A Wide-Field-of-View Monocentric Light Field Camera

Donald G. Dansereau[1], Glenn Schuster[2], Joseph Ford[2], and Gordon Wetzstein[1]

[1]Stanford University, Department of Electrical Engineering
donald.dansereau@gmail.com, gordon.wetzstein@stanford.edu
[2]University of California San Diego, Department of Electrical & Computer Engineering
gschuster@ucsd.edu, jeford@ucsd.edu

Abstract

Light field (LF) capture and processing are important in an expanding range of computer vision applications, offering rich textural and depth information and simplification of conventionally complex tasks. Although LF cameras are commercially available, no existing device offers wide field-of-view (FOV) imaging. This is due in part to the limitations of fisheye lenses, for which a fundamentally constrained entrance pupil diameter severely limits depth sensitivity. In this work we describe a novel, compact optical design that couples a monocentric lens with multiple sensors using microlens arrays, allowing LF capture with an unprecedented FOV. Leveraging capabilities of the LF representation, we propose a novel method for efficiently coupling the spherical lens and planar sensors, replacing expensive and bulky fiber bundles. We construct a single-sensor LF camera prototype, rotating the sensor relative to a fixed main lens to emulate a wide-FOV multi-sensor scenario. Finally, we describe a processing toolchain, including a convenient spherical LF parameterization, and demonstrate depth estimation and post-capture refocus for indoor and outdoor panoramas with $15 \times 15 \times 1600 \times 200$ *pixels (72 MPix) and a 138° FOV.*

1. Introduction

Light field (LF) cameras measure a rich 4D representation of light that encodes color, depth and higher-order behaviours such as specularity, transparency, refraction and occlusion [18]. Post-capture capabilities such as perspective shift, depth estimation, and refocus are well known from consumer-grade LF cameras [24], and these also offer simplification of an expanding range of tasks in computer vision [8–10, 26, 28]. A capability for wide field of view (FOV) LF capture would greatly benefit a wide range of applications from navigation in autonomous vehicles, recognition and tracking, and object segmentation and detection [12, 25, 27].

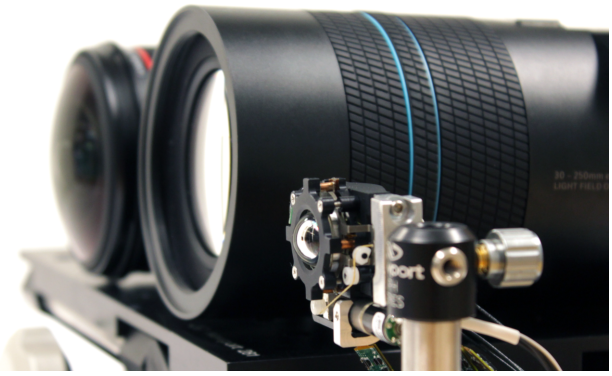


Figure 1. We demonstrate for the first time the feasibility of a compact wide-FOV LF camera, building on prior work on monocentric lenses and LF cameras. Shown here are a monocentric lens (right), a conventional lens with similar FOV and resolution (left), and the Lytro Illum (center). The monocentric lens achieves a large FOV with a smaller f-number than the Illum in a tiny form factor.

Unfortunately, wide-FOV LF imaging is currently restricted to either sequential-capture [3, 13, 37] or camera array [2] approaches. While the former restricts applications to capturing static scenes, multi-camera arrays are usually expensive and bulky. A limiting factor for building small, low-cost, wide-FOV LF cameras is that wide angle lenses are not easily adapted to LF capture. Fisheye lenses are fundamentally limited in their entrance pupil, as depicted in Fig. 2, offering only a very small baseline that prevents effective capture of LF information. Catadioptric systems employing curved mirrors are bulky, and in both catadioptric and fisheye cases resolution is fundamentally limited to that of a single sensor.

To overcome these fundamental limitations, optics designers have recently developed wide-FOV 2D imaging techniques employing monocentric optics [4, 29, 33]. Mono-

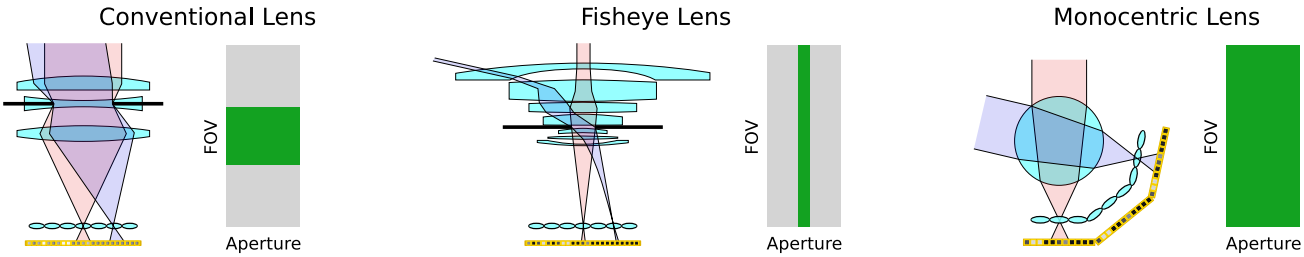


Figure 2. Comparing LF camera lenses: a conventional Cooke triplet, a fisheye lens, and a monocentric lens are shown schematically above. Whereas conventional optics work well, they do not support large FOVs. Fisheye lenses scale to 180° and beyond but have fundamentally limited entrance pupils, making them unsuitable for LF capture. Monocentric lenses support both a wide FOV and a wide aperture but present a curved focal surface. We introduce an LF processing approach to coupling the spherical lens with planar sensor arrays.

centric lenses are concentric glass spheres of differing index of refraction, i.e. multi-shell spherical lenses. These offer rotational symmetry, diffraction-limited resolution and wide FOV in a small form factor.

