High Resolution Image Reconstruction for Plenoptic Imaging Systems using System Response

Sapna A. Shroff and Kathrin Berkner
Ricoh Innovations, Inc., 2882 Sand Hill Rd, Suite 115, Menlo Park, CA 94025-7054
sapna@rii.ricoh.com

Abstract: Plenoptic imaging systems with a microlens array imaging the pupil at the sensor typically produce low resolution images. We introduce an inverse problem solver using the system response to obtain high resolution image reconstructions.

OCIS codes: (070.0070) Fourier optics and signal processing, (110.1758) Computational imaging, (100.1830) Deconvolution, (100.3010), Image reconstruction techniques, (100.3190) Inverse problems, (100.6640) Superresolution

1. Introduction

Plenoptic imaging systems are used for various applications like multi-angle imaging [1], depth imaging, refocusing [1], multimodal imaging [2], etc. A lenslet or pinhole array placed in the traditional image plane re-images the pupil of the main lens at the sensor. This configuration allows the cone of rays from a single point in object space to focus at each lenslet and then propagate beyond the plane of the lenslet, spread out and be sampled by multiple pixels, instead of the traditional sampling of a diffraction spot by 2-3 pixels. This data behind each lenslet is called a superpixel. Each superpixel may comprise of many sensor pixels.

In this configuration, different sensor pixels in a superpixel capture information from rays from different angles starting from a single point in object space. This angle-to-space mapping allows the extraction of angular information about the object. Therefore a single plenoptic image can be used to extract multiple images of the same object from different angles, refocus the images using the concept of multi-view or stereo imaging, etc. Every lenslet images the pupil aperture of the main imaging system in each superpixel. Therefore, a filter array placed at the pupil plane is demagnified and imaged at the sensor behind each lenslet. Separating pixels based on their location corresponding to each filter at the pupil, and binning them into one effective pixel per filter per lenslet, allows single-shot multispectral or multimodal images [2]. Such simple image processing produces images with resolution equal to the number of lenslets in x-y directions [3]. Therefore each spectral image (in multispectral imaging), depth plane (in refocusing) or angle image (for angular imaging) has lower resolution than the image acquired in conventional imaging systems where the sensor, instead of the lenslet array, is placed at the image plane.

Here we present a detailed model for the optical response of the plenoptic system based on wave propagation techniques and propose the use of a linear inverse problem solver to recover high resolution information in a plenoptic image, in an effort to approach the resolution of the main (conventional) imaging system at a given focus plane. We show simulation results that demonstrate our technique.

2. Image formation in a plenoptic system

Fourier optics theory for incoherent image formation in such plenoptic imaging systems has not been discussed in literature. But as plenoptic imaging is applied to systems like microscopes, diffraction analysis becomes important. We first discussed a coherent analysis for these systems in [4] and [5]. Here we have broadened the scope of our analysis to include incoherent objects. The optical response of the plenoptic system for an infinitesimal point in object space is an image of the pupil of the main lens. We call this response the Pupil Image Function (PIF) [5], instead of the Point Spread Function (PSF) to avoid confusion with the traditional airy disk pattern at the sensor. The Pupil Image Function is obtained by Fourier analysis,
PIFv_{ζ,η}^{w} = \left[ \frac{w^{j2\pi \lambda z_{1}}}{j\lambda z_{1}} \exp\left[ \frac{jk}{2z_{1}} \left( i^{2} + w^{2} \right) \right] \sum_{n} \sum_{m} \exp\left[ -\frac{jk}{2f_{2}} \left( z_{m}D_{2} \right)^{2} \left( z_{n}D_{2} \right)^{2} \right] \right] \left\{ du dv \right\}_{2} \left( u - mD_{2}, v - nD_{2} \right) \\
exp\left[ \frac{jk}{2} \left( \frac{1}{z_{2}} + \frac{1}{f_{2}} - 1 \right) \right] \left( u^{2} + v^{2} \right) \exp\left[ -jk \left[ u \left( \frac{t - mD_{2}}{z_{2}} \right) + v \left( \frac{w - nD_{2}}{f_{2}} \right) \right] \right] h_{v}^{'} \left( u - \xi^{*}, v - \eta^{*} \right) \right\}^{2} \tag{1}
\]

where \( h_{v}^{'}(u,v) = FT \left( P_{v}(x\lambda z_{v},y\lambda z_{v}) \exp \left\{ jk/2 \left( 1/z_{2} + 1/z_{3} - 1/f_{3} \right) \right\} \left( x\lambda z_{v} \right)^{2} + \left( y\lambda z_{v} \right)^{2} \right\} \) contains the impulse response of the main lens. \( \lambda \) is the imaging wavelength, \( k = 2\pi /\lambda \). As described in Figure 1(a), \( z_{1}, z_{2}, \) and \( z_{3} \) are the distances between the object and main lens, the main lens and lenslet array, and the lenslet array and sensor, respectively. We denote the coordinates of the object plane as \( (\xi,\eta) \), the pupil plane for the main lens as \( (x,y) \), the lenslet array as \( (u,v) \), and the sensor plane as \( (t,w) \). The magnification from object to the primary image plane is \( M = -z_{2}/z_{1}. \) And \( \xi^{*} \) and \( \eta^{*} \) are reduced object space coordinates where, \( \xi^{*} = M \xi \) and \( \eta^{*} = M \eta. \) The focal lengths and diameters of the main lens and the lenslet array are \( f_{1}, f_{2}, D_{1}, \) and \( D_{2} \) respectively. The generalized pupil function for the main lens and an individual lenslet in the array are given by \( P_{1} \) and \( P_{2} \) respectively. \( M \) and \( N \) are the number of lenslets in the array. \( \tilde{M} \) and \( \tilde{N} \) are indices for the lenslets. Figure 1(b)-(e) shows the simulated system response.

