ISSN 2616-5775 Vol. 5, Issue 11: 97-103, DOI: 10.25236/AJCIS.2022.051115

Research on Speckle Reduction of SAR Image by Convolutional Neural Network

Wei Liu

Tiangong University, Tianjin, 300387, China
liuwei05@foxmail.com

Abstract: *The synthetic aperture radar (SAR) images are contaminated by noise called speckle, making the contaminated images difficult to understand and use. Since SAR images have the advantages of long range and all-weather, which can make up for the poor imaging effect of optical cameras in bad weather conditions, it is meaningful to study the de-speckling algorithm of SAR images. In order to solve the above problem, this paper proposes a method to improve the ability of removing speckles from SAR images using convolutional neural networks. And we use the attention mechanism to further improve the denoising ability based on the proposed denoising network CDNet. After experimental comparison, the speckle removal capability of CDNet proposed in this paper is better than the existing mainstream algorithms.*

Keywords: *Synthetic aperture radar, Convolutional neural network, Attention mechanism*

1. Introduction

Synthetic aperture radar (SAR) is a high-resolution imaging radar that can obtain remote sensing images similar to the imaging effects of optical cameras in environments with poor meteorological conditions. SAR makes up for the fact that traditional optical remote sensing cameras do not have the ability to take clear pictures in the case of insufficient light and occlusion, etc. SAR can obtain good imaging effects at any time and any place, and has good application scenarios in both civil and military fields, such as agriculture, hydrology, forestry, geology, oceanography, etc. At the early stage of radar development, people used real aperture radar (RAR), but since the imaging resolution of RAR is proportional to the length of the radar antenna and inversely proportional to the wavelength and observation distance, it needs to increase the antenna length if it wants to obtain better imaging effect, which limits its application and development, and therefore RAR is gradually replaced by SAR.

The SAR system not only uses the long-range transmission characteristics of radar signals, but also combines with the current high-performance information processing technology to enable the SAR system to provide high-resolution SAR images. Due to the coherence principle of SAR system, there will be echo signals generated after the coherent signals emitted by SAR are irradiated to the relevant targets, but these echo signals will have certain fluctuations, and such random fluctuations act on the imaging system to produce coherent speckle noise. The generation of coherent speckle noise will reduce the signal-to-noise ratio of the image and affect the texture structure of the image, thus reducing the imaging effect of the image, which is incompatible with the current demand for better and better imaging, so it is very meaningful to study how to reduce the effect of coherent speckle noise.

There have been many researchers who have made their contributions in dealing with speckle noise in SAR images. At first some filtering algorithms were proposed based on statistical ideas, such as Lee filter [1], Kuan filter [2], Frost filter [3], SAR-BM3D [4]. Moreover, these algorithms have been improved and developed continuously, and enhanced Lee filtering, enhanced Frost filtering, etc. have been proposed. The spatial filtering algorithm has been proposed for a long time and is simple to operate, but because the spatial filtering algorithm is based on local processing, when the local window selection is small, more information can be retained but the effect of removing speckles is not good, when the local window selection is large, the effect of removing speckles will become better but some information will be lost.

With AlexNet [5] winning the ImageNet competition in 2012, it has caused many scholars to research on convolutional neural networks, which are now achieving good results in many aspects such as image classification, target detection, and speech recognition. There are also many attempts to use convolutional

ISSN 2616-5775 Vol. 5, Issue 11: 97-103, DOI: 10.25236/AJCIS.2022.051115

neural networks in the denoising of SAR images. Because deep learning is trained end-to-end, convolutional neural networks can automatically extract image features, so researchers no longer need to extract features manually. This can improve the network's robustness and enhance the network's denoising ability. In 2017, Zhang et al. proposed the DnCNN [6] network, which simply adds the residual network to VGG and with few modifications makes the DnCNN surpass the traditional method in removing Gaussian noise. In the subsequent research, excellent denoising networks such as SAR-CNN [7], SAR-DRN [8], etc. were born. In this paper, we propose a new denoising network inspired by many previous studies and add an attention mechanism to the network to further improve the denoising effect. The following are the main contributions of this paper.