In this work we combine monocentric optics and LF imaging to construct a single-lens wide-FOV LF camera. We replace the expensive and resolution-limited sensor-coupling fiber bundles conventionally employed in monocentric systems with microlens arrays (MLAs) as employed in lenslet-based LF cameras [24]. This allows the optical coupling to be carried out in software, effectively leveraging the capacity of MLAs to enable light field capture and processing. This reduces cost and complexity while increasing robustness and resolution, and adds LF capture as a significant bonus.

Use cases span 3D reconstruction, autonomous driving, object detection, recognition and tracking, and potentially also cinematic content production for virtual and augmented reality applications. Compact optics and the robustness of LF processing make the device especially appropriate for robotics applications including unmanned aerial vehicles (UAVs), autonomous underwater vehicles (AUVs) and small ground vehicles. Finally, the ability to natively capture depth information through a spherical main optic makes this configuration appealing in underwater and surgical applications where the air-water interface complicates wide-FOV depth sensing.

The chief contributions of this paper are:

- Optical design and prototype of a novel single-lens wide-FOV LF camera
- A parameterization that closely matches the properties of the camera and permits conventional LF processing
- A field flattening method that efficiently couples planar sensors and the spherical focal surface of the monocentric lens
- Panoramic capture in realistic scenarios showing depth estimation, post-capture refocus, and LF resolution enhancement for panoramas with $15 \times 15 \times 1600 \times 200$ pixels (72 MPix) and a 138° FOV

To achieve wide-FOV operation, our system requires multiple sensors. Currently, the prototype only has a single sensor so we demonstrate wide-FOV capture by rotating about the monocentric lens over multiple exposures in a static scene. These results establish the viability of a multi-sensor approach, including validation of the proposed decoding, field flattening and light field processing scheme.

2. Related Work

Panoramic Light Field Imaging Wide-FOV imaging is commonly achieved by stitching multiple images that are recorded from the same center of projection [31]. Parallax is undesirable in these stitching systems because it often leads to artifacts. LF panorama stitching is emerging as a means of overcoming this limitation [3, 13, 37]. All of these approaches rely on multiple, sequentially-captured images which prevents dynamic scenes from being recorded. While we envision the proposed LF imager to operate with multiple sensors simultaneously and discuss clear directions on how to achieve that, the prototype used in this paper also relies on multiple, sequentially-captured LFs.

Wide FOV videos can easily be recorded using fisheye lenses. Unfortunately, these lenses offer only a small entrance pupil (i.e., large f-number), which fundamentally limits the LF baseline and therefore the amount of depth information gathered [22] – see Fig 2. In addition, chromatic and spherical aberrations are often problematic for wide-FOV lenses. Due to their large f-number, conventional fisheye lenses are not suitable for LF camera design.

Catadioptric systems employing curved mirrors [23] are bulkier than the proposed single-lens approach, and limit resolution to a single sensor. Our design employs multiple sensors behind a single main lens, resulting in a compact device with potentially very high pixel count – the monocentric camera on which we build offers images of up to 2.4 GPix [33].

Monocentric Imagers Monocentric lenses are rotationally symmetric shells of different refractive materials that share a common center. The simplest monocentric lens is a glass sphere. Multiple concentric layers with different

refractive indices are often useful to correct for chromatic aberrations [29].

Compared with conventional compound lenses, monocentric lenses offer simplicity, a small form factor, and extremely high, diffraction-limited resolution over the full 360° hemisphere. Referring to Fig. 2, monocentric lenses are the only option that captures a wide-FOV over a large input aperture. For these reasons, monocentric lenses have become increasingly popular for gigapixel imaging within the last few years [4, 5].

The biggest challenge of monocentric imagers, however, is that the focal surface of a spherical lens is curved. Several approaches have recently been proposed to couple this curved surface to one or more planar sensors. Arrays of microcameras, each with their own relay optics, represents the most intuitive approach [4, 33]. Fiber bundles have been proposed as an attractive alternative, removing the need for more complicated relay optics [29]. Finally, curved sensors are difficult to fabricate but can be approximated using piecewise planar segments [36]. In this paper, we propose a computational imaging approach to this problem: instead of optically relaying the curved focal surface onto a set of planar sensors, we record the LF behind a monocentric lens using a combination of planar sensors and MLAs. The focal surface curvature is then corrected in post-processing. Correcting optical aberrations with LFs has been proposed in prior art [14, 34], but we are the first to apply this concept for wide-FOV LF imaging.

Applications of Light Fields Most commonly LF cameras are used for post-capture refocus and depth sensing. Time-of-flight cameras [15] and structured light sensors also offer depth imaging capabilities, but these are active imaging techniques that only provide a limited range, especially for outdoor applications. Scanning LiDAR systems are an attractive alternative, offering long-range depth sensing over 360°, but these systems are currently still several orders of magnitude more expensive than other depth sensors. As a passive depth sensing technology, wide-FOV LF imaging could find applications in robotic vision, autonomous vehicles, and a range of 3D imaging scenarios.

Cinematic virtual reality is another area where wide-FOV cameras have recently been of great interest. Facebook's Surround 360 and Google's Jump [2] systems are only two of many examples that capture an omni-directional stereo panorama. These cameras only capture two views of each scene point, and not an LF. Lytro's Immerge rotates a planar array of 90° FOV cameras over multiple poses to cover a wide FOV, and its complexity precludes its use in size- and budget-constrained applications.

The proposed monocentric imager is compact – see Fig. 1 for size comparison[1] – and unlike fisheye lenses can capture sufficient parallax information to support 3D reconstruction and focus cues for emerging LF VR displays [16]. With a pair of imagers, or extrapolation or reconstruction algorithms, we expect realistic stereo vision to be possible.

