Glare Aware Photography: 4D Ray Sampling for Reducing Glare Effects of Camera Lenses

Ramesh Raskar* Amit Agrawal† Cyrus A. Wilson‡ Ashok Veeraraghavan§
Mitsubishi Electric Research Labs (MERL), Cambridge, MA University of Maryland, College Park, MD

Figure 1: We extract glare components from a *single-exposure* photo in this high dynamic range scene. Using a 4D analysis of glare inside the camera, we can emphasize or reduce glare. The photo in the middle shows a person standing against a sunlit window. We extract reflection glare generated inside lens and manipulate it to synthesize the result shown on the left. On the right we show the glare-reduced component. Notice that the face is now visible with improved contrast.

Abstract

Glare arises due to multiple scattering of light inside the camera's body and lens optics and reduces image contrast. While previous approaches have analyzed glare in 2D image space, we show that glare is inherently a 4D ray-space phenomenon. By statistically analyzing the ray-space inside a camera, we can classify and remove glare artifacts. In ray-space, glare behaves as high frequency noise and can be reduced by outlier rejection. While such analysis can be performed by capturing the light field inside the camera, it results in the loss of spatial resolution. Unlike light field cameras, we do not need to reversibly encode the spatial structure of the ray-space, leading to simpler designs. We explore masks for uniform and non-uniform ray sampling and show a practical solution to analyze the 4D statistics without significantly compromising image resolution. Although diffuse scattering of the lens introduces 4D low-frequency glare, we can produce useful solutions in a variety of common scenarios. Our approach handles photography looking into the sun and photos taken without a hood, removes the effect of lens smudges and reduces loss of contrast due to camera body reflections. We show various applications in contrast enhancement and glare manipulation.

CR Categories: I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Radiometry

Keywords: Glare, Flare, Light Fields, Computational Photography, Masks

1 Introduction

A scene with a bright light source in or near the field of view is difficult to photograph. Glare reduces contrast and causes image fog and large area ghosts. Glare is unavoidable: it disrupts every optical system, including the human eye. Glare can be broadly classified as due to reflection (Fresnel reflection at lens surfaces) and scattering (diffusion in lenses). However, the two are often indistinguishable in the captured 2D photo. In this paper, we analyze glare formation in 4D ray-space. Although glare appears as an additive low-frequency bias in 2D, we show that a significant part of the glare is high frequency noise in 4D ray-space. The key result is that we can remove reflection glare by means of outlier rejection in ray-space and also re-synthesize novel glare effects by manipulating those components.

Our approach involves a minor, inexpensive modification to the camera (Figure 2). We insert a high frequency mask near the camera sensor to act as a sieve that separate spurious rays in ray-space. As we do not explicitly build a light field camera that reversibly encodes the ray-space, the proposed modification leads to simple design choices and requires less precision and minimal calibration. For example, we show glare reduction by non-uniform ray sampling using a pinhole array mask with randomly perturbed pinhole locations.

The procedure is easier to explain using the terminology of a traditional light field camera. A light field camera records the spatial and angular variations of rays incident at each location on the sensor. For an un-occluded Lambertian scene patch in sharp focus, the incoming rays have no angular variations. Reflection glare causes a bright light source in the scene to make a stray contribution to the sensor, but only along a specific angular direction (Figure 3). We eliminate this outlier in the angular dimension and its impact on the recorded luminance of the scene patch. These outliers appear as high frequency noise in 4D although the projection of ray-space onto a 2D sensor creates an apparent low-frequency glare.

Traditional methods use a 2D deconvolution approach to reduce glare in 2D post-processing. However, deconvolution fails due to limited quantization where glare overwhelms the signal. Our outlier rejection approach can handle reflection glare as well as certain types of scattering glare. We reduce the reflection glare sufficiently to permit good results by deconvolution methods.

We sample the ray-space as in a traditional light-field camera. In some scenarios, we can ignore the spatial arrangement of the sub-

*email:[raskar]@media.mit.edu (Currently at MIT Media Lab)
†email:[agrawal]@merl.com, http://www.merl.com/people/agrawal/sig08/
‡email:[cyrus.wilson]@siggraph.org (Currently at Stanford)
§email:[vashok]@umiacs.umd.edu

ACM Reference Format
Raskar, R., Agrawal, A., Wilson, C., Veeraraghavan, A. 2008. Glare Aware Photography: 4D Ray Sampling for Reducing Glare Effects of Camera Lenses. *ACM Trans. Graph. 27*, 3, Article 56 (August 2008), 10 pages. DOI = 10.1145/1360612.1360655 http://doi.acm.org/10.1145/1360612.1360655.

Copyright Notice
Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701, fax +1 (212) 869-0481, or permissions@acm.org.
© 2008 ACM 0730-0301/2008/03-ART56 $5.00 DOI 10.1145/1360612.1360655
http://doi.acm.org/10.1145/1360612.1360655

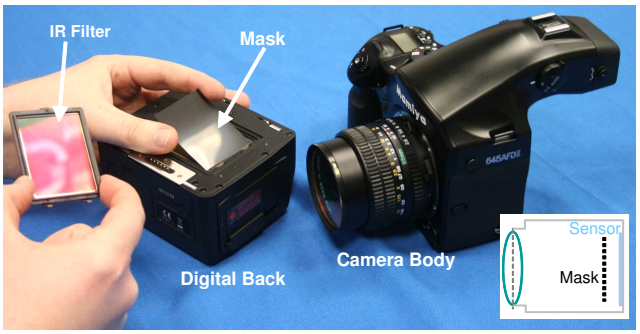

Figure 2: Our hand-held prototype implementation using a medium format camera and a printed film transparency (mask) placed on top of the sensor. Section 6 describes further implementation details.

aperture views recorded on the sensor to remove outliers due to glare. In others, we exploit coherence in neighboring sub-aperture views for clustering the spurious rays due to glare in 4D. To the best of our knowledge, our method is the first attempt to capture and analyze glare in ray-space.

1.1. Contributions

We present a set of techniques to analyze glare by capturing a subset of ray-space while minimizing resolution reduction. Note that our goal is not light field capture but glare reduction and re-synthesis in 2D. Specific technical contributions are as follows

- We explain that glare is a higher dimensional phenomenon and clarify how it manifests as a low frequency contrast reduction in photographs.
- We show a method to decompose glare spread function into reflection (outlier), scattering (bias), body and non-glare components.
- We show a practical method to capture and reduce glare in a *single shot* photo with a portable handheld camera.
- We explore tradeoffs between loss of light, spatial resolution and glare detection by uniform and non-uniform 4D ray sampling using masks.
- We apply these ideas to reduce and manipulate glare in a variety of settings. Our approach handles photography looking into the sun and photos taken without a hood, removes the effect of dust on the lens and reduces loss of contrast due to camera body reflections.

1.2. Benefits and Limitations

We believe that ours is the first method to address glare reduction in a single exposure photo. We are inspired by Talvala et al. [2007] who were first to devise an active method. Unlike light field capture methods, we can recover full resolution information for in-focus parts of the image in several cases. In addition, our method is ideally suited for single-exposure capture of scenes with extremely high dynamic range (HDR). Even without multi-exposure HDR imaging, we can ensure that saturating glare does not overwhelm the scene signal. However, our approach does suffer from several limitations. Our method can handle many types of glare, but not all of them.

- Our approach cannot handle **extended area light sources** in a general way because the resulting glare has low frequency components in angular dimensions.
- Glare due to **large area scattering** (such as foggy or dusty lens, low-quality camera body interior) cannot be reduced.