1) A new end-to-end neural network model is proposed for removing speckle noise from SAR images, and this new neural network model is called CDNet(Cooperative denoising network).

2) An attention mechanism is used in the denoising network of SAR images to further improve the denoising ability.

3) A new loss function is proposed and used in the training.

2. Related Work

2.1. Speckle Noise

SAR is a coherent imaging system. SAR is not the same as the imaging principle of an optical camera. The imaging process of optical camera is to convert the optical signal to electrical signal, then after analog-to-digital conversion finally the converted digital signal will be processed and then stored. The imaging process of synthetic aperture radar system is that the synthetic aperture radar carrying antenna will emit electromagnetic waves, and because the distance between the target and the antenna is different, the echoes will return to the antenna at different times, and finally these echo signals will be coherently superimposed.

That is, the pixel points of the picture of the SAR system coherent imaging is the product of the pixel points of the target object and the probability of noise, the image features contaminated by multiplicative noise is the area where the image pixels change more rapidly, the noise changes more rapidly and the contamination caused is also more serious. The noise model is shown in Equation 1

$$F(\mathrm{x},\mathrm{y}) = R(\mathrm{x},\mathrm{y}) \times N(\mathrm{x},\mathrm{y}) \qquad (1)$$

where F(x,y) and R(x,y) represent the noisy picture and the clean picture, respectively, and N(x,y) represents the noise factor.

It should be noted that the speckle noise in SAR images follows a gamma distribution and its probability density formula is shown in Equation 2

$$\rho(\mathrm{n}) = \frac{L^L n^{L-1} exp(-nL)}{\Gamma(L)} \qquad (2)$$

where $\mathrm{n} \geq 0$, $\mathrm{L} \geq 1$ and Γ represents the gamma function. L represents the average number of looks of SAR images and is calculated as in Equation 3.

$$\mathrm{L} = \frac{mean}{variance} \qquad (3)$$

2.2. Convolutional Neural Network

A convolutional neural network is a deep learning model often used in the field of computer vision. A convolutional neural network mainly consists of a convolutional layer, a pooling layer, an upsampling layer, and an activation function. Convolutional neural networks work by continuously extracting feature map information. A large number of studies have been conducted to use convolutional neural networks for denoising SAR images, and the denoising effect is better than the traditional algorithms using spatial filtering.

In a convolutional neural network, the convolutional layers are the most numerous and occupy the majority of the network. The main role of the convolution layer is to extract information from the input feature map and use the extracted information as input for the next layer. The convolution kernel is the core of the entire convolutional computation, which can be viewed as a sliding window that completes the multiplicative accumulation computation from top to bottom according to the step size from left to

ISSN 2616-5775 Vol. 5, Issue 11: 97-103, DOI: 10.25236/AJCIS.2022.051115

right, thus obtaining the eigenvalues of each part. In order to extract more image features, a large number of convolutional kernels with different parameters are used in the convolutional neural network, and when the size of the convolutional kernel is larger, the receptive field is larger but the computation is also larger, and when the size of the convolutional kernel is smaller, the computation is smaller but the receptive field is also smaller. Most of the current neural networks use a convolutional kernel of size 3. Since the convolutional computation is linear, a nonlinear activation function is added between each two layers of the convolutional computation. The commonly used activation functions are Sigmoid, Relu, Tanh and other functions.

3. Proposed Method

3.1. Network Architecture

In this section, our proposed network architecture for CDNet will be introduced. The details of the proposed CDNet are shown in Figure 1, and the CDNet mainly consists of our proposed CD(Cooperative denoising) layer. The CNN architecture proposed by G. Chierchia et al. for noise reduction of SAR images first processed the input image using the log function, which transformed the SAR image into the log domain space, and then completed the denoising function by learning the relevant image information through the structure of CNN. In our proposed CDNet this approach is not adopted, instead of transferring the input SAR image to the logarithmic domain, the training is done directly and then the learning of speckle noise is completed by multiple CDNet layers. This allows CDNet to learn the speckle noise, and finally the denoised image can be obtained by simply dividing the input SAR image by the learned noise.