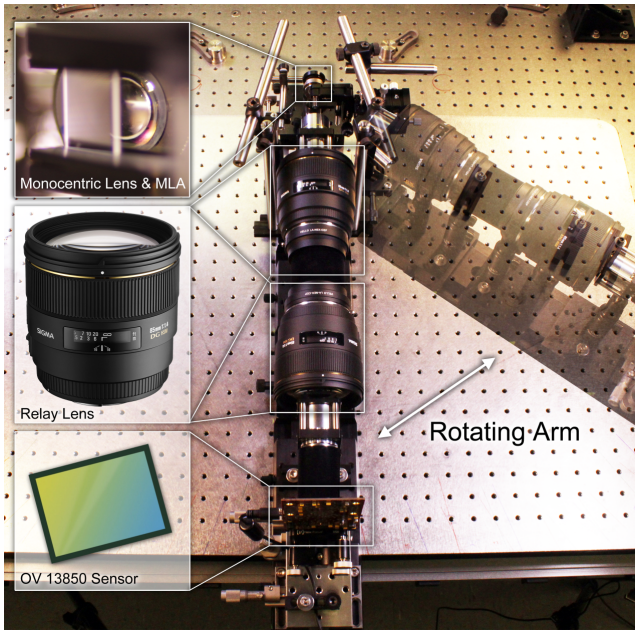


Figure 3. Prototype monocentric LF imager. An MLA is mounted directly behind the monocentric lens (top left). Currently, the sensor is not directly attached to the MLA, but placed conjugate to the MLA by a 1:1 optical relay system. This system is mounted on a mechanical arm that can be rotated around the fixed lens to emulate a multi-sensor LF imager.

3. Optical Design

The critical challenge in high-resolution wide angle imaging is mapping light incident from a wide range of angles onto a flat focal plane. Lenses are naturally well-suited to focusing spheres onto spheres, and so most of the glass in wide-FOV lenses is devoted to focusing planes onto planes.

A monocentric lens, composed of spherically concentric surfaces, "cheats" by forming its image on a spherical focal surface, as opposed to a plane. Most of the classical Seidel aberrations (astigmatism, field curvature, lateral color) are automatically canceled, and as few as two concentric glass shells are needed to form a high-resolution image over a field of 120° or more [29].

Employing a monocentric lens in its natural capacity as a sphere-to-sphere focuser is a key insight of this paper, enabling a wide FOV with the low f-number required for LF imaging. Note that this is a departure from the conventional focal regime of monocentric lenses, in which planes in the scene are focused to shifted spheres within the camera [29].

[1] Part of the bulk of the Lytro's lens is attributed to its variable zoom and focus, however even without these features this lens design is bulky and unable to capture a wide-FOV LF, as seen in Fig. 2.

3.1. Connecting Image Sphere to Tiled Sensors

The monocentric lens presents challenges on both the image and sensor side. On the image side, where we are used to focusing on a plane perpendicular to the optical axis and think about scene depth in terms of distance along the optical axis, we must now focus on a sphere and think in terms of radial distance to the lens center. Indeed, this requires only a shift in thinking, and for wide-FOV imaging makes much more sense than the planar approach.

On the sensor side the issue is deeper, as manipulating the sensor to meet the spherical focal surface is not yet practical. The wide-field image surface requires deeper spherical curvature than is possible by elastic deformation of a continuous CMOS focal plane chip. Spatially segmented CMOS sensor structures [35] limit resolution, and cutting a silicon chip into a lattice of discrete sensing areas interferes with the dense data transfer needed for imaging.

Our approach is therefore to tile sensors in an approximation of the sphere, as depicted on the right in Fig. 2. Similar approaches have been demonstrated in the past: a 2.4 Gigapixel imager used relay imaging of overlapping patches of the scene [4, 33] onto 221 16 Mpixel sensors. A smaller and less expensive imager has been made by coupling through an array of fiber bundle faceplates [30]. This 30 Mpixel prototype used a single row of six 5 Mpixel sensors. Both of these solutions are effective but costly and bulky. Fiber bundles also introduce a granularity that fundamentally limits the resolution of the system.

LF imaging provides a new option for coupling the image sphere onto planar sensors. The introduction of an MLA adjacent to the image sensor is well established as a means of recording both ray position and direction to support LF processing [24]. Here we propose to employ MLAs to take the place of the bulky and expensive couplers discussed above. This then allows an LF processing method, which we derive below, to effect the coupling process in software. This has the added bonus of offering LF capture, allowing post-capture refocus, depth information, and simplification of a range of vision tasks.

3.2. Multiple-Sensor Considerations

In tiling sensors to approximate the spherical focal surface we must consider design tradeoffs between the sensor area, the number of sensors required for given FOV, and the magnitude of the resulting field curvature. Smaller sensors approximate the sphere better, resulting in less severe field curvature, but requiring more sensors and accompanying electronics to cover a given FOV.

Fill factor is another issue: practical sensors cannot be tiled side-by-side due to the constraints of packaging – the current state of the art is a 50% fill factor [33]. One solution is to accept non-contiguous image capture, as would be acceptable in robotic navigation applications, for example. Another approach which we will explore as future work is to interpose a faceted field concentrator adjacent to the focal surface. This acts to divide the continuous image into subfields, consolidating the LF into discrete regions. Such approaches typically necessitate complex and bulky relay optics [33], but again LF processing promises to significantly ease the constraints on the sub-field optics, allowing an inexpensive and compact design.

Finally, focusing a multi-sensor setup will pose unique challenges, as no motion of the main lens can effect focus across all sensors when the FOV is wide. We envision either fixing the focus of a multi-sensor device, using variable focus across the sensors as achieved by adjusting the main lens position, or moving the sensors themselves using micro-actuators.

The key challenges addressed in this work are in coupling planar sensors with the monocentric lens through LF processing, and efficiently representing the data from multiple LF sensors through a shared main lens. Design of a multi-sensor system addressing fill factor and focus is ongoing and will be reported in detail elsewhere.

