# An Analysis of Color Demosaicing in Plenoptic Cameras

Zhan Yu     Jingyi Yu
University of Delaware
Newark, DE 19716, USA
{zyu,yu}@cis.udel.edu

Andrew Lumsdaine
Indiana University
Bloomington, IN 47405, USA
lums@osl.iu.edu

Todor Georgiev
Qualcomm
San Diego, CA 92121, USA
todor@tgeorgiev.net

## Abstract

*A plenoptic camera captures the 4D radiance about a scene. Recent practical solutions mount a microlens array on top of a commodity SLR to directly acquire these rays. However, they suffer from low resolution as hundreds of thousands of views need to be captured in a single shot. In this paper, we develop a simple but effective technique for improving the image resolution of the plenoptic camera by maneuvering the demosaicing process. We first show that the traditional solution by demosaicing each individual microlens image and then blending them for view synthesis is suboptimal. In particular, this demosaicing process often suffers from aliasing artifacts, and it damages high frequency information recorded by each microlens image hence degrades the image quality. We instead propose to demosaic the synthesized view at the rendering stage. Specifically, we first transform the radiance to the desired focal plane and then apply frequency domain plenoptic resampling. A full resolution color filtered image is then created by performing a 2D integral projection from the reparameterized radiance. Finally, we conduct demosacing to obtain the color result. We show that our solution can achieve visible resolution enhancement on dynamic refocusing and depth-assisted deep focus rendering.*

## 1. Introduction

Recent advances in computational photography have given rise to previously unexplored effects in imaging. A notable example is the realization of the plenoptic (or "light field") camera [25, 21, 2, 8], a camera that uses a microlens array to capture 4D radiance about a scene. The acquired radiance information can be post-processed for either synthesizing dynamic depth of field effects or for recovering the 3D scene. There are numerous applications for this emerging camera technology, ranging from entertainment (Lytro [22]) to depth recovery for industrial and scientific applications (Raytrix [26]).

The plenoptic camera, in essence, is a single-shot, multi-view acquisition device. In order to overcome the spatio-angular tradeoff, an ultra-high resolution sensor is commonly used. The resulting images, however, are still at a disappointingly low resolution. The Adobe light field camera [8] captures 20 different views of a scene using a 10 megapixel sensor, resulting in rendered images with visible artifacts at a resolution of $700 \times 700$. Ng [25] proposed a different design with a $296 \times 296$ microlens array covering a 16 megapixel sensor. The dense angular resolution greatly suppressed artifacts with higher refocusing power. Nevertheless, the image resolution is low, equal to the number of microlenses in the camera ($296 \times 296$). The recently released Lytro light field camera uses a 11 megapixel sensor for capturing the radiance. Pictures published on Lytro's website still suffer from a low resolution of 0.7 megapixel, with some visible artifacts around thin objects and sharp edges.

In this paper, we develop a simple but effective technique for improving the image resolution of the plenoptic camera by using a more appropriate demosaicing process. A plenoptic camera, like traditional color cameras, captures color information with a Color Filter Array (CFA) masking the sensor pixels. We first show that the traditional solution [25, 21, 2, 8] that demosaics each individual microlens image and then blends them for rendering is suboptimal. In particular, this demosaicing process damages high frequency information recorded by each microlens image, hence greatly degrading the achievable resolution of the final photograph. We instead perform demosaicing on the synthesized color photograph at each refocusing plane Π. Specifically, we first reparameterize the light field to the desired focal plane and then apply frequency-domain plenoptic resampling. A full resolution color filtered image is then created by performing a 2D integral projection from the reparameterized light field. Demosaicing is performed as a last step to obtain the final color result.

Experiments on synthetic and natural scenes show that our approach generates images with higher resolution and fewer artifacts compared with classical plenoptic rendering. We demonstrate the quality enhancements of our results in
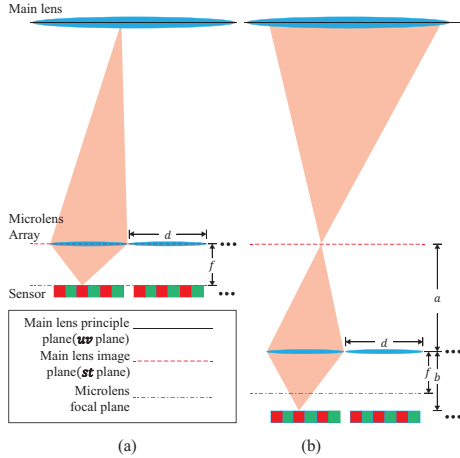
Figure 1. Plenotpic Camera Designs. (a) Ng. (b) Lumsdaine et al.



Figure 2. Projection of the radiance corresponding to focusing at different focal depth with the parameterization of the sampled plane.

several applications such as dynamic refocusing and depth-assisted deep focus rendering.

Without loss of generality, in this paper, a 2D simplification is used for visualization. However, formal mathematical analysis and algorithms are all dealing with 4D space.

## 2. Background

Integral or light field photography has its roots in the methods introduced by Lippmann [19] and Ives [14] over 100 years ago. Recently, it has re-emerged with the introduction of plenoptic cameras. In this section, we briefly review previous work in this field.