If \( I_{w}^{f} \) is a column vector of the sensor data, \( I_{w}^{o} \) is a column vector of the object intensity, \( I_{o,\xi,\eta,\xi',\eta'}^{w} \), and matrix \( PIFv \) contains stacked columns of the optical responses for each object point in a column vector, \( PIFv_{\xi,\eta,\xi',\eta'}^{w} \), then the data captured at the sensor of a plenoptic system can be modeled by a linear system,

\[ \left[ I_{w}^{f} \right] = \left[ PIFv \right] \left[ I_{o} \right] + \text{noise}. \tag{2} \]

We used this analysis to simulate images from a plenoptic system. We assumed a flat 2D object for this simulation. We used an 11 x 11 lenslet array. We assumed a system with focal lengths and diameters of 40mm, 4mm, 4mm, and 0.16mm for the main lens and lenslet array respectively. We have not taken into account any lenslet overflow or cross-talk at the sensor. The images shown in Figure 2 contain no noise or system aberrations. This is consistent with the experimental results shown by [1] and gives us an intuition about focusing such images.

3. High resolution image reconstruction in a plenoptic system

The image formation analysis indicates that the response of a given point in object space is not uniform across the pixels in a superpixel. The form of its distribution across the sensor is a characteristic of the system. The weighting of this distribution is proportional to the object intensity. Therefore, this concept immediately lends itself to the possibility of recovering the original object intensity by solving the inverse problem for the linear system from Eq.(2). Solving this inverse problem is similar to deconvolution, but traditional deconvolution approaches like Wiener filtering cannot be applied directly since the \( PIFv \) matrix in Eq. (2) is not necessarily a convolution matrix. We pose the inversion as computing a solution to the following optimization problem,

\[ \min_{I_{v}} \| PIFv[I_{v}] - I_{f} \|, \text{ subject to constraints on } [I_{v}]. \tag{3} \]

The \( PIFv \) matrix may be obtained by experimental calibration, diffraction analysis or lens design simulations. We used a radially asymmetric amplitude gradient across the pupil of each lenslet to constrain the reconstruction.
Figure 3. (a) Low resolution version of the image obtained by binning the data in a superpixel into one pixel. (b) Reconstructed image for noise-free data, obtained using a pseudoinverse. (c) Reconstructed image for noisy sensor data, SNR = 40db where the reconstruction was obtained using a non-linear iterative least squares algorithm with smoothing.

In the case of refocusing or multi-angle imaging, different pixels within a superpixel are considered to be rays from different angles, emanating from the same point. Here we do not consider refocusing or multi-angle imaging, and only consider the resolution of the object plane that is best in-focus at the lenslet array. This in-focus plane is reconstructed [3] by binning all the data behind a given lenslet to a single pixel. Figure 3(a) shows such a low resolution image that could be obtained from the data captured at the sensor as shown in Figure 2(b). This is the kind of image that could be obtained for a single spectral band in single-shot multispectral imaging with an additional step of filter based segmentation of each superpixel as shown in [2].

In Figure 3(b) we reconstructed the best in-focus image of the object by solving the inversion problem using the pseudoinverse of the matrix $\text{PIFv}$ to compute $\hat{I}_v = [\text{PIFv}]^{-1} [I_v]$ without imposing additional constraints on the object. No noise was added at the sensor for this case. The image shows increased resolution, almost approaching that of the conventional main lens. Next, we added white Gaussian noise at the sensor, SNR = 40 db. Inversion using pseudoinverse did not work very well on noisy data. Therefore, we used an iterative non-linear least squares minimization with smoothing. Other methods such as Tikhonov regularization, etc. may also be used. Figure 3(c) shows our reconstruction which has better resolution than the image shown in Figure 3(a), but retains some low SNR regions between the lenslets. Note that some data in this inter-lenslet spacing is also reconstructed in both Figure 3(b) and (c) because the $\text{PIFv}$ matrix contains the diffraction response, not the geometric optics response.

Note that prior work [6] showed that defocusing the object introduces inter-lenslet aliasing which can be used favorably for high resolution reconstruction. In the results shown here we assumed the object was perfectly focused at the lenslet array and did not incorporate any information from aliasing. But the concept of the $\text{PIFv}$ response matrix can be extended to 3D and our tests with defocus showed that our method can also be used to obtain high resolution reconstruction with and without defocus and, with and without aliasing between lenslets.

4. Conclusions

We propose the use of the response of a plenoptic imaging system to reconstruct a higher resolution image of the object. We show simulations for formation of plenoptic sensor data without noise and with noise (SNR = 40db) and reconstructions of the object which show improved spatial resolution, obtained by solving a linear inverse problem using the system response matrix.

5. References