León-López, Kareth Marcela; Galvis-Carreño, Laura Viviana; Arguello-Fuentes, Henry
Reconstruction of multispectral light field (5d plenoptic function) based on compressive
sensing with colored coded apertures from 2D projections
Revista Facultad de Ingeniería Universidad de Antioquia, núm. 80, septiembre, 2016, pp.
131-141
Universidad de Antioquia
Medellín, Colombia

Reconstruction of multispectral light field (5d plenoptic function) based on compressive sensing with colored coded apertures from 2D projections

Reconstrucción de campos de luz multiespectrales basada en la técnica de muestreo compresivo con aperturas codificadas de color a partir de proyecciones 2D

Kareth Marcela León-López[1*], Laura Viviana Galvis-Carreño[2], Henry Arguello-Fuentes[1]

[1]Escuela de Ingeniería de Sistemas e Informática, Facultad de Ingenierías Fisicomecánicas, Universidad Industrial de Santander. Carrera 27 Calle 9. C. P. 680002. Bucaramanga, Colombia.
[2]Department of Electrical and Computer Engineering, University of Delaware. 19716. Newark, DE, USA.

ABSTRACT: In the last decade, spatio – angular (light field) acquisition systems have advanced due to the inclusion of coded apertures in the optical path. These coded apertures, modulate the light, encoding the information before being captured. Traditionally, these coded apertures are binary, i.e. block and unblock the light rays in the spatial dimension, capturing sparse information of the scene. In this work, the binary coded aperture is replaced by a colored coded aperture which modulates the source not only spatially but spectrally. Thereby, it is possible to capture light fields in multiple wavelengths yielding high spectral resolution. The spectral information provides many features of a scene in different wavelengths, these features are not present in the visible range of the electromagnetic spectrum. In this paper, an algorithm that simulates the light field sampling with colored coded apertures is proposed. The proposed algorithm, exploits the redundant information of the scene based on the compressive sensing theory thus, capturing just a sparse signal. The multidimensional image can be recovered from the underlying signal through a reconstruction algorithm. Several simulations show the quality of the multispectral light field reconstructions. The PSNR (Peak Signal to Noise Ratio) values obtained for the reconstructions are around 25 dB.

RESUMEN: Los sistemas de adquisición de información espacio – angular en 4D o campos de luz, han avanzado en la última década debido a la inclusión de aperturas codificadas en el camino óptico del sistema. Dichas aperturas, modulan la luz codificando la información antes de ser capturada. Tradicionalmente, estas aperturas codificadas son aperturas binarias, es decir, bloquean o dejan pasar los rayos de luz en la dimensión espacial haciendo que se capture información dispersa de la escena. En este trabajo, se reemplazan las aperturas binarias por aperturas codificadas de color, que además de modular la luz espacialmente lo hacen espectralmente. De esta forma, es posible adquirir campos de luz en múltiples longitudes de onda y aumentar su resolución espectral. La importancia de obtener información en el espectro radica en que es posible capturar características que no están presentes en el rango visible del espectro electromagnético. En este artículo, se propone un algoritmo que simula la adquisición de una imagen de campo de luz usando aperturas codificadas de color. El sistema propuesto aprovecha la redundancia de información que se presenta en la escena basándose en la teoría del muestreo compresivo, así, solo se captura una señal dispersa de la escena. La imagen multidimensional es recuperada a partir de dicha señal mediante un algoritmo de reconstrucción. Los resultados de simulación muestran la calidad de reconstrucción del campo de luz multiespectral. La calidad de las reconstrucciones obtenidas es de alrededor de 25 dB en PSNR.

* Corresponding author: Kareth Marcela León-López
e-mail: kareth.leon@correo.uis.edu.co

1. Introduction

The recent theory of compressive sensing (CS) provides a fundamentally new data acquisition approach. CS theory establishes that a signal, e.g. a sound record or an astronomical image, can be sampled at a much smaller rate

than what is commonly prescribed by Shannon – Nyquist [1-3]. This theory has been widely used in different areas as spectral imaging, light field imaging, video acquisition, tomography, among others [4-7] , assuming that the signal under analysis can be sparse in some representation basis. Formally, suppose that a given signal $\mathbf{f} \in \mathbb{R}^N$, is *S*-sparse on some basis $\Psi \in \mathbb{R}^{N \times N}$, such that $\mathbf{f} = \Psi\alpha$ can be approximated by a linear combination of *S* vectors from Ψ, with *S<<N*. The acquired measurements can be expressed as $\mathbf{g} = \mathbf{Hf} = \mathbf{H}\Psi\alpha$, where $\mathbf{H}\Psi$ is known as the sensing matrix. Specifically, in compressive light field, a four dimensional image $\mathbf{F} \in \mathbb{R}^{p_x \times p_y \times p_u \times p_v}$, or its vector form $\mathbf{f} \in \mathbb{R}^{p_x \cdot p_y \cdot p_u \cdot p_v}$, is captured in the detector just taking 2 – dimensional (2D) coded projections of the underlying scene [1]. $\mathbf{H}$ is the measurement matrix, whose structure is determined by the coded aperture entries of the acquisition system and α is a sparse vector that represents the signal $\mathbf{f}$ in the domain Ψ. The inverse problem requires to solve the minimization given by $\hat{\mathbf{f}} = \Psi(\text{argmin}_\alpha \|\mathbf{g} - \mathbf{H}\Psi\alpha\|_2 + \tau\|\alpha\|_1)$, being $\mathbf{f}$ the reconstructed light field from $\mathbf{g}$, τ is a regularization parameter and $\|.\|_2$ and $\|.\|_1$ are the ℓ_2 and ℓ_1 norm operators, respectively.

A light field is a specific parameterization of the plenoptic function. The plenoptic function was introduced as a complete holographic representation of the visual world. This function describes the observable light intensity at every view position and every instant of time (*x, y, z, t*). Adding spectral information intensity in any possible angle $(\theta, \varphi, \lambda)$ results in a 7D function as $P = P(x, y, z, \theta, \varphi, \lambda, t)$ [8], where P is the plenoptic function, *x, y* and *z* are the spatial dimensions, θ and φ represent the angular information, λ represents the wavelength and *t* denotes the time [8]. Plenoptic imaging considers a subset of computational photography approaches; specifically, those that aim at acquiring the dimensions of the plenoptic function with combined optical light modulation and computational reconstruction as CS theory [9].

