DIRECT AND PROGRESSIVE RECONSTRUCTION OF DUAL PHOTOGRAPHY IMAGES

Binh-Son Hua[1] *Imari Sato*[2] *Kok-Lim Low*[1]

[1]National University of Singapore [2]National Institute of Informatics, Tokyo, Japan

ABSTRACT

Dual photography is a well-known application of light transport acquired by a projector-camera system. By applying compressive sensing, compressive dual photography [1] is a fast approach to acquire the light transport for dual photography. However, the reconstruction step in compressive dual photography can still take several hours before dual images can be synthesized because the entire light transport needs to be reconstructed from measured data. In this paper, we present a novel reconstruction approach that can directly and progressively synthesize dual images from measured data without the need of first reconstructing the light transport. We show that our approach can produce high-quality dual images in the order of minutes using only a thousand of samples. Our approach is most useful for previewing a few dual images, e.g., during light transport acquisition. As a by-product, our method can also perform low-resolution relighting of dual images. We also hypothesize that our method is applicable to reconstructing dual images in a single projector - multiple cameras system.

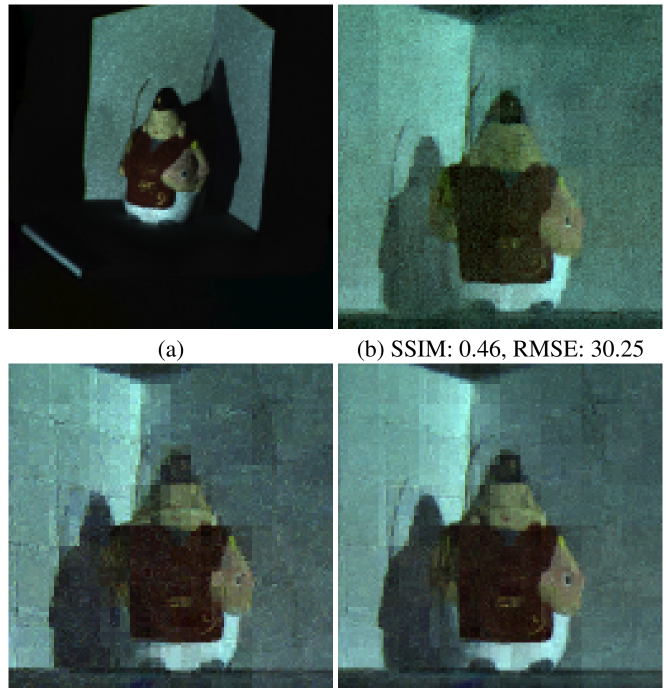
Index Terms— dual photography, compressive sensing

1. INTRODUCTION

Light transport [2] is a mathematical operator that captures how light bounces among surface points in a scene. In computer graphics and computer vision, several applications have been proposed that make use of light transport such as relighting [2], dual photography [3], and radiometric compensation. Among those, dual photography is an interesting and well-known application of light transport thanks to its simplicity and usefulness. Given a light transport of a scene lit by a controlled light source and captured by a camera, dual photography can virtually swap the roles of the light source and the camera to produce dual images. The dual images can be perceived as if the scene is lit by the camera and captured by the light source. Dual photography is also useful in capturing 6D light transport [3].

Traditionally, to obtain dual images of a scene, it is necessary to first acquire and reconstruct the entire light transport matrix of the scene. Several approaches have been proposed to efficiently acquire and reconstruct light transport such as multiplexed illumination [4], compressive sensing of light transport [1, 5], or optical computing of light transport [6]. However, reconstructing the entire light transport matrix from the acquired data can still be very costly since the number of rows and columns of the light transport matrix can be tens or hundreds of thousands. This means a huge amount of computational time is required before the first dual image can be synthesized and ready for display.

In this paper, we present a novel approach to efficiently compute dual images from measured data without reconstructing the light transport. We build our method upon compressive sensing of light transport [1, 5] and propose an approach to directly and progressively reconstruct high-quality dual images using L1-norm optimization. The number of measurement samples needed is comparable to that used for light transport reconstruction in compressive dual photography. Such direct reconstruction allows us to quickly synthesize dual images as soon as the acquisition data is enough. Our method can also generate progressive results while the dual image is being reconstructed. Therefore, our method can be beneficial for previewing a few dual images. Besides, we also demonstrate that our method can be used for low-resolution relighting of dual images. We also hypothesize that our approach is extendable to synthesize dual images in setups that have a single light source and multiple cameras.



