Light Fields and Computational Imaging

Marc Levoy
Stanford University

A survey of the theory and practice of light field imaging emphasizes the devices researchers in computer graphics and computer vision have built to capture light fields photographically and the techniques they have developed to compute novel images from them.

Discoveries in science are frequently triggered by the invention of new instruments, such as the telescope, microscope, or cyclotron. Arguably the most important scientific instrument of the past 50 years is the digital computer. Among its many uses, the coupling of computers with digital sensors has created a powerful new tool called "computational imaging."

From borehole tomography in geophysical exploration to confocal microscopy in the biological sciences, the use of computers during image formation has revolutionized our ability to observe and analyze the natural and manmade worlds. Many of these imaging methods operate at visible wavelengths, and many of those relate to the flow of light through space.

Although the notion that light flows through an environment dates back to ancient times, Michael Faraday was the first to propose, in an 1846 lecture titled "Thoughts on Ray Vibrations," that light should be interpreted as a field. Faraday's proposal, based on his previous work in magnetism, was a good one, but being an experimentalist rather than a mathematician, he couldn't formalize his ideas.

James Clerk Maxwell provided this formalization 28 years later through the equations for which he is famous. Combined with discoveries about the properties of light made by Pierre Bouguer, Johann Lambert, and others, these equations led to an outpouring of theoretical photometry work in the first half of the 20th century. Among the achievements was Subrahmanyan Chandrasekhar's seminal 1950 book, *Radiative Transfer*, about the transport and scattering of light. James Kajiya introduced this work to the computer graphics literature in 1986 in his widely cited paper.[1]

Among the photometry applications deemed useful at the beginning of the age of electricity was the study of surface illumination by artificial lighting. With this application in mind, Arun Gershun defined the light field concept, which gives the amount of light traveling in every direction through every point in space.[2] In his surprisingly readable 1936 paper, Gershun recognized that the amount of light arriving at points in space varies smoothly from place to place (except at well-defined boundaries like surfaces or shadows) and could therefore be characterized using calculus and analytic geometry. Writing before the age of digital computers, Gershun had no way to measure a light field. However, he could derive in closed form the illumination patterns observed on surfaces due to light sources of various shapes positioned above these surfaces.

With the advent of computers, color displays, and inexpensive digital sensors, we can now record, manipulate, and display Gershun's light field. Since light fields were introduced to the computer graphics field 10 years ago,[3,4] researchers have used them to fly around scenes without creating 3D models of them, to relight these scenes without knowing their surface properties, to refocus photographs after they've been captured, to create
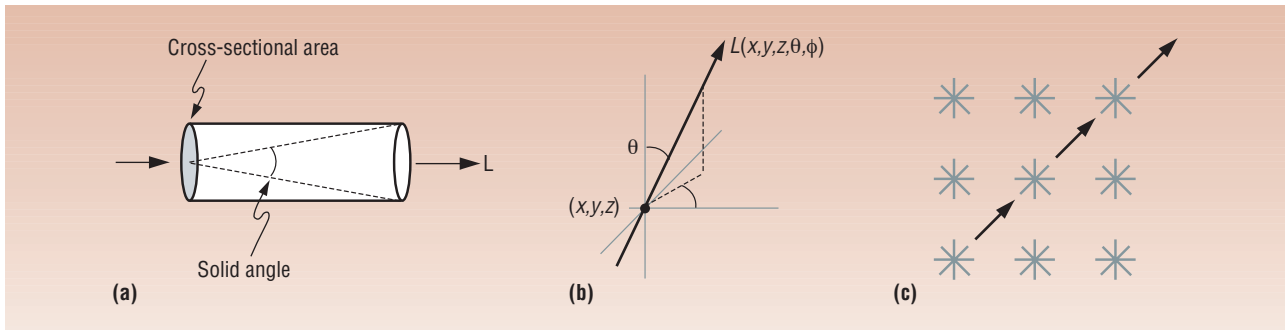
Figure 1. The 5D plenoptic function, representing the flow of light through 3D space. (a) Radiance L along a ray can be thought of as the amount of light traveling along all possible straight lines through a tube whose size is determined by its solid angle and cross-sectional area. (b) Parameterizing a ray by position (x, y, z) and direction (θ, ϕ). (c) Radiance along a ray remains constant if there are no blockers. This leads to redundancy in the plenoptic function.
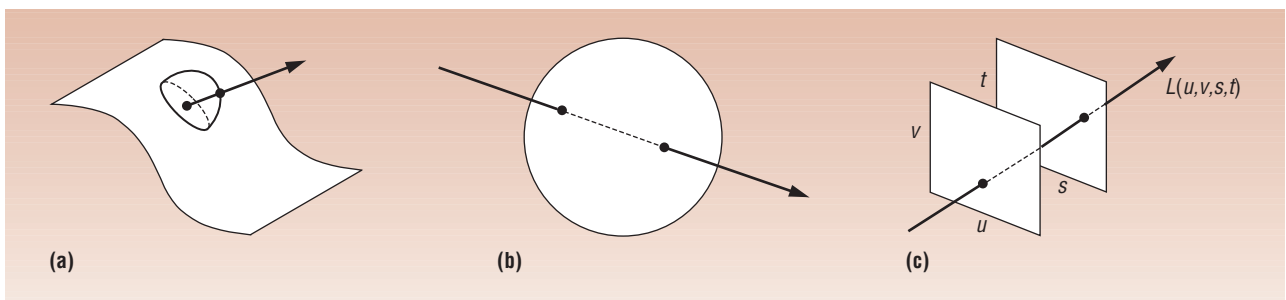