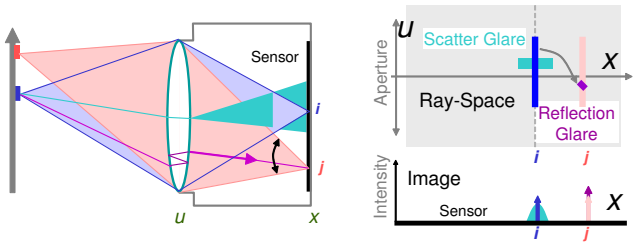

Figure 3: The key idea is based on the observation that reflection glare manifests as outliers in angular samples of ray-space. (Left) A bright pixel imaged at i (blue) contributes a stray reflected ray (purple) to the pixel j (orange). However, among the cone of rays arriving at j, only one ray is spurious, whose contribution can be rejected as an outlier. (Right) Analyzing glare in ray-space and intensity image. Rays due to blue scene patch maps to non-adjacent rays (purple) due to lens inter-reflections, as well as contribute to neighboring spatial samples (scatter glare) due to lens acting as a mild diffuser (cyan).

- We do not remove **diffraction** effects due to finite aperture and aperture diaphragm blades, commonly noticed as streaks in photographs. Diffraction is a wave-based phenomenon and cannot be analyzed with our geometric optic approach. In addition, any sensor-related issues such as **blooming** [Apogee Instruments] or purple-fringing due to sensor pixel microlenses are not addressed.
- We can only address glare created by camera optics and body and not glare due to external elements like haze.

In addition, the proposed method has several disadvantages.

- Adding a mask (or a lenslet) may introduce its own glare and diffraction.
- We reduce the resolution of the sensed image to some extent although we can maintain resolution of in-focus image portions not affected by glare.
- We block significant amount of light necessitating longer exposure time for mask based designs.
- We can handle a glare-inducing light source to appear anywhere in the scene (even outside the field of view), but we assume that the scene of interest is within the depth of field.

1.3 Related Work

Measuring and Removing Glare from Images: ISO standard 9358 [International Organization For Standardization 1994] describes the measurement procedure and defines veiling glare index as the ratio of luminance in the center of a black target to the luminance of the surrounding large area uniform illuminant. McCann & Rizzi [2007] have measured glare in multiexposure HDR imaging. Bitlis et al. [2007] have built a parametric model for stray light effects in digital cameras. In computer graphics, 4D to 8D transport tensors between the light source and sensor have been developed [Sen et al. 2005; Garg et al. 2006] for relighting and view interpolation. These methods can potentially be used to characterize glare. But they are not applicable for reducing or decomposing glare on the camera image sensor. To remove glare, software methods post-process an image that already contains glare via deconvolution [Reinhard et al. 2006]. Similar computational methods are used in X-ray imaging [Seibert et al. 1985].

Glare Prevention in Optics: High-end lenses use novel optical design and materials to reduce glare. Lens-makers' strategies include coating and lens shaping. The 4% to 8% transmission loss

due to reflection at each glass-air interface means that a 5 to 10 element lens can lose half the incident light and instead create significant reflection glare. **Anti-reflective coating** films make use of the light-wave interference effect. Vacuum vapor deposition coats the lens with a $1/4$ wavelength thin film using a $\sqrt{n}$ refractive index substance, where n is the lens glass index. Multilayered coating can bring down the reflection to 0.1%. But this is not sufficient to deal with light sources which are $4+$ orders of magnitude brighter than other scene elements. Ancillary optical elements such as filters also increase the possibility of flare effects. Digital camera sensors are more retro-reflective than film. **Meniscus lenses** with curved profile act as a spherical protective glass in front of the lens assembly and prevent unwanted focused reflections from the sensor. The curved profile defocus creates large area flare rather than ghosts. Lens makers use an electrostatic **flocking process** to directly apply an extremely fine pile to surfaces requiring an anti-reflection finish. The pile stands perpendicular to the wall surfaces acting as Venetian blinds: an effective technique for lenses with long barrel sections. Structural techniques include light blocking grooves and knife edges in lenses to reduce the reflection surface area of lens ends. Hoods or other shading devices are recommended for blocking undesired light outside the picture area.

Comparison with Other Active Approaches: Our work is motivated by the recent work of Talvala et al. [2007]. Our approach follows their lead to prevent glare-producing light from reaching the sensor pixels. As far as we know, theirs is the first and perhaps the only pre-capture method. They used a new direct-indirect separation of lens transport by selectively blocking glare-producing light using a structured occlusion mask. Our approach differs in terms of setup, applications, benefits and resynthesis. Their setup requires a large number of photos. A large sized mask needs to be displaced on an $x - y$ rig and the mask needs to be in focus and close to the scene. The size and focus requirements make it difficult to photograph a scene several meters away from the camera, such as sunlit buildings. In contrast, ours is a handheld setup that can be used like a traditional camera. In terms of applications, our method is suited for isolated bright narrow area light sources (e.g. bright sun or isolated room lights) while their method is best suited for extended area sources and cannot handle point and small area sources. Our method is tolerant of pixel saturation due to glare and hence can work without multi-exposure HDR capture [Debevec and Malik 1997]. We do not require geometric calibration in the scene for different focus settings and it is not necessary to decrease the aperture to increase the depth of field. In terms of analysis and synthesis, we can partition glare into different types providing easy resynthesis opportunities. As mentioned, the disadvantage is that our method does not work well for extended area light sources and it does not directly address situations where lenses are highly scattering. Both methods fail to recover high frequency details near a sharp luminance boundary.

2. Understanding Glare in Ray-Space

We first analyze the sources of glare in ray-space and explain their impact in image space.

2.1. Sources of Glare: Reflection versus Scattering

Reflection glare appears as parasitic images when the sun or another strong light source causes a complex series of reflections among the lens surfaces. Fresnel reflection is the portion of incident light reflected at a discrete interface between two media having different refractive indices (4% to 8% for glass-air interface). For a lens with n surfaces (i.e., glass air interfaces due to $n/2$ lens elements), the number of parasitic ghost images equals $n(n-1)/2$ [Ray 2002]. Ghosts appear as clearly defined aperture-shaped reflections in a position symmetrically opposite the light source. Flare appears as more uniform fogging of a large image area. Flare is most noticeable for large aperture, wider field of view, shorter wavelength and near the center of the image [Ray 2002]. The definition and distinction between ghost and flare varies in the literature. Glare is additionally enhanced by filters, because they have flat profiles perpendicular to the optical axis. Reducing aperture size does not necessarily reduce glare because the aperture diaphragm can in fact contribute to reflections. As an extreme example, imagine an aperture surface made of a mirror or some diffuse reflector.

Scattering glare is created by diffusion at the lenses. The optical elements act as mild diffusers. However, the diffusion angular profile is very narrow and scattering glare falls off very quickly away from the image of a light source. Because reflection glare and scattering glare overlap in a 2D image, they are difficult to automatically identify, classify, and remove. We show that in 4D the distinction between the desired image and the two types of glare is clearer, and that adding a mask to the camera makes image and glare computationally separable.

2.2. Glare Ray Spread Function

The 2D glare point spread function (GPSF) describes the amount of glare intensity created by a point light source as a function of the distance from the center of its ideal image. However, this characterization is restrictive. Rays from the point light source reaching all parts of the lens are assumed to have equal radiance. This GPSF also combines multiple glare effects. We can re-define glare using 4D rays rather than 2D pixels. Glare is a result of the mapping of a given light ray to one or more stray rays due to camera optics. We can define a **glare ray spread function** (GRSF) as a 4D function that describes how much radiance from a single incident ray transfers to every other ray. The GRSF is then composed of *reflection-GRSF*, $\mathbf{R}$, and the *scattering-GRSF*, $\mathbf{S}$. Using the two-plane parametrization of the incident light field $\mathbf{L}(x, y, u, v)$ on the sensor, where (x, y) denotes the sensor plane and (u, v) denotes the main lens aperture plane, the light field $\mathbf{L}$ inside the camera can be written as

$$\mathbf{L} = \mathbf{L}_0 * \alpha(\delta + \mathbf{R} + \mathbf{S}), \tag{1}$$

where $\mathbf{L}_0$ is the original scene light field and α is a normalization constant.