### 2.1. Plenoptic Camera Design

Over the last twenty years, numerous integral cameras have been built [2, 11, 13, 16]. However, it was not until recently that Ng [25] improved the traditional plenoptic camera design and introduced new methods for computational refocusing. This plenoptic camera places the microlens array on the image plane $\Pi$ of the main lens to separate the converging rays onto the sensor behind it (Figure 1(a)). The sensor is located at the focal plane of each microlens so that each microlens is focusing at its optical infinity (main lens principal plane). The F-numbers of the main lens and each microlens are matched to avoid "Cross-Talk" among microlens images. A version of this camera is available from Lytro [22].

Lumsdaine et al. [21] introduced another design by focusing the microlens array on $\Pi$ and correspondingly adjusting the position of the microlens array and the sensor (Figure 1(b)). In this case each microlens image will have samples with more spatial resolution and less angular resolution on $\Pi$. Therefore this design is capable of producing
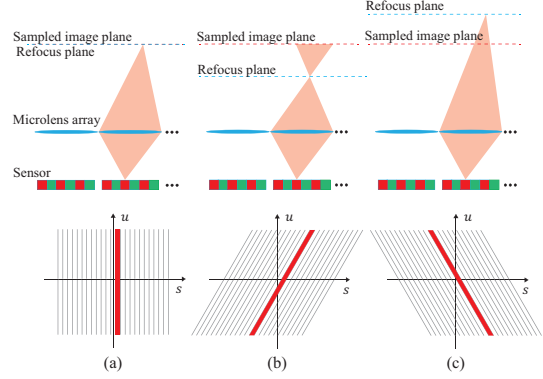
higher resolution results when focusing near the sampled image plane. However, the lower angular resolution may cause ringing artifacts in out of focus regions of the rendered image.

### 2.2. Refocusing as Radiance Transform

In classical plenoptic or light field rendering, a two parallel plane parameterization is commonly used to represent rays, where each ray is parameterized with the coordinates of the mainlens principal plane $\Pi_{uv}$ and the mainlens focal plane $\Pi_{st}$. For simplicity, we use two dimensional vectors $\boldsymbol{s}$ and $\boldsymbol{u}$ to represent positions $(s, t)$ and $(u, v)$ on $\Pi_{st}$ and $\Pi_{uv}$ respectively. The irradiance $I$ at position $\boldsymbol{s}$ can then be computed as:

$$I(\boldsymbol{s}) = \frac{1}{R^2} \int r(\boldsymbol{s}, \boldsymbol{u}) cos^4 \Phi \mathrm{d}\boldsymbol{u}, \qquad (1)$$

where $R$ is the distance between the $\Pi_{st}$ and $\Pi_{uv}$, and $cos^4\Phi$ is *optical vignetting* term. Same as [25], we combine $cos^4\Phi$ into $r(\boldsymbol{s}, \boldsymbol{u})$ for further simplification. Equation 1 shows that any photograph produced by a conventional camera is the result of a 2D integral projection of the angular dimensions of the radiance onto the spatial dimensions, where the slope of the projection relies on the focal depth of the camera. The reparameterization approach of [13] allows us to refocus at different scene depths by transforming the radiance, as shown in Figure 2:

$$I'(\mathbf{s}) = \frac{1}{R'^2} \int r'(\boldsymbol{s}, \boldsymbol{u}) \mathrm{d}\boldsymbol{u}, \qquad (2)$$

where

$$r'(\boldsymbol{s}, \boldsymbol{u}) = r(\boldsymbol{u} + (\boldsymbol{s} - \boldsymbol{u})\frac{R}{R'}, \boldsymbol{u}), \qquad (3)$$

$I'$ is the irradiance at a new focal depth, and $R'$ is the distance between $\Pi_{uv}$ and the target focal plane $\Pi_{s't'}$.

## 2.3. Image Demosaicing

While a significant amount of work on plenoptic cameras has been focusing on improving the image resolution [27, 5, 10], demosaicing remains as an understudied problem. Demosaicing, in essence, converts single-CCD color representations of one color channel per-pixel into full per-pixel RGB. The most popular type of CFA in current use is the Bayer filter [4]. Demosaicing a raw Bayer image requires an underlying image model to guide decisions for reconstructing the missing color channels: at every pixel only one color channel is sampled and therefore we need to use its nearby samples to reconstruct the other two channels. Many sequential methods [17, 15, 1, 12, 24] have been introduced based on the assumption that green channel is less aliased than the other two due to higher sampling frequency. More sophisticated methods impose local gradients [20] or frequency statistics [18, 3, 6, 23] as constraints to improve the performance.

However, by far nearly all demosaicing techniques aim to process images captured by commodity digital cameras and very little work has been focused on developing solutions specifically for plenoptic cameras. Existing plenoptic cameras typically demosaic each individual microlens image and treat the captured plenoptic function as a captured RGB image. One exception is the paper by Georgiev et al. [9] that applies demosaicing after plenoptic rendering to improve plenoptic superresolution. The approach presented in [9] used a straightforward demosaicing that did not resample the lightfield, resulting in significant color artifacts in out-of-focus regions of the rendered images. Other related work is the spatial domain multi-frame demosaicing and super-resolution technique reported in [7]. However, their focus is to combine multiple low resolution images whereas we aim to manipulate demosaicing to improve refocused images produced by plenoptic rendering.

## 3. Image Demosaicing in a Plenoptic Camera

Before proceeding with our analysis, we introduce our notation. Let $I(\boldsymbol{s})$ represent the irradiance of pixel $s$ on the image plane $\Pi_{s't'}$ and $r_i$ represent the RGB radiance of a sample ray in microlens $m_i$. $I_i$ is the *ideal* optical RGB image at $m_i$. In reality, since a color filter is used to separate the colors, we instead consider a color filtered image $I_{fi}$. For each color channel, $I_{fi}$ can be viewed as an undersampled version of $I_i$ in that channel. The demosaicing operator $D$ upsamples $I_{fi}$ to recover $I_i$.