This work studies the plenoptic function in five dimensions (5D), assuming that the air is transparent in a ray space representation and the spectral information is included. The plenoptic function is reduced from 7D to 5D, ignoring the time *t* and the depth z. The specific parameterization of the full plenoptic function can be seen as a light field in multiple wavelengths. This is, every angular 2D image has a manifold of spectrum bands, and this is called *Multispectral Light Field*. The knowledge of the spectrum at various points in the image can be useful in identifying the composition and structure of objects in the scene being observed by the imager. Moreover, spectral images are useful in many areas such as remote sensing, astrophysics, biochemistry, and many others [8]. Five dimensional image acquisition and reconstruction from sparse measurements is possible due to the CS domain [5, 6, 10]. The 5D function that describes a multispectral light field can be written as Eq. (1)

$$P_{5D} = P\left(x, y, \theta, \varphi, \lambda\right).$$

(1)

In general, this paper describes the multispectral light field acquisition using colored coded apertures. More specifically, an algorithm to simulate the multispectral light field

sampling and the robust reconstruction from 2D projections was developed. Sampling is performed by means of a colored coded aperture that modulates not only the spatio – angular dimensions but the spectral information as well. The recovery of the underlying multidimensional signal is performed using the LASSO reconstruction algorithm.

The main contribution of this paper is the development of the mathematical model for multispectral light field acquisition. This paper is organized as follows: the proposed multispectral light field sampling model is developed in section 2, as well as, the multidimensional data cube reconstruction problem. Simulations and results with the proposed acquisition model are detailed in section 3 and finally, some conclusions are presented.
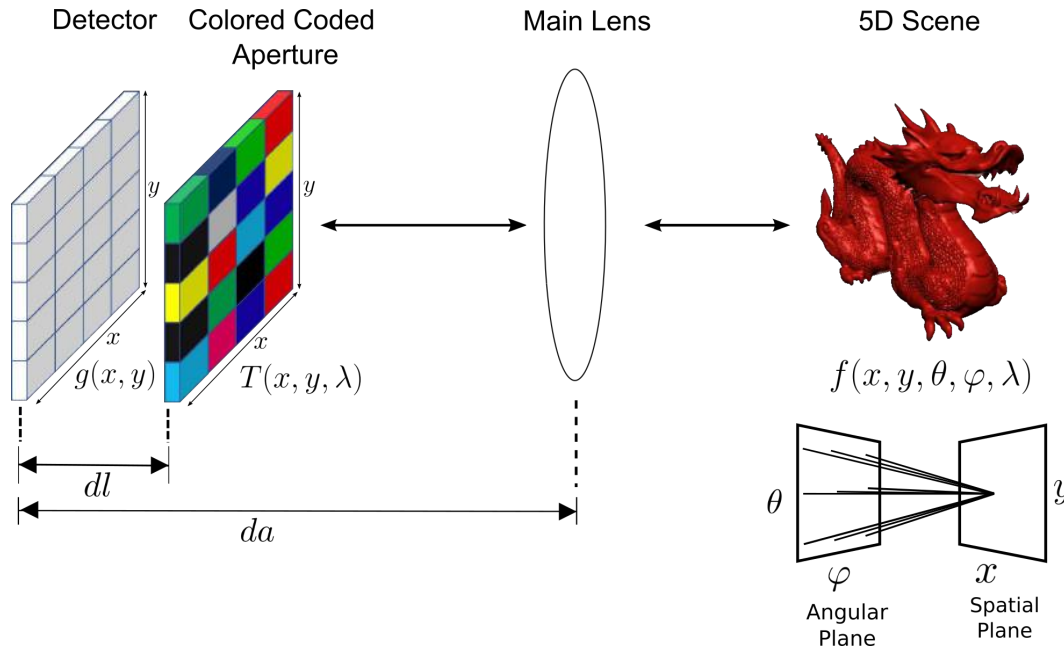
2. Compressive multispectral light field model

This section describes the mathematical model for multispectral light field coded projections acquisition, the development of the spectral overcomplete dictionary as a representation basis, the reconstruction method and a brief description of the colored coded apertures used in this work.

### 2.1. Multispectral light field acquisition model

The CS – based acquisition system is composed by three optical elements: the focal plane array (FPA) or detector, the main lens, and the most important element: the coded aperture. Traditionally, in these compressive light field systems, the light rays are modulated by a binary coded aperture which modulates the spatial and angular features of the scene, blocking and unblocking the light from the scene toward the sensor [5, 11]. In the proposed acquisition system, the binary coded aperture is replaced by a colored coded aperture which modulates both the spatio – angular dimension and the spectral dimension. Figure 1 illustrates a sketch of the proposed acquisition system.

The set of light rays in a free space is denoted as $f(x, y, \theta, \varphi, \lambda)$, where *x* and *y* indicate the spatial coordinates, θ and φ the angular dimensions and λ the spectral dimension (wavelength). The function f is the 5D plenoptic function described in Eq. (1). This set of rays is coded by the colored coded aperture defined by the transmittance as $T(x, y, \lambda)$. Given that the operations along *x* and θ coordinates are equivalent to those along *y* and φ, a simplification in the spatial and angular dimensions, $f(x, \theta, \lambda)$, can be considered and, the continuous measurement *g(x)* on the detector can be expressed by Eq. (2)

$$g(x) = \iint T(x, \lambda) \cdot f(x, \theta, \lambda) d\theta d\lambda.$$

(2)

Figure 1 Compressive multispectral light field acquisition system. A colored coded aperture is placed at a distance dl from the image sensor. The light rays are modulated by the colored coded aperture before they are captured in the detector. The spatio- angular frame reference is illustrated in the bottom-right; each spatial point of the scene can be seen from different angles assuming a two-plane parameterization
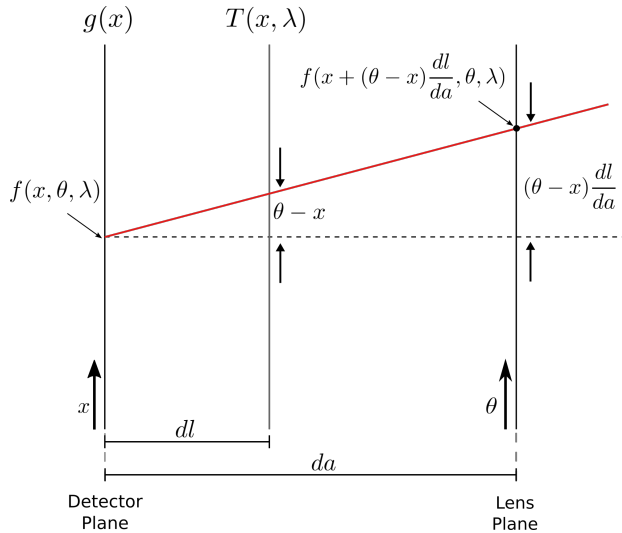
Figure 2 Illustration of the multispectral light field modulation through the colored coded aperture $T(x, \lambda)$. The colored coded aperture modulates the multispectral light field before it is captured in the detector. The parameterization of the light rays is showed using just one light ray for explanation purposes (red line)

In the optical path, the light ray set is once again parameterized by the colored coded aperture $T(x, \lambda)$, producing a defocusing in the trajectory of the light rays. The diagram showed in Figure 2 illustrates this parameterization for a single light ray.