2.3. Higher Dimensional Effects of Glare

The glare spread in higher dimensional ray-space impacts the 2D image intensities after angular integration on the sensor. For the sake of explanation, we will consider a 1D sensor and a 2D light field $\mathbf{L}(x, u)$ as shown in Figure 4. Here the vertical axis represents the aperture coordinate (u) and the horizontal axis represents the sensor pixel coordinates (x). A point light source (blue patch) is in sharp focus at the sensor pixel i, and hence all entries of $L(i, \cdot)$ are equal. Due to the reflection glare, the blue patch maps to a range of rays indicated by the purple lines. The set of rays focused at i also create scattering glare spreading in x-direction, shown in cyan. The rays can also bounce off the camera body (body glare) as shown in green. A different dimmer scene point sharply imaged at pixel j, is shown in orange.

The 1D image formed at the sensor is obtained by the integration of the 2D light field (projection onto the x-axis). Figure 4 compares the image formed by a traditional sensor with that formed by our approach in which a mask is placed in front of the sensor. Notice that luminance values add up sharply at pixel i, but they also contribute to glare shown in purple, cyan and green regions. High frequency reflection glare (purple line) in ray-space emerges as a low frequency profile (purple region) in the 1D image space for a traditional camera. The scattering glare is large near the pixel i but

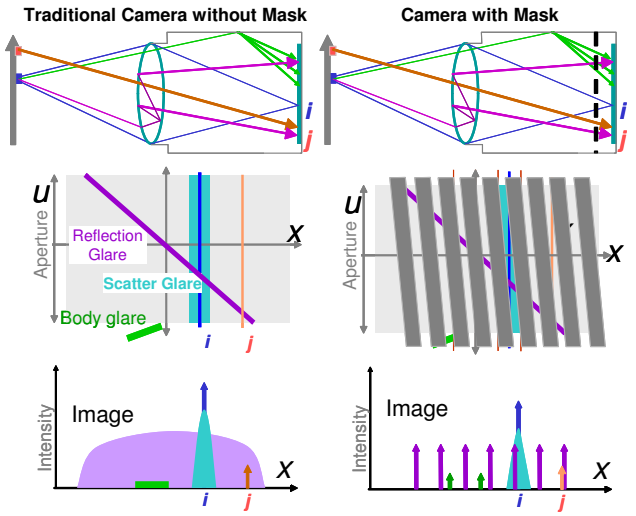

Figure 4: Comparison of glare formation in ray-space and sensor image for a traditional camera and our mask based camera. A focused blue scene patch could contribute to scattering (cyan), reflection (purple) and body glare (green). Since the sensor image is a projection of the ray-space along angular dimensions, the sum of these components creates a low frequency glare for a traditional camera. However, by inserting a high frequency occluder (gray), in front of the sensor, these components are converted into a high frequency 2D pattern and can be separated.

drops off rapidly. The pixel j is simply overwhelmed by the remaining reflection glare. It is difficult to distinguish between the contribution of the reflection or the scattering glare. However, by inserting a high frequency occluder (mask) in front of the sensor, glare in ray-space appears as a high frequency pattern in the captured sensor image. Thus, we can reduce the effects of glare by outlier rejection. Figure 5 shows the photo of a projector in a dark room captured with a pinhole mask inside the camera. Note that the glare appears as high frequency noise in ray-space.

3. Reducing Glare

In the last section, we showed that glare is a 4D phenomenon. In this section, we first describe a basic algorithm for reducing glare. We then outline various design choices and tradeoffs in using masks for ray sampling. We emphasize that it is not necessary to reconstruct the scene light field but by appropriate ray sampling, glare can be reduced.

3.1. A Basic Algorithm

One approach to reduce glare is to explicitly capture and reconstruct the light field using light field cameras based on lenslets [Adelson and Wang 1992; Ng et al. 2005; Levoy et al. 2006; Hanrahan and Ng 2006] or sum-of-cosines masks [Veeraraghavan et al. 2007]. Methods that capture the light field outside the camera, using a camera array [Levoy and Hanrahan 1996; Gortler et al. 1996] or lens-array [Georgiev et al. 2006] are not applicable here. In that case, the algorithm can be summarized as follows:

- Capture a 2D high dynamic range photo with a portable light field camera.
- Reconstruct the corresponding 4D light field $\mathbf{L}(x, y, u, v)$.
- For each spatial sample of the light field, use robust statistics to eliminate the outlier values among its corresponding angular samples.

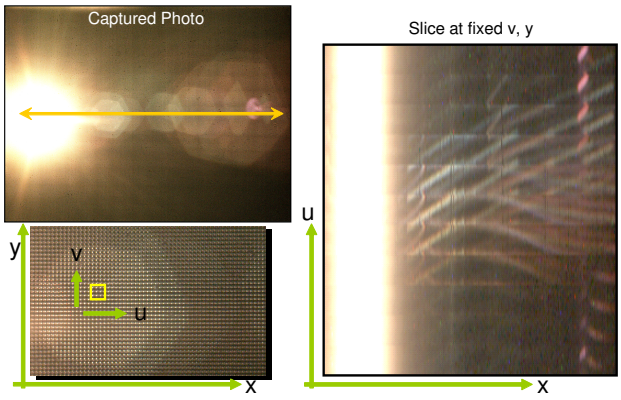

Figure 5: Glare visualization. (Top left) Captured photo of a projector in a dark room using uniform pinhole array mask. The captured photo can be treated as 4D ray-space samples $\mathbf{L}(x, y, u, v)$. (Bottom Left) Close up showing that glare appears as high frequency noise in ray-space. Notice the coherence in the position of glare among neighboring spatial samples of $\mathbf{L}$. (Right) A slice of 4D ray-space for fixed y and v (through the center of the captured photo).

- Reconstruct a *low resolution* 2D image $i(x, y)$ by averaging the remaining angular samples for each spatial sample.

For each spatial sample (x, y), we can also use operators based on order statistics such as min or median to eliminate the outlier angular samples: $i(x, y) = \min_{u,v} \mathbf{L}(x, y, u, v)$ or $i(x, y) = \mathtt{median}_{u,v}\mathbf{L}(x, y, u, v)$. Figure 1 shows an example where we average the bottom 20% of the angular samples for each spatial sample to reduce the glare. Let us refer to this image as i_{gr}. For fair comparison (Figure 1(middle)), we take the average of *all* the angular samples for each spatial sample (without removing the outliers) as this would be equivalent to capturing an image with a traditional low resolution 2D camera. The glare component can be extracted as $\mathtt{glare}(x, y) = i_{max}(x, y) - i_{gr}(x, y)$. The glare enhanced image can be obtained by modifying the glare component and adding it to the glare reduced component (Figure 1(left)). This type of analysis is similar to Nayar et al. [2006] and Vaish et al. [2006].

This strategy works well when the glare outliers are well below 50%. The drawback with this approach is that all light field cameras suffer loss in spatial resolution to achieve angular sampling and computing the outliers could be non-trivial in some cases. Next, we discuss our pinhole array mask based designs that exploit tradeoffs between loss of light, resolution and outlier rejection.

3.2. Design Choices

We choose a pinhole array mask due to the following considerations.

HDR Capture: In a pinhole array mask based design (or a lenslet array design), ignoring blooming, saturation of one pixel does not impact neighboring pixels. Sum-of-cosines mask-based methods sample linear combinations of rays and hence assume a linear response. Under/over exposure makes the linear inversion and decoding unstable. In addition, the decoding process amplifies noise. Thus, a strength of our pinhole array based design is that it could be used without HDR capture, an advantage in outdoor scenarios.

