MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

# Universal Quantization and SAR Compression

Boufounos, P.T.

TR2018-209 March 29, 2019

**Abstract**

As Synthetic Aperture Radar (SAR) technology advances and the resolution and quality of SAR systems improves, there is an increasing need for lightweight compression of SAR raw data. In most satellite-borne SAR systems, due to their limited processing capacity, raw data needs to be transmitted to a ground station for processing. As the resolution and acquisition quality increases, so does the volume of data to be transmitted, making compression necessary. Furthermore, computational constraints on-board such systems impose sever restrictions on the kinds of algorithms that can be implemented, and, therefore, on the compression quality. This report proposes a novel lightweight compression approach, based on the principles of universal quantization, which allows the compression system to exploit the structure of the signal in hindsight, i.e., during the decompression stage. This approach shifts the computational complexity to the decoder, which needs to impose the appropriate image model to recover the data. Thus, the heavy lifting in this approach is performed by the decoder at the ground station, which has significantly more computational resources available.

*MERL website*

# Universal Quantization and SAR Compression

Petros T. Boufounos

March 15, 2019

**Abstract**

As Synthetic Aperture Radar (SAR) technology advances and the resolution and quality of SAR systems improves, there is an increasing need for lightweight compression of SAR raw data. In most satellite-borne SAR systems, due to their limited processing capacity, raw data needs to be transmitted to a ground station for processing. As the resolution and acquisition quality increases, so does the volume of data to be transmitted, making compression necessary. Furthermore, computational constraints on-board such systems impose sever restrictions on the kinds of algorithms that can be implemented, and, therefore, on the compression quality. This report proposes a novel lightweight compression approach, based on the principles of universal quantization, which allows the compression system to exploit the structure of the signal in hindsight, i.e., during the decompression stage. This approach shifts the computational complexity to the decoder, which needs to impose the appropriate image model to recover the data. Thus, the heavy lifting in this approach is performed by the decoder at the ground station, which has significantly more computational resources available.

## 1 Introduction

In radar imaging systems, image resolution is a function of the array aperture. Thus to improve resolution it is necessary to increase the array aperture size. SAR systems, in particular, acquire high-resolution images by exploiting the motion of a moving platform on which an antenna is mounted, thus creating a synthetic aperture, with size significantly larger than the physical antenna. Thus, it is possible to obtain high-resolution images even with a small physical antenna.

The main principle of SAR systems is simple: as the moving platform, e.g., a satellite, is moving along a trajectory, the imaging radar transmits pulses at regular intervals towards a region of interest and receives their reflections. These reflections are recorded and used to synthesize the SAR image of the region of interest, through a process known as image formation. There are several SAR modes of operation, depending on how the antenna beam is directed as the platform moves along its track. For example, in stripmap mode, the antenna beam direction remains fixed with respect to the moving platform, i.e., as the platform moves, the beam slides along the region of interest. In spotlight mode, instead, the beam is always directed to the same spot on the ground, thus obtaining more views of this spot from a larger effective aperture.

In typical SAR systems, stripmap mode has lower resolution than spotlight mode, while spotlight mode images a smaller area. Other modes have also been developed and might be available, such as sliding mode—which has flexible trade-off between resolution and area size, operating in-between stripmap and spotlight modes—and scan mode—which allows imaging and even larger areas than stripmap mode, albeit at even lower resolution. An overview of SAR principles can be found in [1,2]

Despite the variety of SAR modes, the fundamental structure of SAR data remains the same. For each transmitted pulse, a reflection of fixed length, in samples, is recorded by sampling the received signal in baseband, i.e., acquiring a sequence of complex-valued samples. Thus, the SAR data comprise of several lines of complex numbers of fixed length. The index of each line is often referred to as the azimuth dimension, the along-track dimension, or the slow time. The index of each sample within a line is often referred to as the range dimension, the across-track dimension, or the fast time. Thus, for all SAR modes, the raw SAR data can be expressed as a two-dimensional complex-valued array of numbers.

Once the data is acquired, the SAR image is formed using one of many algorithms, according to the SAR mode of operation. These algorithms include the range-Doppler algorithm (RDA), the chirp scaling algorithm (CSA), and the Omega-$K$ algorithm ($\omega$-$k$), among others. While the problem can be expressed as a delay-and-sum operation, these algorithms explore several trade-offs and simplifying assumptions to achieve significant gains in computational efficiency, by exploiting the structure of the problem and fast algorithms, such as the fast Fourier transform (FFT). Different assumptions and different approximations are appropriate for each mode of operation and each application, according to system, application, and processing requirements.