Figure 2. Alternative parameterizations of the 4D light field, which represents the flow of light through an empty region of 3D space. (a) Points on a plane or curved surface and directions leaving each point. (b) Pairs of points on the surface of a sphere. (c) Pairs of points on two planes in general (meaning any) position.

nonperspective panoramas, and to build 3D models of scenes from multiple images of them.

This survey of the theory and practice of light field imaging emphasizes the devices researchers in computer graphics and computer vision have built to capture light fields photographically and the techniques they've developed to compute novel images from them.

PLENOPTIC FUNCTIONS AND LIGHT FIELDS

This article focuses on geometrical optics—that is, spatially incoherent illumination—and on objects significantly larger than the wavelength of light. In geometrical optics, rays are the fundamental light carrier. The amount of light traveling along a ray is radiance, denoted by L and measured in watts (W) per steradian (sr) per meter squared (m^2). Steradians measure a solid angle, and meters squared are used here as a measure of cross-sectional area, as Figure 1a shows.

The radiance along all such rays in a region of 3D space illuminated by an unchanging arrangement of lights has been dubbed the *plenoptic function*.[5] Since rays in space can be parameterized by coordinates, x, y, and z and angles θ and ϕ, as Figure 1b shows, it is a 5D function.

If the region of interest contains a concave object (think of a cupped hand), then light leaving one point

on the object can travel only a short distance before another point on the object blocks it. We know of no device that can measure the plenoptic function in such regions. However, if we restrict ourselves to locations outside the object's convex hull (think shrink-wrap), we can measure the plenoptic function easily using a digital camera.

In this case, the function contains redundant information, because the radiance along a ray remains constant from point to point, as Figure 1c shows. In fact, the redundant information is exactly one dimension, leaving us with a 4D function that Parry Moon called the photic field[6] and Pat Hanrahan and I call the *4D light field*.[3] Formally, the 4D light field is defined as radiance along rays in empty space.

This 4D set of rays can be parameterized in a variety of ways, which Figure 2 shows. One option is to parameterize rays by their intersection with two planes in general position, as Figure 2c shows. While this parameterization can't represent all rays (for example, rays parallel to the two planes if the planes are parallel to each other), it relates closely to the analytic geometry of perspective imaging. Indeed, a simple way to think about a two-plane light field is as a collection of perspective images of the *st* plane (and any objects that may lie beyond it), each taken from an observer position on the
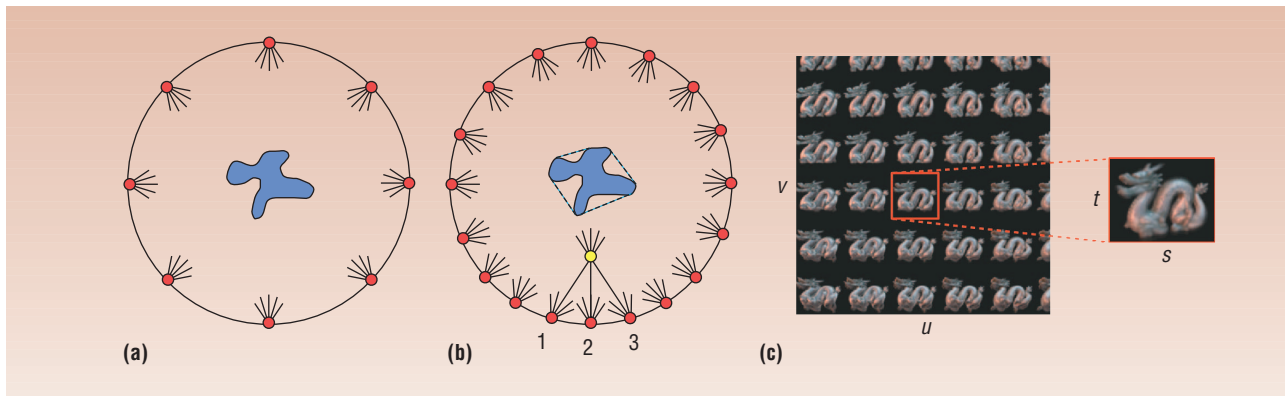
Figure 3. QuickTime VR versus light field rendering. (a) QuickTime VR's object-movie function lets the user fly around an object (blue shape) by flipping among closely spaced photographs of it (red dots). (b) If the dots are spaced closely enough, the user can re-sort the pixels to create new perspective views without having stood there (yellow dot); this is light field rendering. (c) A light field can be interpreted as a 2D collection of 2D images, each taken from a different observer position.