By similar triangles, the light ray that intersects the *x* plane (detector plane) and the *T* plane, also intersects the θ plane (lens plane) at $(\theta - x)\frac{dl}{da}$. Given that the parameterization is conducted by the coded aperture plane, Eq. (2) can be rewritten as in Eq. (3)

$$g(x) = T \iint \left(x + (\theta - x)\frac{dl}{d\alpha}, \lambda \right) \cdot f(x, \theta, \lambda) d\theta d\lambda.$$

(3)

This parameterization shows that multispectral light field imaging can be thought of as shearing the 5D multispectral light field, and then projecting it down to 2D [12].

The discretization of the colored coded aperture is defined through its transmittance function as Eq. (4),

$$T(x, \lambda) = \sum_{i,k} T_{i,k} \cdot rect\left(\frac{x}{\Delta_T} - i, \frac{\lambda}{\Delta_d} - k\right),$$

(4)

where $T_{i,k}$ is a binary value accounting for the block and unblock operation at the (i,k) $(i,k)-th$ voxel of the 5D coding, Δ_T and Δ_d denote the pixel size of the colored coded aperture and the sensor, respectively.

Using Eq. (4), the discrete detector measurement can be expressed as Eq. (5)

$$G_i = \sum_{u,k} T_{i,k} \mathcal{F}_{i,u,k} + \omega_i,$$

(5)
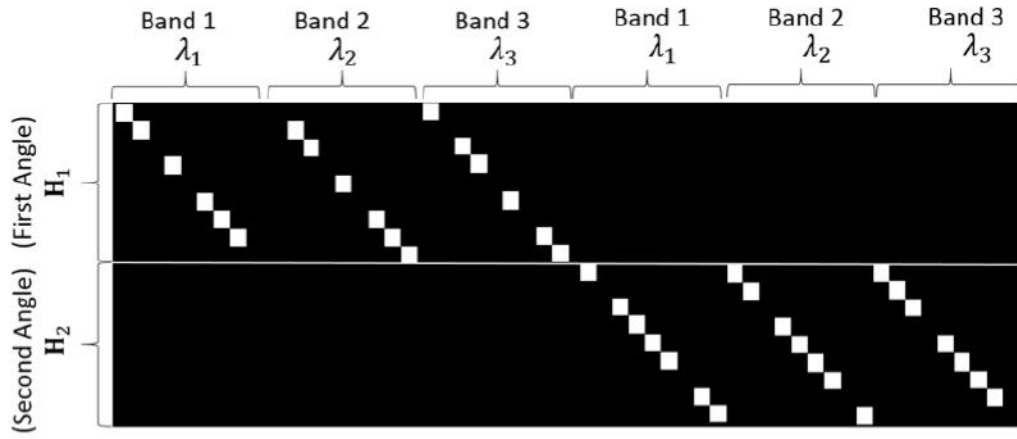
Figure 3 Illustration of the **H** matrix structure for L=3, $p_u = 1, p_v = 2, p_x = 3$ and $p_y = 3$. The black pixels on the diagonal block the light ($T_{i,u,k} = 0$) and white pixels let the light pass through ($T_{i,u,k} = 1$)

where G_i is the intensity at the (i)–*th* position of the detector (considering just one dimension of the 2D detector), $\mathcal{F}$ is a spatio – spectro – angular data cube and, ω_i is the noise of the capturing process. The multispectral light field coded projection at the detector G_i can be also written in matrix form as Eq. (6)

$$\mathbf{g} = \mathbf{Hf}, \quad \mathbf{H} = \left[\mathbf{H_1^1\ H_1^2\ H_1^3 \dots H_1^L \dots H_j^1 \dots H_{p_up_v}^L} \right], \quad (6)$$

where $\mathbf{f} \in \mathrm{R}^N \left(N = p_x \cdot p_y \cdot p_u \cdot p_v \cdot L \right)$ is the vectorized multispectral light field $\mathbf{g} \in \mathrm{R}^M (M = p_x \cdot p_v)$, is the corresponding measurement, and $\mathbf{H} \in \mathrm{R}^{M \times N}$ is the sensing matrix for the whole multispectral light field. Each submatrix $\mathbf{H_j^k} \in \mathrm{R}^{M \times M} (j = 1, \dots, p_u p_v, k = 1, \dots, L)$ in Eq. (6) is a sparse matrix that contains the structure of the colored coded aperture $T(x, \lambda)$ on its diagonal, for each angular view *j* and for each spectral band *k*. Figure 3 shows an example of the **H** matrix structure for $p_x = 3$ and $p_y = 3$ pixels in the spatial dimension, $p_u = 1$ and $p_v = 2$ angular views and $L = 3$ spectral bands. The top half of the **H** matrix in Figure 3 corresponds to the first angular view and the bottom half corresponds to the second angular view.

The **H** matrix, mathematically, represents the spatio – spectro – angular modulation element in the acquisition system.

The sampling and reconstruction processes for the multispectral light field are developed using small patches of the image instead of the full image. This approach allows to simultaneously solve many small and independent problems, instead of attempting to solve a single and large optimization problem [5]. The notion of patches will be exposed in following sections. Algorithm 1 in Figure 4, developed in this work, shows the sampling process of a multispectral light field.

The input parameters of the sampling algorithm 1 are: the **H** matrix, the multispectral light field **f**, and the patch sizes. For each iteration, a set of patches is sampled using the Eq. (6) and each measurement is stacked in the measurement vector **g**.

Algorithm 1: Sampling Multispectral Light Field

Input :

- $\mathbf{H} \in \mathbb{R}^{M \times N}$ (Measurement Matrix)
- $\mathbf{f} \in \mathbb{R}^N$ (Multispectral Light Field)
- $s \leftarrow [p_x, p_y, p_\theta, p_\varphi, L]$ (Patch size)
- $w_1 \leftarrow x/p_x, w_2 \leftarrow y/p_y$ (Number of patches in each 2D spatial image (x, y)

Output: Multispectral light field measurement vector **g**

1: $n \leftarrow 1$
2: $k \leftarrow 1$
3: **for** $i = 1\,to\,w_1$ **do**
4:    **for** $j = 1\,to\,w_2$ **do**
5:       $\mathbf{g}_{n:k \cdot p_x \cdot p_y} \leftarrow \mathbf{H}_k \cdot \mathbf{f}_{i,j}(:)$
6:       $n \leftarrow k \cdot p_x \cdot p_y + 1$
7:       $k \leftarrow k + 1$
8:    **end for**
9: **end for**

Figure 4 Algorithm 1. Process to sample a multispectral light field

2.2. Reconstruction of Multispectral Light Field Projections

The inverse problem to recover the multispectral light field f from the measurements g requires to invert the linear system presented in Eq. (6). Since the number of measurements g is significantly smaller than the number of columns of H, i.e. $M \ll N$, it is necessary to use sparse coding techniques to solve the underdetermined problem.