### 3.1. Classical Rendering

The classical plenoptic rendering approach first applies demosaicing to each individual microlens image and then applies integral projection as given in Equation 2 for focused image formation. Let $b$ denote the distance from the
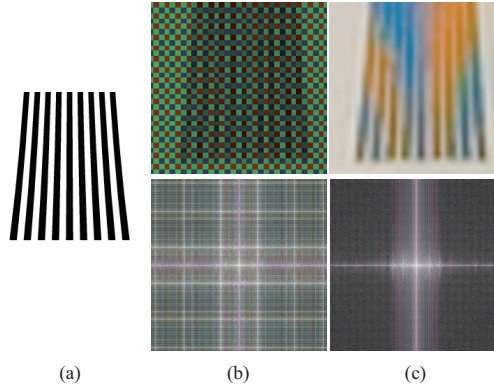


Figure 3. Artifacts on the captured radiance introduced by classical demosaicing. (a) Ground Truth. (b) Raw microlens image and its frequency spectrum. (c) Demosaiced microlens image and its frequency spectrum.

sensor to the microlens array and $\boldsymbol{s}_i$ denote the location of the optical center of $m_i$. In the discrete case, if we focus at $\Pi_{s't'}$ with distance $a$ to the microlens array, we can rewrite the irradiance of Equation 2 as:

$$I'(\boldsymbol{s}) \approx \sum_i D(I_{fi}((\boldsymbol{s}_i - s)\frac{b}{a} + \boldsymbol{s}_i), \qquad (4)$$

Let $\omega_i$ denote the highest frequency of $I_i$ and $\omega$ denote the sampling frequency of $I_{fi}$. In the trivial case $(\forall i)[2\omega_i \leq \omega]$, we can completely recover the full frequency microlens images $I_i$ and hence the refocused image $I'$. In the general case when $(\exists j)[2\omega_j > \omega]$, the spectrum of $I_{fj}$ exhibits aliasing due to undersampling as shown in Figure 3(b). In this case, the demosaic operator $D$ is used to eliminate undersampling artifacts. However, $D$ generally behaves as a low pass filter, indiscriminately removing high frequencies, thereby degrading the image sharpness of the final refocused image. Finally, if $I_{fi}$ is severely undersampled, demosaicing (such as that performed by Adobe Photoshop Camera Raw) can introduce inconsistent color interpolation and cause color blending in the refocused image as shown in Figure 3(c) (black and white to colorful).

### 3.2. Resolution on the Refocus Plane

Unlike the classical approach, which directly applies demosaicing to the microlens images, we first project $I_{fi}$ onto the focused plane $\Pi_{s't'}$ and then perform demosaicing.

In this section, we provide a theoretical analysis to show that the projected image $I_f$ on plane $\Pi_{s't'}$ has a higher sampling frequency than any of the microlens images, hence performing demosaicing on $I_f$ could greatly improve the image resolution. For simplicity, we model each microlens as a pinhole camera and only analyze rays passing through each optical center. Also for simplicity, and without loss of
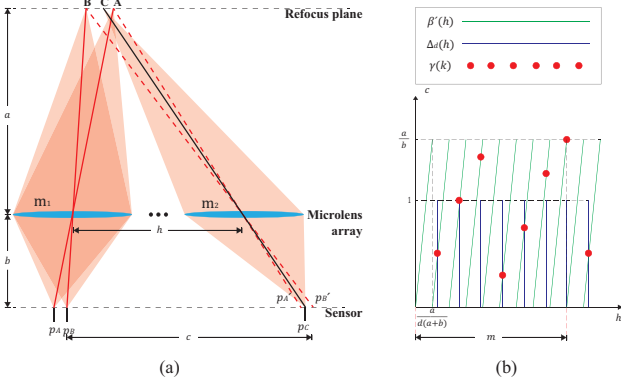
Figure 4. (a) Possible resolution enhancement on the refocus plane by projecting multiple microlens images. (b) Plots of function $\Delta_d(h)$, $\beta(h)$, and $\gamma(k)$.

generality, we show only one spatial dimension $s$. Consider two adjacent pixels $p_A$ and $p_B$ ($p_A < p_B$) in a specific microlens $m_1$ that map to two points $A$ and $B$ on the target focal plane $\Pi_{s't'}$. Assume the distance between $p_A$ and $p_B$ is 1, the distance between two adjacent microlenses is $d$, $\Pi_{s't'}$ lies at distance $a$ to the microlens array, the sensor lies at distance $b$ to the microlens array, and the spacing between $m_1$ and $m_2$ is $h$, as shown in Figure 4(a). Note that since the pixel distance is vanishingly small compared with $a$, $b$, and $d$, we simply treat these latter quantities as integers.

Our goal is to study how many rays (pixels) from other microlenses would fall between $A$ and $B$ on $\Pi_{s't'}$. This number approximates the factor of resolution enhancement compared with the classical demosaicing followed by rendering approach. In order to approximate this number, we first introduce a function $\gamma$ which maps the index of a given microlens to its sampling point between $A$ and $B$. Since all the microlenses out of the minimum period $T$ of $\gamma$ are duplications of samples within $T$, we find out $T$ of $\gamma$ and use it as the upper bound of the resolution enhancement.

