ONBOARD PROCESSING CAPABILITIES OF AN EARTH OBSERVATION COMPRESSIVE SENSING PAYLOAD

Tiziano Bianchi, Martina Cilia, Enrico Magli, Andrea Migliorati, Nicola Prette, Diego Valsesia

Politecnico di Torino - Torino, Italy

ABSTRACT

In this paper, we explore the onboard processing capabilities of an optical Earth observation instrument operating under the principles of compressed sensing, currently under preliminary study. In particular, we focus on two main aspects for onboard operations: i) how to process measurements in a computationally-efficient way to obtain previews of the reconstructed image that can be easily used by downstream inference algorithms; ii) the possibility of having simultaneous compression and encryption by proper management of the pseudorandom patterns used for the sensing matrix and measurements.

Index Terms— Compressed sensing, onboard processing.

1. INTRODUCTION

Since its introduction more than a decade ago, compressive sensing (CS) has established itself as a radically different imaging paradigm that combines compression and sensing. CS relies on the hypothesis that real images have a sparse nature, i.e., they can be compactly represented with few nonzero coefficients in some transform domain, and this allows to sample them at rates lower than what the Nyquist criterion would dictate. The single-pixel camera [1] has demonstrated the idea that imaging hardware exploiting CS principles may require much fewer detectors than conventional designs. This has recently raised interest for the development of a novel generation of payloads for Earth observation missions [2]. Key to the CS theory is the acquisition of measurements of the light field obtained via spatial light modulation (SLM) with pseudorandom masks. Programmable micromirror devices are typically used to implement this behavior by driving each micromirror by means of the corresponding value of the pseudorandom mask.

Measurements are then used to estimate the imaged scene by means of a non-linear reconstruction process. Early reconstruction methods relied on solving optimization problems, typically minimizing the ℓ_1 norm of the transform coefficients of the reconstructed image [3], or the total variation norm [4], or, alternatively, using greedy techniques such as Orthogonal Matching Pursuit [5]. A common issue with those early techniques was the need to define a handcrafted prior about the image to be reconstructed. For instance, total variation minimization implicitly assumes that the image is well-described by a spatially piecewise-smooth signal. Such handcrafted prior have limited the performance of reconstruction algorithms due to their simplistic modelization of the complexity of real images. More recently, deep learning techniques [6] have been used for compressive sensing reconstruction, significantly outperforming optimization-based techniques thanks to the more sophisticated priors that can be learned by neural networks in a data-driven fashion.

Nevertheless, the reconstruction process remains a computationally expensive operation. In the context of an Earth observation mission, the satellite would directly acquire compressive measurements by means of a suitable optical instrument and then transmit them to a ground station for reconstruction. However, there is a growing need in current and future missions for onboard processing capabilities. In fact, the total latency of the acquisition-transmission-reconstruction pipeline can be in the order of hours to days depending on many design factors. This hampers a whole range of time-sensitive applications such as monitoring for environmental disasters, surveillance and many more, which call for the rapid solution of inference problems directly onboard the satellite. This is not trivial to achieve in a design employing compressive sensing because only measurements rather than image data are available to the satellite.

In this paper, we explore the onboard processing capabilities of a compressive Earth observation instrument currently under preliminary study. In particular, we focus on two main aspects: i) how to process measurements in a computationally-efficient way to obtain previews of the reconstructed image to be easily used by downstream inference algorithms for some problems of interest; ii) the possibility of having simultaneous compression and encryption by proper management of the pseudorandom mask patterns and measurements.

This work is funded by the SURPRISE project. The SURPRISE project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 870390.

2. PROPOSED METHOD

In this section, we are presenting the two main operations of interest to be performed onboard, namely fast reconstruction methods, and simultaneous sensing and encryption.

2.1. Fast Onboard Reconstruction

CS reconstruction is typically driven by highly non-linear methods, ranging from optimization-based techniques to deep learning. Such methods are needed in order to exploit the most sophisticated image priors for the regularization of the inverse reconstruction problem. However, their computational complexity is generally high. While some progress has been made in recent years, thanks to the shift from iterative optimization methods to neural networks, which may only require one single forward pass, the amount of floating point operations required is still too large to suit low-complexity onboard implementations. At the same time, providing an estimate of the reconstructed image directly onboard would be beneficial to address a number of inference problems. As a few examples, problems of interest can range from the detection of fires, ships, extreme atmospheric events, etc. In fact, it is desirable to reuse highly optimized and validated algorithms for such problems, but this poses the requirement of producing an image as input to the inference method, rather than compressive measurements.