4. Light Field Processing

Here we introduce a parameterization and toolchain for interpreting the camera's images. As in conventional lenslet-based LF cameras, the raw data consists of closely packed lenslet images which multiplex the 4D LF onto the 2D sensor.

Approaches for rendering and filtering directly from raw lenslet images apply [11, 20]. These are efficient and benefit from irregular sampling patterns associated with real-world optics. However, many LF filtering and processing algorithms are more naturally described and applied in 4D, and so we present here a method for decoding the camera's raw sensor data into a 4D LF. This does not preclude later projection of algorithms back onto the 2D sensor for more direct operation.

The proposed approach requires only a flat-field image for calibrating the lenslet image centers, and metric calibration is left as future work. Here we discuss the parameterization and processing in two dimensions, with straightforward generalization to four.

4.1. Light Field Parameterization

Because the conventional two-plane parameterization (2pp) [19] does not handle wide FOVs well, we turn to spherical representations as a natural fit to what the camera captures. Several such parameterizations have been proposed [17, 32], but none addresses our needs exactly. Because they are scene-centric these representations do not efficiently describe the rays entering the camera, and for common tasks such as refocus or depth estimation more complex algorithms are required than with 2pp LFs.

We therefore propose the camera-centric relative spherical parameterization depicted in Fig. 4a. Whereas past work has conventionally enclosed the scene with a reference surface, here we enclose the camera, centering a reference sphere of radius R on the monocentric lens. Each ray is parameterized by its two points of intersection with the reference sphere. The first point, blue in Fig. 4a, is described as an absolute angle θ_u. The image of this point across the sphere is given by $\tilde{\theta}_u = \theta_u + \pi/2$, and the second point, in green, is measured relative to this $\tilde{\theta}_u$, as shown in the figure.

It is a critical insight that for small θ_s and θ_u this representation is well approximated by a local absolute 2pp with an $\hat{s}$ plane through the origin and a $\hat{u}$ plane at distance R, as depicted in orange in Fig. 4a. The $\hat{}$ denotes a local representation that is rotated with θ_u, yielding an LF that locally behaves as though it were planar, but that is radially symmetric and elegantly accommodates a wide FOV.

As in a conventional 2pp LF, a point on a Lambertian surface appears as a constant-valued plane in the LF, with slope related to the object's depth. The geometry of this is depicted in Fig. 4b. From the small-angle approximation and the similar triangles highlighted in yellow, we have

$$\theta_u \approx \hat{u}, \qquad \theta_s \approx 2\hat{s}. \tag{1}$$

For an object at radial distance r and the similar triangles highlighted in cyan,

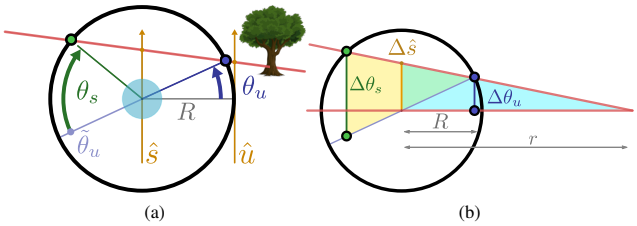
$$\frac{\Delta\theta_u}{r - R} = \frac{\Delta\hat{s}}{r}. \tag{2}$$

Substituting θ_s from (1) and rearranging yields the slope

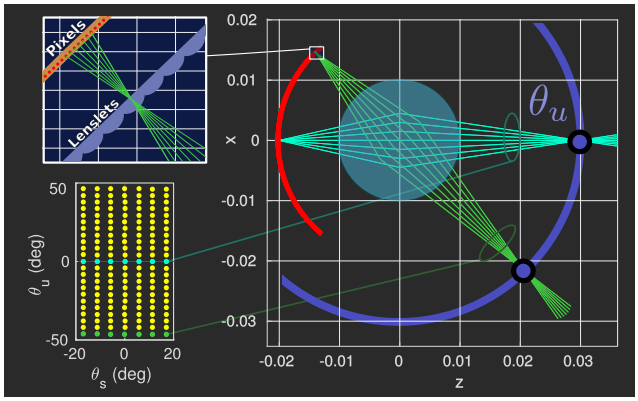
$$\frac{\Delta\theta_u}{\Delta\theta_s} = \frac{1}{2}(1 - R/r), \tag{3}$$

which is identical to the 2pp counterpart up to the factor $1/2$. Because of the strong local behavioural similarity, many conventional LF algorithms including depth estimation and refocus will apply directly to the camera-centric relative spherical representation, provided the factor of 2 in the $\hat{s}$ dimension is accounted for. The only functional difference is that depth in the spherical LF is expressed as a radial distance to the center of the lens, rather than along a fixed vector z.

Another important feature of the proposed representation is that it fits our optical setup very well: an ideal monocentric camera with spherical sensor and MLA measures a regular grid of samples in θ_s, θ_u, centered at $[0, 0]$. If we set the reference sphere radius to match the main lens' focus distance, the samples will be on a rectangular grid. This is demonstrated with a 2D raytracer in Fig. 4c, for which each pixel in an ideal spherical camera casts a single ray into the scene and is registered into the proposed parameterization. The figure depicts only two bundles of rays in the spatial



(a) (b)



(c)

Figure 4. (a) The camera-centric relative spherical parameterization, with monocentric lens at the center of the reference sphere, absolute entry angle θ_u, and relative exit angle θ_s; (b) Geometry behind the point-plane correspondence (3); (c) Raytracer demonstrating the sampling pattern of an ideal spherical camera to be very close to a rectangular grid in the proposed parameterization; for clarity rays from only two lenslets are shown in the diagram at right, and θ_s is not depicted.

diagrams, in cyan and green, but all pixels are shown in yellow in the ray-space diagram on the bottom-left, demonstrating an ideal rectangular sampling grid.

