SHADOW TRACKING OF MOVING TARGET BASED ON CNN FOR VIDEO SAR SYSTEM

Yun Zhang, Shiyu Yang, Hongbo Li, ZhenHua Xu
School of Electronics and Information Engineering, Harbin Institute of Technology, China
drbobo@hit.edu.cn

ABSTRACT

Fast Moving targets always are shifted or smeared outside the scene in different images sequence to make video by Circle Synthetic Aperture Radar (SAR).In this paper, a novel moving target tracking approach with the shadow detection and tracking (SDT) is presented based on Convolution Neural Network. Based on the shadow characteristic of moving target in SAR imagery, CNN tracking classification is employed on potential moving target candidates extracted from a sequence of temporal and spatial sub-aperture SAR images to detect and track the moving targets. By the simulation experiments and performance analysis, the validity of the proposed algorithm can be demonstrated. Real data set processing results are provided to demonstrate the effectiveness of the proposed approach.

***Index Terms*—**Shadow detection, Moving target tracking, Video SAR, Convolution Neural Network.

1. INTRODUCTION

Synthetic aperture radar is widely used to observe the scene of interest under all-weather and day-and-night. However, moving targets detection in SAR system is a challenging task due to the displacement or smearing of moving targets as well as strong ground clutter in SAR imaging. With the longer integration time, the resolution in azimuth is improved, but the frame rate of SAR imaging is far lower than the need for target tracking. To solve it, Video-SAR or Circle AR is proposed by Lars Well of Sandia National Lab in 2003 [1], which can obtain multi frame sequence of target in a movie-like form. A wide-angle's influence on car detection is researched by Kerry E. Dungan's about [2], DRDC use BP algorithm for detection [3] and Tomoya Yamaoka presented a low complexity method to observe by Radar video [4]. Moving detect target on PCNN and track target by KCF (Kernelized Correlation Filter) is presented by Xu [5]. Approaches for multi-target tracking in spotlight SAR system based on Gaussian mixture PHD filtering were presented in [6].A novel shadow of moving target tracking algorithm based on CNN in SAR system is presented in this paper. CNN has been established as a powerful class of models, which handles massive training samples[7], and features learned from the network automatically instead of manual ones. In order to use CNN to get the initial shadows of targets, pre-process of shadow of moving target is applied to form a sequence of sub-aperture SAR images and generate a set of track proposals for CNN inputs. Subsequently, the observation set is generated utilizing the moving target candidates and the azimuth shift of moving targets. The moving target model based on the SAR image sequence is built and the CNN shadow detector and tracker is applied. The performances of the proposed algorithm are demonstrated using both simulated and SAR raw.

2. TARGET TRACKING IN VIDEO SAR

Shadow extraction of moving target in the first Video SAR frame is done, which always worked as circular SAR. The stationary targets show different reflection by different incident, and moving targets show the different location and focusing. To keep tracking a target or monitoring an interested area, the processing of video on circular SAR data can improve the change detection on imagery sequence. Consider Fig. 1 that depicts the geometry of a Circular SAR to make video imaging.

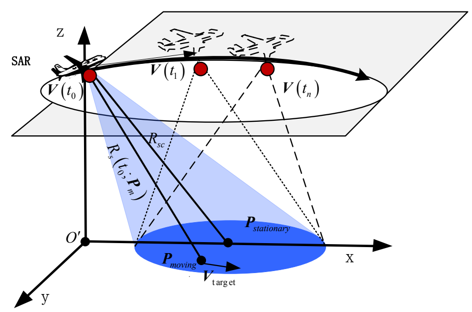


Fig. 1 The geometry of Circular SAR

In ViSAR of multi frame images, the motion of target results in the position shift or defocus, and the scattering characteristics of the moving target (car, vehicle, truck) is completely different indifferent frames. It's hard to detect and track the fast moving targets, which Doppler energy may be shifted or smeared outside the scene. However the shadow of the target is created by the incident angle of look light, which located around the real area. So target shadow is observed, which depend on its physical dimension. The shadow is moving followed the mobility of a target, which is shown in the figure2(a) and (b). The deformation of Shadow is related to the downward angle of view, the velocity of the target and the three-dimensional structure of the target.