The main idea of sparse coding techniques is that underdetermined problems can be solved if the unknowns are sparse or can be represented in a sparse basis or overcomplete dictionary. Assuming that the signal f is S-sparse on some overcomplete dictionary $D \in R^{N \times d}$, f can be approximated by a linear combination of S columns of the dictionary $\mathbf{D}$ as $\mathbf{f} = \mathbf{D}\alpha$, where most of the coefficients in $\alpha \in \mathbb{R}^d$ $\left(S \ll d \right)$ have values close to zero. Thus, the coded projections in Eq. (6) can be rewritten as $\mathbf{g} = \mathbf{HD}\alpha$. Then, the reconstruction consists on recovering α such that the $\ell_2 - \ell_1$ cost function is minimized as Eq. (7)

$$\hat{\mathbf{f}} = \mathbf{D}(\text{argmin}_{\alpha}\|\mathbf{g} - \mathbf{HD}\alpha\|_2 + \tau\|\alpha\|_1), \quad (7)$$

Being τ a regularization parameter and $\hat{\mathbf{f}}$ the reconstructed multispectral light field from $\mathbf{g}$.

Algorithm 2 (Figure 5) shows the reconstruction process where M – long vectors $\mathbf{Y}$, from the measurements vector g are proccesed. For each subproblem, the optimization in Eq. (7) is solved. The main algorithm inputs are the measurements vector $\mathbf{g}$, the measurement matrix $\mathbf{H}$ and the overcomplete dictionary $\mathbf{D}$. The output of the reconstruction algorithm is the reconstructed multispectral light field $\mathbf{f}$. Notice that the reconstructed multispectral patches are stacked in the vector $\mathbf{f}$, and it is necessary to rearrange the vector to obtain the 5D multispectral image.

The quantification of the reduction in the data representation size is the compression ratio and is given by Eq. (8)

$$r = \frac{M}{N} = \frac{1}{p_u p_v p_L}, \quad (8)$$

where M is the *compressed size*, in this case, it is the size of the spatial dimension (detector size) and N, the *uncompressed size*, is the size of the full data (multispectral light field size). The compression ratio r can be used as a comparison measure.

2.3. Learning an Overcomplete Dictionary

Previous works in compressive light field have demonstrated high compression and a good data representation using an overcomplete dictionary as a representation basis. An overcomplete dictionary is composed of a set of fundamental building blocks of natural light fields or light field atoms [5].

A light field atom is a 4D spatio – angular patch, or a set of pixels extracted from a training light field.

Algorithm 2: Reconstruction of a Multispectral Light Field from the measurements vector **g**

Input :

- $\mathbf{g} \in \mathbb{R}^M$ (Measurements Vector)
- $\mathbf{H} \in \mathbb{R}^{M \times N}$ (Measurement Matrix)
- $\mathbf{D} \in \mathbb{R}^{N \times d}$ (Overcomplete Dictionary)
- $s \leftarrow [p_x, p_y, p_\theta, p_\varphi, L]$ (Patch size)
- $w_1 \leftarrow x/p_x, w_2 \leftarrow y/p_y$ (Number of patches in each 2D spatial image (x, y))

Output: Reconstructed Multispectral light field $\hat{\mathbf{f}}$

1: $r \leftarrow p_x \cdot p_y \cdot p_\theta \cdot p_\varphi \cdot L$
2: $m \leftarrow 1$
3: $q \leftarrow 1$
4: **for** $i = 1 \, to \, w_1$ **do**
5:     **for** $j = 1 \, to \, w_2$ **do**
6:         $\mathbf{Y} \leftarrow \mathbf{g}_{m:q \cdot p_x \cdot p_y}$
7:         $\alpha = \min_{\alpha}\|\mathbf{Y} - \mathbf{H}_q \mathbf{D}\alpha\| + \lambda\|\alpha\|$
8:         $\hat{\mathbf{f}}_{r \cdot (q-1)+1:r \cdot q} = \mathbf{D}\alpha$
9:         $m \leftarrow q \cdot p_x \cdot p_y + 1$
10:         $q \leftarrow q + 1$
11:     **end for**
12: **end for**

Figure 5 Algorithm 2. Reconstruction process of a multispectral light field

In order to obtain a good representation of the 5D multispectral light field in the dictionary domain, the 4D light field atoms presented in [5] are modified adding the spectral information. Let $\{c_i\}_{i=1}^{q}$ be a set of q training signals or 5D patches, where the size of each 5D patch is $N = p_x \times p_y \times p_u \times p_v \times L$. Adding these signals as the columns of a matrix $\mathbf{C}$, the general dictionary learning problem is to find a sparsifying dictionary $\mathbf{D}$ by solving Eq. (9)

$$\underset{\{D, A\}}{\text{minimize}}\|\mathbf{C} - \mathbf{DA}\|_F$$

$$\text{subject to } \forall_i \, \|\alpha_i\|_0 \leq S, \quad (9)$$

where $C \in R^{N \times q}$ is a training set that contains q signals, $D \in R^{N \times d}$ is the overcomplete dictionary, $A = [\alpha_1, ..., \alpha_q] \in R^{d \times q}$ is a set of S-sparse coefficient vectors, and S is the sparsity level or the number of nonzero entries. The non – convex ℓ_0 pseudo – norm, $\|.\|_0$, counts the amount of nonzero elements in a vector. In general, the minimization in Eq. (9) looks for the best possible dictionary for the sparse representation of the set $\mathbf{C}$ [5, 13]. Algorithm 3 shows the process to train a dictionary (Figure 6).

The optimization problem of learning a dictionary can be seen as a search for sparse representations with coefficients summarized in the matrix $\mathbf{C}$. The dictionary is iteratively updated together with the sparse coefficients. Assuming that both $\mathbf{C}$ and $\mathbf{D}$ are fixed, each dictionary column $\mathbf{d}_k$ and the corresponding k_{th} row in $\mathbf{C}$ are denoted as α_R^k. For each column of $\mathbf{C}$, a group of examples ω_k are defined using the atom $\mathbf{d}_k$. Then, the overall representation error matrix is calculated in $\mathbf{E}_k$. The value E_k^R is built choosing only the columns corresponding to ω_k in $\mathbf{E}_k$, implying a selection of error columns that correspond to examples that use the atom $\mathbf{d}_k$. Applying SVD (Singular Value Decomposition) in $\mathbf{E}_k^R$ such as $\mathbf{E}_k^R = \mathbf{U}\Delta\mathbf{V}^T$, the dictionary column $\mathbf{d}_k$ is updated choosing the first column of the matrix U, and the coefficient vector α_R^k is updated choosing the first column of $\mathbf{V}$ multiplied by $\Delta(1,1)$. The best dictionary is obtained to represent the 5D patch signal as sparse composition [13].