(a) (b) SSIM: 0.46, RMSE: 30.25

(c) SSIM: 0.47, RMSE: 25.93 (d) SSIM: 0.55, RMSE: 27.32

Fig. 1. Dual photography. (a) Camera view. (b) Dual image directly reconstructed from 16000 samples, which is not practical. (c) Dual image progressively reconstructed from only 1000 samples using our method with 64 basis dual images. (d) Dual image reconstructed with settings as in (c) but from 1500 samples. Haar wavelet is used for the reconstruction.

2. RELATED WORKS

Recently, several approaches to efficiently acquire and reconstruct light transport have been proposed [4, 3, 1, 5, 6]. In the seminal work about dual photography [3], Sen et al. proposed a hierarchical approach to detect projector pixels that can be turned on simultaneously in a single light pattern. This greedy-like approach can reduce the number of light patterns in the acquisition to the order of thousands. In the worst case when most of projector pixels conflict to each other and can only be scheduled to be turned on sequentially, this approach can be as slow as brute-force acquisition.

In compressive dual photography [5, 1], the authors proposed to use rows of measurement matrices in compressive sensing as light patterns, thus turns light transport acquisition into a compressive sensing problem that allows the light transport to be reconstructed using the well-known L1-norm optimization. This approach works well for high-rank and sparse light transport matrix which is often seen in a projector-camera system. In this work, we also build our approach based upon compressive sensing. We provide a simple reformulation of compressive dual photography that allows us to directly and progressively reconstruct dual images using L1-norm optimization.

Recently, O'Toole and Kutulakos [6] proposed to use Arnoldi iterations to determine eigenvectors of a light transport using optical computing. While their method only requires less than a hundred of images, it is more suitable for dense and low-rank light transport where the light source is diffuse. In this work, we target sparse and high-rank light transport.

While compressive sensing of light transport is designed to minimize the number of images to acquire, it often results in long computation time needed to reconstruct the light transport in the post-processing step. This is an issue for dual photography, especially when we only need to see a handful number of dual images. Therefore, it is necessary to have an approach that can compute dual images from measured data as fast as possible. In this work, we fill in this gap by proposing such an approach based on compressive sensing and L1-norm optimization.

Sen et al. [1] also discussed about single pixel imaging and how it is related to compressive dual photography. This is probably most closely related to direct reconstruction of dual images which we proposed in this work. The authors noticed that directly recovering the reflectance function of this single pixel, which is equivalent to directly computing the dual image under floodlit lighting, is rather troublesome because the dual image is more complicated and therefore a lot more samples are needed. In this work, we solve this problem by presenting a simple basis so that dual images can be progressively reconstructed from a small amount of measurement samples.

Finally, while it is not closely related to dual photography, we note that the idea of direct reconstruction using compressive sensing was also exploited to obtain the inverse light transport [7].

3. COMPRESSIVE DUAL PHOTOGRAPHY

Let $\mathbf{T}$ be the light transport matrix of a scene captured by a projector-camera system. Suppose that the light source emits pattern $\mathbf{l}$. The image $\mathbf{c}$ of the scene captured by the camera can be represented by the light transport equation:

$$\mathbf{c} = \mathbf{T}\mathbf{l}. \tag{1}$$

In dual photography, by utilizing Helmholtz reciprocity, the dual image can be computed as

$$\mathbf{c}' = \mathbf{T}^\top \mathbf{l}', \tag{2}$$

where $\mathbf{l}'$ is the dual light pattern virtually emitted by the camera and $\mathbf{c}'$ is the dual image virtually captured by the light source.

By projecting a set of N light patterns $\mathbf{L} = [\mathbf{l}_1 \dots \mathbf{l}_N]$ and capturing images of the scene $\mathbf{C} = [\mathbf{c}_1\ \mathbf{c}_2 \dots \mathbf{c}_N]$ lit by this set of patterns, we can rewrite the light transport equation as

$$\mathbf{C}^\top = \mathbf{L}^\top \mathbf{T}^\top, \tag{3}$$

which suggests an elegant way to measure light transport $\mathbf{T}$ using compressive sensing. Each row of $\mathbf{T}$ can be measured by letting $\mathbf{L}^\top$ be a measurement matrix such as Bernoulli or Gaussian matrix that satisfies the restricted isometry property [8]. Each row $\mathbf{t}_i$ of light transport matrix $\mathbf{T}$ can be independently reconstructed by minimizing

