangle is also equivalent to the frequency compactness of the interval, i.e., the frequency alienated by the width of the interval. The whole area of the histogram is equivalent to the amount of data. A histogram might also be normalized displaying comparative frequencies. It then shows the proportion of cases that fall into each of several categories, with the entire area totaling 1. The categories are usually stated as successive, non-overlapping intervals of an adjustable. The categories must be contiguous, and frequently are chosen to be of the similar size. The rectangles of a histogram are drawn so that they trace each other to designate that the original variable is continuous.

SVM Based Classification

In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces.An SVM classifies data by finding the best hyper plane that separates all data points of one class from those of the other class. The best hyper plane for an SVM means the one with the largest margin between the two classes. Margin means the maximal width of the slab parallel to the hyper plane that has no interior data points.

The SVM classifier involves a nonlinear mapping from the input parameter space to the feature space. The nonlinear mapping was performed using a radial basis function. Then tried linear kernels first, but the obtained classification rates were lower than those obtained with the radial function.

Bootstrap resembling methods were used both to select test and training data to evaluate the classification accuracy of our method and to estimate the SVM parameters. For this latter purpose, alternative approaches have been proposed in the literature, for instance, cross-validation methods such as leave-one-out orn-fold cross-validation [7].

IV. Experimental Results

Due to the lack of standard data sets of HR remote sensing images for object detection, evaluate the RIPBM on Quick-Bird for aircraft detection with the resolution of 60 cm/pixel. The size of the training set is 100 patches with aircraft and 150 patches with background. In the training phase, the parts of the aircraft are manually labeled, i.e., one head, two wings, and one empennage. In the testing phase, take 40images independent from the training set, and there are several aircrafts with different types and directions in each image. One labeled example and some samples from the training set and the positives are aligned.

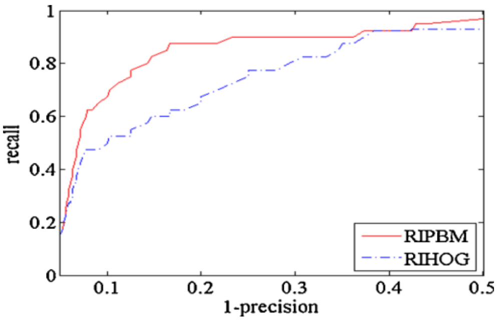

Fig.3 RPC of the RIPBM and RIHOG

Then assume a detection result is correct if more than 75% of an aircraft is detected in qualifying the performance. Therecall-precision curve (RPC) is also plotted as in to tradeoff between Precision and Recall.

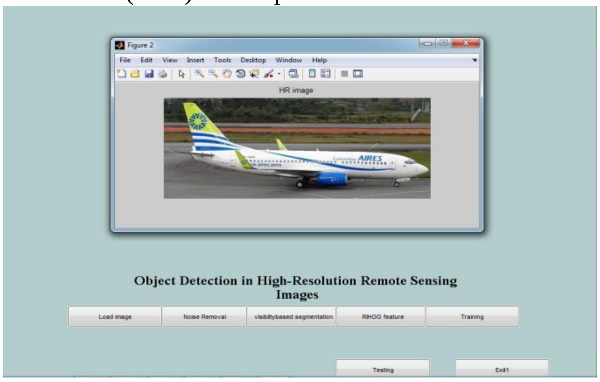

Fig.4 Load image GUI


Fig.5 Various steps of noise reduction

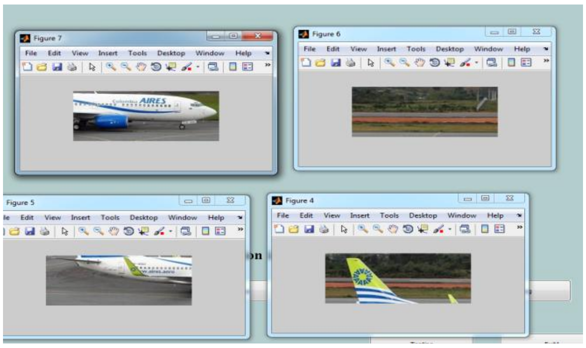

Fig.6 Parts based Segmentation process

For aircraft detection, the precision and recall change with the sizes of cell and block. 2 × 2 cell blocks of 8 × 8 pixelscells perform best with the 98.7% precision and 98.3% recall,with 50% overlapping between nearby blocks, and the 4 × 4cell blocks of 8×8 pixels cells are a close second.