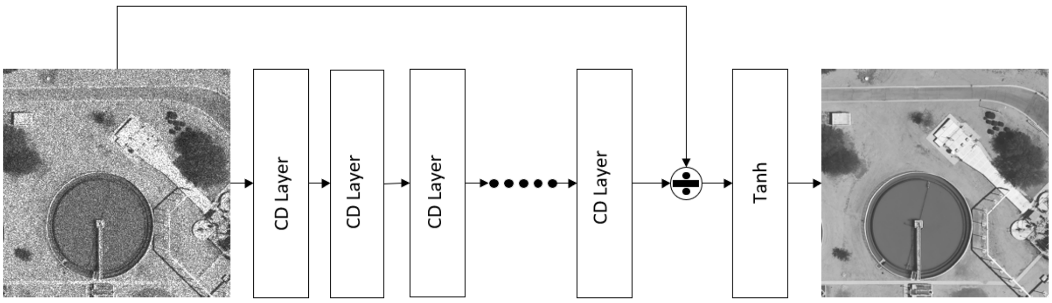

Figure 1: Network structure of CDNet.

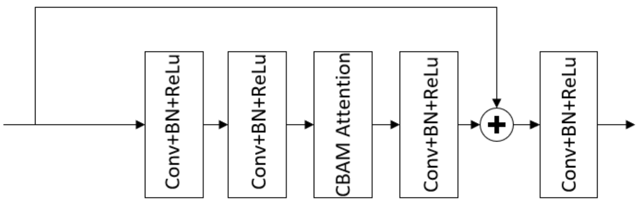

Figure 2: Structure of the CD layer.

The structure of the CD layer is shown in Figure 2. The whole CD layer consists of 4 convolutions, BatchNormal, ReLu and a CBAM(Convolutional Block Attention Module) [9] attention mechanism. In CDNet, multiple CD layers are used, so here we do not specify the number of input and output channels of the convolution in the CD layer, in practice you can choose the number of input and output channels of the convolution in the CD layer according to your needs, it should be noted that the size of the convolution kernel in the CD layer is 3, but the number of input and output channels need to be recalculated when using.

The number of channels of the convolutional kernels in the CD layer is calculated in such a way that the input and output channels of the first convolutional kernel are the same, the input channel of the second convolutional kernel is the same as the output channel of the first convolutional kernel, the output channel of the second convolutional kernel is twice the input channel, the input channel of the third convolutional kernel is the same as the output channel of the second convolutional kernel, the output channel of the third convolutional kernel is one-half the input channel, the input channel of the fourth convolutional kernel is the same as the output channel of the third convolutional kernel, and the output

ISSN 2616-5775 Vol. 5, Issue 11: 97-103, DOI: 10.25236/AJCIS.2022.051115

channel of the fourth convolutional kernel is the same as the input channel of the next CD layer.

Humans always care about the parts they want to care about, and ignore the parts they don't want to care about, so they can get the information they want faster and easier. The attention mechanism mimics this human mechanism by allowing the AI to learn the important and unimportant parts of the data so that the trained AI model can pay more attention to the information it wants to focus on. Therefore, we also use the CBAM attention mechanism in the proposed CD layer to further improve the denoising ability of CDNet. The CBAM attention mechanism consists of two parts, one is the CAM, which is the channel attention mechanism, and the other is the SAM, which is the spatial attention mechanism, which are independent of each other, thus saving computational resources and ensuring that the CBAM module can be easily integrated into other networks.

Although a deeper neural network means that the network is able to extract more information, as the depth of the neural network increases it also introduces the problem of gradient dispersion. In the ResNet published by He [10] et al. a residual network is used to solve the gradient dispersion problem caused by deep neural network training. This residual network approach is used in our proposed CD layer, where the output feature map of each CD layer is obtained by first convolving the input feature map several times and then adding the input feature map to this result and finally by another convolution.

3.2. Loss Function

The loss function is an important part of the deep learning model. When the difference between the predicted result and the correct result of the deep learning model is large, the value of the loss function will be larger, and when the prediction of the deep learning model is more accurate, the value of the loss function will become smaller. A good loss function is helpful for the convergence of convolutional neural networks. The loss function may be different in different application scenarios, and there has been a large amount of research on the loss function within the field of SAR image denoising. In this network of SAR-DRN, the authors used the more popular mean squared error(MSE) loss function. And in ID-CNN [11] the authors used a combination of euclidean loss(L2) function and total variance loss function in order to better recover the features of the images.

