# Bayesian SAR Amplitude Image Compressive Sensing Based on Directional Lifting Wavelet Transform

Xingsong Hou [1], Lan Zhang [1], Chen Gong [2], Lin Xiao [1], Jinqiang Sun [1] and Xueming Qian [1]

1：Xi'an Jiaotong University, 2：Qualcomm Research San Diego, USA, Email: houxs@mail.xjtu.edu.cn

## ABSTRACT

Compared with the traditional Discrete Wavelet Transform (DWT), DLWT (Directional Lifting Wavelet Transform) features better compressibility, interscale attenuation and intrascale directional clustering property for SAR (Synthetic Aperture Radar) amplitude images. In this work, a new Bayesian SAR amplitude image compressive sensing （CS） algorithm based on DLWT (DLWT_TDC: DLWT_Tree_ Directional Clustering) is proposed which fully exploits interscale attenuation and intrascale directional clustering property of the DLWT coefficients. To exploit the intrascale clustering property more accurately, a new directionally and locally adaptive prior probability model is proposed for the three high frequency subbands. Experimental results show that DLWT_TDC can achieve the best reconstruction performance for high sampling rates 0.5 to 0.9. We also observed that, compared with the DWT-based CS reconstruction algorithm, DLWT-based CS reconstruction algorithm achieved better reconstruction performance at the typical 0.5 sampling rate.

***Index Terms*— SAR, CS, DLWT, Bayesian**

## 1. INTRODUCTION

Compressive Sensing for SAR amplitude image has attracted significant interest in recent years. According to the fundamental theory of compressive sensing, assuming a P-dimensional non-sparse signal $f$, which is compressible in a wavelet basis represented by the matrix $\Psi$ :

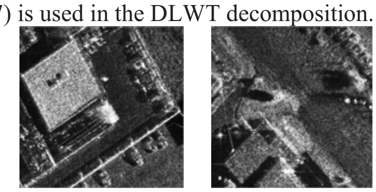
$$\boldsymbol{f} = \sum_{i=1}^{P} x_i \Psi_i \tag{1}$$

The measurement process can be represented as: $\boldsymbol{y} = \Phi \boldsymbol{f}$ and $\Phi \in R^{N \times P} (N << P)$. When measurement matrix $\Phi$ and sparse basis $\Psi$ satisfy the Restricted Isometry Property (RIP) condition, it can be measured randomly and reconstructed with deterministic method, and Bayesian statistical inference method[1]. He et al. [1] applies the CS reconstruction for natural images by exploiting interscale dependency of DWT coefficients. However, these algorithms have not fully exploited the dependency of DWT coefficients, e.g., the intrascale dependency. Moreover, compared with DWT, DLWT coefficients achieve better compressibility, interscale attenuation and intrascale clustering property. DLWT has already been used to represent the complex SAR images and achieved high performance for complex SAR images compression [2]. However, to the best of our knowledge, DLWT has not been used for SAR amplitude image CS reconstruction based on our survey. Therefore, in the paper, a new DLWT-based Bayesian CS reconstruction algorithm (DLWT_TDC) is proposed by fully exploiting intrascale and interscale correlation of DLWT coefficients.

The rest of the paper is organized as follows: in Section 2, DLWT and DWT coefficients' property are illustrated for SAR amplitude images. In Section 3, the proposed algorithm is presented in detail. Experimental results and performance analysis are presented in Section 4. Finally, conclusion is shown in Section 5.

## 2. DLWT and DWT COEFFICIENTS' PROPERTY ANYLISIS FOR SAR AMPLITUDE IMAGES

For SAR amplitude images, we compare the differences in compressibility, interscale and intrascale dependency between DWT and DLWT. We test two $512 \times 512$ with 16 bpp (bits per pixel) SAR amplitude images(as shown in Fig. 1) obtained by the modulo operation with the real part and imaginary part of the complex SAR images downloaded from the Sandia National Laboratories of the United States[3]. Five scalar wavelet decomposition (including both DWT and DLWT) are performed and the biorthogonal 9/7 filter (B9/7) is used in the DLWT decomposition.

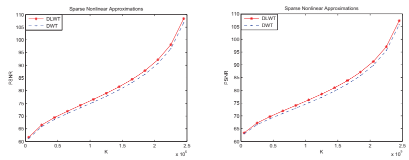


(a) Img1 (b) Img2

**Fig.1.** $512 \times 512$ test images.

### 2.1. K-term Nonlinear Approximation

We employ K-term nonlinear approximation to show the sparse representation ability. Fig.2 plots the PSNR of the reconstructed image against the number of the sorted large amplitude coefficients for both DLWT and DWT. It is seen that DLWT outperforms DWT in terms of larger PSNR for the same number of the sorted large amplitude coefficients.