Note that for each microlens $m_2$ different from $m_1$, we can have at most 1 point between $AB$ that maps to a pixel to $m_2$ as the length $AB$ is preserved in all microlenses. Assume $A$ and $B$ map to points $p'_A$ and $p'_B$ in $m_2$, as shown in Figure 4(a). Note that $p'_A$ and $p'_B$ may not be pixels. In the first case, $p'_A$ and $p'_B$ fall exactly on the pixels position. In that case, no additional rays (pixels) from $m_2$ would intersect the segment $AB$ on plane $\Pi_{s't'}$. Therefore, $m_2$ would not contribute to enhancing the resolution between $AB$. Under similitude relationship, the conclusion holds for any pair of adjacent pixels in $m_1$ and $m_2$, i.e., $m_2$ would not contribute to enhancing the resolution to $m_1$'s image.

In the second case, $A$ and $B$ do *not* coincide with pixels in $m_k$ and there is exactly one point $C$ between $A$ and $B$ that maps to a pixel in $p_C$ in $m_k$. We call $C$ a super-pixel

as it will increase the resolution between $AB$. We can then compute $p_C = \frac{a+b}{a}h$ and the distance $\beta$ between $A$ and $C$ on the focal plane as:

$$\beta(h) = (\frac{a+b}{a}h - \lfloor \frac{a+b}{a}h \rfloor)\frac{a}{b}. \qquad (5)$$

Note that function $\beta(h)$ is a periodic function with a minimum period of $\frac{a}{a+b} < 1$. For each microlens, we can substitute its distance $h$ into $m_1$ and compute the location of this super-pixel. If the super-pixels in some $N$ microlenses have identical $\beta$ values, then these microlenses only contribute 1 rather $N$ super-pixels for enhancing the resolution between $AB$.

To finally compute the exact resolution enhancement, recall that in the microlens array setting, $h = kd$ from $m_1$, where $k$ is some positive integer and $d > 1$. We can then concatenate the microlens sampling function (a Dirac comb) $\Delta_d(h)$ with the distance function $\beta(h)$ as: $\Delta_d(h) \cdot \beta(h)$. To further simplify, we can factor $d$ into $\gamma(h)$ so that $\gamma(k) = \Delta(k) \cdot \beta'(k)$. where $\beta'(k)$ has period $\frac{a}{d(a+b)}$ and $\Delta(k)$ has period 1.

Clearly $\gamma(k)$ has minimum integer period equal to the least common integer multiple of $\frac{a}{d(a+b)}$ and 1. We rewrite $\frac{a}{d(a+b)}$ as an irreducible fraction two integers $\frac{m}{n}$. Thus, $S'(k)$ has minimum integer period $m = \frac{a}{gcd(a,d(a+b))}$, where $gcd$ denotes the greatest common divisor operator (Figure 4(b)).

Note that the number of microlenses sharing a field of view also constrains the number of distinct samples between $p_A$ and $p_B$. Since the shift from one microlens image to another for any point $p$ on $\Pi_{s't'}$ is $\Delta = d\frac{b}{a}$, we can compute the number of microlens covering $p$ as:

$$n_p = \lfloor \frac{d}{\Delta} \rfloor = \lfloor \frac{a-f}{f} \rfloor. \qquad (6)$$

Combining with Equation 5 we obtain that the resolution enhancement factor from microlens image $m_i$ to $\Pi_{s't'}$ is equal to $min(m, n_p)$. Since $b = \frac{af}{a-f}$, this factor is controlled only by $a$ and $f$, namely, the depth of the scene and the camera optics.

### 3.3. Our Approach

Projecting samples of each microlens to the refocus plane $\Pi_{s't'}$ gives us a higher resolution image $I_f$. However, as shown in Figure 5(a), when the captured radiance is transformed to $\Pi_{s't'}$ for projection, as proposed by Georgiev et al. [9], the spacing of each color component is not uniform on $I_f$, resulting in random RGB patterns (Figure 6(a)). This issue creates trouble for demosaicing $I_f$. Therefore a crucial step of our approach is to resample the radiance with the parameterization of $\Pi_{s't'}$ to achieve constant spacing on each dimension (Figure 5(b)).
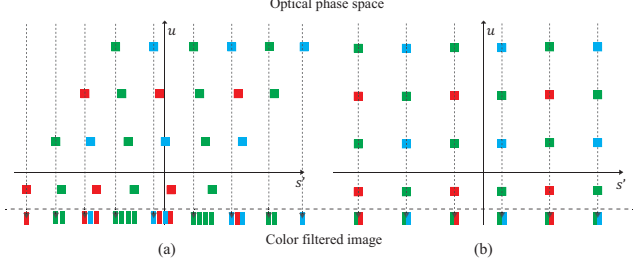
Figure 5. Optical phase space illustration of resampling the captured radiance. (a) Directly projecting the captured radiance onto the refocus plane. (b) Projecting the resampled radiance onto the refocus plane.

**Resampling** We adopt a similar approach to that in [28], which was originally developed for multi-frame single channel image restoration. We use a frequency-domain approach to resample the 4D color filtered radiance. This simplifies to reconstructing a higher resolution color image by perfect registration with an array of low resolution color images taken at the same time in a 2D image restoration case.