In the field of image restoration, image quality is an important evaluation criterion, and we also tend to pay more attention to whether the denoised image is more in line with the intuitive visual perception. The SSIM (Structural Similarity Index Measure) measures the degree of distortion of an image and also the degree of similarity between two images. SSIM includes three aspects: luminance, contrast and structure. The value of SSIM ranges from 0 to 1. When the value of SSIM is larger, it means that the two images are structurally similar, and Equation 4 indicates the use of SSIM as the loss function. Where μ_x is the mean of image x, μ_y is the mean of image y, σ_x^2 is the variance of image x, σ_y^2 is the variance of image y, and σ_{xy} is the covariance of image x and image y, c_1 and c_2 are a constant set to avoid the denominator being zero.

$$L_{ssim} = 1 - \frac{(2\mu_x\mu_y+c_1)\times(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)\times(\sigma_x^2+\sigma_y^2+c_2)} \tag{4}$$

In image restoration tasks (e.g., image denoising), noise points can affect image quality, and many denoising algorithms cannot reduce the effect of noise. For both noisy and noiseless images, the total variation of the noisy image is significantly larger than the total variation of the noiseless image, so reducing the value of the total variation increases the denoising ability. The full variance loss is defined as:

$$L_{TV} = \sum_{w=1}^{W}\sum_{h=1}^{H}\left[\left(\hat{X}^{w+1,h} - \hat{X}^{w,h}\right) + \left(\hat{X}^{w,h+1} - \hat{X}^{w,h}\right)\right]^{\frac{1}{2}} \tag{5}$$

where W and H represent the length and width of the image, and $\hat{X}$ represents the image output from the network.

Models using MSE as the loss function are likely to force-fit singularity data in order to reduce the value of the loss function, which has an impact on the prediction results. The huber loss function requires a pre-given parameter δ. It uses a squared error when the prediction deviation is less than δ, and a linear function when the prediction deviation is greater than δ. Huber loss function as

$$L_\delta = \begin{cases} \frac{1}{2}a^2, & for\ |a| \le \delta \\ \delta\left(|a| - \frac{1}{2}\delta\right), & otherwise \end{cases} \tag{6}$$

ISSN 2616-5775 Vol. 5, Issue 11: 97-103, DOI: 10.25236/AJCIS.2022.051115

where a denotes the actual value minus the predicted value.

Based on the above mentioned SSIM loss function, total variance loss function and huber loss function we propose new loss functions such as

$$\mathrm{L} = L_{\delta} + \lambda_{ssim} L_{ssim} + \lambda_{TV} L_{TV} \tag{7}$$

where λ_{ssim} and λ_{TV} represent the weights of the two loss functions, and the proportion of these two loss functions to the total loss function can be changed by mediating the size of the weights.

4. Experiment

4.1. Data Set

Because the scattering coefficients fluctuate randomly during the imaging process of SAR causing noise in SAR images, clean SAR images do not exist in reality, so there is a lack of labels for network training. So to train the network we construct the dataset by adding speckle noise to the optical images, constructed as

$$\mathrm{Y} = \mathrm{FX} \tag{8}$$

where X is the original optical image, Y is the image after adding noise, and F is the random speckle noise obeying the gamma distribution.

The optical dataset we used is the UCMerced_LandUse dataset used for classification. This dataset includes 21 types of data such as airplane, forest, port, and river, covering a large number of scenes that are similar to the real SAR image scenes.

4.2. Evaluation Indicators

The evaluation of speckle removal in SAR images is usually based on several aspects such as the degree of speckle removal, the degree of preservation of image edge texture and the degree of smoothness of the image area. The current evaluation of spot removal is based on two parts: subjective and objective evaluation.

Subjective evaluation is to judge the effectiveness of removing speckle noise from your own speckle noise removal images by visual observation. Although subjective evaluation can evaluate the effectiveness of speckle noise removal, this approach is influenced by the evaluator's emotion, experience, and other aspects, and lacks a uniform evaluation index.