(a) Img1 (b) Img2

**Fig.2.** K-term nonlinear approximation performance.

### 2.2. Interscale Attenuation Property

The wavelet coefficients show the following interscale attenuation property: 1) the energy is concentrated to the low frequency subbands; 2) if a parent coefficient is insignificant, its children coefficients tend to be insignificant. We use energy distribution percentage of wavelet coefficients to compare the interscale dependency between DWT and DLWT.

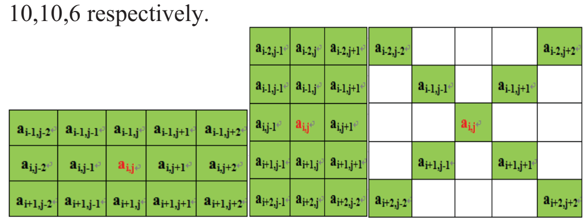
| Test images | Represen-tation | Scale | | | | | |
|---|---|---|---|---|---|---|---|
| | | LL | LL | LL | LL | LL | LL |
| Img1 | DLWT | 0.4002 | 0.0857 | 0.1104 | 0.1423 | 0.1779 | 0.084 |
| | DWT | 0.3609 | 0.0569 | 0.0723 | 0.1380 | 0.2413 | 0.130 |
| Img2 | DLWT | 0.2538 | 0.1438 | 0.1774 | 0.1430 | 0.1887 | 0.0934 |
| | DWT | 0.2818 | 0.0490 | 0.1150 | 0.1598 | 0.2390 | 0.1553 |

**Table 1.** Energy distribution percentage in each scale

LL represents the low frequency coefficients and H5,H4,H3,H2 and H1 represent high frequency coefficients from scale 5 to scale 1, respectively. TABLE 1 shows that energy in LL, H5, and H4 subbands for DLWT is higher than that of DWT coefficients. In other words, for DLWT the coefficients are more concentrated in low frequency subbands, which shows that DLWT has better interscale attenuation property.

### 2.3. Intrascale Clustering Property Analysis

Wavelet coefficients have the following intrascale clustering property: 1) if the number of the significant coefficients in its neighboring $c \times d$ block is bigger than threshold K, we impose the belief that wavelet coefficient is likely to be also significant whereas the opposite; 2) coefficients in different subband have different directional correlation. It is known that the HL, LH, HH subbands have better horizontal correlation, vertical correlation, and diagonal correlation, respectively. Thus, we employ $3 \times 5$ , $5 \times 3$ and $5 \times 5$ neighboring block for HL,LH and HH subbands, respectively (as shown in Fig. 3). Take HL $3 \times 5$ neighboring block(Fig.3(a)) for example, there are 14 wavelet coefficients among the preserved significant coefficient $a_{i,j}$ , where i and j denote the rows and columns of the image, respectively. Then, if the number of significant coefficients for the 14 coefficients is bigger than a setting threshold 10, we regard it as a cluster. The threshold is determined on the basis of a large number of experiments. In the proposed algorithm, the threshold for $3 \times 5$ 、 $5 \times 3$ 、 $5 \times 5$ neighboring block is set as 10,10,6 respectively.



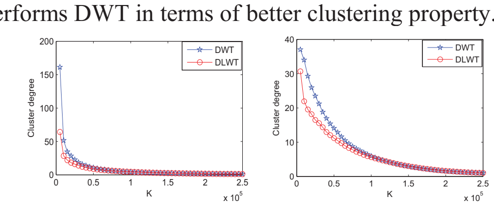
(a) HL $3 \times 5$ (b) LH $5 \times 3$ (c) HH $5 \times 5$

**Fig.3.** Neighboring block.

We count the number of clusters in wavelet-domain with 5-level transform of DLWT and DWT for various numbers of significant coefficients K in the K-term nonlinear approximation, for the 4 images in Fig.1. Then the cluster degree of the two transforms is defined as follows:

$$cluster\ degree = \frac{K}{C} \tag{2}$$

where $C$ is the number of clusters. For the same $K$ , smaller cluster degree means the coefficients have more clustering property. Fig.4 plots the cluster degree against the number of preserved significant coefficients K for the 2 images in Fig.1, for both DLWT and DWT. It is seen that DLWT outperforms DWT in terms of better clustering property.