Here we only consider the green rays. The other two channels can be computed in a similar manner. Suppose we have $q$ microlenses. Each microlens captures a low resolution radiance with $N_s$ and $N_u$ samples on each dimension. Let $r_o(s, u)$ be the original green rays parameterized by $\Pi_{st}$ and $\Pi_{uv}$. Given the distance $a$ from $\Pi_{st}$ to the microlens array, the registration of a recorded sub-radiance $r_i$ can be computed accurately as offsets $\sigma_s, \sigma_u$ on each dimension respectively. Therefore, the sampled rays by microlens $m_i$ is $r_i(s, u) = r_o(s + \sigma_{si}, u + \sigma_{ui})$. In frequency domain, this yields:

$$R_i(\boldsymbol{S}, \boldsymbol{U}) = e^{j2\pi(\sigma_{si}\boldsymbol{S} + \sigma_{ui}\boldsymbol{U})} R_o(\boldsymbol{S}, \boldsymbol{U}) \qquad (7)$$

where $R_o(\boldsymbol{S}, \boldsymbol{U})$ and $R_i(\boldsymbol{S}, \boldsymbol{U})$ are CFT of $r_o(s, u)$ and $r_i(s, u)$ respectively. Let pixels under $m_i$ capture $r_i$ with a uniform spacing $(T_s, T_u)$, and $R_{d_i}(\Omega)$ be the discrete Fourier transform (DFT) of the rays recorded by $i^{th}$ microlens at frequency $\Omega = (\omega_s, \omega_u)$. From the aliasing relationship between CFT and DFT, $R_{d_i}(\Omega)$ satisfies the following equation:

$$R_{d_i}(\Omega) = K \sum_{m_s} \sum_{m_u} (R_i(\frac{\omega_s}{N_s T_s} + m_s f_s, \frac{\omega_u}{N_u T_u} + m_u f_u)), \qquad (8)$$

where $K = \frac{1}{T_s T_u}$, and $f_s$, $f_u$ are sampling frequencies on each dimension of all micro images. All $\sum$ operators range from $-\infty$ to $\infty$ and $m_s, m_u$ are integers. Substituting $R_i$ from Equation 7 to Equation 8 yields:

$$V_\Omega = M_\Omega R_\Omega, \qquad (9)$$

where $V_\Omega$ is a $q$ dimensional column vector with $i^{th}$ element equal to $R_{d_i}(\Omega)$; Let $B_S, B_U$ be periodic boundaries of $R_o$ such that $R_o(\boldsymbol{S}, \boldsymbol{U}) = 0$ for any condition of
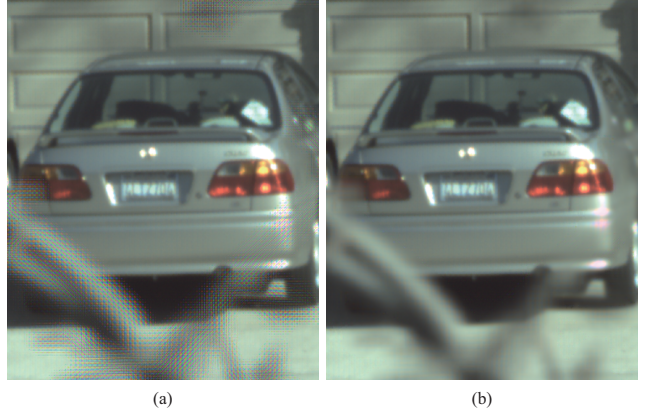


Figure 6. Rendered results using (a) the approach proposed by Georgiev et al. [9] and (b) our approach. The out of focus foreground objects exhibit RGB patterns in (a) due to non-uniform spacing of color components after integral projection.

$|\boldsymbol{S}| > B_S f_s, |\boldsymbol{U}| > B_U f_u$ satisfies; $R_\Omega$ is a $4B_S B_U$ dimensional column vector with the $k^{th}$ element $R_o(\frac{\omega_s}{N_s T_s} + \gamma_s f_s, \frac{\omega_u}{N_u T_u} + \gamma_u f_u)$, and $\gamma_s = k \bmod(2B_S) - B_S, \gamma_u = \lfloor \frac{k}{2B_S} \rfloor - B_U$, and $M_\Omega$ is a $q \times 4B_S B_U$ matrix with $(i, k)^{th}$ element

$$\frac{1}{T_s T_u} \exp\{j2\pi[\sigma_{si}(\frac{\omega_s}{N_s T_s} + \gamma_s f_s) + \sigma_{ui}(\frac{\omega_u}{N_u T_u} + \gamma_u f_u)]\}.$$

Since we know the locations of $\Pi_{st}$ and of each microlens $m_i$, $\sigma_{si}$ and $\sigma_{ui}$ can be accurately computed. $R_{d_i}(\Omega)$ can be acquired by performing the 4D DFT on the sampled radiance by each microlens. Therefore Equation 9 is solvable for unknown $R_\Omega$, which contains $2B_S$ and $2B_U$ frequency samples of $R_o(\Omega)$ on each dimension respectively. Combining all $R_\Omega$ provides an estimate of $R_o$ with $2N_s B_S, 2N_u B_U$ samples ranging from $(-B_S f_s, -B_U f_u)$ to $(B_S f_s$ and $B_U f_u)$ with spacing $(\frac{1}{N_s T_s}, \frac{1}{N_u T_u})$ on each dimension respectively. We then use it to estimate $r_o(s, u)$ from $(0, 0)$ to $((N_s - 1)T_s, (N_u - 1)T_u)$, with spacing $(\frac{T_s}{2B_S}, \frac{T_u}{2B_U})$. Hence the resolution of the resampled radiance is increased by $2B_S, 2B_U$ on $s$ and $u$ compared with that of each original microlens image. An optimized process of solving for $R_\Omega$ is presented in Appendix A.