$$\mathbf{t}_i = \arg\min_{\mathbf{u}} \|\mathbf{c}_i^\top - \mathbf{L}^\top \mathbf{u}\|_2^2 + \lambda \|\mathbf{W}^\top \mathbf{u}\|_1 \tag{4}$$

where $\mathbf{c}_i^\top$ denotes column i of $\mathbf{C}^\top$, $i \in [1 \dots |\mathbf{T}|]$, $|\mathbf{T}|$ the number of rows of matrix $\mathbf{T}$, $\mathbf{W}$ the basis of the space where each row of the transport matrix can be sparse. However, since $|\mathbf{T}|$ can be tens of thousands, e.g., $|\mathbf{T}| = 128 \times 128$ which represents a rather low-resolution camera view, the reconstruction of $\mathbf{T}$ can take several hours to complete [1].

To speed up, it is possible to further exploit coherency among pixels in each column of matrix $\mathbf{T}$ by using another compression basis $\mathbf{P}$ as in [5]. We get:

$$\mathbf{P}^\top \mathbf{C} = (\mathbf{P}^\top \mathbf{T} \mathbf{W})(\mathbf{W}^\top \mathbf{L}). \tag{5}$$

We capture images as before but transform them into basis $\mathbf{P}$ in the post-processing. As before, compressive sensing can be applied to reconstruct each row of the compressed matrix $\mathbf{P}^\top \mathbf{T} \mathbf{W}$ independently, but this time the number of rows needed to reconstruct can be less. However, in our observation, the number of non-zero rows of $\mathbf{P}^\top \mathbf{C}$ is still in the order of thousands because the captured images $\mathbf{C}$ lit by measurement patterns $\mathbf{L}$ can contain a lot of complex blocky patterns that are difficult to compress by basis $\mathbf{P}$.

4. DIRECT AND PROGRESSIVE RECONSTRUCTION

4.1. Direct reconstruction

We are now ready to present our approach to directly reconstruct dual images, which we build on top of compressive dual photography [1]. We start by showing that dual image can be directly computed from the acquired images and light patterns. By multiplying the dual light pattern $\mathbf{l}'$ to both sides of Equation 3, it is easy to get:

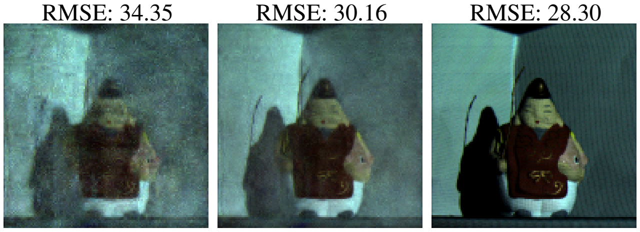
$$\mathbf{C}^\top \mathbf{l}' = \mathbf{L}^\top \mathbf{c}'. \tag{6}$$

By letting $\mathbf{L}^\top$ be a measurement matrix and pre-computing the left part $\mathbf{C}^\top \mathbf{l}'$, we can view dual image synthesis as a compressive sensing problem. Therefore, the dual image can be directly reconstructed by L1-norm optimization:

$$\mathbf{c}' = \arg\min_{\mathbf{u}} \|\mathbf{C}^\top \mathbf{l}' - \mathbf{L}^\top \mathbf{u}\|_2^2 + \lambda \|\mathbf{W}^\top \mathbf{u}\|_1. \tag{7}$$

Theoretically, this approach should be able to reconstruct the dual image $\mathbf{c}'$. Unfortunately, in practice, in order to obtain a high-quality dual image, almost tens of thousands number of measurement samples, or camera images and light patterns, are necessary. This is because dual image is not as sparse as reflectance functions stored in rows of light transport $\mathbf{T}$, thus it requires more samples in the reconstruction.

(a) 4000 samples. SSIM: 0.43 RMSE: 34.35 | (b) 8000 samples. SSIM: 0.44 RMSE: 30.16 | (c) 16000 samples. SSIM: 0.44 RMSE: 28.30



(d) 1000 samples. SSIM: 0.41 RMSE: 31.86 | (e) 2000 samples. SSIM: 0.57 RMSE: 28.82 | (f) Ground truth.

Fig. 2. Comparison between direct and progressive reconstruction. Dual image (a), (b), and (c) are from direct reconstruction. Dual image (d) and (e) are from progressive reconstruction with 64 basis dual images. (f) Ground truth is generated from light transport from 16000 samples by inverting the circulant measurement matrix. Daubechies-8 wavelet is used for the reconstruction.