For these reasons, we study a fast method to generate "previews", i.e. coarse estimates of the reconstructed image, directly onboard. This is done via the use of an optimized linear reconstruction operation. This operator acts directly on the measurements vector to recover an estimate of the image. Instead of relying on the simple pseudoinverse of the sensing matrix, which would minimize the least squares criterion in the measurements domain, we seek to optimize the linear reconstruction operator in a data-driven fashion so that it can learn the typical autocorrelation pattern of the class of images of interest. In particular, the linear reconstruction operator $\mathbf{Q}$ is computed as:

$$\mathbf{Q} = \arg\min_{\mathbf{Q}} \mathbb{E}_{\mathbf{x}\sim\mathcal{D},\mathbf{y}=\Phi\mathbf{x}}[\|U(\mathbf{Q}\mathbf{y}) - \mathbf{x}\|_2^2], \qquad (1)$$

being $\mathbf{x}$ a vectorized image sampled from the training dataset, $\mathcal{D}$ and U a fixed upsampling operation (e.g., bilinear interpolation), while $\mathbf{y}$ and Φ respectively denote the measurements vector and the sensing matrix. The fast preview is then generated from the measurements as

$$\hat{\mathbf{x}} = \mathbf{Q}\mathbf{y}. \qquad (2)$$

It is interesting to notice that we can further reduce the computational complexity of this operation by targeting a lower resolution than the one of the original $\mathbf{x}$ that has generated the measurements, thus reducing the number of rows of matrix $\mathbf{Q}$, at the cost of a degraded quality of image $\hat{\mathbf{x}}$.

2.2. Onboard Encryption

Let us now discuss how the CS acquisition process as described in Equations (1) and (2) can effectively be seen as a symmetric-key encryption system. Specifically, given a vectorized image $\mathbf{x}$ and a sensing matrix Φ such that $\mathbf{y} = \Phi\mathbf{x}$, where $\mathbf{y}$ denotes the measurements vector, it is possible to achieve secrecy by randomizing the matrix Φ, assuming it is kept secret and known only to the hardware used for the CS acquisition. In such a fashion, the image reconstruction can be seen as the process of encrypting a payload $\mathbf{x}$ by means of an encryption key Φ to obtain the encrypted message $\mathbf{y}$. In the proposed framework, we employ binary sensing matrices generated independently and changed for each different CS acquisition. In particular, we employ very efficient,cryptographically-secure pseudo-random generators to generate the binary entries for the sensing matrices. As demonstrated in [7] and further illustrated in Section 4, despite the fact that random binary sensing matrices do not offer perfect secrecy, they nonetheless ensure that, when the size of Φ grows, the advantage of a hypothetical attacker in attacking the proposed system compared to a theoretically secure one is negligible. Hence, practical security is achieved.

3. EXPERIMENTAL RESULTS

3.1. Onboard reconstruction

We investigate the performance of the proposed onboard reconstruction method by comparing the reconstruction error achieved by the fast preview method and that of a full reconstruction method, namely the ISTA-Net neural network. For this experiment, we used a subset of the DFC2020 dataset [8] composed of Sentinel 2 multispectral images. A training partition has been used to optimize the linear reconstruction operator and to train ISTA-Net, while a disjoint test partition is used for numerical tests. The CS acquisition process uses binary random matrices with ± 1 entries with a block size of 32×32 pixels. We study three compression ratios, i.e., the number of measurements acquired for each block, namely 75%,50%,25% (768,512,256 measurements, respectively). Table 1 shows the root mean square error of the test images generated by the fast preview method and the full reconstruction by means of the ISTA-Net neural network [6] for various system design parameters. On the other hand, Table 2 shows the computational complexity of the methods, measured in FLOPs. It can be noticed that the fast preview is an order of magnitudes less expensive than the full reconstruction, and could reasonably be implemented on dedicated hardware onboard. At the same time, the degradation in quality with respect to ISTA-Net is not too severe and enables the solution of inference problems. Fig. 1 shows a visual comparison of the reconstructions.