**Integral Projection and Demosaicing** As shown in Figure 5 (b), with the previous resampling process, we can achieve an evenly-sampled radiance on the target focal plane $\Pi_{s't'}$. The integral projection is immediately applied to get $I_f$. An example of the green channel of $I_f$ is shown by Figure 7(c). However, due to the higher sampling rate of the green channel, a demosaicing process is still needed for red and blue channels of $I_f$ to render a full RGB image with the resolution of the green channel.
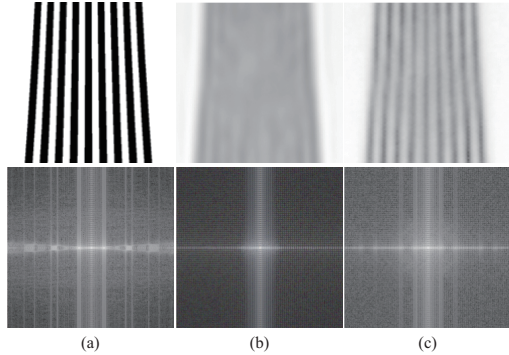
Figure 7. From (a)-(c), we compare the ground truth, the result using classical approach, and the result using our approach. The frequency spectrums are shown in the bottom row.



Figure 8. Our plenoptic demosaicing and rending pipeline.

Traditional sequential demosaicing frameworks first recover a full resolution green channel and subsequently use that green channel to facilitate the recovery of red and blue channels. In our case, the full resolution green channel is already known after the integral projection. Based on this green channel, the red and blue channels are reconstructed by applying the state-of-the-art anisotropic adaptive filtering [18] in the frequency domain. Figure 6(b) shows that by employing the resampling scheme, the demosaicing can be performed on the integral projection result and the final image is free of RGB patterns.

Suppose the resampled radiance has highest frequency $\omega'$. The most common situation is $(\exists i)[\omega' > 2\omega_i > \omega]$. In this case the new demosaicing process preserves more high frequency information of the radiance, hence producing a higher resolution image (Figure 7). In other cases such as $(\forall i)[\omega' > \omega > 2\omega_i]$ (very smooth regions such as places with constant color), both processes recover the full radiance and the resolution of the resultant images are the same. If $(\forall i)[2\omega_i > \omega' > \omega]$ (texture rich regions or sharp edges), the final images are both over-smoothed.

As illustrated by column (a) and (b) of Figure 7, with the classical approach, significant losses in high frequency components occur in texture-rich regions and the rendered result suffers from over-smoothing compared with the ground truth. Column (c) shows our method preserves much more high frequency information of the ground truth, therefore capable of producing a higher resolution image.

## 4. Implementation and Applications

Figure 8 shows the pipeline for implementing our proposed plenoptic demosaicing and rendering scheme. We first resample the radiance, then integral project it onto the spatial domain, and finally demosaic the color filtered result.

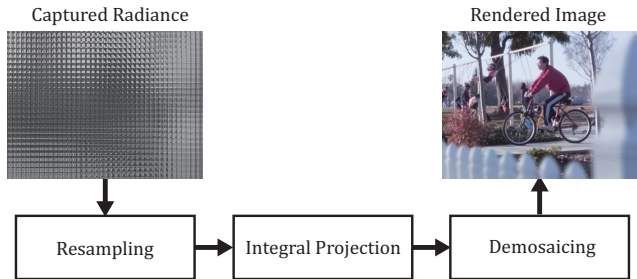Our experimental data is captured by a plenoptic camera similar to that described in [21]. We use a 39-megapixel sensor with pixel size 6.8 $\mu$m. The main lens is mounted on the camera with a 13mm extension tube, which provides the needed spacing to establish an appropriate distance from the main lens focal plane to the microlens array. The focal length of the main lens and of each microlens are 80mm and $1500\mu$m respectively. The microlens pitch is 500 $\mu$m, which makes it work with the F-number of the main lens. The distances between microlenses are 74 pixels.

### 4.1. Enhanced Dynamic Refocusing

We first test our resolution enhancement performance by synthesizing photographs with a shallow depth of field. Figure 9 shows the comparison of our approach (b) and classical rendering (a) on a resolution chart scene. The bottom rows of (a) and (b) compare the demosaiced and raw microlens images of three highlighted regions. Note that severe aliasing effects appear on each raw microlens image and the structure of the resolution chart is not visible. If demosaicing is performed directly on each microlens image, colorful artifacts are introduced, damaging the high frequency information and over-smoothing microlens images. As a result, these regions could not be successfully reconstructed in the final image, as shown in (a). On the contrary, our approach utilizes each aliased microlens image to resample a high resolution radiance before demosaicing is performed. Thus preserving a larger portion of high frequency information and producing a higher resolution image, as shown in (b). Also note that low frequency regions such as the left bottom part of the chart are equally clear in both cases, and very high frequency regions such as the bottom of the red highlighted region are both blurry.

The top row of Figure 10 shows an outdoor scene. Apparently, the numbers on the licence plate in (b) are not visible but readable in (c). Another visible artifact of the classical framework here is that small regions of specular highlight appear less shiny due to over-smoothing on each microlens image.