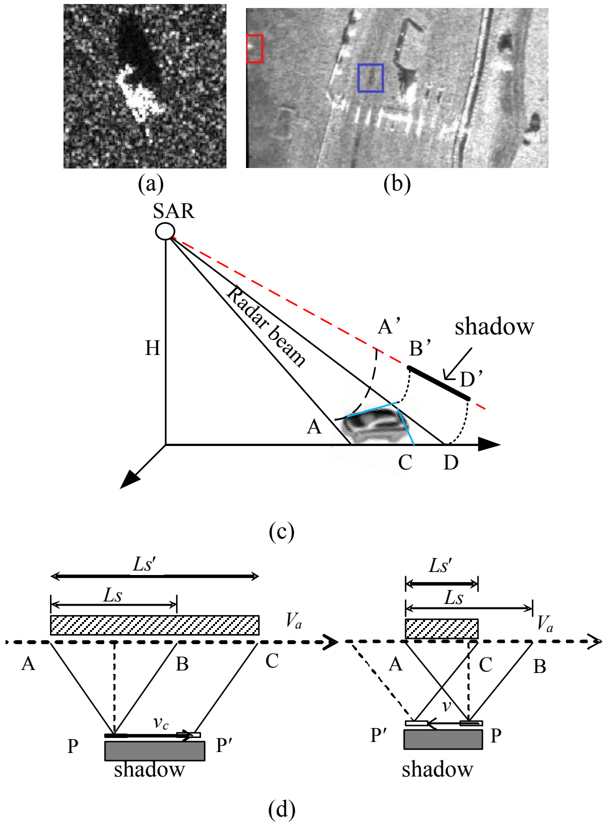


Fig. 2 The geometry of target shadow, (a)shadow of static target, (b) shadow of moving target, (c) The geometry of target, (d) the synthetic aperture time of shadow of moving target

Tangential motion causes the target to be different from the radar exposure time, that is, the effective accumulation time is different. The following analysis is made to analyze the synthetic aperture time of the moving target shadow, as shown in Figure 2(d), the synthetic aperture time T_s (L_s=T_s× V_a), the SAR carrier moves from point A to B, the moving target is from point P to point P ', and L_s is the length of the synthetic aperture of the system, and the relation between the two is as $V_a/(V_a - v_c) = T_s'/T_s = L_s'/L_s$.Because of the motion of radar platform, the azimuth of the radar beam reaching the target is constantly changing, and the whole time of synthetic aperture is moved by the shadow region of the stationary target. Due to the motion of target, the shadow area is soon covered by clutter, which is not conducive to shadow detection, which is blurring to increase the width.

3. TARGET TRACKING ALGORITHM ON CNN

3.1. Target track information extraction

Multi frame image of maneuvering object and static clutter is transient, especially because of the moving target, radial velocity, tangential velocity variation, the target echo position offset, serious defocusing even by clutter and submerged. The moving target echoes and background clutter are instantaneous in the multi frame SAR image sequence. Due to the maneuvering of the platform, the tangential and radial motion components of the moving target are changed in real time, causing serious defocus and migration of the imaging results, and the defocus and offset may be different in each frame. Although video SAR can obtain image sequences and have continuous observation advantages, the traditional tracking method is faced with the problem of high tracking heel rate and low target retention because of the factors mentioned above. The road information can guide to track the shadow of moving target more effectively, and train it together in neural network shown in the figure 3. The tracking processor is based on radar information (platform attitude, lower angle of view, motion parameters, etc.) and target information (position, defocus degree, contour, etc.) Shadow, background and so on complete the tracking of moving targets.

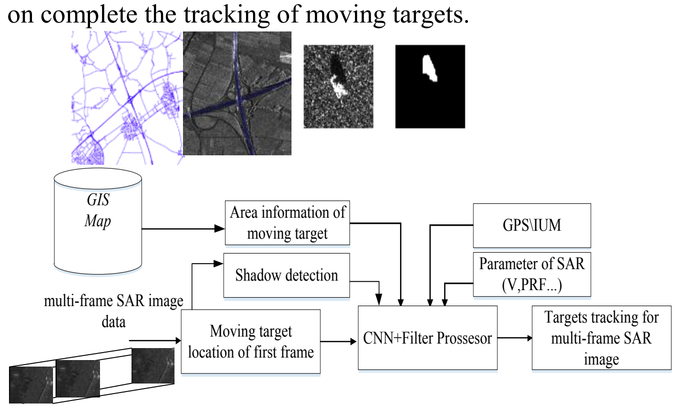


Fig. 3 the procedure of target shadow detection and target tracking

3.2. Convolution Neural Network

Typically, CNN consists of convolutional layers and sub-sampling layers. The former is usually interspersed with the latter to reduce computation time. However, in the paper, we set the sub-sampling factors to be 1, which means that the sub-sampling layers do not work as the size of feature maps are extremely small. In forward computation, the feature maps from previous layer are convolved by learnable kernels and the results will be put through the activation function to form the output maps.

$$x_j^l = f\left(\sum_{i\in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right) \qquad (1)$$

where M_j represents a selection of input maps x_i in the layer $l-1$ to form the output map x_j in layer l . k_{ij}^l refers to the kernel convolving input map x_i^{l-1} to form output map x_j^l , for each input map will be convolved with distinct kernels. Each output map is given an additive bias b_j^l . $f(\square)$ is the activation function.