uv plane. This interpretation brings us into the realm of photography, which in turn brings us to consider some of the uses for photographically captured light fields.

Light field rendering

One use falls under the paradigm of *image-based rendering*, a family of techniques invented primarily during the 1990s for conveying an object's shape on a computer display using previously captured images instead of a 3D geometric model. Consider the situation diagrammed in Figure 3a. We place an object, let's say a terra-cotta dragon, at the center of a sphere 6 feet in diameter. We then move a camera to 100 positions distributed across the sphere's surface, and photograph the dragon at each position. As long as the sphere is large enough to not intersect the dragon's convex hull, the collection of images is a 4D light field, albeit coarsely sampled. Flipping quickly among these images gives the impression of orbiting around the dragon, or of standing in one place while the dragon is turned every which way.

Proposed by Eric Chen in 1995,[7] this idea provided the basis for the object-movie function in Apple's proprietary QuickTime VR system. With this function, the user can fly around, but not toward, the dragon, and magnify the images without a change in perspective. Neither the relative sizes of features nor the occlusions (what blocks what) change.

If the collection of positions is denser, perhaps a thousand distributed across the sphere's surface, then we can generate enough pixels to fly toward the dragon. For example, while standing at the yellow dot in Figure 3b, the central pixel (or equivalently, ray) in this view of the dragon is the same as the central pixel in Photograph 2. More interestingly, the rightmost pixel in this view is identical to some pixel in Photograph 1, and the leftmost pixel is identical to some pixel in Photograph 3.

Thus, if the set of original observer positions is dense enough, by selecting among the pixels, possibly with interpolation among nearby pixels, you can construct new, perspectively correct views from observer positions where you never stood. In fact, you can stand anywhere you want, as long as you stay outside the dragon's convex hull (the dashed lines around the blue shape in the figure).

This idea is called *light field rendering*.[3] Stated more formally, a light field can be interpreted as a 2D collection of 2D images of a scene—hence, a 4D array of pixels, as Figure 3c shows. Computing a novel perspective view of the scene can be interpreted as extracting an appropriately positioned and oriented 2D slice from this 4D array.

How many images does light field rendering require? A light field we captured in 1999 of Michelangelo's *Night* in Florence's Medici Chapel is the largest I know of, containing 24,000 1.3-megapixel images (http:// graphics.stanford.edu/projects/mich/lightfield-of-night). In our lab, we routinely capture light fields of 1,000 megapixel images, and we use our camera array to capture light field videos, each frame of which contains 100 VGA-resolution images.

At a deeper level, the answer to this question depends on where you'd like to stand after capturing these images. If you want to walk completely around an opaque object, then you need to photograph its back side. Less obviously, if you want to walk close to the object, you need images taken at finely spaced positions on the sphere's surface (which is now behind you), and these images need to have high spatial resolution. The number and arrangement of images and the resolution of each image are together called the "sampling" of the 4D light field.