Despite the existence of fast algorithms, image formation is a too complex a task to be performed on-board a satellite. Thus, the raw SAR data needs to be transmitted to a ground station and processed there to obtain a SAR image. In most modern systems, considering the data size and the communication link capacity, it is necessary to compress the data before transmitting it. Unfortunately, at first inspection, raw SAR data do not seem to exhibit any exploitable structure and resemble uncorrelated noise or noise with low correlation properties [3]. In contrast, formed SAR images exhibit significant structure which can be used to achieve greater compression rates. Thus, it is, in principle, possible to compress the raw SAR data by first performing image formation, which is an invertible process, and then compressing the formed image. Of course, this is not possible in practice, since it is prohibitively expensive to perform on-board a satellite, but it demonstrates that raw SAR data are, in principle, compressible. The main goal of on-board compression algorithms is to be able to exploit some structure but at a very low complexity and computational cost.

# 2 Background

## 2.1 SAR as an Inverse Problem

While there are different modes of SAR operation, mathematically they can all be represented by the same high-level linear model. In particular, given a two-dimensional region of interest (ROI) denoted $X$, the acquired two dimensional data $Y$ are obtained through a linear operator $A(\cdot)$:

$$Y = A(X). \tag{1}$$

Depending on the mode of operation, the forward operator $A(\cdot)$ describes how the complex image $X$ is converted to raw SAR data, according to the operating mode of the SAR system.

In order to be able to acquire a SAR image without significant artifacts, such as spatial aliasing, the typical assumption is that $A$ is invertible. Significant literature has been devoted to inverting $A$ efficiently, which provides good algorithms for implementing $A$, its adjoint $A^*$, and its inverse $A^{-1}$ with low computational complexity. Depending on the system design, appropriate algorithms include the RDA, CSA, $\omega$-$k$, all of which exhibit different trade-offs between efficiency, accuracy, and applicability to different SAR modes.

This general characterization of SAR as a linear system has spawned significant work on considering SAR in the context of general inverse problems, and applying recent advances in this area. For example, there is significant work on applying compressive sensing techniques to reduce the required number of SAR pulses [4–6], to provide robustness to acquisition error, such as saturation [7], or to improve the operating characteristics, such as ROI size and resolution [8–12]. A significant component of these efforts relies on appropriate signal models for $X$, such as sparsity under a basis or a learned dictionary. It should be noted, however, that SAR images are not as sparse as regular images, primarily because the phase component does not have easily exploitable structure, but also because they exhibit significant speckle noise.

## 2.2 SAR Raw Data Compression

Typical SAR raw data compression approaches include Block-Adaptive Quantization (BAQ) [13, 14] and Flexible Block-Adaptive Quantization (FD-BAQ) [15, 16] and their variants, such as [3] and references within. These algorithms operate on raw data blocks and do not attempt to take an image model into account. In particular, image blocks are modeled as an i.i.d. process of a certain distribution—typically Gaussian or uniform—and an optimal quantizer is designed for this process, trading-off the target data rate and desired quality. The output is then coded using an entropy coder, the complexity of which varies, depending on the computational power available on-board.

Typical variations include lightweight pre-processing of the data, for example filtering, de-chirping, or using a small low-complexity fast Fourier transform (FFT) to transform the data to the frequency domain. In addition, FD-BAQ further exploits prior information about the level of noise in each block to adapt the bit-rate used to encode the block. However, even with those variations, none of the existing popular approaches is able to exploit the structure of the image generated by the image formation algorithm.

## 2.3 Universal Quantization

Universal quantization, first introduced in [17], has been proven a promising approach in a number of applications requiring lightweight compression of different aspects of a signal, although not the signal itself. The main tenet of universal quantization is that higher order bits often contain redundant information and can be discarded. In particular, when prior information about the signal can be exploited, a signal prediction of reasonable accuracy can be estimated. If the signal is transformed using a randomized incoherent measurement process, e.g., using a Gaussian random projection matrix, then it is possible to determine higher order bits with very high probability, whereas the lower-order bits are more difficult to determine.

