Light Field Compressive Sensing

Mahdad Hosseini Kamal, Mohammad Golbabaee and Pierre Vandergheynst
Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
E-mail:{mahdad.hosseinikamal, mohammad.golbabaei, pierre.vandergheynst}@epfl.ch

***Abstract*—This paper presents a novel approach to capture light field in camera arrays based on the compressive sensing framework. Light fields are captured by a linear array of cameras with overlapping field of view. In this work, we design a redundant dictionary to exploit cross-cameras correlated structures in order to sparsely represent cameras image. We show experimentally that the projection of complex synthetic scenes into the designed dictionary yield very sparse coefficients.**

***Index Terms*—Redundant Dictionary, Compressive Sensing, Light Fields, l_1-minimization.**

I. Introduction

The growing flood of data, particularly due to the emergence of image processing systems with multiple visual sensors, causes an increasing need to quickly process large data sets in compressed domain. Newly introduced light field cameras are one of the widely used class of computational cameras. In order to tackle the large amount of data, we employ the Compressive Sensing (CS) technique [1] that has superiority over traditional schemes because of low complexity in the encoder and universality with respect to scene model.

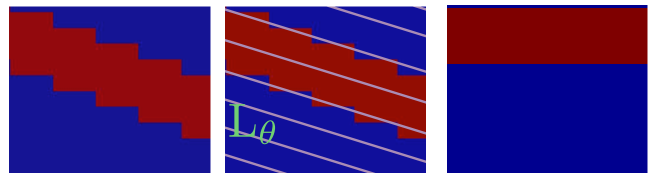
II. Light Field Recovery Scheme

The CS technique holds for signals that are sparse either in their standard coordinate or in any orthogonal basis. Although bases such as Fourier and wavelet can provide a good representation of signals, they are generic and not specific enough to very restrictive class of signals. An alternative signal representation is to consider a redundant dictionary.

The geometric features of a signal are the heart of dictionary design. A piecewise constant function is sparsely represented in the 1D wavelet domain. For epipolar images (EPI) similar to fig. 1(a), it would be best to consider the reordered version of the image grid to have a piecewise-constant-like images [2]. The reordering process is described by selecting a direction η, which is as parallel as possible to the real geometry of the curve. As it is shown on Fig. 1(b), we select grid points L_η and reorder the image grid according to the indices of image samples on these lines Fig. 1(c). Afterwards, we make a piecewise smooth 1D discrete function f from the reconstructed image, which can be sparsely represented using 1D wavelet domain. Selecting a proper direction η for reordering lines is a crucial task. In the case of EPIs with different directions, we do not have a preferential orientation. Therefore, we can benefit from a redundant dictionary, which consists of the concatenation of several reordered wavelet transform $\mathbf{\Phi}^r$ with different directions η. Hence, the designed dictionary $\mathbf{\Psi} = \left[\ \mathbf{\Phi}_1^r, \mathbf{\Phi}_2^r, \cdots, \mathbf{\Phi}_\gamma^r\ \right]$ benefits from different reordering directions to exploit the geometry induced by the natural correlations within light field images. For the camera array, we stack all the cameras image to make an image volume $\mathcal{X} \in \mathbb{R}^{i \times j \times k}$. The image volume would have EPI structure along (i,k)-planes. In addition, a suitable 1D wavelet transform can be applied to sparsify $\mathcal{X}$ along the remaining dimension. To achieve an efficient representation, we reshape X into a matrix $\widehat{\mathbf{X}} \in \mathbb{R}^{ik \times j}$ whose columns contain the information of (i,k)-planes. Following the discussion above, there exists a *sparse* matrix of coefficients $\mathbf{\Theta} \in \mathbb{R}^{\gamma ik \times j}$ such that $\widehat{\mathbf{X}} = \mathbf{\Psi}\mathbf{\Theta}\mathbf{\Gamma}^T$ where $\mathbf{\Psi} \in \mathbb{R}^{ik \times \gamma ik}$ is the previously defined dictionary transform along (i,k)-plane and $\mathbf{\Gamma} \in \mathbb{R}^{j \times j}$ denotes the 1D wavelet basis along j dimension. Thus, if we rewrite $\widehat{\mathbf{X}}$ and $\mathbf{\Theta}$ matrices in vectorial format, we will have $\widehat{X}_{vec} = \mathbf{\Omega}\Theta_{vec}$ with $\mathbf{\Omega} = \mathbf{\Psi} \otimes \mathbf{\Gamma} \in \mathbb{R}^{nk \times \gamma nk}$, where $\otimes$ denotes the Kronecker product between two matrices and $\mathbf{\Omega}$ is the dictionary that is applied to encode the whole image volume into a sparse vector Θ_{vec}. The following convex problem can be applied to reconstruct X from the compressive measurements,

$$\underset{\Theta_{vec} \in \mathbf{R}^{\gamma nk}}{\operatorname{argmin}} \ \|\Theta_{vec}\|_1 \quad \text{subject to} \quad \|Y - \widehat{\mathbf{A}}\mathbf{\Omega}\Theta_{vec}\|_2 \leq \epsilon. \quad (1)$$

Here $\widehat{\mathbf{A}}$ contains the same elements as $\mathbf{A}$ (a block diagonal measurement matrix for the camera array), and is reshaped with respect to $\widehat{X}_{vec} = \mathbf{\Omega}\Theta_{vec}$ so that $\widehat{\mathbf{A}}\widehat{X}_{vec} = \mathbf{A}X$. This optimization can be solved iteratively using Douglas-Rachford splitting method [3]. The experimental results demonstrate the superiority of our methods by about $3\ dB$ in compare with the state-of-the-art 3D wavelet transform.



(a) (b) (c)

Fig. 1: The reordering example of EPI. (a) Original E image. (b) Parallel reordering lines L_η to capture regularity along direction η. (c) Reordered 2D EPI along L_η.

References

[1] D. Donoho, "Compressed Sensing," *IEEE Trans. Inform. Theory*, 2006.
[2] G. Peyré and S. Mallat, "Surface Compression with Geometric Bandelets," *ACM SIGGRAPH*, 2005.
[3] P. L. Combettes and J. C. Pesquet, "Proximal Splitting Methods in Signal Processing," *arXiv/0912.3522*, 2009.