Algorithm 3: Learning an Overcomplete Dictionary with K - SVD

Input :

- $\mathbf{D}^0 \in \mathbb{R}^{N \times d}$ (Initialization of the dictionary matrix)
- $\{c_i\}_{i=1}^q$ (Set of q training signals)
- S (Sparsity level)

Output: The best dictionary to represent the data samples $\{c_i\}_{i=1}^q$ as sparse composition

1: **for** $j = 1\,to\,stopping\,rule$ **do**
2:     Using any pursuit algorithm to compute:
$$\tilde{\alpha}_i = \operatorname*{argmin}_{\alpha_i}\{||c_i - \mathbf{D}\alpha_i||_2^2$$
$$s.t.\ ||\alpha_i||_0 \leq S, i = 1, 2, ..., q$$
3:     **for** $k = 1\,to\,d$ in $\mathbf{D}^{J-1}$ **do**
4:         $w_k = \{i | 1 \leq i \leq q, \alpha_T^k(i) \neq 0\}$
5:         $\mathbf{E}_k = \mathbf{C} - \sum_{j \neq k} \mathbf{d}_j \alpha_T^j$ (Error matrix)
6:         $\mathbf{E}_k^R \leftarrow$ Restriction of $\mathbf{E}_k$ choosing the columns corresponding to w_k
7:         SVD in $\mathbf{E}_k^R$ as $\mathbf{E}_k^R = \mathbf{U}\triangle\mathbf{V}^T$
8:         Choose the updated column $\tilde{\mathbf{d}}_k$ as the first column of $\mathbf{U}$
9:         Update α_R^k as the first column of $\mathbf{V}\triangle(1,1)$
10:     **end for**
11: **end for**

Figure 6 Algorithm 3. The dictionary learning algorithm

2.4. Colored Coded Apertures

The colored coded apertures are a recent advance that allows to modulate the wavelengths of interest. These apertures have been used in applications such as artificial compound eyes [14], multi-focusing and depth estimation [15], deblurring and matting [16], and compressive spectral imaging [17].

A colored coded aperture can be seen as an array of binary coded apertures where each voxel blocks or unblocks a specific wavelength set of the spectral data cube pixel (Figure 7).

The colored coded aperture pixels can modulate the spectral data as low – pass ($\mathcal{L}$) or high – pass ($\mathcal{H}$) optical filters. Each colored coded aperture pixel is one of many possible optical filters whose spectral response can be selected. For instance, a low pass filter $\mathcal{L}$ (Figure 7 bottom) permits to pass the wavelengths λ_0, λ_1 and λ_2 in the spectral data cube pixel. A high pass filter $\mathcal{H}$ permits to pass the wavelengths λ_2, λ_3 and λ_4. The use of this optical element has demonstrated a high probability of correct reconstruction (high PSNR) in previous researches [17, 18].

3. Results

In order to test the proposed model, a set of compressive measurements is simulated using the forward model in Eq. (6) and the algorithm 1. These measurements are constructed using a synthetic multispectral light field data set. The multispectral light field was built employing a light field data set published by the MIT Media Lab [5]. The original light field has 593×840 pixels of spatial resolution, 3×3 angular views and L=3 bands in the visible spectrum (RGB). This light field was modified adding two new spectral bands to each angular image. These matrices are located between the existing bands and a cubic interpolation (Cubic spline) defines the value of the new pixels. Thus, the resultant synthetic dataset is a multispectral light field with 5 spectral bands, 593×840 pixels of spatial resolution, and 3×3 angular views (Figure 8(a)). The spectral bands range between 450nm and 650nm.

A reduced version with 405×360 pixels, 3×3 angular views and 5 spectral bands of the data set is used to test the proposed algorithm (Figure 8(b)). The underlying signal is reconstructed solving the optimization problem described in Eq. (7), using the LASSO algorithm [19] briefly summarized in Algorithm 2. The reconstruction quality was quantified with the Peak signal-to-noise ratio (PSNR), which is measured in decibels (dB). For the simulations, a patch size N=9×9×3×3×5 was used i.e. 9×9 pixels in spatial resolution, 3×3 angular views and 5 spectral bands.

The learned dictionary was created choosing 5D random patches from four training multispectral light field data sets show in Figure 9. The patch size was N=9×9×3×3×5, i.e. 9×9 pixels of spatial resolution, 3×3 angular views and 5 spectral bands. Algorithm 3 was employed to learn the overcomplete dictionary [13]. All simulations were conducted using a dual Intel Xeon E5-2670v2, 2.5 GHz processor and 64 GB RAM.
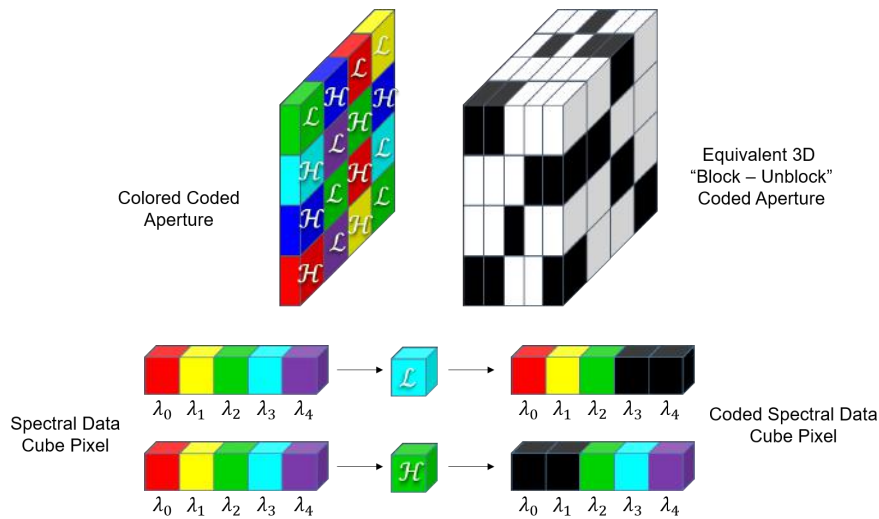
Figure 7 Illustration of a colored coded aperture and its equivalent binary coded aperture equivalence. Each colored coded aperture voxel can operate as a low – pass ($\mathcal{L}$) or high – pass ($\mathcal{H}$) optical filter