4.2. Tiled Planar Sensors

Our optical design tiles planar sensors and MLAs to approximate the ideal spherical elements described above. This causes the camera's sampling pattern to deviate from the ideal following two distinct effects: first, the planar geometry shifts lenslet images outward on the sensor compared to the spherical counterpart, and affects their scaling, following the tangent of the angle to the optical axis. This effect is easy to model and correct during 2D processing of the raw imagery, and is not strong for small sensors.

The more important effect is Petzval field curvature, which dramatically warps the LF sampling geometry. Fig. 5 depicts the issue, in which the focal surface in the scene is nearly parabolic, rather than the desired sphere – note that this differs from the usual interpretation of field curvature, in which the ideal is a plane, not a sphere. Starting at the center of a sensor, the main lens at distance a_0 from the centermost lenslet brings a point in the scene in focus at a distance b_0, as desired. Moving some distance x towards

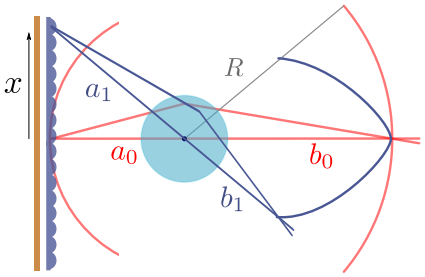


Figure 5. Geometry of field curvature: the ideal camera has a spherical MLA and focal surface (red), while a planar MLA results in a curved focal surface resembling a parabola described by (4) (blue).

the edge of the sensor, the lenslets-to-main lens distance increases to a_1, resulting in a reduced focal distance b_1. Tracing through x reveals a parabola-like shape, and this impacts both depth of field and the sampling pattern of the LF.

To describe this focal surface we set the reference sphere radius $R = b_0$, so that objects at the center of the FOV at distance b_0 are sampled ideally at a slope $\Delta\theta_u/\Delta\theta_s = 0$ as shown in Fig. 4c. Moving some distance x towards the edge of the FOV, the focus distance decreases, giving the sampling pattern a negative slope. This causes objects that should exist at slope 0 to take on positive apparent slopes. The sampling pattern slope can be derived by applying Pythagoras' theorem, the thin lens equation applied radially, and the point-plane correspondence (3),

$$a(x)^2 = x^2 + a_0^2, \qquad b(x)^{-1} = f^{-1} - a(x)^{-1}, \tag{4}$$

$$m(x) = \frac{1}{2}\left(1 - \frac{b_0}{b(x)}\right) = \frac{f}{2}\left(\frac{\sqrt{(1 + x^2/a_0^2)^{-1}} - 1}{a_0 - f}\right), \tag{5}$$

where f is the focal length of the main lens, a and b are the sensor-side and object-side focal distances as depicted in Fig. 5, and m is the sampling pattern slope described above. Then the actual sampling locations θ'_u can be expressed in terms of the ideal locations following

$$\theta'_u = \theta_u + m(x)\theta_s. \tag{6}$$

Note that in generalizing to 4D, x must be replaced with the radial distance from the center of the sensor.

4.3. Decoding and Field Flattening

We now have the tools required to coregister the multiple sensors of our optical design into a single, convenient representation. Our decoding process closely follows that of a conventional lenslet-based LF camera, building on the Light Field Toolbox for MATLAB [7]. That process consists of locating lenslet centers using a flat-field image, slicing the LF into a 4D structure, then applying a 1D interpolation to correct for hexagonal packing of the MLA.

We accommodate variation in lenslet image spacing across the FOV by fitting the lenslet grid model to the central region of the flat-field image. Then we allocate extra space in the sampled θ_s, θ_t dimensions to allow the lenslet images to drift outward towards the edge of each sensor. To register images into a common LF it is most convenient to fix the grid model spacing between sensors, ensuring a consistent sampling pattern.

The geometry of field curvature affects the slope of the sampling pattern independently of the lenslet image shift. As such, we can address the field curvature without further processing. To this end we correct the sampling pattern by interpolating the ideal locations θ_u given the warped locations θ'_u. Note that this is ideally done in the desired θ_u direction, but a good and computationally efficient approximation is to interpolate along the existing θ'_u dimension. Thus we have reduced field curvature correction to 2D interpolations applied across the θ_u, θ_v slices.

Higher order effects impact the lenslet image locations and the sampling pattern of our design. The monocentric lens is close to ideal, and so we neglect for now higher-order aberrations it might cause in the sampling pattern. Inclusion of an external iris and relay optics in our optical prototype have a stronger impact, chiefly in shifting lenslet image centers onto a nonuniform grid. Future prototypes will omit the relay and external iris, and so we address the lenslet image shifts with a pre-rendering step that fits a data-driven smoothed grid model to the flat-field image. This is then used to shift the lenslet images onto a common center prior to rendering.

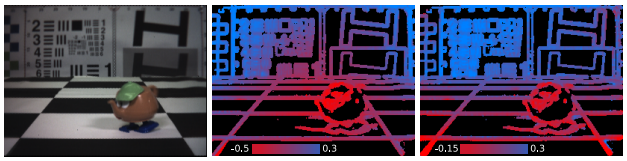
5. Results

5.1. Optical Prototype

We constructed a prototype monocentric LF camera consisting of a monocentric lens and a single sensor and MLA. Although we envision eventually bonding the MLA directly to the sensor, optical bonding is a precision mechanical task which we did not attempt for this prototype. As such, we employ an optical relay system, as illustrated in Fig. 3.