In another real scene shown in the second row of Figure 10. In column (b), the first line of characters are barely
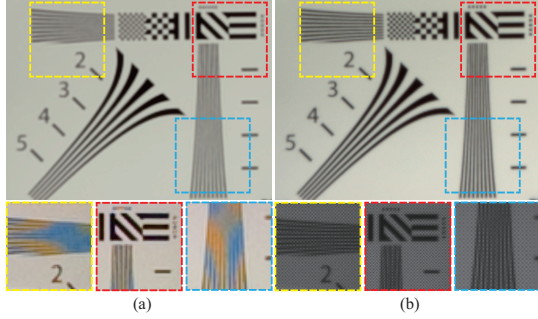
Figure 9. Comparison of rendered image employing classical approach and our approach. (a) Classical approach. Top row: Rendered image. Bottom Row: Demosaiced microlens image. (b) Our approach. Top row: Rendered image. Bottom row: Raw microlens image.

readable using the classical rendering. Nevertheless, they are clearly rendered with our approach. Note that colorful artifacts introduced by demosaicing each microlens image remain on positions of "nf" and "ffi" in (b) and ringing artifacts also appear around the edges of the characters. Furthermore, the lower characters are totally blurry in (b) while still readable in (c).

### 4.2. Extended Depth of Field

Another popular application of our method is the extended depth of field photography. Our approach precomputes the depth of the sampled radiance and renders each pixel by choosing its own depth among samples automatically. We present our depth estimation algorithm designed for the captured radiance in Appendix B.

The third row of Figure 10 shows our extended depth of field application on the same data as the second row. Note that the original out of focus regions such as the face and hair of the person are brought into focus, as if the photograph is captured by a pinhole aperture camera. However, with our framework, shown in (c), the rendered result preserves more high frequency information than the classical approach shown in (b), therefore produces a much more detailed look.

### 5. Discussions and Limitations

We have presented a well-principled plenoptic demosaicing and rendering framework, which preserves more high frequency information from the captured radiance and generate less aliasing artifacts compared with the classical approach.

Our framework does not apply demosaicing directly to the image captured by the plenoptic camera. Instead, with a resampling scheme which helps achieve constant spacing on each dimension, it dynamically performs demosaicing after integral projection. Extensive experiments show that

this framework could produce photographs with commercially acceptable resolution.

As analyzed in Section 3.2, the resolution enhancement of each plane in the scene achieved by our algorithm varies according to the depth of the plane. This could cause unpleasant results if the resolution enhancements are low on planes of interests. In the extreme case, the resolution could be as low as the classical framework. Like classical plenoptic photography, our approach assumes the captured radiance are thin rays in order to reconstruct a refocused image. This is also our assumption for theoretical resolution enhancement analysis.

### 6. Acknowledgments

### References

[1] J. Adams and J. H. Jr. Adaptive color plan interpolation in single sensor color electronic camera, patent us 5506619, 1996. 3

[2] E. Adelson and J. Wang. Single lens stereo with a plenoptic camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:99–106, 1992. 1, 2

[3] D. Alleysson, S. Susstrunk, and J. Herault. Linear demosaicing inspired by the human visual system. *IEEE Transactions on Image Processing*, 14(4):439–449, apr. 2005. 3

[4] B. E. Bayer. Color imaging array. US Patent 3,971,065, 1976. 3

[5] T. E. Bishop, S. Zanetti, and P. Favaro. Light field superresolution. In *IEEE ICCP*, 2009. 3

[6] E. Dubois. Filter design for adaptive frequency-domain bayer demosaicking. In *IEEE International Conference on Image Processing*, oct. 2006. 3

[7] S. Farsiu, M. Elad, and P. Milanfar. Multiframe demosaicing and super-resolution of color images. *IEEE Transactions on Image Processing*, 15(1), jan. 2006. 3

[8] T. Georgeiv, K. C. Zheng, B. Curless, D. Salesin, S. Nayar, and C. Intwala. Spatio-angular resolution tradeoff in integral photography. In *Proceedings of Eurographics Symposium on Rendering*, pages 263–272, 2006. 1

[9] T. Georgiev, G. Chunev, and A. Lumsdaine. Superresolution with the focused plenoptic camera. In *Proc. SPIE 7873*, 2011. doi:10.1117/12.872666. 3, 4, 5

[10] T. Georgiev and A. Lumsdaine. Focused plenoptic camera and rendering. *Journal of Electronic Imaging*, 19, 2010. 3

[11] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen. The lumigraph. In *Proceedings of ACM SIGGRAPH*, pages 43–54, 1996. 2

Figure 10. Comparison of three results with classical approach and our approach. First and second row show shallow depth of field rendering. The third row shows extended depth of field rendering. (a) Our rendered result. (b) and (c) are enlarged highlighted regions in (a) with classical approach and our approach respectively.