Universal quantization-based compression approaches shift the complexity to the decoder. In a distributed source coding setup, such as [18–20], the decoder will try to estimate a good approximation of the signal, incorporating all known models and side information, attempting a prediction of the measurements. Since the least significant bits are more difficult to predict, the encoder should encode correction information at a higher rate, often transmitting them as is. On the other hand, higher order bits are easier to predict, so the correction information requires a lower rate to transmit. Thus, similarly to classical distributed coding, the correction information becomes the compression of the signal. However, this approach requires a good prediction of the signal to be available at the decoder, which is usually originating from side information.

Unfortunately, in the case of compression of SAR raw data, side information is not available. In this case, hierarchical universal quantization [21] can be used to code the signal. This approach, however, is not competitive in practice and requires significant parameter tuning. Still, it demonstrates that it is possible to use a few measurements of the signal to provide the signal prediction and recover bits of higher order. The remainder of this report exploits this insight and explores approaches to apply methods based in universal quantization to lightweight compression of SAR data, in which side information is not available. We qualitatively compare with BAQ and FDBAQ approaches, and describe their comparative advantages and disadvantages.

# 3 Universal Quantization and SAR Measurements

## 3.1 Baseline Image

In order to generate a good prediction of the measurements, we first need to obtain a good prediction of the signal. In particular, it is by now well-understood that SAR images exhibit structure that can be exploited using compressive sensing principles [4–12, 22–25]. Thus, given the linear SAR acquisition system (1), it is possible to reconstruct, with reasonable error, the SAR image from subsampled quantized data

$$Y = Q(S(A(X))), \tag{2}$$

where $S(\cdot)$ is a linear subsampling operator, designed such that $S(A(\cdot))$ is incoherent with the signal model for $X$[1], and $Q(\cdot)$ is an element-wise scalar quantizer designed to satisfy the reconstruction quality specs for the end-to-end system. This initial reconstruction can serve as a prediction of the signal that can be used to recover higher order bits of the remaining measurements.

## 3.2 Hierarchical Reconstruction

The principle above can be generalized in a hierarchical fashion with several levels of image recovery, followed by prediction of the higher order bits. In particular we consider a sequence of operators $S_i(\cdot)$, $i = 1, \ldots, I$ such that their range is mutually disjoint and span the range of $A(\cdot)$. In other words, $\text{range}(S_i) \cap \text{range}(S_j) = \{0\}$, for all $i \neq j$ and $\bigcup_i \text{range}(S_i) = \text{range}(A)$. This implies that their direct sum spans the space of measurements: $\bigoplus_i \text{range}(S_i) = \text{range}(A)$. Furthermore, as above, these operators should satisfy the same incoherence properties as $S$ above. We denote the measurements using

$$Y_i = Q(S_i(A(X))), \tag{3}$$

[1] Technically, the term incoherence applies to $X$ exhibiting a sparsity model under some basis $X$. However, there are several alternative models possible for $X$, including manifolds,We abuse the term incoherence to mean that using the signal model and the appropriate reconstruction algorithm, it is possible to estimate the image $X$ from the measurements.