4.2. Progressive reconstruction

We propose a simple approach in order to overcome the above issue. Suppose that we can project the dual light pattern $\mathbf{l}'$ into a basis $\mathbf{Q} = [\mathbf{q}_1\ \mathbf{q}_2 \dots \mathbf{q}_{|\mathbf{Q}|}]$:

$$\mathbf{l}' = \mathbf{Q}\mathbf{w} = \sum_i w_i \mathbf{q}_i, \tag{8}$$

where $|\mathbf{Q}|$ is the number of basis vectors in $\mathbf{Q}$, $i \in [1 \dots |\mathbf{Q}|]$, $\mathbf{w}$ the coefficient vector of $\mathbf{l}'$ in basis $\mathbf{Q}$. Therefore, the dual image can be computed by

$$\mathbf{c}' = \sum_i w_i \mathbf{c}'_i \tag{9}$$

where $\mathbf{c}'_i$ is the *basis dual image* which satisfies

$$\mathbf{C}^\top \mathbf{q}_i = \mathbf{L}^\top \mathbf{c}'_i. \tag{10}$$

Each basis dual image can be found independently by optimizing

$$\mathbf{c}'_i = \arg\min_{\mathbf{u}} \|\mathbf{C}^\top \mathbf{q}_i - \mathbf{L}^\top \mathbf{u}\|_2^2 + \lambda \|\mathbf{W}^\top \mathbf{u}\|_1. \tag{11}$$

The intuition behind this formulation is that we can split the reconstruction of the dual image into several passes, and reconstruct each basis dual image that forms a part of the dual image in each pass. It is significant to guarantee that each basis dual image should be sufficiently sparse so that it can be successfully reconstructed using Equation 11 without using too many measurement samples. As shown in Figure 1 and 2, the number of samples needed to reconstruct basis dual images is comparable to that required to reconstruct the entire light transport in traditional compressive dual photography, which is more practical than direct reconstruction. Figure 3 shows a few examples of the progressive reconstruction.

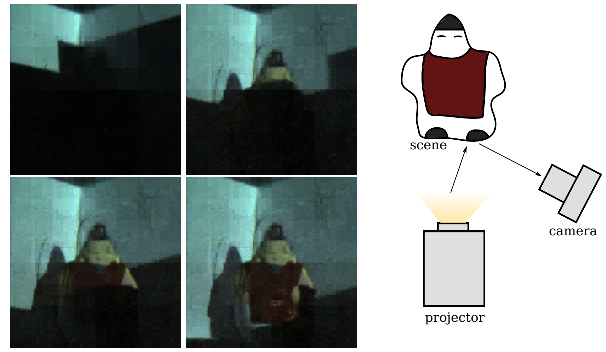


Fig. 3. Progressive results of the dual image in Figure 1(d) by accumulating those reconstructed basis dual images. Our projector-camera setup to acquire light transport is shown in the diagram.

We choose basis $\mathbf{Q}$ based on two following criteria. First, the dimension of space $\mathbf{Q}$ should be as low as possible. It is best to choose $\mathbf{Q}$ of which the dimension is about tens or hundreds. Second, the basis dual images $\mathbf{c}'_i$ obtained by setting dual lighting to basis vectors of $\mathbf{Q}$ should be sparse so that high quality reconstruction can be achieved.

Based on such criteria, we propose a simple and easy to implement basis $\mathbf{Q}$ as follows. We subdivide the dual lighting pattern $\mathbf{l}'$ into a grid and let each patch in the grid be a basis vector $\mathbf{q}_i$. Therefore, the weight w_i is simply set to one. It is easy to see that smaller patch size tends to produce sparser coefficients of basis dual images in the wavelet domain. This can yield higher accuracy in the reconstruction but result in longer computational time.

An advantage of choosing basis $\mathbf{Q}$ as above is that we can display progressive results of the dual image by accumulating existing basis dual images while other remaining basis dual images are pending for reconstruction, which is useful for previewing applications.

5. IMPLEMENTATION

We use a projector-camera system to acquire the light transport. The projector is a Sony VPL-DX11. The camera is a Sony DXC-9000 of which the response curve is linear.

The light patterns to compressively acquire the light transport are obtained from a circulant matrix of which the first row is an i.i.d Bernoulli distribution with value -1 and 1 [9]. An advantage of using a circulant measurement matrix is that its multiplication with a vector can be quickly computed using fast Fourier transform. Also, circulant matrix requires very little memory storage as only the first row needs to stored.