Ray Sampling versus Light Field Reconstruction: All previously described light field cameras are targeted towards explicitly reconstructing the scene light field. Our goal is to obtain a glare reduced 2D image as opposed to *reversibly re-bin or encode* the projection of 4D ray-space onto the 2D sensor. Pinhole array

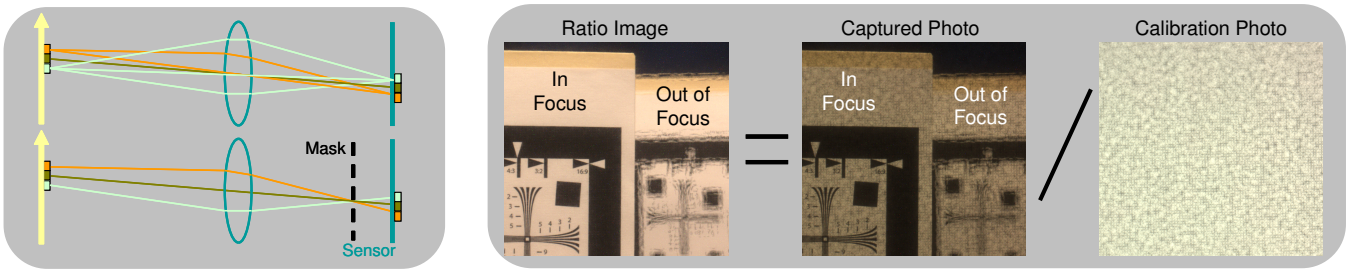

Figure 6: Recovering a full resolution image. (Left) Without mask, three scene patches create a focused image. With mask, for scene parts in focus, at least a subset of rays reach the sensor, creating the same (but less bright) image of the three scene patches. (Right) By dividing the captured photo with a calibration photo, we can recover the scene parts in focus at full resolution. Out of focus scene parts, however, exhibit aliasing as shown by the resolution chart in the back.

masks allow us to explore non-uniform ray sampling (using randomized pinhole array masks) and larger aperture size (which leads to aliased light fields). Non-uniform ray sampling also provides visually pleasing results as opposed to uniform ray sampling.

Loss of Spatial Resolution: The biggest win in using masks is that we can avoid the loss of spatial resolution common to all light field cameras. We attempt to get a full resolution image for the *in-focus* parts of the captured photo. In our setup, we keep the scene in sharp focus on the sensor as a photographer would do, compared to Ng et al. [2005] where the scene is focused on the lenslet array.

Consider a scene point in focus at the sensor. The cone of rays emerging from the scene point are refracted through the lens and fall on a single sensor pixel. Introducing a mask simply blocks some of these rays without changing the pixel position. As long as a part of any sub-aperture contributes rays to this pixel, the sensor can record it. Thus, for in-focus scene parts, the captured image is dimmer but full spatial resolution can be recovered. However, for out of focus scene parts, every neighboring scene point is captured by a slightly different center of projection. Thus, out of focus parts exhibit artifacts. Figure 6 shows proof of concept on in-focus and out of focus resolution charts. Note that the out of focus resolution chart shows a stair-step like aliasing effect, but for the in-focus chart, full resolution can be obtained. In practice, one needs to divide the captured photo with a calibration photo of a uniform Lambertian light box to account for variations in the mask.

As stated earlier, our assumption is that the entire scene lies within the depth of field of the lens. Thus, for the purpose of removing glare, we can avoid the loss of spatial resolution. However, when the number of glare rays is 'sufficiently large' (as described later), we revert back to a low resolution 2D output.

4. Exploring Masks for Glare Reduction

In this section, we describe several approaches for glare reduction using uniform and randomized pinhole array masks in the camera. We first outline the terminology (see Figure 7) for describing the subsequent approaches. We discretize the captured ray-space into a finite number of 'sub-aperture' bins denoted by (u, v) and a finite number of sensor pixels (x, y). We will refer the union of the rays passing through a given sub-aperture and contributing a single sensor pixel as one 'ray' of the system. A 'tile' refers to the maximum size of the non-overlapping aperture image formed by a single pinhole. The size of the tile depends on the distance of the mask from the sensor and restricts the maximum aperture size allowing non-overlapping aperture images on the sensor. In our setup, this maximum aperture size corresponds to $f/8$. Smaller $f/\#$ (larger aperture) will lead to aliased light fields but are useful for glare reduction as shown later. The tiles are uniform for uniform pinhole

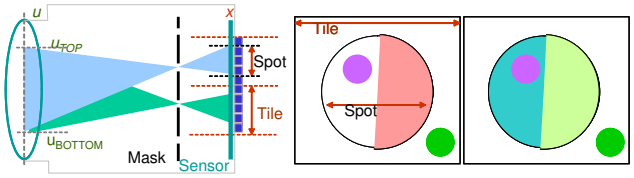

Figure 7: Capturing a photo by placing a pinhole mask near the sensor. 'Tile' refers to the maximum size of the non-overlapping aperture image formed by a single pinhole. For small apertures ($u_{top} \leftrightarrow u_{bottom}$), the image of the aperture is restricted to a 'spot' within the tile. (Middle) An orange colored vertically striped object is focused on the sensor. The purple dot indicates an outlier corresponding to the reflection glare. (Right) Scatter-glare (cyan) creates bias making the orange colored scene object look more greenish. Body glare (green dot) is contributed from a ray outside the view cone towards the aperture.

mask and the notion of a tile remains the same for the randomized pinhole mask, while being non-uniform in spacing. Depending on the $f/\#$ of the main lens, the image under each pinhole may not cover the entire tile and is referred to as an aperture-shaped 'spot'. The image created by a sub-aperture is referred to as a 'sub-view'. Figure 7 also shows the representation of a spot under a pinhole for different types of glare. Reflection glare appears as a bright outlier and scatter glare appears as a bias, reducing contrast. For larger $f/\#$ (smaller lens aperture), body glare appears *outside* the spot. Reflection glare is caused by stray rays and can be removed using outlier analysis. Scattering glare adds a bias which is handled by deconvolution after removing the reflection glare. Body glare formed by scattering of rays after the last lens element appears outside the spot and can be removed by using a smaller aperture.

Our approach can also handle saturation. In a traditional camera, glare contributes to all the pixels in a neighborhood potentially saturating or biasing all of them. In our camera, reflection glare contributes to a *few* pixels in a tile. If those pixels are saturated, their intensity is clamped at the maximum without affecting the neighbors. By using a non-linear response camera, such as a logarithmic sensing camera, one might achieve even better noise and quantization behavior despite saturating glare.

4.1. 2D Image Analysis

We first describe a method that works entirely in the 2D image domain without any ray-space analysis. This approach is useful for both uniform and non-uniform pinhole array masks and can minimize the loss of resolution if the reflection glare contributes to less than 50% of the pixels in each spot. Figure 8 shows an example, where we captured a photo with a randomized pinhole array mask

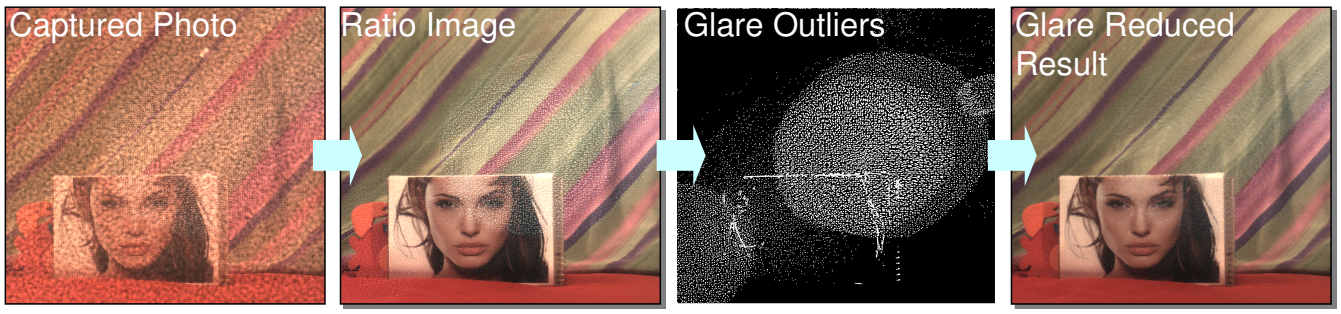

Figure 8: Glare removal without loss of spatial resolution. We captured the photo shown on the left using a randomized pinhole array mask in the camera at $f/4$. A projector on the top-right corner in the scene (not shown) causes glare. By dividing the captured photo with a calibration photo, the mask effects are removed. Note, however, that the mask makes glare appear as high frequency 2D noise. A glare outlier stencil is obtained by comparing the ratio image with its median filtered output. The glare is reduced by interpolating the pixels in outlier stencil from neighboring pixels. All processing is done in 2D image domain.