Fig.7 Classified output image

Our resultis different from the best performance with 3 × 3 cell blocksof 6×6 pixels cells for human detection, because thewidth of a part in aircraft is about 5m, about 8 pixels in theHR remote sensing images.

V. CONCLUSION

The RIPBM for object detection in HR remote sensing images model solved the rotation invariance of parts in rotated objects by defining the displacement deformation cost and the rotation cost for parts. In encoding the features of rotated parts, extended HOG to RIHOG by evaluating the dominant orientations similar with SIFT descriptors. The resolution of all parts is identical with the original image, twice the spatial resolution of the full object. In representing the configuration of objects, propose the RIHOG feature to represent the rotated objects and the parts, and the RIPBM to define a new deformation cost and a rotation cost. The deformation cost here describes the relative displacements among parts, and the rotation cost defines the pose variance of each part relative to a detection window. During the searching of parts, a clustering method was introduced with our model to reduce the search pace of parts and fuse the final detection results. The final experiments showed the robustness of this model.

References

[1] S. Xu, T. Fang, D. R. Li, and S. W. Wang, "Object classification ofaerial images with bag-of-visual words," IEEE Geosci. Remote Sens.Lett., vol. 7, no. 2, pp. 366–370, Apr. 2010.

[2] H. Sun, X. Sun, H. Wang, Y. Li, and X. Li, "Automatic target detectionin high-resolution remote sensing images using spatial sparse codingbag-of-words model," IEEE Geosci. Remote Sens. Lett., vol. 9, no. 1,pp. 109–113, Jan. 2012.

[3] Z. Lei, T. Fang, H. Huo, and D. Li, "Rotation-invariant object detectionof remotely sensed images based on Textonforest and Hough voting,"IEEE Trans. Geosci. Remote Sens., vol. 50, no. 4, pp. 1206–1217, Apr.2012.

[4] M. Weber, M. Welling, and P. Perona, "Towards automatic discovery ofobject categories," in Proc. IEEE CVPR, vol. 2. Jun. 2000, pp. 101–108.

[5] M. Burl, M. Weber, and P. Perona, "A probabilistic approach to objectrecognition using local photometry and global geometry," in Proc. IEEEECCV, Jun. 1998, pp. 628–641.

[6] R. Fergus, P. Perona, and P. Zisserman, "Object class recognition byunsupervised scale-invariant learning," in Proc. IEEE CVPR, vol. 2. Jun. 2003, pp. 264–271.

[7] R. Fergus, M. Weber, and P. Perona, "Efficient methods for objectrecognition using the constellation model," California Inst. Technol.,Tech. Rep., 2001.

[8] R. Fergus, P. Perona, and A. Zisserman, "A visual category filter forGoogle images," in Proc. IEEE ECCV, vol. 3021. May 2004, pp.242–256.

[9] T. Kadir and M. Brady, "Scale, saliency and image description," IEEEIJCV, vol. 45, no. 2, pp. 83–105, Nov. 2001.

[10] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan,"Object detection with discriminatively trained part based models," IEEEPattern Anal. Mach. Intell., vol. 32, no. 9, pp. 1627–1645, Oct. 2010.

[11] P. Felzenszwalb, D. McAllester, and D. Ramanan, "A discriminativelytrained, multiscale, deformable part model," in Proc. IEEE CVPR, Jun.2008, pp. 1–8.

[12] P. Felzenszwalb "Pictorial structures for object recognition," IEEECVPR, vol. 61, no. 1, pp. 55–79, Jan. 2005.

[13] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A.Zisserman, (2007) The PASCAL VOC2007 Results [Online]. Available:http://www.pascalnetwork.org/challenges/VOC/voc2007/

[14] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A.Zisserman, (2008) The PASCAL VOC2008 Results [Online]. Available:http://www.pascalnetwork.org/challenges/VOC/voc2008/

[15] T.Maheswari and P.Sukumar, R.K.Gnanamurthy(2015) "Minimizing MCU in SRAM Cell to Enhance Memory Reliability Using DMC" International Journal of Advanced and Innovative Research (2278-7844) / # 21 / Volume 4 Issue 1.

[16] T.Maheswari and P.Sukumar(2015) " Error Detection and Correction in SRAM Cell Using Decimal Matrix Code" IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 5, Issue 1, Ver. II (Jan - Feb. 2015), PP 09-14 e-ISSN: 2319 – 4200, p-ISSN No. : 2319 – 4197.