Objective evaluation metrics are mainly done by computers and do not require human intervention. The objective evaluation method evaluates the denoising effect by a uniform formula and compares it with other algorithms. The commonly used evaluation metrics are peak signal-to-noise ratio (PNSR) and structural similarity index measure (SSIM).

4.3. Comparative Analysis of Results

Table 1: Comparison with other methods.

Method	L=1		L=2		L=4		L=8	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Lee	22.78	0.3735	23.29	0.4132	24.84	0.4740	25.84	0.5192
PPB	22.85	0.5385	24.85	0.6268	26.79	0.7110	28.66	0.7841
SAR-BM3D	24.63	0.6668	26.28	0.7088	28.51	0.7980	30.24	0.8493
SAR-CNN	25.86	0.6795	27.04	0.7348	29.29	0.8142	30.74	0.8564
SAR-DRN	25.94	0.6911	27.49	0.7500	29.10	0.8071	30.91	0.8585
CDNet	26.69	0.7366	28.44	0.7964	29.64	0.8425	31.42	0.9024

To illustrate that our proposed CDNet denoising algorithm outperforms other denoising algorithms, we have done tests on noise levels 1, 2, 4, and 8, we did comparison experiments with BM3D, Lee, SAR-CNN and other algorithms. The experimental results are shown in Table 1.

From Table 1, it can be seen that CDNet shows good performance compared to several other denoising methods in the above four noise levels. The values of both PNSR and SSIM increase as the noise level level increases, while the PNSR and SSIM values of CDNet are better than other methods at

ISSN 2616-5775 Vol. 5, Issue 11: 97-103, DOI: 10.25236/AJCIS.2022.051115

the same noise level. The SSIM value reflects the similarity of luminance, contrast, and structure between two images or two signals. PNSR is more concerned with the relationship between pictures and noise. It can be seen from Table 1 that the CDNet proposed in this paper has high SSIM and PNSR values, which shows that CDNet has good denoising ability and at the same time has good ability to preserve image detail texture.

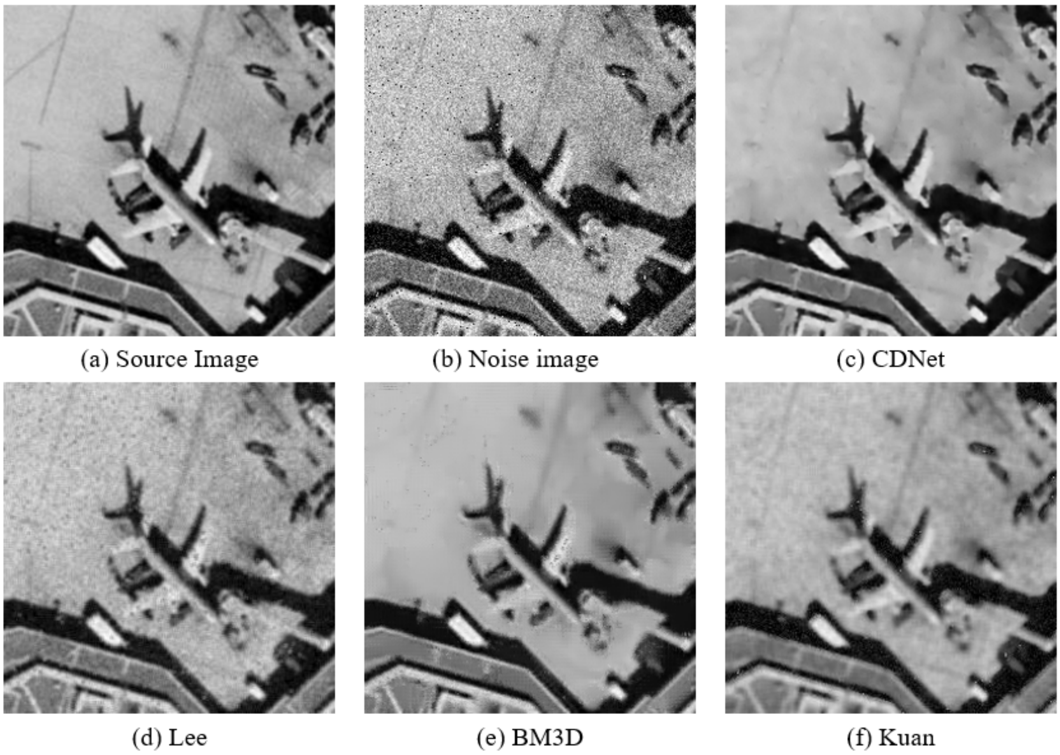

Figure 3: Simulated SAR image denoising comparison.

