Compressive spectral video by optimal 4D-sphere packing

Nelson Díaz[1], Alejandro Alvarado[1], Pablo Meza[2], and Esteban Vera[1,*]

[1] *School of Electrical Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile*
[2] *Department of Electrical Engineering, Universidad de la Frontera, Temuco, Chile*
[*] *esteban.vera@pucv.cl*

Abstract: This work presents a mathematical framework to design 4D-coded aperture (CA)s for compressive snapshot spectral video (SV) exploiting 4D-sphere packing (SP). Simulation results using state-of-the-art datasets and metrics show a promising performance of the proposed approach to capture high-dimensional datacubes with limited sensing resources.

OCIS codes: 110.0110, 100.3020, 110.4234

1. Introduction

Sensing spectral video (SV) involves capturing a sequence of spectral datacubes over time, where each spectral datacube hyperspectral images (HI) has two spatial dimensions and one spectral dimension. SV may find outstanding applications in environmental sciences or medicine, tracking dynamic processes that have not been properly studied using traditional spectral imaging devices.

Nevertheless, capturing SV is a challenging problem in the 4D-Euclidean space, where a conventional approach to acquire the SV relies on scanning to complete the sequence of datacubes, giving up on temporal resolution. Since scanning is a time-consuming process, compressive sensing (CS) has emerged as a sensing protocol to capture a compressed version of the signal, where the underlying signal can be later recovered by solving an ill-posed inverse problem. Traditionally, the encoding patterns known as CA are random realizations. In contrast, novel approaches have shown outstanding performance using uniform patterns, and those approaches are known as SP-driven methods [1, 2], inspired by early methods based on Fourier analysis [3]. SP is a problem in discrete geometry that consists of packing congruent balls in a container such that only the spheres touch each other in a tangential manner [4]. In detail, optimal SP in 1D is trivial because the points are separated by a corresponding diameter whose resulting density is $\eta_1 \approx 1$, SP in 2D is circular packing with density $\eta_2 \approx 0.90$. 3D-SP or the Kepler conjecture packs 3D spheres, the resulting density is $\eta_3 \approx 0.74$. Recently, 3D-SP has been successfully applied to temporal sensing [1] and MSFA design [2]. Specifically, in SV the container is a hypercube with size $K \times K \times K \times K$ and the number of balls is $L = K^3$ because the number of spheres is proportional to the sensor's size and the number of frames. In this work, we introduce 4D-SP to design CA, and our approach exploits the best-known 4D-sphere packing density $\eta_4 = \frac{\pi^2}{16} = 0.61\ldots$ in the 4D-Euclidean space of the $\mathbf{D}_4 \in \mathbb{Z}^{4\times 4}$ lattice.

2. Methods

The corresponding compressive measurement corresponds to the following discrete model:

$$\mathbf{Y} = \sum_{t=0}^{T-1} \sum_{k=0}^{K-1} \mathcal{X}_{(:,:,k,t)} \odot \mathcal{C}_{(:,:,k,t)} + \mathbf{\Omega}, \tag{1}$$

where K is the number of bands and T is the number of frames, $\mathcal{X}_{(:,:,k,t)} \in \mathbb{R}^{M\times N}$ denotes the spectral temporal scene at the k^{th} band, and the t^{th} frame, the coded aperture is given by $\mathcal{C}_{(:,:,k,t)} \in \mathbb{R}^{M\times N}$, let $\odot$ be the Hadamard product, and $\mathbf{\Omega} \in \mathbb{R}^{M\times N}$ is the additive noise. In 4-dimensional Euclidean space the best-known packing density corresponds to the following lattice $\mathbf{D}_4 = [-1,-1,0,0;1,-1,0,0;0,1,-1,0;0,0,1,-1]$, and the corresponding Gram matrix is $\mathbf{A}_4 = \mathbf{D}_4^T\mathbf{D}_4$. The density of a 4D-lattice is given by

$$\frac{\text{Vol}(S_r^4)}{\text{Vol}(\mathbb{R}^4/\mathbf{D}_4)} = \frac{\frac{\pi^2}{2} r^4}{\sqrt{\det(\mathbf{A}_4)}} = 0.61\ldots \tag{2}$$

where $\text{Vol}(S_r^4) = \frac{\pi^2}{2} r^4$ is the volume of the 4−dimensional sphere and $\text{Vol}(\mathbb{R}^4/\mathbf{D}_4) = \sqrt{\det(\mathbf{A}_4)}$ is the volume of the lattice, the radius is $r = 1/\sqrt{2}$.