Since our patterns contain both positive and negative values, we project positive and negative patterns separately and combine the corresponding camera images in the post-processing by the formula $\mathbf{c} = \mathbf{T}(\mathbf{l}^+ - \mathbf{l}^-) = \mathbf{c}^+ - \mathbf{c}^-$, where superscript $+$ and $-$ denote positive and negative patterns and images, respectively. For simplicity, we also crop and downsample camera images to the same size as the light patterns so the light transport is a square matrix.

We implement our system in MATLAB. We implement split Bregman iterations [10] for L1-norm optimization in Equation 11. We let $\lambda = 0.001$ for all progressive reconstruction. We let $\lambda = 0.05$ for direct reconstruction to further suppress noise. We test the reconstruction with Haar wavelet and Daubechies-8 wavelet provided by the Rice Wavelet Toolbox [11].

During progressive reconstruction, we discard basis dual images of which the absolute maximum value of their corresponding left-hand side vector $\mathbf{C}^\top \mathbf{q}_i$ is less than 10^{-4}. In fact, this corresponds to regions that can be lit by the projector but are out of field of view of the camera so zero solutions for basis dual images are appropriate.

6. EXPERIMENTS

The results of our method are shown in Figure 1. The resolution of the dual image is 128×128. As can be seen, our method is able to reconstruct a good-quality dual image without first obtaining the light transport. We provide quantitative comparisons between our results of direct and progressive reconstruction and the ground truth shown in Figure 2 using both structural similarity index (SSIM) [12] and root-mean-square error (RMSE).

In Figure 1, by using basis $\mathbf{Q}$ with patch size set to 16 pixels, only 1000 samples are needed to reconstruct total 64 basis dual images and a high-quality final dual image. Figure 3 shows some of the progressive results of the dual image during reconstruction. In contrast, directly reconstructing the dual image without basis $\mathbf{Q}$ requires 16000 samples in order to reach similar image quality, which is far less practical. In fact, given 16000 samples, it is often more preferable to reconstruct the entire light transport in the post-processing by inverting the circulant measurement matrix using FFT, which is fast. Here we use this approach to generate the ground truth dual image as shown in Figure 2(f). We do not opt to reconstruct the light transport from only 1000 measurement samples since it takes tens of thousands of L1-norm optimization which is too time consuming to perform.

Figure 2 further demonstrates how our method works with different number of samples for both direct and progressive reconstruction using Daubechies-8 wavelet for compression. As expected, more samples allows more details of the dual image to be revealed.

As a by-product, we demonstrate a relighting application by linearly combining basis dual images by setting the weight vector to a low-resolution lighting pattern. Figure 4 shows our relit images. The new lighting has resolution 8×8 since our basis vectors are derived from 8×8 grid patches.

We measured the running time of our progressive reconstruction on an Intel Core 2 Quad processor clocked at 2.8 GHz with 8 GB of RAM. Our MATLAB implementation output the direct result (b) of Figure 1 in 10 minutes and the progressive result (c) in 40 minutes. While progressive reconstruction is a few times slower, it saves a large amount of acquisition time as it requires far less number of samples to reach similar image quality. With the same number of samples, progressive reconstruction is also faster than reconstructing the entire light transport when only a few images are needed.

6.1. Running time analysis

We provide a simple analysis to estimate how much and when progressive reconstruction is better than traditional light transport reconstruction in terms of running time as follows. We assume the following model to predict the running time of progressive reconstruction:

$$t = 2\alpha N + k\rho|\mathbf{Q}|, \tag{12}$$

where t is the running time in seconds, N the number of samples acquired, α the time to acquire a single image, ρ the time to reconstruct a basis dual image, k the number of dual images we are interested in in total. The constant 2 represents the need to capture two images per sample due to positive and negative entries of the measurement

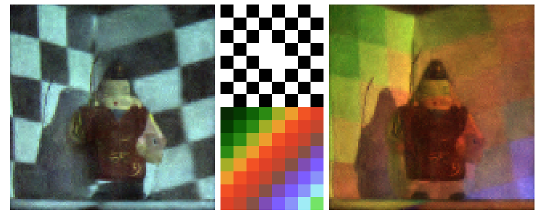


Fig. 4. Relighting of the dual image in Figure 2(e).

matrix. Similarly, the running time of traditional light transport reconstruction can be predicted by:

$$t' = 2\alpha N' + \rho'|\mathbf{T}|, \tag{13}$$

where N' is the number of samples needed to acquire for light transport reconstruction, ρ' the time to reconstruct a row of $\mathbf{T}$.