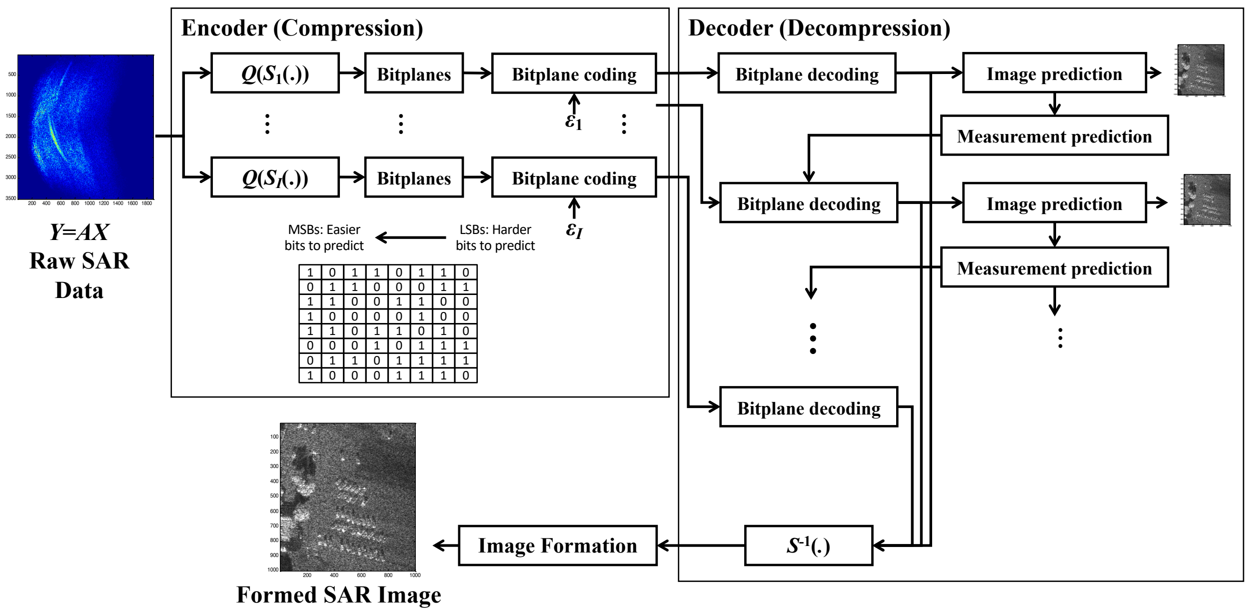


Figure 1: Overview of the proposed system architecture and operation

where $Q(\cdot)$ is designed such that the reconstruction using all $Y_i,\ i = 1, \ldots, I$ using the preferable image formation algorithm satisfies the quality requirements of the system.

In practice, the easiest approach to generate the sequence of $S_i$ is to measure using a randomized full-rank operator of the same dimensionality as $A$ and then partition the measurements to $I$ partitions. Typical examples are matrices generated with random Gaussian entries, or a randomized fast transform. In particular for SAR systems, SAR measurements are already significantly incoherent with the SAR image, and, therefore, $S_i$ could be just selecting a randomized non-uniform subsampling of the raw SAR data.

Given the sequence of measurement operators, reconstruction can be performed in a hierarchical sequence, where at hierarchy level $j$ all sets of measurements $Y_i, i = 1, \ldots, j$ are used to reconstruct an estimate $\widehat{X}_j$ using the appropriate reconstruction algorithm for the model. Since the reconstruction at every level of the hierarchy uses more data than the previous level, we expect the estimation error, denoted using

$$\epsilon_j = \left\| \widehat{X}_j - X \right\|_2, \tag{4}$$

to decrease, i.e., $\epsilon_1 \geq \epsilon_2 \geq \ldots \geq \epsilon_I$. The remainder of this development assumes that the sequence of $\epsilon_j$, or an upper bound for it, is known during encoding time. A drawback of this approach is that this sequence is not straightforward to compute, especially when non-linear reconstruction methods are considered.

## 3.3 Hierarchical Coding/Decoding

Using this hierarchy of operators, a hierarchical encoding/decoding strategy, summarized in Fig. 1, follows. The encoder, at each level of the hierarchy starting at $i = 1$, encodes the measurements $Y_i =$

$Q(S_i(A(X)))$ in sequence, protecting bits of higher order less than bits of lower order, i.e., using higher rate to encode lower order bits. The protection of the higher order bits decreases as the level of the hierarchy increases. In particular, at level $i = 1$, all bits are equally protected, i.e., are transmitted as is. As the level $i$ increases, lower order bits of $Y_i$ are transmitted as is, while higher might be dropped, to be fully recovered by the decoder.

There are two approaches possible for encoding correction bits.

- Using the approach in [18, 19], we can recognize that bits higher than a certain order can be almost exactly recovered with very few negligible errors. These bits are discarded and bits of lower order are transmitted as is. This approach is not as efficient, but requires less precise knowledge of the sequence of estimation errors $\epsilon_i$.
- Using the approach in [20], for each bitplane of the measurements we can compute a probability there will be an error using [20, Thm. 1]. This probability can then be used to derive the rate necessary to encode this bitplane, i.e., to protect it from errors in estimation, using distributed coding. Some bitplanes will be completely discarded, to be filled by the decoder, and other will be transmitted as is. Intermediate bitplanes will be encoded with varying rate, increasing as the bitplane order decreases. This approach provides much better compression rates, but has the drawback that it requires more accurate knowledge of $\epsilon_i$ at the encoder.

The decoder works iteratively, starting with level $i = 1$. After the decoding of each level, i.e., at the end of each iteration, the decoder produces two output items: the decoded measurements $Y_i$ for this level, and an estimate of the image $\widehat{X}_i$. The assumption is that at the end of each decoding iteration the measurements $Y_i$ are exactly recovered and available to the next level, which is only possible because $Y_i$ is quantized and taking values in a finite set.