2.1. Coding optimization strategy

The sampling of SV can leverage the following solution to 3DN^2QP [5] to place the spheres within a 3D-container $\mathcal{B}$ as follows:

$$\mathcal{B}_{(:,:,t)} = ((\alpha \odot \mathbf{V} + \beta \odot \mathbf{H} + \gamma \odot q_t) \mod K + \mathbf{1}), \tag{3}$$

where K is the number of bands, the $\mathbf{V}$ and $\mathbf{H}$ are the vertical and horizontal translation matrices, respectively. The vertical translation matrix is denoted by $\mathbf{V} = \mathbf{f}^T \otimes \mathbf{q}$, where $\mathbf{V} \in \mathbb{N}^{K\times K}$, $\mathbf{f}$ is a vector of all ones given by $\mathbf{f} \in \{1\}^K$, and $\mathbf{q} = [1, \ldots, K]^T$ where $\mathbf{q} \in \mathbb{N}^K$, $\otimes$ is the Kronecker product, and $\odot$ is the Hadamard product. The horizontal translation matrix is $\mathbf{H} = \mathbf{V}^T$, and matrix $\mathbf{1} \in \mathbb{N}^{K\times K}$. The parameters, α, β, and $\gamma \in \mathbb{N}$ and must be selected such that the minimum distance between spheres is maximized. The positions of the Multispectral Filter Array (MSFA) at the t^{th} frame is given by $\mathcal{G}_{(:,:,t)} = \mathbf{A} \otimes \mathcal{B}_{(:,:,t)}$, where $\mathbf{A}$ is a matrix of all ones such that $\mathbf{A} \in \{1\}^{\delta\times\epsilon}$, where $\delta = \lfloor \frac{M}{K} \rfloor$, and $\epsilon = \lfloor \frac{N}{K} \rfloor$, the successive t^{th} frame is computed using the tensor $\mathcal{B}_{(:,:,t)}$. The multispectral pattern $\mathcal{G}$ can be reorganized as CA

$$C_{(i,j,k,t)} = \begin{cases} 1 & \text{if } k = \mathcal{G}_{(i,j,t)} \\ 0 & \text{if } k \neq \mathcal{G}_{(i,j,t)}, \end{cases} \tag{4}$$

the resulting tensor $\mathcal{G} \in \mathbb{R}^{M\times N\times T}$ can be reorganized as $\mathbf{p}_l = [i, j, \mathcal{G}_{(i,j,t)}, t]$, where $\mathbf{P} = [\mathbf{p}_1, \ldots \mathbf{p}_l \ldots \mathbf{p}_L] \in \mathbb{R}^{4\times L}$, with indexes $i, j \in \{1, \ldots, K\}$ and $k \in \{1, \ldots, K\}$, where $L = K^3$ is the number of spheres. Thus, the distance function of L spheres is $d^*(L) = \max(\min_{1\leq l_1 < l_2 \leq L}, D_{l_1,l_2})$, where $D_{l_1,l_2} = \|\mathbf{p}_{l_1} - \mathbf{p}_{l_2}\|_2^2$ is the all pairwise distance matrix, $l_1, l_2 \in \{0, \ldots, L-1\}$ index the $l_1{}^{\text{th}}$ and $l_2{}^{\text{th}}$ spheres.

3. Results

To prove the performance of our approach, we use a state-of-the-art SV dataset [6]. The spatial resolution is 256×256, the spectral resolution is 8 bands and the number of frames is 8. Moreover, state-of-the-art metrics are used. To evaluate the spatial fidelity, the peak-signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were used; to test the spectral accuracy, the spectral angle mapper (SAM) was used. The reconstruction algorithm is interpolation. Fig. 1 depicts the groundtruth and reconstruction of frames 1st, 3rd, 5th, and 7th, whose corresponding PSNR is 25.5807 [dB], SSIM is 0.69274, and SAM is 0.40898.

Fig. 1: Compressive measurement and comparison of image reconstruction quality. (a) Compressive measurement, (b) the first row shows 4 groundtruth frames. The second row depicts the recovered image using interpolation.

4. Conclusions

This paper introduces 4D-CA to reconstruct the SV. Our approach leverages the best-known 4D-packing density of 4D congruent SP to design the spectral filter for different instants of time. Future work will explore reconstruction algorithms driven by deep learning.

Funding

This work has been supported by ANID ANILLOS ATE220022, ANID FONDECYT 1201081, ANID FONDECYT 1221883, and ANID FONDECYT Postdoctorado 3230489.

References

1. E. Vera, F. Guzmán, and N. Díaz, "Shuffled rolling shutter for snapshot temporal imaging," Opt. Express **30**, 887–901 (2022).
2. N. Diaz, A. Alvarado, P. Meza, F. Guzman, and E. Vera, "Multispectral Filter Array Design by Optimal Sphere Packing," IEEE Transactions on Image Process. (2023).
3. K. Hirakawa and P. J. Wolfe, "Spatio-spectral color filter array design for optimal image recovery," IEEE Transactions on Image Process. **17**, 1876–1890 (2008).
4. J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, vol. 290 (Springer Science & Business Media, 2013).
5. L. Allison, C. Yee, and M. McGaughey, *Three-dimensional Queens Problems* (Monash University, Department of Computer Science, 1989).
6. F. Xiong, J. Zhou, and Y. Qian, "Material based object tracking in hyperspectral videos," IEEE Transactions on Image Process. **29**, 3719–3733 (2020).