# Compressive Light Field Photography

Kshitij Marwah[1] Gordon Wetzstein[1] Ashok Veeraraghavan[2] Ramesh Raskar[1]
[1]MIT Media Lab [2]Rice University

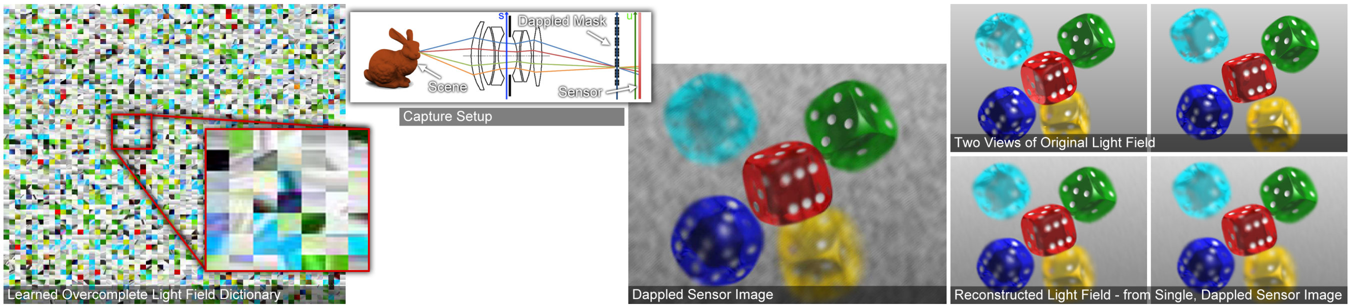


**Figure 1:** *Compressive light field acquisition using joint optical light coding and compressive computational reconstruction. Using machine learning techniques, we extract the essence of natural light fields from archives and store them in overcomplete dictionaries (left). The same optimization procedure also generates optimal attenuation mask patterns that are physically mounted, at a slight offset, in front of a camera sensor (center left). These masks create a "dappled" pseudo-random pattern in captured photographs (center). From a precomputed dictionary and a single sensor image, we then reconstruct a full-resolution light field that contains slightly different viewpoints of a scene (right) and allows a photograph to be refocused after the fact (see supplement).*

**Introduction and Overview** Light field cameras, e.g. [Lytro 2012; Veeraraghavan et al. 2007], have ushered a new direction in photography allowing consumers to synthesize photographs with novel viewpoints or varying focus after the actual recording. Unfortunately, current light field camera designs impose a fixed tradeoff between spatial and angular resolution — spatial resolution is reduced to capture angular light variation on the sensor. We introduce a principled computational framework and a new camera design to acquire and reconstruct light fields at full spatial and angular resolution from a single exposure. Our framework introduces a high-dimensional sparse basis for light fields learned from millions of light fields patches. The same optimization procedure also allows for the synthesis of optimal mask patterns that are mounted at a slight offset in front of the sensor and optically attenuate the light field before it is recorded by the sensor. Finally, a weighted compressive sensing-style reconstruction is performed to recover the light field. We demonstrate, in theory and with simulations, how our compressive approach to light field photography outperforms state-of-art techniques.

**Sparse Coding of Light Fields** Light fields can be thought of as a collection of slightly different perspectives of the same scene that vary over the size of the camera aperture. As these viewpoints have a very narrow baseline, they contain a large amount of redundancy. This redundancy can be computationally exploited using a variety of different bases commonly used for redundant or sparse coding, including data-independent bases such as Fourier transforms, contourlets, bandelets, and Hadamard codes as well as data-dependent functions such as principal components or overcomplete dictionaries. In the signal processing literature, learned overcomplete dictionaries are commonly considered the best approach for compressive sensing [Mairal et al. 2009]. We consider millions of light field patches from various light field archives that each encode the essence of depth variations and occlusions in natural scenes. This basis is about $100\times$ overcomplete in which each light field patch is represented as a linear combination of a few dictionary elements.

**Compressive Reconstruction** By placing a mask on the sensor we optically modulate the light rays incident on the sensor; the mask creates a dappled pattern in the sensor image. The pattern of modulation is learned simultaneously with the dictionary to be the most incoherent physically probable sensing matrix for capture. To reconstructing the light field back we employ a weighted L1-minimization (see [E. Candes and Y. Eldar and D. Needell and P. Randall 20010] for more details) on the dappled sensor image as:

$$\begin{aligned} &\arg\min_{\alpha} \; \|\alpha\|_1 \\ &\text{s.}t. \; \|I_{sensor} - \Phi_M D\alpha\|_2 \leq \epsilon, \end{aligned} \tag{1}$$

where $\Phi_M$ is the measurement matrix modeling the attenuation mask and angular integration of the sensor, $I_{sensor}$ is the captured sensor image, $D$ is the overcomplete dictionary, and $\alpha$ is the unknown vector with sparse coefficients. The light field is represented as the combination of a sparse set of coefficients and corresponding elements in the dictionary $D\alpha$.

**Results and Discussion** We learned an overcomplete dictionary for patch size $16 \times 16$ in space and $5 \times 5$ viewpoints from synthetic light fields rendered with a raytracer and also reparameterized light fields from the Stanford Light Field Archive. Our dictionary contains about 100,000 elements with each patch being represented in no more than ten coefficients. We tested our reconstruction algorithm using a new light field that was not part of the training dataset, which is simulated to be captured with the learned mask pattern in front of the sensor (Fig. 1). The sensor image is then used to recover the light field, which includes multiple viewpoints and can be used to refocus a photograph, at a higher resolution than conventional methods, after it is captured (see supplement).

Our approach has the potential to significantly improve the resolution of next-generation computational cameras, which are already emerging on the consumer market.

## References

E. Candes and Y. Eldar and D. Needell and P. Randall. 20010. Compressed sensing with coherent and redundant dictionaries. *Harmonic Analysis*.

Lytro, I., 2012. Lytro Light Field Camera.

Mairal, J., Bach, F., Ponce, J., and Sapiro, G. 2009. Online dictionary learning for sparse coding. *International Conference on Machine Learning*.

Veeraraghavan, A., Raskar, R., Agarwal, A., Mohan, A., and Tumblin, J. 2007. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Siggraph*.


SIGGRAPH 2012, Los Angeles, California, August 5 – 9, 2012.
ISBN 978-1-4503-1435-0/12/0008