at $f/4$. The randomized pinhole array mask is obtained by perturbing the location of each pinhole randomly compared to a uniform grid. A projector in top-right corner of the scene causes reflection glare. We first divide the captured photo with the photo of a uniform intensity Lambertian lightbox (referred to as the calibration photo). Any brightness variation in overlapping spots is thus normalized. Then we use a tile-sized moving window and compute the median at every pixel. If the current pixel value is brighter and differs from the tile median by some threshold, we mark it for removal as an outlier. For the removed pixels, we interpolate to fill in the values from neighboring pixels and recover a full resolution glare reduced image. Glare free pixels remain untouched and the resolution is decreased only by the percentage of glare pixels removed.

Note that a larger aperture ($f/4$) is useful in this case, as it leads to overlapping spots, which avoid issues of dark pixels at the corner of the tiles if the aperture is small. A larger aperture makes the calibration photo non-zero at all pixels, ensuring that the division by the calibration photo is stable. Also, such large apertures would have lead to aliasing if a light field was recovered from the captured photo. Thus, appropriate ray sampling (even non-uniform) is sufficient for glare reduction without explicit light field reconstruction. An additional advantage of using a randomized pinhole array is that visually pleasing results are obtained after division by the calibration photo. In case of a uniform pinhole array, such a division leads to visually strong vertical and horizontal grid like pattern due to uniformity of tiles.

4.2. 4D Ray-Space Analysis

The 2D processing fails when a single tile contains multiple outlier elements covering more than 50% of the tile. Flare and ghosts have a characteristic structure in 4D because they are created by highly symmetric elements of optics, resulting in high coherence among glare rays in 4D ray-space. We exploit this spatio-angular coherence of glare in neighboring tiles for better segmentation of glare rays. Figure 9 shows an example. Note that since we utilize the spatial structure of tiles, this analysis only applies to uniform ray sampling.

For 4D ray-space analysis, we *treat* the captured photo as a 4D lattice $\mathbf{L}(x, y, u, v)$. The captured photo may not correspond to an un-aliased light field due to the large aperture size (small $f/\#$) of the main lens. Each tile at (x, y) corresponds to the angular samples of $\mathbf{L}(x, y, u, v)$. We perform a 4D segmentation via a 2-way clustering to partition the 4D ray-space into glare and non-glare parts as follows.

- Create a 4D graph from $\mathbf{L}(x, y, u, v)$. The nodes of the graph correspond to the samples of $\mathbf{L}(x, y, u, v)$. Each node (except on the boundary) is connected to $3^4 - 1 = 80$ neighbors[1].
- If nodes i and j are connected, assign edge weights $w_{ij} = \exp(-(\mathbf{L}(x_i, y_i, u_i, v_i) - \mathbf{L}(x_j, y_j, u_j, v_j))^2)/(2\sigma^2))$, where σ is a constant.
- Assign a few bright pixels as source nodes of the graph and perform 2-way clustering using a maxflow (mincut) algorithm [Boykov and Kolmogorov 2004] to segment $\mathbf{L}(x, y, u, v)$ into glare and non-glare components[2].
- Since each pixel in the captured 2D photo corresponds to a sample in $\mathbf{L}(x, y, u, v)$, mark the glare pixels in 2D photo.

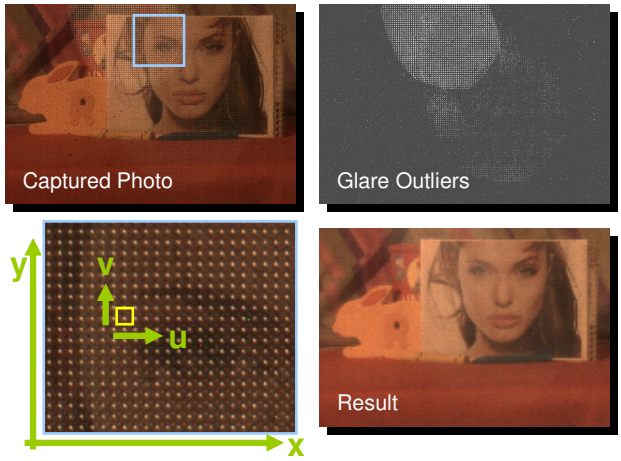

Figure 9: Glare reduction using 4D ray-space analysis. Top left shows captured photo with a uniform pinhole array at $f/8$. After 4D clustering, we mark the 4D samples corresponding to one cluster as white (glare) and other cluster as black (glare-free). Note that since each 4D sample corresponds to a 2D pixel, we can visualize the segmentation in 2D as shown in top right. The aperture rings are clearly segmented although they are not seen strongly in the captured photo. We recover a low resolution result by averaging non-glare pixels in each tile.

The 4D analysis allows us to separate the glare pixels in 2D photo. These could be interpolated as before to recover glare reduced out-

[1]This is similar to 2D planar graphs where each node is connected to $3^2 - 1 = 8$ neighbors.

[2]We modified the graph cut software from http://www.cs.cornell.edu/People/vnk/software.html.

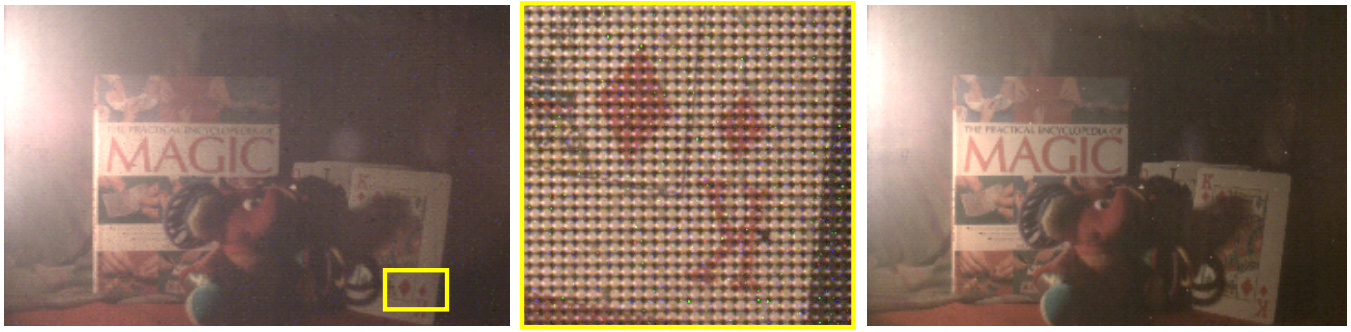

Figure 10: Reflection glare in regions far from the light source. (Left) Captured photo using uniform pinhole array at $f/8$. (Middle) Notice the dominance of outliers indicating reflection glare even at regions away from the light source. (Right) Glare reduced low resolution result showing improved contrast, especially in shadows.