In particular, at each iteration $i$ the decoder uses the image estimate from the previous iteration $\widehat{X}_{i-1}$ to obtain an estimate of the unquantized measurements $\overline{Y}_i = S_i(A(\widehat{X}_{i-1}))$. Using this estimate, and the approach described in [18–20], the decoder can obtain $\widetilde{Y}_i$, the best uncorrected estimate of $Y_i$. However, using the correction bits transmitted by the encoder for this level, as described above, the decoder can produce $\widehat{Y}_i$, a corrected estimate of $Y_i$. Of course, at level $i = 1$ we assume that $\widehat{X}_0 = 0$, and that all bitplanes for $Y_1$ are transmitted as is.

Once all the decoding iterations are finished and $Y$ has been fully recovered, the image formation algorithm of choice can be used on $Y' = S^{-1}(Y)$, where $S(\cdot) = [S_1(\cdot), \ldots, S_I(\cdot)]$ is the combination of all the $S_i(\cdot)$, and $Y = [Y_1, \ldots, Y_I]$ is the combination of all the measurements, i.e., $Y' = A(X) + e$, where $e$ is bounded noise due to the quantization of $S_i(A(X))$.

# 4 Discussion

When comparing with conventional SAR compression approaches, there are several advantages and several disadvantages in this approach.

The first key advantage is the very low complexity at the encoder. Encoding only requires straightforward linear operations followed by quantization for the measurement of the $Y_i$, and simple bit operations for computing the protection bits, either dropping some and transmitting other, or applying simple binary field operations to produce a syndrome using standard coding approaches.

The second key advantage is the ability of the algorithm to exploit the structure of the signal at the decoding, rather than the encoding. In particular, a key ingredient in the success of the algorithm is the signal model used in the decoding stage to recover each $\widehat{X}_i$, which allows better recovery with more accurate signal models, and, therefore, lower values for $\epsilon_i$. This, in turn, translates to improved coding rates for the corresponding $Y_i$. Thus, as expected, better signal models reduce the required rate necessary at the encoder, assuming the encoder is aware of the improvement in $\epsilon_i$ or its upper bounds.

Unfortunately, in SAR systems, the signal models do not perform as well as signal models for more conventional modalities, such as natural images. In particular, speckle and the random phase associated with the image is difficult to take into account in a SAR model. One approach is to only apply the model in the magnitude image [7]. Another is to try to learn SAR-specific image dictionaries. Still, there is significant scope for improvement in this front. Any gains automatically imply gains in the achievable compression rate.

The main disadvantage of this approach is the requirement that an estimate of the error at each hierarchy level is available at the time of encoding. This might be possible to estimate off-line using experience with sample training data. However, this estimate has to be conservative and can be critical. Depending on the encoding choice, an incorrect estimate could lead to transmitting a smaller number of correction bits, and failing to decode the correction at the decoder.

In contrast, conventional approaches, such as BAQ and FDBAQ completely ignore the structure of the image and only rely on the statistical properties of the SAR raw data. The advantage of these approaches is that, for the most part, the encoder does not need any additional information other than what exists in the measurements. The disadvantage is that most of the exploitable structure for SAR systems is in the formed image, not on the measurements. To improve the rate, these approaches rely on entropy coding, which only looks at the statistics of the bitstream. Depending on the entropy coder choice, this can be a computationally intensive component of the encoder, which can be quite costly when implemented for a satellite system.

## References

[1] J. C. Curlander and R. N. McDonough, "Synthetic aperture radar–systems and signal processing," *New York: John Wiley & Sons, Inc, 1991.*, 1991.

[2] I. G. Cumming and F. H. Wong, "Digital processing of synthetic aperture radar data: Algorithms and implementation [with cdrom](artech house remote sensing library)," *Boston, MA, USA: Artech House*, 2005.

[3] S. Rane, P. Boufounos, A. Vetro, and Y. Okada, "Low complexity efficient raw sar data compression," in *Proc. SPIE Defense and Security, Algorithms for Synthetic Aperture Radar Imagery XVIII*, (Orlando, FL), April 25-29 2011.

[4] V. M. Patel, G. R. Easley, D. M. Healy, and R. Chellappa, "Compressed sensing for synthetic aperture radar imaging," in *2009 16th IEEE international conference on Image Processing (ICIP)*, pp. 2141–2144, IEEE, 2009.