(a) Img1 (b) Img2

**Fig.4.** Comparison of clustering property between DWT and DLWT.

### 3. BAYESIAN SAR AMPLITUDE IMAGE CS BASED ON DLWT

In this section we propose a Bayesian CS algorithm for SAR amplitude images. We first establish a prior probability model which fully exploits wavelet coefficients' dependency. In this work, we employ the spike-and-slab prior model [4], given as follows,

$$x_i \sim (1-\pi_i)\delta(0)+\pi_i\mathcal{N}(0,\alpha_s^{-1}) \qquad i=1,2,\cdots,P \tag{3}$$

$$\alpha_s \sim Gamma(c_0,d_0), \quad s=1,2\cdots L \tag{4}$$

$$\boldsymbol{y}|\boldsymbol{x},\alpha_n \sim \mathcal{N}(\boldsymbol{\Phi x},\alpha_n^{-1}\mathbf{I}) \tag{5}$$

$$\alpha_n \sim Gamma(a_0,b_0) \tag{6}$$

where $x_i$ denotes the $i$th wavelet coefficient and P is the number of the wavelet coefficients. In the following we elaborate the model described by equation (3). $\delta(0)$ is a point mass concentrated at zero modeling insignificant coefficients; the second component is a zero-mean Gaussian distribution with variance $\alpha_s^{-1}$ . $\pi_i$ denotes the probability of the significant coefficients. $\alpha_n^{-1}$ represents the unknown noise variance. *Gamma*(.,.) and *Beta*(.,.) are the *Gamma* and *Beta* functions, respectively, and $a_0$, $b_0$, $c_0$ and $d_0$ are the hyper-parameters.

The TSW-CS reconstruction algorithm [1], which uses DWT as sparse representation for natural images, only exploits the interscale attenuation property of the wavelet coefficients. In this paper, we propose the following model for $\pi_i$ based on whether the parent is significant and the number of significant coefficients in the neighboring block. More specifically, we have the following model,

$$\pi_i = \begin{cases} \pi_r & if \quad s=1 \\ \pi_{s,i,b}^{0,0} & if \quad 2 \le s \le L \quad x_{pa(s,i)} \ insignificant \quad x_{s,i,b}^{c\times d}=0 \\ \pi_{s,i,b}^{0,1} & if \quad 2 \le s \le L \quad x_{pa(s,i)} \ insignificant \quad x_{s,i,b}^{c\times d}\neq 0 \\ \pi_{s,i,b}^{1,0} & if \quad 2 \le s \le L \quad x_{pa(s,i)} \ significant \quad x_{s,i,b}^{c\times d}=0 \\ \pi_{s,i,b}^{1,1} & if \quad 2 \le s \le L \quad x_{pa(s,i)} \ significant \quad x_{s,i,b}^{c\times d}\neq 0 \end{cases} \quad (7)$$

$$\pi_r \sim Beta\left(e_0^r, f_0^r\right), s=1 \quad (8)$$

$$\pi_{s,i,b}^{0,0} \sim Beta\left(M_b^{0,0}, N_b^{0,0}\right) \quad 2 \le s \le L \quad (9)$$

$$\pi_{s,i,b}^{0,1} \sim Beta\left(M_b^{0,1}, N_b^{0,1}\right) \quad 2 \le s \le L \quad (10)$$

$$\pi_{s,i,b}^{1,0} \sim Beta\left(M_b^{1,0}, N_b^{1,0}\right) \quad 2 \le s \le L \quad (11)$$

$$\pi_{s,i,b}^{1,1} \sim Beta\left(M_b^{1,1}, N_b^{1,1}\right) \quad 2 \le s \le L \quad (12)$$

where $x_{pa(s,i)}$ and $x_{s,i,b}^{c\times d}$ denote the corresponding parent coefficients in the next level and the neighboring block coefficients of the subband $b \in \{HL, LH, HH\}$ at the same scale, respectively; $\pi_r$ denotes the probability of the significant coefficients $x_i$ in the biggest scale.