Empirically, we set $\alpha = 1$ second, $N = 1000$ samples, $\rho = 75$ seconds, according to the examples in the previous figures. Since the reflectance function stored in each row of light transport $\mathbf{T}$ can be more sparse than dual images, we pessimistically assume that $N' = 500$ which means our progressive reconstruction requires twice the number of samples. We also set $\rho' = 2$ to assume that each row of $\mathbf{T}$ can be reconstructed much faster. We also have $|\mathbf{Q}| = 64$ and $|\mathbf{T}| = 16000$.

As a result, in order to guarantee $t < t'$, we need to bound $k \leq 6$. This indicates the maximum number of dual images we can reconstruct before our method cannot offer any time savings. When only a dual image is needed, or $k = 1$, the speed up is about $5\times$.

7. DISCUSSION

Conventionally, in order to compute a dual image of light transports of a scene captured by a single projector and multiple cameras, the light transport matrix between each pair of projector-camera needs to be reconstructed. In such case, for quick reconstruction, our method is still applicable. In the case of two cameras, we have:

$$\begin{bmatrix} \mathbf{C}_1^\top & \mathbf{C}_2^\top \end{bmatrix} \begin{bmatrix} \mathbf{l}_1' \\ \mathbf{l}_2' \end{bmatrix} = \mathbf{L}^\top \mathbf{c}'. \tag{14}$$

It is natural to extend the formulation to the case of multiple cameras. We leave the implementation of such a system for future works.

8. CONCLUSIONS

In this paper, we presented an approach based on compressive sensing to directly and progressively reconstruct dual photography images without the need of reconstructing the entire light transport. Our method can be useful for previewing of dual images. We are also able to perform low-resolution relighting of dual images.

There are a few limitations in our approach. First, our reconstructed dual images tend to be noisier than those produced by the full light transport. This can be explained by the dot product between the camera images and the dual lighting pattern, which sums up the variance of each camera pixel. Second, our method may fail when the basis dual images are not sparse enough.

It is interesting to extend this work further in the future. First, it can be useful to have a careful noise analysis of dual images obtained by our method. Second, it can be exciting to seek a more optimal basis than our grid basis in order to reconstruct dual images in higher quality.

9. REFERENCES

[1] Pradeep Sen and Soheil Darabi, "Compressive dual photography," *Comput. Graph. Forum*, vol. 28, no. 2, pp. 609–618, 2009.

[2] Ren Ng, Ravi Ramamoorthi, and Pat Hanrahan, "Triple product wavelet integrals for all-frequency relighting," *ACM Trans. Graph.*, vol. 23, pp. 477–487, August 2004.

[3] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch, "Dual photography," in *ACM SIGGRAPH*, 2005, pp. 745–755.

[4] Yoav Y. Schechner, Shree K. Nayar, and Peter N. Belhumeur, "A Theory of Multiplexed Illumination," in *IEEE International Conference on Computer Vision (ICCV)*, 2003, vol. 2, pp. 808–815.

[5] Pieter Peers, Dhruv K. Mahajan, Bruce Lamond, Abhijeet Ghosh, Wojciech Matusik, Ravi Ramamoorthi, and Paul Debevec, "Compressive light transport sensing," *ACM Trans. Graph.*, vol. 28, no. 1, pp. 1–18, 2009.

[6] Matthew OToole and Kiriakos N. Kutulakos, "Optical computing for fast light transport analysis," in *SIGGRAPH Asia*, 2010.

[7] Xinqi Chu, Tian-Tsong Ng, Ramanpreet Pahwa, Tony Q. Quek, and Thomas Huang, "Compressive inverse light transport," *Proceedings of the British Machine Vision Conference*, 2011.

[8] Richard Baraniuk, "Compressive sensing," *IEEE Signal Processing Mag*, pp. 118–120, 2007.

[9] Wotao Yin, Simon Morgan, Junfeng Yang, and Yin Zhang, "Practical compressive sensing with toeplitz and circulant matrices," *Proceedings of Visual Communications and Image Processing (VCIP)*, 2010.

[10] Tom Goldstein and Stanley Osher, "The split bregman method for l1-regularized problems," *SIAM J. Img. Sci.*, vol. 2, pp. 323–343, April 2009.

[11] Richard Baraniuk et al., "Rice wavelet toolbox," 2002.

[12] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, april 2004.