Table 1. Quality of Fast Preview methods - Root Mean Squared Error

	Compression ratio		
	25%	50%	75%
Full reconstruction	66.85	43.56	26.91
Full-resolution preview	79.01	51.33	30.47
Half-resolution preview	84.30	69.69	64.61
Quarter-resolution preview	104.53	100.96	100.09

Table 2. Complexity of Fast Preview methods - FLOPs

	Compression ratio		
	25%	50%	75%
Full reconstruction	$\sim 4 \times 10^9$	$\sim 4 \times 10^9$	$\sim 4 \times 10^9$
Full-resolution preview	5.2×10^5	1.0×10^6	1.6×10^6
Half-resolution preview	1.3×10^5	2.6×10^5	3.9×10^5
Quarter-resolution preview	3.3×10^4	6.5×10^4	9.8×10^4

4. SECURITY EVALUATION

As explained in Section 2, we employ binary sensing matrices generated independently and changed at each different CS acquisition. A perfectly secure encryption system ensures that $\mathbf{x}$ and $\mathbf{y}$ are statistically independent, i.e. $p(\mathbf{x}|\mathbf{y}) = p(\mathbf{x})$, or, in other words, the mutual information between $\mathbf{x}$ and $\mathbf{y}$ is zero ($I(\mathbf{x}, \mathbf{y}) = 0$). Under the assumption that the sensing matrices are generated independently and changed for each reconstruction, as known from [7], perfect secrecy can be obtained if two conditions are met: ***(i)*** the sensing matrix is generated from Gaussian variables identically distributed and independent of each other; ***(ii)*** the energy of the acquired signal is constant. As said, our design employs random binary matrices. Hence, theoretical security is not guaranteed. However, it is possible to precisely measure the advantage of an attacker in attacking the proposed system compared to a perfectly secure one. In particular, we can show that the advantage is negligible for growing block sizes in the CS acquisition process. For this measurement, we employ the concept of *θ-distinguishability*. Given two signals x_1 and x_2, a ciphered signal y, and a decision function $D(y)$, we measure the capacity of $D(y)$ to correctly understand whether y has been obtained from x_1 or x_2. Hence, given P_c and P_w, respectively the probability of a correct decision and the probability of making a wrong one, x_1 and x_2 are defined as *θ-distinguishable* if, for every $D(y)$:

$$P_c - P_w \leq \theta. \tag{3}$$

The θ parameter measures the advantage for the decision function $D(y)$ in picking the correct signal, i.e. effectively decipher y, with respect to a random guess between the two signals x_1 and x_2. Specifically, the probability that $D(y)$ picks the correct signal is $P_c \leq 1/2 + \theta$. The smaller the θ value is, the smaller the advantage of the attacker becomes. Straightforwardly, $\theta = 0$ in the case of a perfectly secure encryption framework. Fig. 3 collects the *θ-distinguishability* results as a function of the macro-pixel size n. The curves are obtained under the assumption that an independent CS acquisition is done for contiguous macro-pixels. We evaluate two scenarios in which the acquired signals are k-sparse in the DCT-2D domain and in the original pixel domain, respectively. The sparsity is set to $1/8$ of the number of micro-pixels in the block. Considering for example $n = 32$, corresponding to a macro-pixel composed of 1024 micro-pixels, the probability for an attacker to decipher the encrypted signal y is greater than the probability of randomly guessing the correct payload x_1 or x_2 only by a negligible 10^{-5}.

It is also worth noticing that our security analysis refers to the single macro-pixel attack, while in practice the attacker would try to infer the entire original image. For this reason, the complexity of the proposed encryption framework is comparable to the complexity of a brute-force attack to obtain the encryption key. As a matter of example, let us consider a 128×128 micro-pixels image composed of 1024 macro-pixels of size 4×4. In this scenario, the probability for an attacker to infer the original image by observing the encrypted measurements y can be estimated as $(1/2 + 0.01)^{1024} = 3.5634e^{-300}$. To put things in perspective, the probability of brute-force guessing a 256-bit encryption key is equal to $1/2^{256} = 8.6362e^{-78}$. Hence, it would be more advantageous for an attacker to try to infer the encryption key used to generate the random binary sensing matrix, instead of devising an attack based on the specific properties of the binary matrices employed in our secure framework.