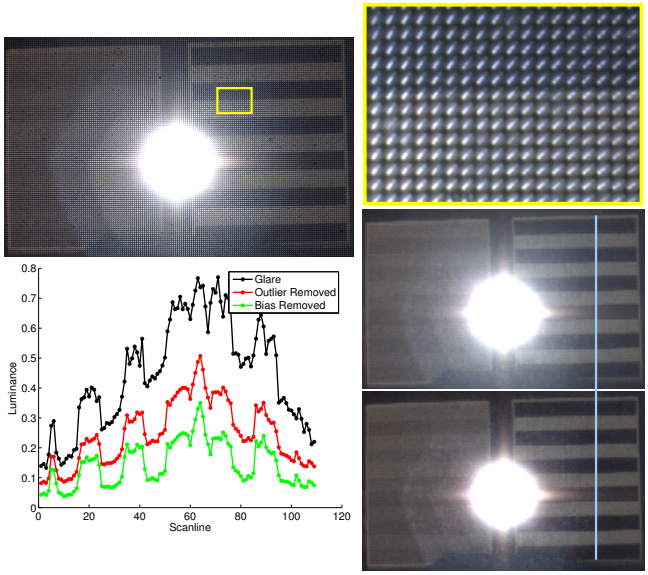

Figure 11: Contrast improvement. (Top Left) Photo of glare obscured contrast charts with inset shown in bottom left (uniform pinhole array at $f/8$). (Right) Improvement in contrast after outlier rejection (top) and further reduction in fog via 2D deconvolution (bottom). The plots show the luminance profile along the contrast chart, before (red) and after (green) 2D deconvolution.

put. However, when the number of glare pixels are large, we obtain a more reliable low resolution result as follows. For each tile, we ignore the glare pixels and simply average the remaining pixels to recover one intensity value per tile. The resolution is thus reduced by the number of angular samples as in a traditional light field capture.

4.3 Reducing Scattering Glare

Glare depends highly on the position of the light source and modeling it as a 2D point spread function is difficult. Thus, convolution based techniques fail in the presence of strong reflection glare. But after removing the outliers corresponding to the reflection glare, the remaining low frequency scattering glare behaves in a more predictable way. Similar to [Talvala et al. 2007], we model the scatter glare PSF in 2D as a mixture of 2D Gaussians and deconvolve the reflection glare reduced output to remove the scattering glare. Figure 11 shows an example.

5. Applications

We show several examples of contrast enhancement including outdoor scenes, strong aperture ghosts, and multiple lights in the scene.

Contrast Enhancement: The sunlit scene in Figure 1 spans a wide dynamic range. We took just one photo with uniform pinhole array at $f/8$. Parts of the scene, such as the sun and the sky are saturated. Although this example involves a relatively extended light source, we note that the majority of glare on the indoor portions of the image is reflection glare. Another example is shown in Figure 10. However, the visible aperture rings overlap so that several tiles have more than 50% glare values. We use the basic algorithm explained in Section 3.1 to recover low resolution glare reduced output.

Lens Smudges: Dust specks, fingerprints and water droplets on lenses or filters introduce unwanted diffusion and/or refraction. This causes a localized reduction in contrast, appearing as a low frequency smudge. If the distorting element is limited to a narrow area of the lens or filter, we can eliminate its effect via ray-space analysis. There are no bright lights sources or saturation, in this case. An example is shown in Figure 12. Imagine replacing the distorting element by an opaque black surface. Given the finite aperture, blocking a part of the lens will only nullify certain rays making the image dimmer. We want to eliminate those rays in software so that we can remove the artifact and recover a slightly dimmer image. This elimination is a blind estimation problem because we do not have information to model the distorting element. We assume that the distorting element (on the lens) is highly out of focus while the scene is in sharp focus. We compute the sub-aperture views and reconstruct a low resolution image by choosing pixels from views that maintain the highest spatial contrast in a small neighborhood.

Body Glare: Body glare is caused by reflections occurring in the space between the last lens element and the sensor. Rays from objects outside the field of view, rays blocked by the anti-reflective coating or aperture eventually end up on the camera body. High-end cameras use baffles to prevent their impact. To demonstrate body glare effects, we intentionally placed a diffusing reflector inside the lens and used a small aperture ($f/11$) so that the rays due to body glare fall outside the spots. Figure 13 shows an example, where the red cloth causes glare on the green region on the left as evident by the close up of tiles. To reduce body glare, we eliminate pixels outside the spot to remove the contribution of red cloth to the green region. Our method is also applicable to more ad hoc imaging arrangements where it may be difficult to create high quality baffles to attenuate stray light. By giving up a small percentage of resolution, the stray light can be trapped outside the spot.

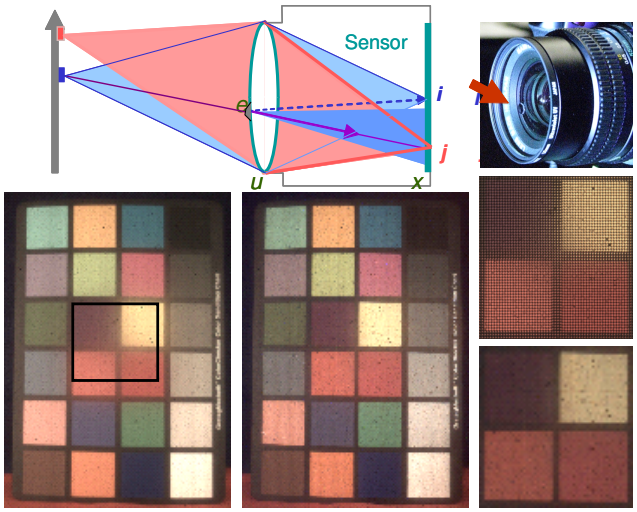

Figure 12: A water drop on the lens creates localized smudging. (Top Left) Rays passing through a distorter at e diffuses or refracts light in multiple directions. (Top Right) A water drop on the lens. (Bottom Left) The color chart before and after glare removal.

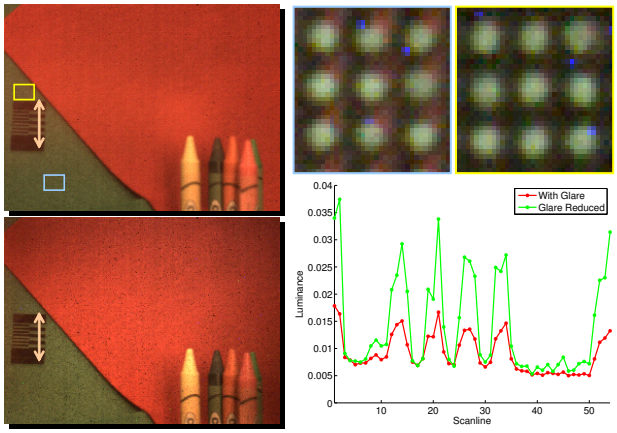

Figure 13: Reducing body glare. (Left) Photo of reddish wash on green cloth before and after outlier removal. (Top Right) Insets show the red values *outside* the spots, which we remove. (Bottom Right) The intensity profile of contrast chart shows dramatic improvement.

Rendering and Synthesis: Photographers intentionally use glare for artistic effects and to create illusion of bright light sources. Glare is rendered via software image filters or raytracing to create visually compelling flare effects [Knoll Light Factory ; Kakimoto et al. 2005; Spencer et al. 1995]. But, at the capture time, photographers do not have much control over the glare formation. In addition, glare is difficult to examine through optical or digital viewfinder. We use the decomposed glare components in several interesting and physically realistic ways for digital glare manipulation. For example, the rejected outlier component in a tile provides an estimate of reflection glare. In Figure 1, we show enhanced glare. We also easily change the color of glare. From a single photograph, we create animations of an apparently moving light source generating dramatic flare (video).