(a) (b)
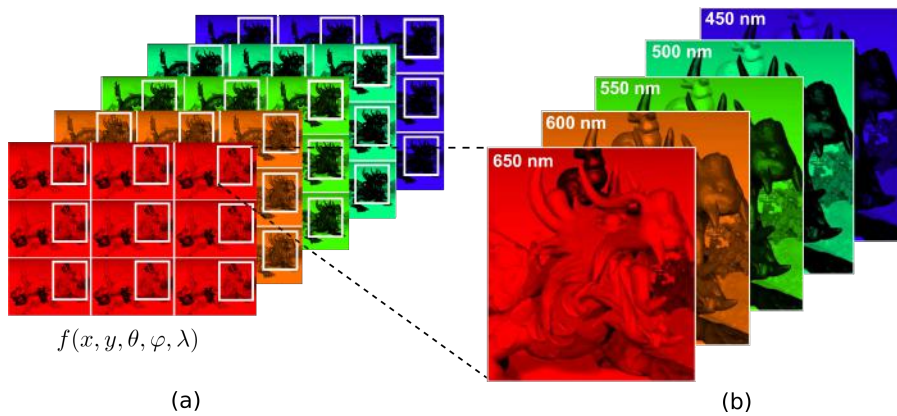
Figure 8 Synthetic data set, dragon's scene. (a) Multispectral light field with spatial (x,y), angular (θ,ϕ) and spectral information (λ). The spatial dimension of the data set is reduced for each angular view in each spectral band. (b) The reduced version of the data set is used to test the proposed algorithm, the size of the reduced data set is 405×360 pixels, 3×3 angular views and 5 spectral bands


**Figure 9** Synthetic training light fields mapped to RGB. Available online at the MIT Media Lab webpage

3.1. Reconstruction from the measurements vector

The notion of using patches helps in speeding up the reconstruction process. The 5D overlapped patches are reconstructed separately. A sliding window was used to merge the reconstructed overlapping patches with a median filter. The reconstructions are shown in Figures 10 and 11, and the PSNR achieved is exhibited for each reconstructed slide. Figure 10 (top) shows the first angular view of the multispectral light field ground truth, mapped to a RGB profile and its five spectral bands between 450nm and 650nm. The bottom row shows the original RGB profile and its five reconstructed spectral bands. Figure 11 shows the reconstruction of the third spectral band for all angular views. The 5D multispectral light field was reconstructed from just 2D coded projections. It can be noticed that the reconstructions exhibit a good spatial and angular approximation of the synthetic scene. Figure 12 presents the spectral signatures of the original image compared to the reconstructed signatures for three points of a specific
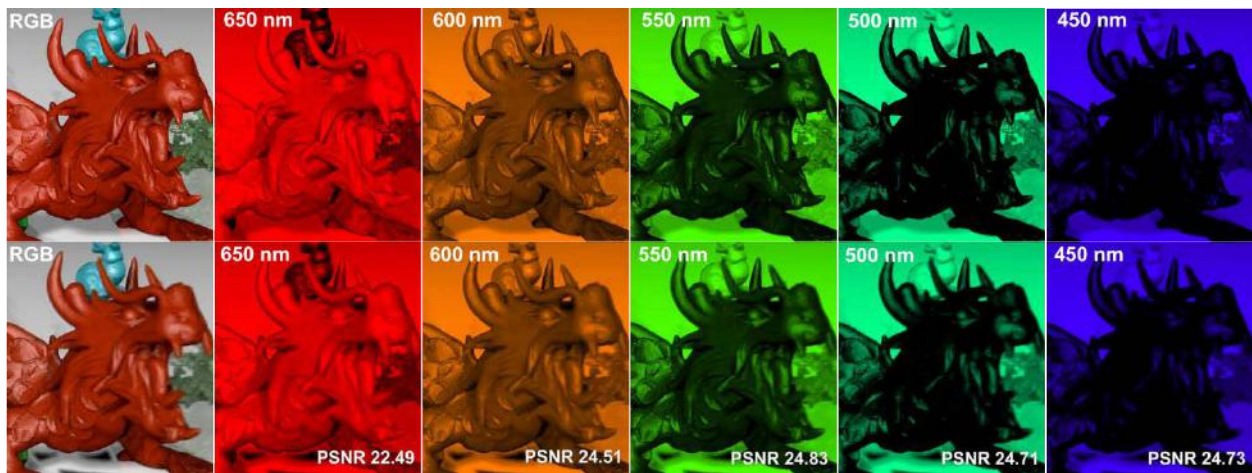
Figure 10 (Top) The original and (Bottom) reconstructed multispectral light fields for the first angular view and its five spectral bands. The PSNR is presented for each reconstructed spectral band