Many researchers have analyzed light field sampling.[8,9] According to their findings, if the images don't have enough pixels, the light field renderings will be blurry, especially as you move away from the original observer positions. If you don't take enough images, the renderings will
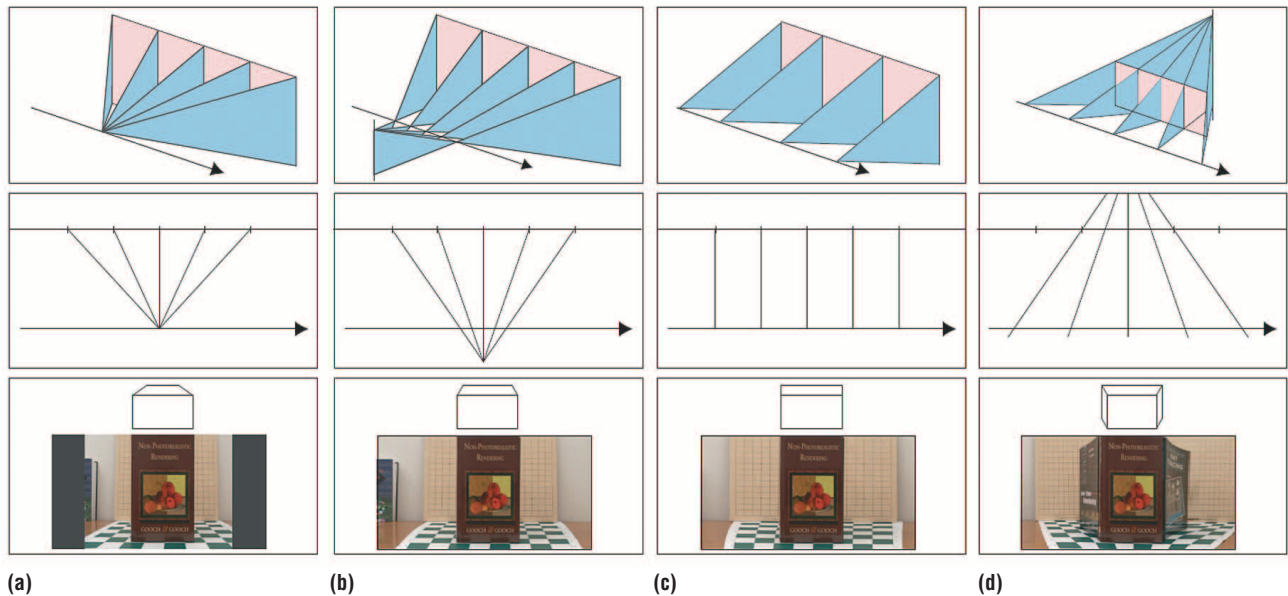
Figure 4. Crossed-slits projections, formed by the intersection with a picture plane of the set of rays passing through two lines in general position. From top to bottom are diagrams in 3D and 2D of the lines of sight, thumbnail drawings of the perspective induced on a simple cube, and renderings computed from a light field of three books arranged in a square and standing on a checkerboard. (a) Perspective. If the two slits are coincident, we obtain an ordinary perspective view. (b) Crossed slits. Moving the slits apart, we obtain a view in which the perspective is different in the horizontal and vertical directions. (c) Pushbroom. Moving one slit to infinity produces a pushbroom panorama, which is perspective vertically but orthographic horizontally. (d) Inverted slits. Placing the slits on opposite sides of the picture plane, and placing the object astride the picture plane, produces an inverted crossed-slits projection. Note the unnatural appearance of the checkerboard. Note also that we can see the books on both sides of the square at once.

contain ghosts arising from blending different views of an object. If, however, you augment the light field with an approximate 3D geometric model of the object, many fewer images are needed.[4]

Taken to an extreme, you can reconstruct an accurate model of the object from a handful of images, then fly around the object by rendering this model.[10] That approach and light field rendering represent two ends of a continuum of rendering techniques, which is indexed by the amount of geometric information known about the scene.

Multiperspective panoramas

Why limit ourselves to perspective views? Linear perspective, invented in 1413 by Renaissance artist Filippo Brunelleschi, is defined as the intersection of a plane and the set of rays passing through a point. The rays are the lines of sight, the point is the observer position, and the plane is the surface of the canvas, or more generally the "picture plane." In photography, the picture plane is the film or sensor chip, and the effective observer position lies at the center of the first principal plane of the camera's optical system, usually buried somewhere inside the lens system.

A simple variant on linear perspective is to move the observer infinitely far from the scene—a sort of supertelephoto view. The lines of sight become parallel, there is no perspective distortion, and occlusions don't change as the observer moves sideways relative to these lines. This is called an *orthographic projection*. While it's unusual to find optical systems other than the microscope that capture orthographic projections, it's easy to compute one using light field rendering and an input light field with an assortment of available lines of sight.