When CNN is well built with specific size of kernels and number of layers, weights of the network will be initialized. In the backward computation, those weights will be corrected iteratively with the training samples to learn features from the train dataset. According to BP algorithm, gradient decent is implemented to update the weights, when a new definition called sensitivity is used to simplify the derivation. The sensitivity of map j in convolution layer l is δ_j^l, and the next layer $l+1$ is subsampling layer whose map j has the sensitivity of δ_j^{l+1}.

$$\delta_j^l = \beta_j^{l+1}\left(f'\left(u_j^l\right)\circ up\left(\delta_j^{l+1}\right)\right) \tag{2}$$

where β_j^{l+1} is the weight defined in subsampling layer. $up(\square)$ is the operation of up-sampling. δ_j^{l+1} is multiplied by the derivative of activation function evaluated at the current layer's pre-activation input. The gradient of bias b_j^l and weights k_{ij}^l can be derived from

$$\frac{\partial E}{\partial b_j^l} = \sum_{u,v}\left(\delta_j^l\right)_{u,v} \tag{3}$$

$$\frac{\partial E}{\partial k_{ij}^l} = \sum_{u,v}\left(\delta_j^l\right)_{u,v}\left(p_i^{l-1}\right)_{u,v} \tag{4}$$

where $\left(p_i^{l-1}\right)_{u,v}$ is the patch in x_i^{l-1} that multiplied element by k_{ij}^l during convolution in order to compute the element at (u,v) in the output convolution map x_i^l.

Although neural networks can theoretically improve the overall tracking performance, the neural network is a supervised learning process, which has a huge demand for training samples, and the large training sample of shadow is obviously not provided by the multi frame SAR image sequence. To some extent, it limits the application of neural network in this field. Therefore, tracking multi frame moving targets through CNN neural network, the first thing is to solve the lack of training samples. After learning the sample set, a small number of tagged samples of the current task are used to fine tune the parameters, thus the purpose of training CNN for small samples is achieved. The first frame of the actual tracking task is tuned according to the first frame of the actual tracking task, and then the second frames to the N frame data are processed with the learned CNN network to achieve the target tracking effect.

The CNN improvement for the training problem with tags in the lack of training samples, first supervise and pre training, the training process in the offline, can be used to contain a large number of training sample data set (but the data set may not apply to the current task), and then according to the current task, fine-tuning with a small amount of training samples in order to achieve the expected goal. Maneuver trajectory obtained by SAR Video video tracking can learn from this idea, first extracted from SAR images in the existing large amounts of labeled training samples, pre training CNN network, first let CNN to learn the target feature, to fine tune the network parameters and image data of the first frame, so as to realize better tracking. The implementation process is shown in Figure 4.

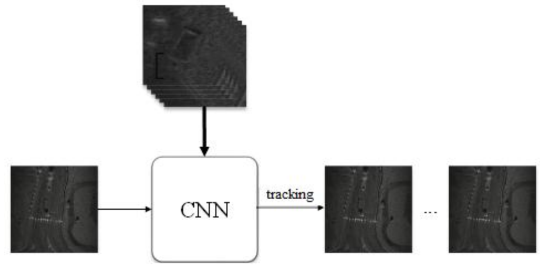


Fig. 4 shadow feature extraction process of multi frame SAR images based on depth learning

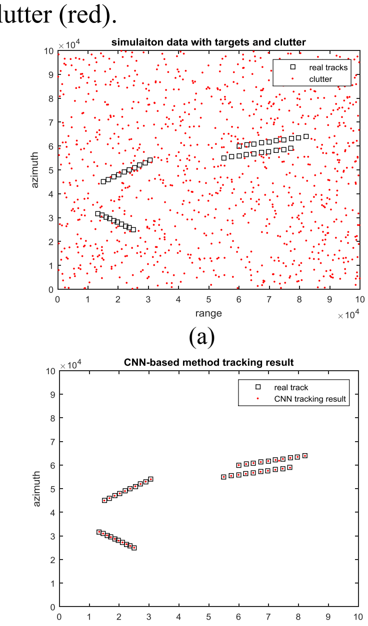
Pre training process

Unlike the traditional CNN, the training output here is no longer a single value, but a graph (a matrix), which becomes a probability graph. For example, the size of the input image is 100 ×100, and the size of the probability map is 50 × 50, so a pixel in the probability map corresponds to a 2× 2 area corresponding to the input image. Therefore, the training label is also a 50 ×50 image, here we set in a position with a goal of 1, set the position without a goal of 0, using logistic regression model to complete the parameter adjustment process through the minimum cost function of J, which is defined as the cost function

$$J = \min_{p_{ij}}\sum_{i=1}^{50}\sum_{j=1}^{50} -(1-t_{ij})\log(1-p_{ij}) - t_{ij}\log(p_{ij}) \tag{5}$$