Central View

Spatio - Angular view
of the third spectral band (550nm)
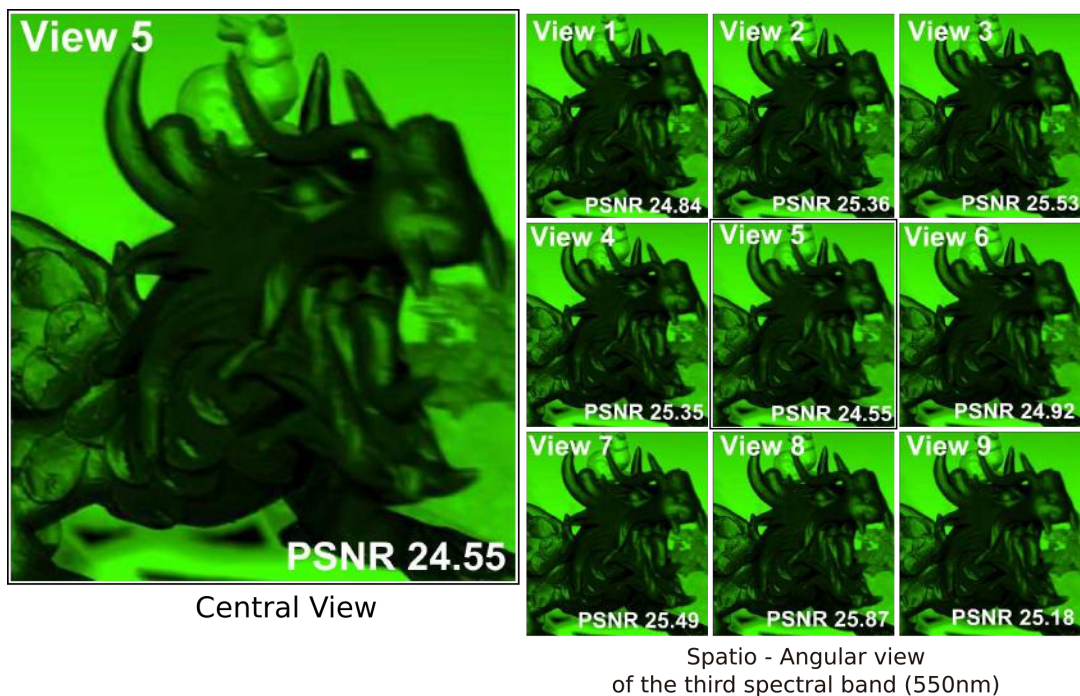
Figure 11 Reconstruction of the third spectral band (L=550 nm). All the angular views are depicted. The range of achieved PSNR is between 24 dB and 25 dB for the different views
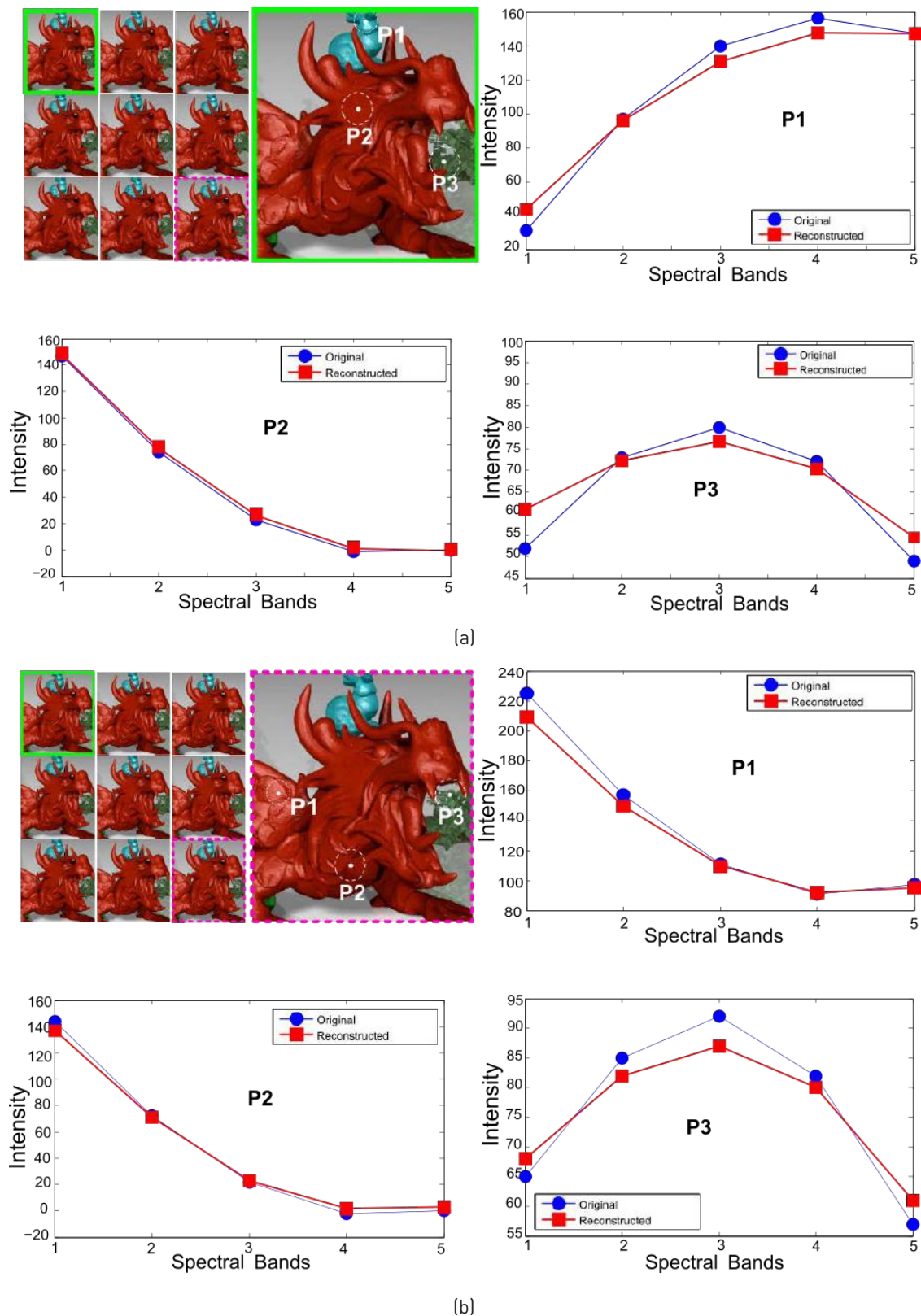
(a)

(b)

Figure 12 Spectral signature comparison for three spatial locations indicated by the points P1, P2 and P3 for the first angular view (a) and the last angular view (b)

angular view. The resultant curves are close to the originals showing a good estimation of the multispectral light field information.

The achieved compression ratio in the reconstruction of the multispectral light field is calculated in Eq. (10) using Eq. (8)

$$r = \frac{1}{3 \times 3 \times 5} = \frac{1}{45} \approx 0.022$$

(10)

The percentage of achieved compression is 97.8%. In other words, the multispectral light field reconstruction obtained with the proposed model requires only 2.2% of the whole

multidimensional image compared with a traditional acquisition model that requires 100% of the information. Figure 13 shows an overview of the reconstruction quality for the dragon's scene. The achieved PSNR is between 22 dB and 25 dB.