The hyper-parameters $a_0$ , $b_0$ , $c_0$ , $d_0$ , $e_0^r, f_0^r, M_b^{0,0}, N_b^{0,0}, M_b^{1,0}, N_b^{1,0}, M_b^{0,1}, N_b^{0,1}, M_b^{1,1}, N_b^{1,1}$ are set respectively as follows:

$$a_0=b_0=c_0=d_0=10^{-6} \quad \left[e_0^r, f_0^r\right]=\left[0.9, 0.1\right]\times P_1 \quad (13)$$

$$[M_b^{0,0}, N_b^{0,0}] = \left[1/P, 1-1/P\right] \times P_S \quad (14)$$

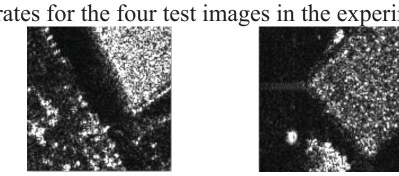
$$[M_b^{1,0}, N_b^{0,1}] = [M_b^{1,0}, N_b^{0,1}] = \left[0.5, 0.5\right] \times P_S \quad (15)$$

$$[M_b^{1,1}, N_b^{1,1}] = \left[1-1/P, 1/P\right] \times P_S \quad (16)$$

where $p_s$ and $p$ represent the number of wavelet coefficients in scale s( $1 \le s \le L$ ) and the total number of wavelet coefficients ,respectively.

## 4. EXPERIMENTAL RESULTS

We test the performance of the DLWT based CS algorithms for real SAR amplitude images. We compare the efficiency of DLWT_TDC with that of DLWT_T which only uses interscale dependency of DLWT coefficients and that of the DLWT_IID which assumes the DLWT coefficient obeying the independently identical distribution (IID). Moreover, we compare performance with MS-BCS-SPL algorithm[5] which uses a dual-tree DWT (DDWT) as the sparsity transform and blocks of sizes $B_l$ = 2,4,8,16 for decomposition levels $l$ = 1, 2, 3,4 respectively ($l$ = 4 is the highest-resolution level). To accelerate the experimental speed, we cut the images in Fig.1 into two 16 bpp test SAR amplitude images with size $128\times128$ , as shown in Fig.5. Four-level DLWT with B9/7 filter is performed at different sampling rates for the four test images in the experiments.

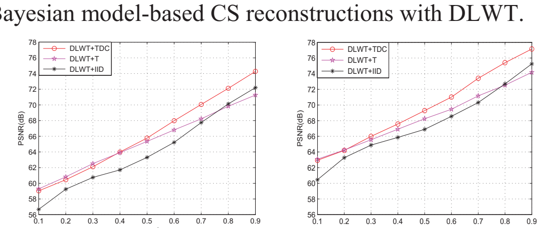


(a)Img1 (257:384,1:128) (b)Img2(385:512,1:128)

**Fig.5.** $128\times128$ test images

### 4.1. Performance of DLWT-based CS Algorithm at Different Sampling Rates

Fig.6 plots the PSNR of the reconstructed image against the sampling rates 0.1 to 0.9 for the two images, for the three Bayesian model-based CS reconstructions with DLWT.




(a) Fig.5(a) (b)Fig.5(b)

**Fig.6.** *PSNR – percent* diagram at different sampling rate.

From Fig.6, it is seen that DLWT_TDC shows the best reconstructed image quality among all DLWT-based CS schemes for high sampling rate 0.5 to 0.9. Let us see the sampling rate of 0.7. Compared with DLWT_T algorithm, DLWT_TDC can improve PSNR up to 2.25dB. Compared with DLWT_IID algorithm, DLWT_TDC can improve PSNR up to 3.11dB. Compared with MS-BCS-SPL algorithm, DLWT_TDC can improve PSNR up to 2.39dB.

This is because DLWT_TDC fully considers the inter- and intra-scale dependency of DLWT coefficients.

Note that, from Fig.6, when the sampling rate is lower than 0.4, the PSNR of DLWT_TDC and DLWT_T almost coincide with each other. This is due to the fact that at the lower sampling rate, the number of the reconstructed wavelet coefficients is small and thus more independent, and thus the improvement of the DLWT_TDC is negligible. The performance gain of the DLWT_TDC is more significant for the sampling rate higher than 0.4. Among the three DLWT-based Bayesian CS reconstruction schemes, DLWT_TDC and DLWT_T, the algorithms exploiting the dependencies, show better reconstruction performance compared with that of DLWT_IID.

To intuitively show the impact on the estimation of the wavelet coefficients probability distribution when we add inter- and intra-dependency into reconstruction algorithm, the probability distributions of the wavelet coefficients for test image Fig.5(a) is shown in Fig.8 at 0.7 sampling rate.