4. EXPERIMENTS AND PERFORMANCE ANALYSIS

In this section, both simulated SAR data and Gotcha data set are used to valid the effectiveness and superiority of the proposed algorithm. In simulation experiments, there are 4 targets with velocity from 20m/s to 30m/s, which include ten frame SAR image. Fig.5shows tracks (black) with a certain degree of clutter (red).



(a)

(b)

Fig. 5 Simulation data with targets and clutter. (a) Targets in the Clutter background, (b) Targets tracking by CNN

To demonstrate the effectiveness of our approach, we apply it to the publicly released videoSAR data set. The data was collected by a circular SAR system in an urban environment. In this part, we mainly show the simulation results and relative analysis. Some results are still remained to be optimized in the future so better results can be expected. The tracking result is a series of images with interested targets tracked by the templates, and the tracking process with target tracked at the same time on a normal laptop. Results are shown in Fig.6.

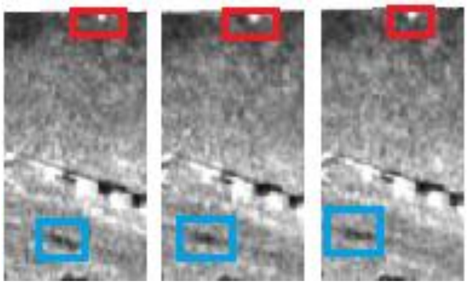


Fig. 6 Frames extracted from tracking results. Target and its shadow are well tracked but the top target(red template) in the 22th 25th and 28th frame

To obtain targets' speed and moving routes, the tracking templates' location in every frame. Also the records of tracking templates can be very helpful in our further method optimization. The results of 5 moving target tracking is shown in the Figure7.

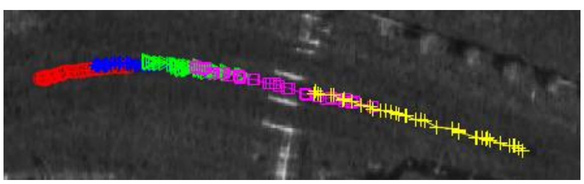


Fig. 7 The routes of tracking templates

5. CONCLUSION

In this paper, a new moving target tracking approach based on CNN in video SAR system has been presented. CNN is adopted to detect and track multiple moving targets in heavy clutter environment, using the background information. It can provide estimates for the states and the number of moving targets. Multiple-target simulation and experiments have been conducted to show that the proposed algorithm is effective to obtain an improved detection and tracking performance.

ACKNOWLEDGMENTS

Thanks for the National Natural Science Foundation of China (61201308 and 61201304).And it was supported by Key Laboratory of Marine Environmental Monitoring and Information Processing, Ministry of Industry and Information Technology

REFERENCES

[1] Wells L, Sorensen K, Doerry A, et al. Developments in SAR and IFSAR systems and tech-nologies at sandia national laboratories[C]//Aerospace Conference, 2003.

[2] Dungan K E. Vehicle detection in wide-angle SAR]//SPIE Defense+ Security. International Society for Optics and Photonics, 2016: 98430H-98430H-6.

[3] Damini A, Balaji B, Parry C, et al. A videoSAR mode for the X-band wideband experimental airborne radar[C]//SPIE Defense, Security, and Sensing. International Society for Optics and Photonics, 2010: 76990E-76990E-11.

[4]Yamaoka T, Suwa K, Hara T, et al. Radar video generated from synthetic aperture radar image[C]//Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International. IEEE, 2016: 6509-6512.

[5]Zhenhua Xu, Yun Zhang , Hongbo Li, Huilin MU, Yuan ZhuangA New Shadow Tracking Method to Locate the Moving Target in SAR Imagery Based on KCF.2017 International Conference on Communications, Signal Processing, and Systems (CSPS'17), Harbin,2017.07

[6] N. Daryasafar, R.A. Sadeghzadeh, M. Naser-Moghadasi,A technique for multitarget tracking in synthetic aperture radar spotlight imaging mode based on promoted PHD filtering approach,*Radio Science*, 52(2), pp. 248-258, February, 2017.

[7] Yun Zhang,Shiyu Yang, Hongbo Li, Huilin Mu.A Novel Multi-target Track Initiation Method Based On Convolution Neural Network The International Workshop on Remote Sensing with Intelligent Processing(RSIP) Shanghai.2017.05