6. Analysis and Implementation

Using a probabilistic analysis, we can analyze the number of pixels in a tile and the pinhole spacing for removing glare as outliers. For uniform sampling, let us assume that the probability that a pixel is not affected by glare is β. Let M be the number of pixels in a tile. Then the probability that r out of M pixels in a tile are not affected by glare is given by $\binom{M}{r}\beta^r(1-\beta)^{(M-r)}$. For outlier detection to work well, $r \geq M/2$. Thus, the probability that at least 50% of pixels are not affected by glare is given by

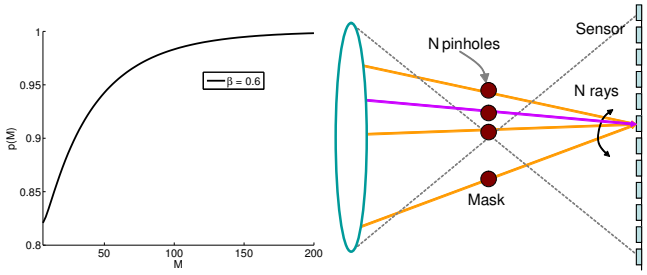

Figure 14: (Left) Plot showing the probability that at least 50% of the sensor pixels are not corrupted by glare as a function of the number of pixels within each tile for $\beta = 0.6$. (Right) Multiple pinholes contribute to a single pixel if the spacing between the pinholes is reduced.

$$p(M) = \sum_{r=M/2}^{r=M} \binom{M}{r} \beta^r (1-\beta)^{(M-r)}. \qquad (2)$$

Figure 14 shows the plot of $p(M)$ versus M assuming $\beta = 0.6$. For reasonable fidelity ($p(M) = 0.98$), we need $\approx$ 95 pixels within each tile. In our implementation, we used a tile size of 11×11, accounting for modeling errors.

As discussed earlier, we also used non-uniform sampling and larger apertures so that the tiles overlap. This means that compared to uniform sampling, here every pixel could receive contribution from more than one glare ray. Let us analyze how much overlap can be allowed. Let α be the probability that a ray is affected by glare and N be the average number of rays contributing to a pixel. Then the probability that a pixel does not receive contribution from any glare affected rays is $(1-\alpha)^N$. For this to be greater than β, we could increase N to $\log(\beta)/\log(1-\alpha)$. Assuming a 5% probability of a ray getting affected by glare ($\alpha = 0.05$), N could be increased to 9.959. Thus, in 2D, on an average the spacing between the pinholes could be reduced by $\sqrt{N} = 3.16$ times, or the size of the aperture could be increased by the same factor keeping the same spacing.

Implementation: We used a Mamiya 645ZD medium format digital camera with a 22 mega-pixel sensor digital back having a $36 \times 48mm^2$ Dalsa CCD imaging sensor. The sensor resolution is 5344×4008 pixels. A 1.2mm thick glass protects the sensor. We printed a pinhole array mask of the same size and simply dropped it on top of the sensor protective glass. We used an additional glass piece to push and flatten the mask to hold it in place. We used 50mm, 80mm and 210mm fixed focal length lenses. The maximum aperture size which avoids the overlap between tiles is $f/8$, but we used larger apertures as explained earlier. Figure 2 shows the medium format camera, lens, mask and digital back. We printed the transparency mask at a resolution of $25\mu m$ with pinhole size $25\mu m \times 25\mu m$. For the randomized mask, the location of the pinholes were uniformly distributed within a range of $50\mu m$. We obtain 242×182 tiles and 22×22 angular samples per tile. Instead of Bayer interpolation, we down-sampled the raw Bayer pattern by two to obtain a color image resulting in 11×11 angular samples in each tile. A single $A4$ sized transparency holding 20 masks can be printed for $100, making the additional cost of our setup just $5.

A disadvantage of the pinhole array based design is that it allows

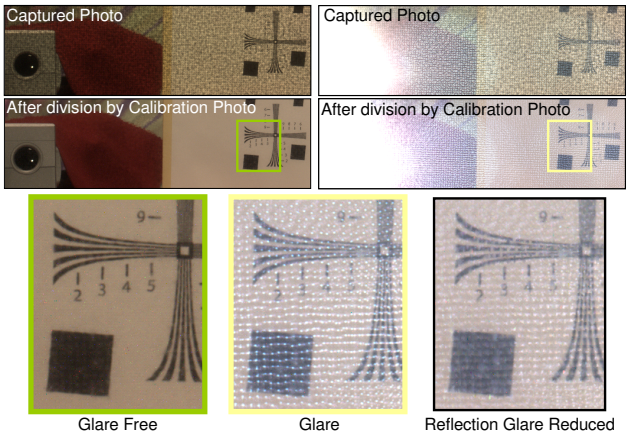

Figure 15: Reducing reflection glare and recovering full resolution image. The projector in the scene causes reflection glare on the resolution chart which is removed by identifying outlier pixels and interpolating from neighboring values. The bias due to scatter glare, however, remains.

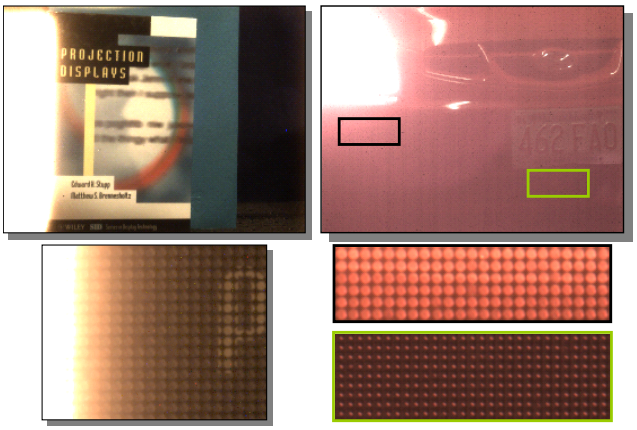

Figure 16: Failure cases: (Left) Extended area light source can cause multiple reflection glare components to mix so as to create a bias at each pixel. (Right) Overwhelming light bleeds through the mask leaving no unpolluted pixels.

$\approx 2\%$ average light throughput, forcing longer exposure times: 0.5sec in sun-light scenes (Figure 1) and up to 30sec indoors. One may improve the light efficiency using a sum-of-cosines pattern [Veeraraghavan et al. 2007]. Smarter patterns that require linear inversion will not work for a single shot approach in presence of saturation and will require HDR capture. Although each exposure time will be short and arguably one can take several exposure bracketed photos in the same time, we take only one photo with pinhole array masks.

Contrast and Resolution: Figure 11 shows a glare-obscured contrast chart. The plot shows the 1D intensity profile without and with glare correction. In our case the ratio between white and black region of the chart improves from $2.1:1$ to $4.6:1$, an improvement of a factor of two. For in-focus portions, there is a minimum impact on resolution. Figure 15 shows resolution chart captured using randomized pinhole array at $f/5.6$ and recovered glare reduced result at full resolution. Glare impacts certain pixels and their values can only be hallucinated. At $25\mu m$ pinhole opening, the diffraction for green light is 0.0240 radians, and after 1.2mm propagation, the diffraction blur size is $28\mu m$, which is roughly 3 pixels.

Failure Cases: Our method is not suited for extended area light sources. Reflection glare is convolved with the area of the light source and each tile receives stray light from multiple directions, which appears as scattering glare. Figure 16 shows the failure cases. The tiles appear washed out. With bright car-headlights, the contrast ratio of our mask is also a limitation. Our mask has a contrast ratio of only $1000:1$ for blocking light in opaque mask regions, while the dynamic range close to the light source is $3 \times 10^6 : 1$. Bleeding through the mask creates significant saturation close to the light source and we fail to reliably reduce reflection glare.

7. Future Directions

4D Analysis: Our 4D analysis of glare offers a more detailed representation of glare based on glare *ray* spread function. This can support improved lens designs as well as physically valid filters for glare deconvolution and rendering. It can support analysis of camera body or lens barrel reflections. We can decompose glare components generated by a single light source, but more powerful techniques may allow us to decompose multiple light glare contributions from a single image. We used the 4D analysis for distortions inside a camera. A similar statistical approach may be interesting for studying and minimizing lens aberrations such as chromatic aberrations.