The lens used in our prototype system is a 12 mm focal length F/1.35 achromat. It is made of a 3.6 mm radius sphere of S-LAL13 glass, surrounded by a symmetric shell of S-LAH79 glass with 7.15 mm outer radius. The elements are ground and polished as hemispheric pairs, and assembled with a 4.4 mm aperture at the center. The resulting lens appears to be a solid ball of glass 14.3 mm in diameter. The lens is characterized in [30], showing MTF25 of 300 lp/mm over a full 120° FOV. Such performance has never been achieved with a conventional lens.

We use an OmniVision OV13850 CMOS image sensor with 13 MPix and 1.12 µm pixel pitch. The MLA is the same as that used for the Lytro Illum LF camera: an $f =$



(a) (b) (c)

Figure 6. Validating field flattening: (a) example input, (b) slope estimate before correction showing noticeable radial bulging, and (c) field-flattened image showing the correct depth behaviour.

40 µm F/2.0, hexagonally packed array with 20 µm lenslet diameter.

Note that the f-number of the experimental system is currently limited by the microlenses, not the main lens. An external iris is therefore introduced on the scene side of the lens to effectively match the monocentric lens f-number to that of the MLA. This is mounted as close as possible to the main lens so as to limit its impact on the LF geometry.

The MLA, optical relay, external iris and sensor are mounted on a mechanical arm that rotates around the stationary monocentric lens to emulate a multi-sensor setup. The arm swings through a 120° arc, and so with a per-frame FOV of about 24° the device can capture over 140° panoramas, neglecting vignetting. By allowing the external iris to rotate with the rest of the arm, we prevent it from causing vignetting. The external iris emulates the lens design in [1], in which total internal reflection from an internal convex surface provides a perfectly angle-invariant aperture and response. We plan to adopt such a lens in follow-on work.

5.2. Field Flattening

To confirm the efficacity of the field flattening method we tested it on single-frame LFs collected over a range of focal depths between 0.5 and 4 m. Distortion due to field curvature is most easily seen when animating a perspective shift, appearing as a bulging near the center of the frame. In print, it is more clearly seen in slope estimates, appearing as a bulging towards the camera near the center of the frame.

Fig. 6a depicts an example input scene, focused at a depth of 0.8 m, and visualized at the center of the s, t sample range. Following (6), the distortion is not visible in this central view, but the impact on the slope estimates in the uncorrected image, shown in Fig. 6b, is clear. The corrected image Fig. 6c shows the expected slope behaviour – we are correcting to a spherical isoslope, not planar, thus the slight nonplanarity on the back surface.

The slope estimates shown here are taken from the direction of the gradient of the LF [6,21], and are related to depth via the point-plane correspondence (3). Estimating metric depth requires calibration, which is left as future work.

5.3. Panoramas

We used the optical prototype to capture LF panoramas in indoor and outdoor scenarios – examples are shown in Figs. 7 and 8. These were captured by rotating the sensor and MLA about the monocentric lens, as described in Sect. 3, to mimick a multi-sensor setup. For the panoramas shown here we captured 11 frames with the arm angle spanning 120° in 12° increments. Each frame has about a 24° FOV, and so this resulted in a 144° panorama, but because of vignetting near the edges this was reduced to 138° – the vignetting is visible in Fig. 7.

In an ideal multi-sensor setup, contiguous sensors ensure full coverage of the scene. Because of the difficulty in repeatably achieing contiguous capture with our single-sensor setup, we captured overlapping views and employed the overlap to render seamless panoramas.

The panoramas have an effective pixel count of $15 \times 15 \times 1600 \times 200$, i.e. 72 MPix. The lenslets nominally have a diameter of 17.5 pix, and so fine tuning of the iris and relay optics should allow closer to 92 MPix in future experiments.

The indoor panorama shown in Fig. 7 occupies a depth range of 0.4–6 m, with the back wall at 2 m from the camera near the center of the panorama, and the camera focused at 0.8 m. The slope estimate (center) shows strong depth information throughout the LF, and the refocus results (bottom) demonstrate clear depth selectivity.

The outdoor panorama shown in Fig. 8 has similar geometry, but was captured over a variable focal depth, shifting from 4 m to 1 m to capture the stuffed toy to the right of the scene. A true multi-sensor setup could handle such a shift in focus by translating individual sensors and lenslets relative to the main lens. Some blending seams are evident, especially near the stuffed animal. This is due in part to defocus blur near the edges due to the closer focal setting for that frame. Per-sensor exposure is also enabled in a multi-sensor setup, and we emulate this by varying exposure over 4 settings spanning 2–6 ms over the panorama.

The insets in Fig. 8 demonstrate depth selectivity for refocus (left- and right-most insets), and also resolution enhancement through interpolation (center insets). The later is the same operation as focus, applied on a finer sampling grid to benefit from geometric information spread over the LF. Ihrke et al. [18] refer to this as geometric superresolution, to distinguish it from methods employing aliased information – we anticipate superior results are possible using more sophisticated methods.

6. Discussion

We presented an optical design and toolchain for collecting wide-FOV panoramas from a compact, single-lens camera. Constructed from concentric glass spheres, the monocentric lens offers the radial symmetry required for this to