[12] R. Hibbard. Apparatus and method for adaptively interpolating a full color image utilizing luminance gradients, patent us 5506619, 1995. 3

[13] A. Isaksen, L. McMillan, and S. Gortler. Dynamically reparameterized light fields. In *Proceedings of ACM SIGGRAPH*, pages 297–306, 2000. 2

[14] F. Ives. Parallax stereogram and process of making same, patent us 725567, 1903. 2

[15] R. Kakarala and Z. Baharav. Adaptive demosaicing with the principal vector method. *IEEE Transactions on Consumer Electronics*, 48(4):932–937, nov. 2002. 3

[16] M. Levoy and P. Hanrahan. Light field rendering. In *Proceedings of ACM SIGGRAPH*, pages 31–42, 1996. 2

[17] X. Li and M. Orchard. New edge-directed interpolation. *IEEE Transactions on Image Processing*, 10:1521–1527, 2001. 3

[18] N. Lian, L. Chang, Y. Tan, and V. Zagorodnov. Adaptive filtering for color filter array demosaicking. *IEEE Transactions on Image Processing*, 16(10):2515 –2525, oct. 2007. 3, 6

[19] G. Lippmann. La photographie intgrale. *Comptes-Rendus, Acadmie des Sciences*, 146:446–451, 1908. 2

[20] W. Lu and Y. Tan. Color filter array demosaicking: new method and performance measures. *IEEE T-IP*, 12:1194–1210, oct. 2003. 3

[21] A. Lumsdaine and T. Georgiev. The focused plenoptic camera. In *In Proc. IEEE ICCP*, 2009. 1, 2, 6

[22] Lytro. www.lytro.com. 1, 2

[23] D. Menon and G. Calvagno. Demosaicing based on wavelet analysis of the luminance component. In *IEEE International Conference on Image Processing*, volume 2, pages 181–184, 2007. 3

[24] D. Muresan and T. Parks. Demosaicing using optimal recovery. *IEEE Transactions on Image Processing*, 14(2), feb. 2005. 3

[25] R. Ng, M. Levoy, M. Brdif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Stanford University Computer Science Tech Report*, 2(2005-02):1–11. 1, 2

[26] Raytrix. www.raytrix.com. 1

[27] J. Stewart, J. Yu, S. J. Gortler, and L. McMillan. A new reconstruction filter for undersampled light fields. EGSR '03, pages 150–156, 2003. 3

[28] R. Tsai and T. Huang. Multi-frame image restoration and registration. *Advances in Computer Vision and Image Processing*, 1, 1984. 5

# Appendices

## A. Solving $R_\Omega$

If we separate $M_\Omega$ into components related and unrelated to $\Omega$, it can be further decomposed into $D_\Omega M_v$, where the $D_\Omega$ is a diagonal matrix with $i^{th}$ diagonal element equals to:

$$\frac{1}{T_s T_u}\exp\{j2\pi[\sigma_{si}(\frac{\omega_s}{N_s T_s} - B_S f_s) + \sigma_{tu}(\frac{\omega_u}{N_u T_u} - B_U f_u)]\},$$

and $M_v$ is a $p \times 4B_S B_U$ matrix with $[M_v]_{ik}$ equals to:

$$\exp\{j2\pi[\sigma_{si}k\mathrm{mod}(2B_S) + \sigma_{ui}\lfloor\frac{k}{B_S}\rfloor]\}.$$

therefore we can rewrite Equation 9 as:

$$M_v R_\Omega = D_\Omega^{-1} V_\Omega. \tag{10}$$

Since $M_v$ is independent of $\Omega$, therefore we only need to solve it once for all sampled frequencies. The rows of $M_v$ are linearly independent if the shifts on dimensions satisfies $\sigma_{si} \neq \sigma_{sj} + n_s T_s$, $\sigma_{ui} \neq \sigma_{uj} + n_u T_u$, where $i \neq j$ and $n_s, n_u$ are some integers. If $q \geq 4B_S B_U$, Equation 10 is consistent. Otherwise, $M_v$ is singular and we obtain the minimum-norm, least-square-error estimate of $R_\Omega$ by first computing the generalized inverse of $M_v$. This yields:

$$R_\Omega = (M_v^T M_v)^{-1}(M_v^T M_v)D_\Omega^{-1}V_\Omega. \tag{11}$$

Note that for each micro image, the sampling frequency on $t, v$ dimensions out of $s, u$ is doubled for the green channel, therefore the resampled radiance has a two times higher resolution on $t, v$ dimensions of the green channel.

## B. Depth Estimation on the Radiance

Recall that each microlens image captures part of the scene from a different view point. Therefore, a multi-view depth estimation is feasible to retrieve the depth information from the entire captured radiance. Out algorithm iteratively computes the depth of the pixel based on Graph Cuts. Given a certain depth, we utilize the variance of the corresponding pixels among microlens images as the data term. Since our framework does not perform demosacing on the raw data. Therefore, when computing the data term for each pixel, only the pixels with the same color channel in different microlens images are used.

One common problem with depth estimation is that even though the correct depth is assigned to one pixel, the data term could still be large due to occlusion. This problem is especially severe in the radiance case since many views are involved in the computation. To resolve the this issue, a gradually increasing confidence threshold is given in each iteration, so that the pixels with lower depth will converge first and the pixels with larger depth may still generate data terms bigger than the threshold due to occlusion. Once the depth of one pixel is decided, it will not be involved in the computation anymore. Thus, the pixels with large depth will avoid these low depth pixels and produce a low data term if assigned the correct depth in the later iterations. The complete algorithm is given in Algorithm 1.

---

**Algorithm 1** Calculate depth of each pixel $d_i$

---

**Require:** Captured radiance $r$
  **for** $j = 1 \rightarrow M$ **do**
    **for** $i = 1 \rightarrow N$ **do**
      **if** $p_i$ has been assigned with any depth value **then**
        Continue;
      **end if**
      Compute data term $e_d$ of pixel $p_i$;
      **if** $e_d > k_j$ **then**
        {Data term is bigger than current threshold}
        Continue;
      **else**
        Add $p_i$ to graph
      **end if**
    **end for**
    Perform graph cuts;
    Assign depths to pixels in the graph;
  **end for**

---