We visualize the denoising effect of many algorithms in Figure 3. From Figure 3, it can be seen that the Lee filter is the worst, and the denoised image still has a lot of speckle noise. The BM3D and Kuan algorithms remove the speckle noise but lose much of the image information and the edges of the image become blurred. And the denoising effect of our proposed CDNet is superior to other algorithms, while retaining the detailed information of the image and the edges of the image are well preserved.

5. Conclusion

In this paper we propose an algorithm for speckle noise removal from SAR images, using the proposed CD layer to form the backbone feature extraction network of CDNet, using residual connectivity in the CD layer for solving the gradient dispersion problem, and using the CBAM attention mechanism to improve the feature extraction capability. Compared with other methods for removing speckle noise, our proposed CDNet has a good ability to remove speckle noise while preserving the detailed information of the image. In future work we will explore the use of transformer's network structure for better removal of speckle noise from SAR images.

References

[1] J. -S. Lee, "Digital Image Enhancement and Noise Filtering by Use of Local Statistics," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-2, no. 2, pp. 165-168, March 1980, doi: 10.1109/TPAMI. 1980. 4766994.
[2] D. T. Kuan, A. A. Sawchuk, T. C. Strand and P. Chavel, "Adaptive Noise Smoothing Filter for Images with Signal-Dependent Noise," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-7, no. 2, pp. 165-177, March 1985, doi: 10.1109/TPAMI.1985.4767641.
[3] V. S. Frost, J. A. Stiles, K. S. Shanmugan and J. C. Holtzman, "A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise," in IEEE Transactions on Pattern

ISSN 2616-5775 Vol. 5, Issue 11: 97-103, DOI: 10.25236/AJCIS.2022.051115

Analysis and Machine Intelligence, vol. PAMI-4, no. 2, pp. 157-166, March 1982, doi: 10.1109/TPAMI.1982.4767223.

[4] S. Parrilli, M. Poderico, C. V. Angelino and L. Verdoliva, "A Nonlocal SAR Image Denoising Algorithm Based on LLMMSE Wavelet Shrinkage," in IEEE Transactions on Geoscience and Remote Sensing, vol. 50, no. 2, pp. 606-616, Feb. 2012, doi: 10.1109/TGRS.2011.2161586.

[5] Krizhevsky, Alex, Ilya Sutskever and Geoffrey E. Hinton. "ImageNet classification with deep convolutional neural networks." Communications of the ACM 60 (2012): 84 - 90.

[6] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," in IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3142-3155, July 2017, doi: 10.1109/TIP. 2017. 2662206.

[7] G. Chierchia, D. Cozzolino, G. Poggi and L. Verdoliva, "SAR image despeckling through convolutional neural networks," 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2017, pp. 5438-5441, doi: 10.1109/IGARSS.2017.8128234.

[8] Q. Zhang, Q. Yuan, J. Li, Z. Yang, and X. Ma, "Learning a Dilated Residual Network for SAR Image Despeckling," Remote Sensing, vol. 10, no. 2, p. 196, Jan. 2018, doi: 10.3390/rs10020196.

[9] Woo, Sanghyun, Jongchan Park, Joon-Young Lee and In-So Kweon. "CBAM: Convolutional Block Attention Module." ECCV (2018).

[10] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[11] P. Wang, H. Zhang and V. M. Patel, "SAR Image Despeckling Using a Convolutional Neural Network," in IEEE Signal Processing Letters, vol. 24, no. 12, pp. 1763-1767, Dec. 2017, doi: 10.1109/LSP.2017.2758203.