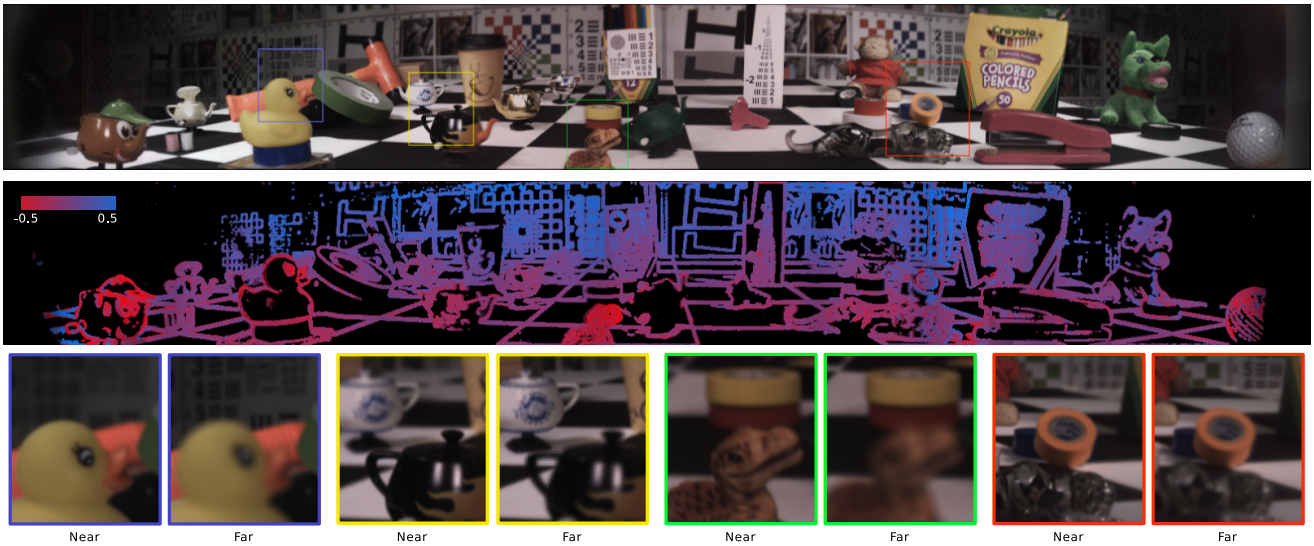


Figure 7. A 138° 72 MPix ($15 \times 15 \times 1600 \times 200$) LF panorama captured using our system: (top) A 2D slice of the 4D LF shown with (center) depth estimate and (bottom) detail views of near and far focus.

Figure 8. An outdoor scene employing per-sensor exposure and focus settings. This 138° LF is generated from 11 images over 4 exposure and 2 focal settings – the rightmost part of the scene places the toy animal in focus, resulting in defocus blur near the edges of that image's field. Insets show refocus and resolution enhancement using LF interpolation.

work efficiently. We showed that planar sensors can be coupled with the spherical focal surface of this lens in software, using LF processing techniques. We proposed a parameterization appropriate to the camera, and reduced field curvature correction to a set of 2D interpolations in this space.

We demonstrated an optical prototype capable of capturing 138° $15 \times 15 \times 1600 \times 200$ (72 MPix) panoramas, emulating multiple-sensor operation by rotating a sensor and lenslet array around a fixed moncentric lens. Indoor and outdoor examples demonstrate relative depth estimation, post-capture refocus, and LF superresolution. We also demonstrated extended dynamic range and depth of field by varying exposure and focus settings across exposures, emulating per-sensor control in a multi-sensor system.

This work presents the first steps towards practical, compact, wide-FOV LF capture. We anticipate applications spanning embedded vision including virtual and augmented reality, autonomous vehicles, and robotics in general.

The next step in this development is creating a compact multi-sensor rig similar to the one depicted in Fig. 1. It is difficult to achieve contiguous image capture by tiling sensors, and we anticipate simplifying this challenge by again employing low-cost optics paired with LF processing. For metric operation we anticipate exploring a suitable camera model and calibration scheme.

Acknowledgments We thank Kurt Akeley and Lytro for a hardware donation that enabled this work. This work is supported by the NSF/Intel Partnership on Visual and Experiential Computing (Intel #1539120, NSF #IIS-1539120). The authors thank Google ATAP for providing the Omnivision sensor interface, and Julie Chang and Sreenath Krishnan for their help with early optical prototypes. The monocentric lenses used in this work were fabricated within the DARPA SCENICC research program.

References

[1] I. P. Agurok and J. E. Ford. Angle-invariant imaging using a total internal reflection virtual aperture. *OSA Applied Optics*, 55(20):5345–5352, 2016. 7

[2] R. Anderson, D. Gallup, J. T. Barron, J. Kontkanen, N. Snavely, C. Hernández, S. Agarwal, and S. M. Seitz. Jump: virtual reality video. *ACM Transactions on Graphics (TOG)*, 35(6):198, 2016. 1, 3

[3] C. Birklbauer and O. Bimber. Panorama light-field imaging. In *Eurographics*, 2014. 1, 2

[4] D. J. Brady, M. E. Gehm, R. A. Stack, D. L. Marks, D. S. Kittle, D. R. Golish, E. M. Vera, and S. F. Feller. Multiscale gigapixel photography. *Nature*, 486:386–389, 2012. 1, 3, 4

[5] O. S. Cossairt, D. Miau, and S. K. Nayar. Gigapixel computational imaging. In *Computational Photography (ICCP)*, pages 1–8. IEEE, 2011. 3

[6] D. G. Dansereau and L. T. Bruton. Gradient-based depth estimation from 4D light fields. In *Intl. Symposium on Circuits and Systems (ISCAS)*, volume 3, pages 549–552. IEEE, May 2004. 7

[7] D. G. Dansereau, O. Pizarro, and S. B. Williams. Decoding, calibration and rectification for lenselet-based plenoptic cameras. In *Computer Vision and Pattern Recognition (CVPR)*, pages 1027–1034. IEEE, Jun. 2013. 6

[8] D. G. Dansereau, S. B. Williams, and P. I. Corke. Simple change detection from mobile light field cameras. *Computer Vision and Image Understanding (CVIU)*, 145C:160–171, 2016. 1

[9] F. Dong, S.-H. Ieng, X. Savatier, R. Etienne-Cummings, and R. Benosman. Plenoptic cameras in real-time robotics. *The Intl. Journal of Robotics Research*, 32(2):206–217, 2013. 1

[10] C. U. S. Edussooriya, D. G. Dansereau, L. T. Bruton, and P. Agathoklis. Five-dimensional (5-D) depth-velocity filtering for enhancing moving objects in light field videos. *IEEE Transactions on Signal Processing (TSP)*, 63(8):2151–2163, April 2015. 1