[5] S. I. Kelly, C. Du, G. Rilling, and M. E. Davies, "Advanced image formation and processing of partial synthetic aperture radar data," *IET signal processing*, vol. 6, no. 5, pp. 511–520, 2012.

[6] J. Fang, Z. Xu, B. Zhang, W. Hong, and Y. Wu, "Fast compressed sensing sar imaging based on approximated observation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 1, pp. 352–363, 2014.

[7] D. Wei and P. T. Boufounos, "Saturation-robust sar image formation," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, (Prague, Czech Republic), May 22-27 2011.

[8] D. Liu and P. T. Boufounos, "High resolution sar imagingc using random pulse timing," in *Proc. IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*, (Vancouver, Canada), July 24-29 2011.

[9] D. Liu and P. T. Boufounos, "Random steerable arrays for synthetic aperture imaging," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, (Vancouver, Canada), May 26-31 2013.

[10] D. Liu and P. T. Boufounos, "Synthetic aperture imaging using a randomly steered spotlight," in *Proc. International Geoscience and Remote Sensing Symposium (IGARSS)*, (Melbourne, Australia), July 21-26 2013.

[11] D. Liu and P. T. Boufounos, "High resolution scan mode sar using compressive sensing," in *Proc. Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, (Tsukuba, Japan), September 23-27 2013.

[12] D. Liu and P. T. Boufounos, "Compressive sensing based 3d sar imaging with multi-prf baselines," in *Proc. International Geoscience and Remote Sensing Symposium (IGARSS)*, (Quebec, Canada), July 13-18 2014.

[13] R. Kwok and W. T. Johnson, "Block adaptive quantization of magellan sar data," *IEEE Transactions on Geoscience and remote sensing*, vol. 27, no. 4, pp. 375–383, 1989.

[14] S. Peskova and S. Vnotchenko, "Analysis of complex sar raw data compression," in *SAR workshop: CEOS Committee on Earth Observation Satellites*, vol. 450, p. 619, 2000.

[15] P. Snoeij, E. Attema, A. M. Guarnieri, and F. Rocca, "Fdbaq a novel encoding scheme for sentinel-1," in *2009 IEEE International Geoscience and Remote Sensing Symposium*, vol. 1, pp. I–44, IEEE, 2009.

[16] E. Attema, C. Cafforio, M. Gottwald, P. Guccione, A. M. Guarnieri, F. Rocca, and P. Snoeij, "Flexible dynamic block adaptive quantization for sentinel-1 sar missions," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 4, pp. 766–770, 2010.

[17] P. T. Boufounos, "Universal rate-efficient scalar quantization," *IEEE Trans. Info. Theory*, vol. 58, pp. 1861–1872, March 2012.

[18] D. Valsesia and P. T. Boufounos, "Multispectral image compression using universal vector quantization," in *2016 IEEE Information Theory Workshop (ITW)*, pp. 151–155, September 11-14 2016.

[19] D. Valsesia and P. Boufounos, "Universal encoding of multispectral images," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, (Shanghai, China), pp. 4453–4457, March 20-25 2016.

[20] M. Goukhshtein, P. T. Boufounos, T. Koike-Akino, and S. C. Draper, "Distributed coding of multispectral images," in *IEEE International Symposium on Information Theory (ISIT)*, (Aachen, Germany), pp. 3230–3234, June 25-30 2017.

[21] P. T. Boufounos, "Hierarchical distributed scalar quantization," in *Proc. Int. Conf. Sampling Theory and Applications (SampTA)*, (Singapore), May 2-6 2011.

[22] V. M. Patel, G. R. Easley, D. M. Healy Jr, and R. Chellappa, "Compressed synthetic aperture radar," *IEEE Journal of selected topics in signal processing*, vol. 4, no. 2, pp. 244–254, 2010.

[23] M. Çetin, Ö. Önhon, and S. Samadi, "Handling phase in sparse reconstruction for sar: Imaging, autofocusing, and moving targets," in *EUSAR 2012; 9th European Conference on Synthetic Aperture Radar*, pp. 207–210, VDE, 2012.

[24] G. Rilling, M. Davies, and B. Mulgrew, "Compressed sensing based compression of sar raw data," in *SPARS'09-Signal Processing with Adaptive Sparse Structured Representations*, 2009.

[25] A. Soğanlı and M. Cetin, "Low-rank sparse matrix decomposition for sparsity-driven sar image reconstruction," in *2015 3rd International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, pp. 239–243, IEEE, 2015.