Mask Pattern and Hardware: Talvala et al. [2007] used direct-global separation to partition direct rays from glare causing reflected and scattered rays. We have instead used a 4D analysis. Both methods use high frequency masks close to either the scene or the sensor. By combining the two, we may be able to deal with high as well as low frequency glare. Further, a hybrid of inside/outside mask could be a promising approach to deal with a range of other direct-global separation problems. We have demonstrated the use of planar masks for 4D ray-space sampling. However, other types of masks such as volumetric masks may allow further control over accepting and rejecting a subset of 4D samples. One can use high-contrast spatial light modulators or LCDs to create dynamic masks. Since the mask resolution is close to pixel resolution, Bayer mosaics used for color filtering could be combined with 4D sampling. One might imagine a consumer camera with an aperture shape bracketing mode to support intelligent reshaping over multiple photos. The effective aperture shape could then be varied on a pixel-by-pixel basis to form a composite image in which glare and distortion effects are reduced.

8. Conclusions

We believe that we have presented the first system to exploit statistical properties of 4D ray-space distortions inside a camera. Ours is the only single-shot optical mechanism for glare reduction. It works with minimal changes to the camera, is portable and produces results in many practical challenging scenarios. Glare-producing scenes inherently span a large dynamic range. Even with unavoidable saturation, our method works in a single shot without a need for multi-exposure HDR capture. Appearance of high frequency 4D glare as a low frequency 2D image pattern has traditionally confounded glare analysis. We have studied the phenomenon in ray-space allowing us to classify, visualize and, possibly for the first time, decompose glare into meaningful sub-components. We have shown results on traditional problems due to bright illumination as well as problems that have not been explored such as lens smudges and body-induced glare.

Light field applications using uniform 4D sampling have so far been limited to digital refocusing, view interpolation and shape recovery.

We have extended it to non-uniform sampling and have presented glare analysis as a new opportunity. Glare is a common problem in several other fields, such as X-ray based tomography [Goszczynska et al. 2000; Skinner 1988], lithography, fluorescent microscopy and in astronomy for isolating dim light sources appearing near bright stars. In the majority of the cases, noise is removed by deconvolution in software. We hope that our approach exploiting higher dimensional sieving will inspire new designs and support better recording probes.

Acknowledgements

We thank the anonymous reviewers and several members of MERL for their suggestions, and John McCann and Douglas Lanman for helpful discussions. We thank Clifton Forlines and Takafumi Aoki for help in preparing the video. We also thank Keisuke Kojima, Jay Thornton, Joseph Katz and John Barnwell, along with Haruhisa Okuda and Kazuhiko Sumi, Mitsubishi Electric, Japan for their help and support.

References

ADELSON, T., AND WANG, J. 1992. Single lens stereo with a plenoptic camera. *IEEE Trans. Pattern Anal. Machine Intell. 14*, 99–106.

APOGEE INSTRUMENTS. Blooming vs. anti-blooming by ccd university. http://www.ccd.com/ccd102.html.

BITLIS, B., JANSSON, P. A., AND ALLEBACH, J. P. 2007. Parametric point spread function modeling and reduction of stray light effects in digital still cameras. *Computational Imaging V, SPIE 6498*, 29–31.

BOYKOV, Y., AND KOLMOGOROV, V. 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Anal. Machine Intell. 26* (Sept.), 1124–1137.

DEBEVEC, P. E., AND MALIK, J. 1997. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97*, 369–378.

GARG, G., TALVALA, E.-V., LEVOY, M., AND LENSCH, H. P. A. 2006. Symmetric photography: Exploiting data-sparseness in reflectance fields. In *Rendering Techniques 2006: 17th Eurographics Workshop on Rendering*, 251–262.

GEORGIEV, T., ZHENG, C., NAYAR, S., CURLESS, B., SALASIN, D., AND INTWALA, C. 2006. Spatio-angular resolution tradeoffs in integral photography. In *Eurographics Symposium on Rendering*, 263–272.

GORTLER, S., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. 1996. The lumigraph. In *SIGGRAPH '96*, 43–54.

GOSZCZYNSKA, H., DORIA-DERNALOWICZ, J., PURZYCKI, Z., AND LESZCZNSKI, L. 2000. Estimation of the x-ray scatter and veiling glare rate in coronarographic densitometric measurements. In *Engineering in Medicine and Biology Society, 2000. Proc. 22nd Annual Intl. Conf. IEEE*, vol. 3, 1714–1716.

HANRAHAN, P., AND NG, R. 2006. Digital correction of lens aberrations in light field photography. In *International Optical Design*, Optical Society of America, WB2.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. 1994. Optics and optical instruments veiling glare of image forming systems definitions and methods of measurement. *ISO 9358:1994*.

KAKIMOTO, M., MATSUOKA, K., NISHITA, T., NAEMURA, T., AND HARASHIMA, H. 2005. Glare generation based on wave optics. *Computer Graphics Forum 24*, 2 (June), 185–194.

KNOLL LIGHT FACTORY. Glare creation software. http://www.digitalanarchy.com/knollPS/ (Photoshop Plugin).

LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *SIGGRAPH 96*, 31–42.

LEVOY, M., NG, R., ADAMS, A., FOOTER, M., AND HOROWITZ, M. 2006. Light field microscopy. *ACM Trans. Graph 25*, 924–934.

MCCANN, J. J., AND RIZZI, A. 2007. Veiling glare: The dynamic range limit of hdr images. In *Human Vision and Electronic Imaging XII, SPIE*, vol. 6492.

NAYAR, S. K., KRISHNAN, G., GROSSBERG, M. D., AND RASKAR, R. 2006. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Trans. Graph. 25*, 3 (July), 935–944.

NG, R., LEVOY, M., BRDIF, M., DUVAL, G., HOROWITZ, M., AND HANRAHAN, P. 2005. Light field photography with a hand-held plenoptic camera. Tech. rep., Stanford Univ.

RAY, S. F. 2002. *Applied Photographic Optics: Lenses and Optical Systems for Photography*, 3rd edition ed. Focal Press.

REINHARD, E., WARD, G., PATTANAIK, S., AND DEBEVEC, P. 2006. *High Dynamic Range Imaging - Acquisition, Display and Image-based Lighting*. Morgan Kaufman Publishers, 500 Sansome Street, Suite 400, San Francisco, CA 94111.

SEIBERT, J. A., NALCIOGLU, O., AND ROECK, W. 1985. Removal of image intensifier veiling glare by mathematical deconvolution techniques. *Medical physics 12, 3*, 281288.

SEN, P., CHEN, B., GARG, G., MARSCHNER, S., HOROWITZ, M., LEVOY, M., AND HENDRIK, P. A. L. 2005. Dual photography. *ACM Trans. Graph. 24*, 745–755.

SKINNER, G. K. 1988. X-Ray Imaging with Coded Masks. *Scientific American 259* (Aug.), 84.

SPENCER, G., SHIRLEY, P. S., ZIMMERMAN, K., AND GREENBERG, D. P. 1995. Physically-based glare effects for digital images. In *SIGGRAPH '95*, 325–334.

TALVALA, E.-V., ADAMS, A., HOROWITZ, M., AND LEVOY, M. 2007. Veiling glare in high dynamic range imaging. *ACM Trans. Graph. 26,3* (July), 37:1–37:9.

VAISH, V., LEVOY, M., SZELISKI, R., ZITNICK, L., AND KANG, S. 2006. Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures. In *Proc. Conf. Computer Vision and Pattern Recognition*, vol. 2, 2331–2338.

VEERARAGHAVAN, A., RASKAR, R., AGRAWAL, A., MOHAN, A., AND TUMBLIN, J. 2007. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph. 26*, 3 (July), 69:1–69:12.