The Structural SIMilarity (SSIM) index is calculated in order to estimate the image quality in terms of: luminance, contrast and structure [20]. SSIM provides a better consistency with the perceived quality measurement. The SSIM range is between -1 and 1, where 1 is the identical image. Table 1 shows the SSIM calculated for the third spectral band and a comparison with the calculated PSNR. Similar results were obtained for the remaining 4 spectral bands.
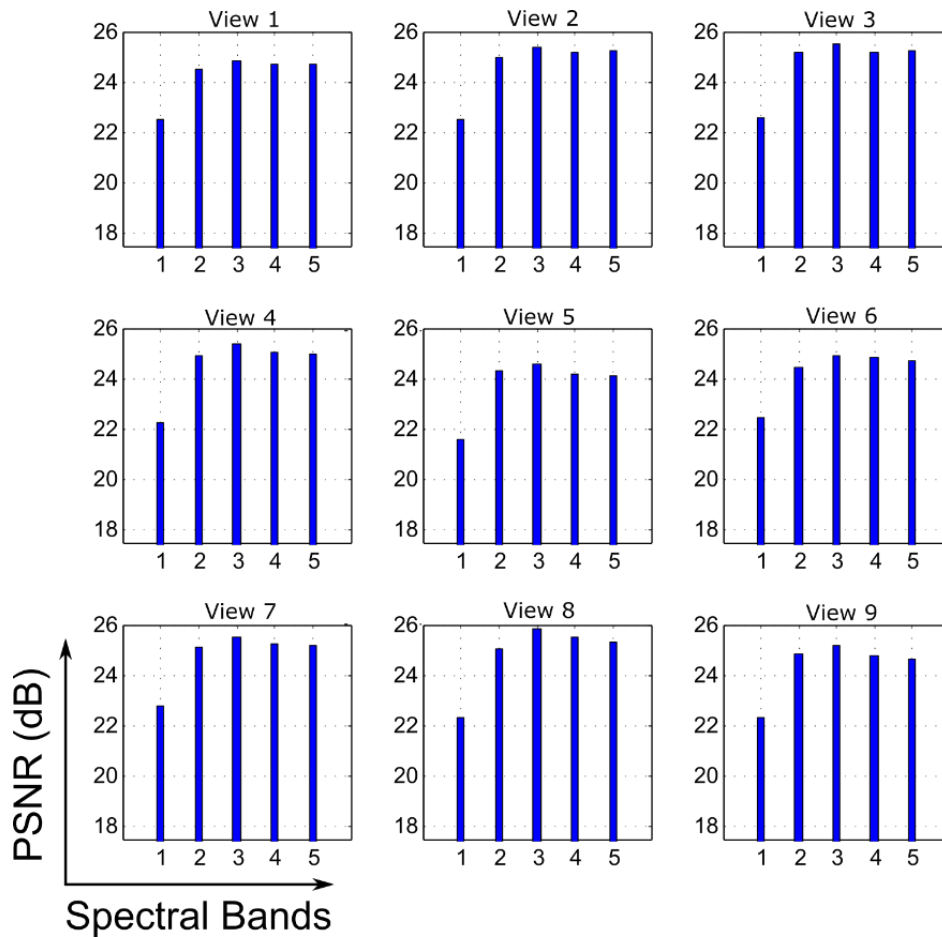
Figure 13 Reconstruction quality (PSNR) achieved for each angular view of each spectral band measured in dB

Table 1 Reconstruction quality comparison in terms of SSIM and PSNR for the third spectral band and all angular views

Angular View	1	2	3	4	5	6	7	8	9
SSIM	0.85	0.88	0.88	0.88	0.84	0.86	0.88	0.89	0.87
PSNR [dB]	24.83	25.35	25.53	25.35	24.55	24.92	25.49	25.86	25.18

4. Conclusions

A mathematical model of the sampling and reconstruction multispectral light field based on compressive sensing with colored coded apertures has been developed. The proposed model includes the spectral information in a light field without increasing the measurements size captured in the detector. Simulations show the reconstruction of a multispectral light field using just the 2.2% of information of the scene. The full spatio –spectro– angular data cube is reconstructed exhibiting PSNR values between 22 dB and 25 dB. In a future work, this model could be extended to the optimization of the colored coded aperture pattern in order to improve the reconstruction quality.

5. Acknowledgments

6. References

1. E. Candès, "Compressive sampling", in *International Congress of Mathematicians*, Madrid, Spain, 2006, pp. 1433-1452.

2. M. Fornasier and H. Rauhut, "Compressive sensing", in *Handbook Mathematical Methods in Imaging*, O. Scherzer (ed). New York, USA: Springer, 2011, pp. 187-228.

3. D. Donoho, "Compressed sensing", *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289-1306, 2006.

4. G. Shaw and H. Burke, "Spectral imaging for remote sensing", *Lincoln Lab. J.*, vol. 14, no. 1, pp. 3-28, 2003.

5. K. Marwah, G. Wetzstein, Y. Bando and R. Raskar, "Compressive light field photography using overcomplete dictionaries and optimized projections", *ACM Trans. Graph.*, vol. 32, no. 4, 2013.

6. A. Mian and R. Hartley, "Hyperspectral video restoration using optical flow and sparse coding", *Optics Express*, vol. 20, no. 10. p. 10658-10673, 2012.

7. K. Choi and D. Brady, "Coded aperture computed tomography", in *SPIE 7468: Adapt. Coded Aperture Imaging, Non-Imaging, Unconv. Imaging Sens. Syst.*, San Diego, USA, 2009.

8. E. Adelson and J. Bergen, "The plenoptic function and the elements of early vision", in *Computational Models of Visual Processing*, M. Landy and J. Movshon (eds). Cambridge, USA: MIT Press, 1991, pp. 3-20.

9. G. Wetzstein, I. Ihrke, D. Lanman and W. Heidrich, "Computational plenoptic imaging", *Comput. Graph. Forum*, vol. 30, no. 8, pp. 2397-2426, 2011.

10. H. Rueda, H. Arguello and G. Arce, "Compressive spectral imaging based on colored coded apertures", in *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), Florence, Italy, 2014, pp. 7799-7803.

11. S. Babacan, R. Ansorge, M. Luessi, R. Molina and A. Katsaggelos, "Compressive Sensing of Light Fields", in *IEEE International Conference on Image Processing* (ICIP), Cairo, Egypt, 2009, pp. 2337-2340.

12. R. Ng, "Digital light field photography", Ph.D. dissertation, Standford University, Stanford, USA, 2006.

13. M. Aharon, M. Elad and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation", *IEEE Trans. Signal Process.*, vol. 54, pp. 4311-4322, 2006.

14. A. Brückner, "Microoptical Multi Aperture Imaging Systems", Ph.D. dissertation, Friedrich Schiller University Jena, Jena, Germany, 2011.

15. S. Kim, E. Lee, M. Hayes and J. Paik, "Multifocusing and depth estimation using a color shift model-based computational camera", *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4152-4166, 2012.

16. Y. Bando, B. Chen and T. Nishita, "Extracting depth and matte using a color-filtered aperture", *ACM Transactions on Graphics*, vol. 27, no. 5, 2008.

17. H. Arguello and G. Arce, "Colored coded aperture design by concentration of measure in compressive spectral imaging", *IEEE Trans. Image Process.*, vol. 23, pp. 1896-1908, 2014.

18. H. Arguello, H. Rueda, Y. Wu, D. Prather and G. Arce, "Higher-order computational model for coded aperture spectral imaging", *Appl. Opt.*, vol. 52, no. 10, pp. 12-21, 2013.

19. E. van den Berg and M. Friedlander, "Probing the Pareto Frontier for Basis Pursuit Solutions", *SIAM Journal on Scientific Computing*, vol. 31, no. 2. pp. 890-912, 2009.

20. Z. Wang, A. Bovik, H. Sheikh and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity", *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, 2004.