[11] J. Fiss, B. Curless, and R. Szeliski. Refocusing plenoptic images using depth-adaptive splatting. In *Computational Photography (ICCP)*, pages 1–9. IEEE, 2014. 4

[12] T. Gandhi and M. M. Trivedi. Vehicle mounted wide FOV stereo for traffic and pedestrian detection. In *IEEE International Conference on Image Processing 2005*, volume 2, pages II–121. IEEE, 2005. 1

[13] X. Guo, Z. Yu, S. B. Kang, H. Lin, and J. Yu. Enhancing light fields through ray-space stitching. *IEEE Transactions on visualization and computer graphics*, 22(7):1852–1861, 2016. 1, 2

[14] P. Hanrahan and R. Ng. Digital correction of lens aberrations in light field photography. In *OSA Intl. Optical Design*. OSA, 2006. 3

[15] M. Hasard, S. Lee, O. Choi, and R. Horaud. *Time of Flight Cameras: Principles, Methods, and Applications*. Springer, 2012. 3

[16] F.-C. Huang, K. Chen, and G. Wetzstein. The light field stereoscope: immersive computer graphics via factored near-eye light field displays with focus cues. *ACM Transactions on Graphics (TOG)*, 34(4):60, 2015. 3

[17] I. Ihm, S. Park, and R. K. Lee. Rendering of spherical light fields. In *IEEE Pacific Graphics*, pages 59–68. IEEE, 1997. 4

[18] I. Ihrke, J. Restrepo, and L. Mignard-Debise. Principles of light field imaging. *IEEE Signal Processing Magazine*, 1053(5888/16), 2016. 1, 7

[19] M. Levoy and P. Hanrahan. Light field rendering. In *SIGGRAPH*, pages 31–42. ACM, 1996. 4

[20] C.-K. Liang and R. Ramamoorthi. A light transport framework for lenslet light field cameras. *ACM Transactions on Graphics (TOG)*, 34(2):16, 2015. 4

[21] J. Luke, F. Rosa, J. Marichal, J. Sanluis, C. Dominguez Conde, and J. Rodriguez-Ramos. Depth from light fields analyzing 4D local structure. *Display Technology, Journal of*, 2014. 7

[22] K. Miyamoto. Fish eye lens. *Journal of the Optical Society of America*, 54(8):1060–1061, 1964. 2

[23] S. K. Nayar. Catadioptric omnidirectional camera. In *Computer Vision and Pattern Recognition (CVPR)*, pages 482–488. IEEE, 1997. 2

[24] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Stanford University Computer Science Technical Report CSTR*, 2, 2005. 1, 2, 4

[25] D. Scaramuzza and R. Siegwart. Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles. *IEEE transactions on robotics*, 24(5):1015–1026, 2008. 1

[26] A. Shademan, R. S. Decker, J. Opfermann, S. Leonard, P. C. Kim, and A. Krieger. Plenoptic cameras in surgical robotics: Calibration, registration, and evaluation. In *Robotics and Automation (ICRA)*, pages 708–714. IEEE, 2016. 1

[27] C. Silpa-Anan, R. Hartley, et al. Visual localization and loopback detection with a high resolution omnidirectional camera. In *Workshop on Omnidirectional Vision*. Citeseer, 2005. 1

[28] K. A. Skinner and M. Johnson-Roberson. Towards real-time underwater 3D reconstruction with plenoptic cameras. In *Intelligent Robots and Systems (IROS)*, pages 2014–2021. IEEE, 2016. 1

[29] I. Stamenov, I. P. Agurok, and J. E. Ford. Optimization of two-glass monocentric lenses for compact panoramic imagers: general aberration analysis and specific designs. *OSA Applied Optics*, 51(31):7648–7661, 2012. 1, 3

[30] I. Stamenov, A. Arianpour, S. J. Olivas, I. P. Agurok, A. R. Johnson, R. A. Stack, R. L. Morrison, and J. E. Ford. Panoramic monocentric imaging using fiber-coupled focal planes. *OSA Optics Express*, 22(26):31708–31721, 2014. 4, 6

[31] R. Szeliski. *Computer vision: algorithms and applications*. Springer, 2010. 2

[32] S. Todt, C. Rezk-Salama, A. Kolb, and K. Kuhnert. Fast (spherical) light field rendering with per-pixel depth. Technical report, University of Siegen, Germany, 2007. 4

[33] E. J. Tremblay, D. L. Marks, D. J. Brady, and J. E. Ford. Design and scaling of monocentric multiscale imagers. *OSA Applied Optics*, 51(20):4691–4702, 2012. 1, 2, 3, 4

[34] L.-Y. Wei, C.-K. Liang, G. Myhre, C. Pitts, and K. Akeley. Improving light field camera sample design with irregularity and aberration. *ACM Transactions on Graphics (TOG)*, 34(4):152, 2015. 3

[35] T. Wu, S. Hamann, A. Ceballos, O. Solgaard, and R. Howe. Design and fabrication of curved silicon image planes for miniature monocentric imagers. In *Intl. Conference on Solid-State Sensors, Actuators and Microsystems*, pages 2073–2076. IEEE, 2015. 4

[36] T. Wu, S. S. Hamann, A. C. Ceballos, C.-E. Chang, O. Solgaard, and R. T. Howe. Design and fabrication of silicon-tessellated structures for monocentric imagers. *Microsystems & Nanoengineering*, 2, 2016. 3

[37] Z. Xue, L. Baboulaz, P. Prandoni, and M. Vetterli. Light field panorama by a plenoptic camera. In *IS&T/SPIE Electronic Imaging*, pages 90200S–90200S. Intl. Society for Optics and Photonics, 2014. 1, 2