5. CONCLUSIONS

This paper presented a study of the onboard processing capabilities of an optical payload working under the principles of compressed sensing. We showed how it is possible to obtain high-quality reconstructions directly onboard with mod-

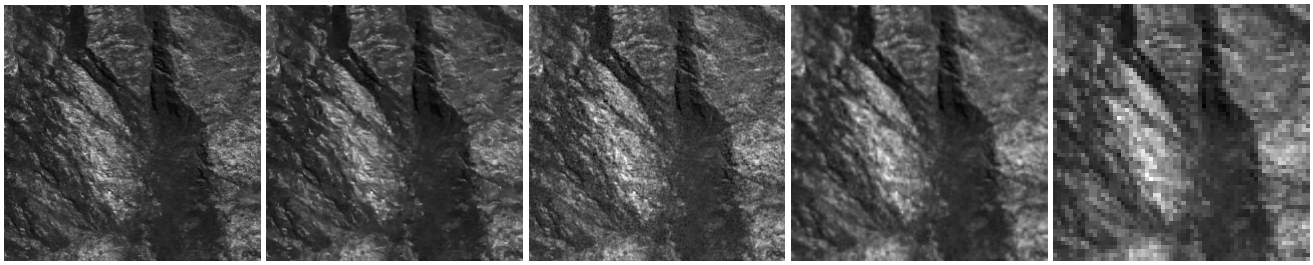

Fig. 1. Reconstructions comparisons. Left to right: ground truth, ISTA-Net, full-resolution preview, half-resolution preview, quarter-resolution preview.

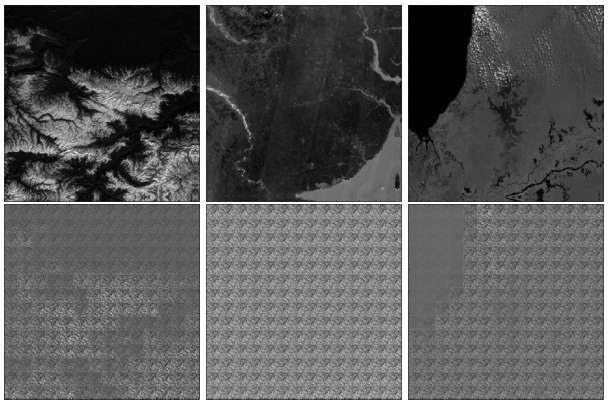

Fig. 2. Examples of CS reconstructions when the sensing matrix Φ, i.e. the encryption key, is not known.

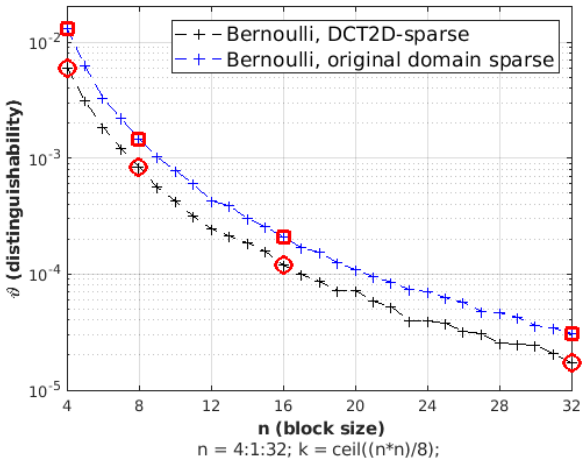

Fig. 3. *θ-distinguishability* as a function of the macro-pixel size n. Even for small values of n, i.e. $n = 4$, θ takes significantly small values ($\theta \simeq 0.01$).

est computational requirements and how the sensing process can be leveraged to achieve simultaneous encryption.

6. REFERENCES

[1] Marco F Duarte, Mark A Davenport, Dharmpal Takhar, Jason N Laska, Ting Sun, Kevin F Kelly, and Richard G Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 83–91, 2008.

[2] Valentina Raimondi, Luigi Acampora, Gabriele Amato, Massimo Baldi, Dirk Berndt, Alberto Bianchi, Tiziano Bianchi, Donato Borrelli, Valentina Colcelli, Chiara Corti, et al., "Spatial light modulator-based architecture to implement a super-resolved compressive instrument for earth observation," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 7864–7867.

[3] Emmanuel J Candès and Michael B Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[4] Chengbo Li, Wotao Yin, Hong Jiang, and Yin Zhang, "An efficient augmented lagrangian method with applications to total variation minimization," *Computational Optimization and Applications*, vol. 56, no. 3, pp. 507–530, 2013.

[5] Joel A Tropp and Anna C Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on information theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[6] Jian Zhang and Bernard Ghanem, "Ista-net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1828–1837.

[7] Matteo Testa, Diego Valsesia, Tiziano Bianchi, and Enrico Magli, *Compressed Sensing for Privacy-Preserving Data Processing*, Springer, 2019.

[8] Michael Schmitt; Lloyd Hughes; Pedram Ghamisi; Naoto Yokoya; Ronny Hänsch, "2020 ieee grss data fusion contest," 2019.