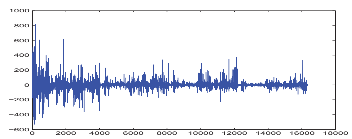
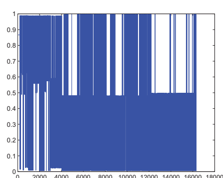
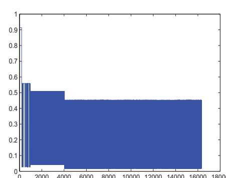


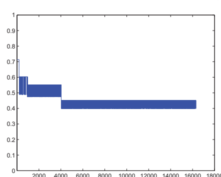
**Fig.7.** The original DLWT wavelet coefficients.



(a) DLWT_TDC



(b) DLWT_T



(c) DLWT_IID

**Fig.8.** The posterior probability distributions of three Bayesian CS reconstruction algorithms.

Fig.7 shows the original wavelet coefficients of the test image shown in Fig.5 (a). Fig.8 shows the posterior probability distribution where x axis denotes the location of wavelet coefficients and y axis denotes the probability of the significant coefficients.

As illustrated in Fig.8 (c), the significant wavelet coefficients' probability is concentrated on 0.5. This is due to the fact that we assume the wavelet coefficients obey the independently identical distribution. From Fig.8 (b) we can see that if the probability for significant parent coefficients is small, the probability for its significant children coefficients is much smaller when we only consider the interscale attenuation property in reconstruction algorithm. However, from Fig.7 we can see that although the parent coefficient is small, its children coefficients are not always small. Thus if we only consider the interscale attenuation property in CS reconstruction algorithm, we cannot give accurate model of the wavelet coefficients. Whether a child coefficient is significant or not is not only determined by its parent coefficient but also by the neighboring block coefficients when we fully exploit interscale attenuation and intrascale directional clustering. Compared with the original wavelet coefficients distribution in Fig.7, we can see that the probability distribution in DLWT_TDC can describe the location of wavelet coefficients more accurately as illustrated in Fig.8(a).

### 4.2. Comparison between DWT and DLWT

Finally we compare the DLWT- and DWT-based CS algorithm. TABEL 2 shows the PSNR values for the three CS reconstruction algorithms based on DWT and DLWT at the sampling rate 0.5. From TABEL 2 we can see that the DLWT-based CS algorithms show higher reconstructed image PSNR than the DWT-based CS counterparts just as illustrated in Section 2.

| Test image | IID | | T | | TDC | |
|---|---|---|---|---|---|---|
| | DLWT | DWT | DLWT | DWT | DLWT | DWT |
| Fig5(a) | 63.306 | 62.5353 | 65.3473 | 64.0339 | 65.774 | 64.7702 |
| Fig5(b) | 66.8935 | 66.2486 | 68.2425 | 67.2035 | 69.2834 | 68.1942 |

**Table 2.**Comparison between DWT- and DLWT- based algorithms

## 5. CONCLUSION

We have adopted DLWT as sparse representation for SAR amplitude images to achieve better compressibility, interscale attenuation property and intrascale clustering property. We have proposed a novel DLWT-based Bayesian CS reconstruction algorithm, which fully exploits the interscale attenuation property and intrascale directional clustering property. The experimental results show correctness and effectiveness of the proposed algorithm.

## 6. ACKNOWLEDGES

This work was supported by 2007 National Natural Science Foundation of China (project no. 60602024) and 2009 Campus Research Foundation of Xi'an Jiaotong University (project no. xjj2012023).

## REFERENCES

[1] L.H.He and L.Carin, "Exploiting Structure in Wavelet-Based Bayesian Compressive Sensing," *IEEE Trans.Signal Process.*, vol.57, no.9, pp. 3488-3497,Sep.2009.

[2] X.S.Hou, J.Yang and G.Jiang, "Complex SAR Image Compression Based on Directional Lifting Wavelet Transform With High Clustering Capability," *IEEE Trans.Geoscience and Remote Sensing,* vol.51, no.1, pp. 527-538,2012.

[3] Sandia National Laboratories, Sandia SAR Data. [Online]. http://www.sandia.gov/radar/sar-data.html

[4] H.Ishwaran and J.S.Rao, "Spike and Slab Variable Selection: Frequentist and Bayesian Strategies," *Ann.Statics.,*vol.33, no. 2, pp. 730-773, Sep. 2005.

[5] E.F.Janmes, M.Sungkwang, and W.T.Eric, " Multiscale Block Compressive Sensing with Smoothed Projected Landweber Reconstruction, " *Proc. EUSIPCO,* 2011.