Moving away from projections in which all rays pass through a single point, suppose we replace the lens in a digital camera with a pair of masks, one containing a horizontal slit and the other containing a vertical slit. With this arrangement, the camera records a view in which the lines of sight for each column of pixels converge to a point on the horizontal slit and the lines of sight for each row of pixels converge to a point on the vertical slit.

Invented in 1888 by color photography pioneer Louis Ducos du Hauron, images like this are called *crossed-slits projections*.[11] As Figure 4 shows, moving the slits produces a variety of unusual perspectives. Even wilder camera models have been proposed, but building a complete taxonomy is an open problem and outside the scope of this article.

All these projections can be computed by extracting slices from a light field. As an example, suppose you drove down a city street, pointed a video camera out the side window, and recorded the storefronts you passed.
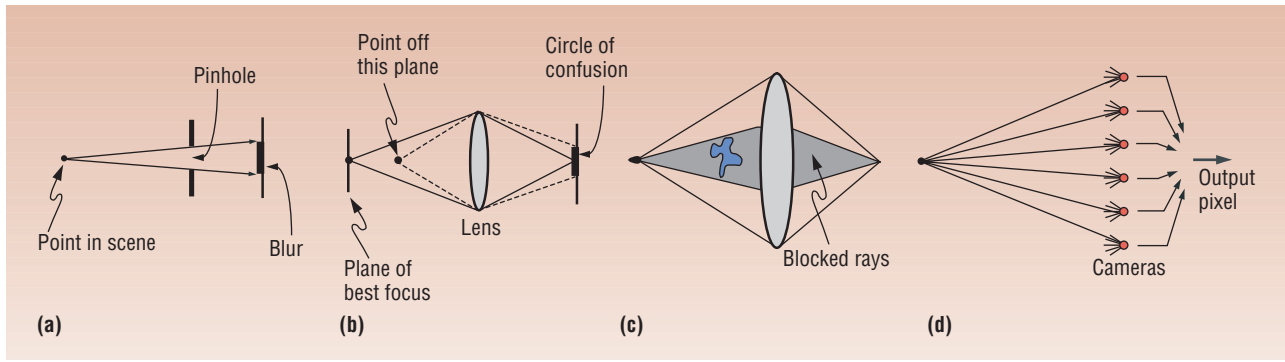
Figure 5. The principle of synthetic aperture photography. (a) A pinhole camera creates a blur on the picture plane. (b) Adding a lens admits more light and focuses it, but only points lying on one plane are sharply focused; points off this plane image as a blur called the circle of confusion. (c) If the lens is larger than an occluding object (in blue), although some rays are blocked, the object doesn't completely obscure your view of points on the plane of best focus. (d) Discretely approximating a large aperture by adding rays extracted from the views seen by an array of cameras.

If you extract the center column of pixels from each frame of video and abut these columns together horizontally, you obtain a pushbroom panorama in which one slit is the path of the camera down the street and the second slit is vertical and placed opposite the storefronts and infinitely far away. You don't actually need a 2D collection of images to construct such a panorama; since you're limiting the vertical perspective to converge on the camera's path, a 1D collection suffices.

Unfortunately, pushbroom panoramas compress objects that are close to the camera and stretch distant objects. To minimize these distortions, two or more different projections can be combined in a single image. This produces a *multiperspective panorama*. In the Stanford CityBlock Project, we computed an orientation for each fan of rays (blue triangles in Figure 4) that locally minimizes this distortion.[12]

In work concurrent to our own, Aseem Agarwala addresses the same problem by semiautomatically segmenting the panorama into regions, each of which is extracted from a different input image.[13] Automatically generating multiperspective panoramas is challenging, so this will undoubtedly be an active research area for a long time.

Synthetic aperture photography

In the previous two sections, each pixel in a computed image represented a unique line of sight—hence, a single sample extracted from the light field. Of course, real cameras don't work this way, or they would capture infinitely little light. A real camera has a finite-size aperture.

This rule even applies to the pinhole cameras that elementary school students make by poking a hole in the side of a shoebox and looking inside for the image it formed. Named *camera obscura* by Johannes Kepler in 1604, the pinhole camera has been known since antiquity. As Figure 5a shows, the larger the pinhole, the more light it admits, but the blurrier the image becomes.

This problem can be addressed by placing a lens in the pinhole, as Figure 5b shows. The device's light-gathering power is unchanged, but now objects at one particular distance from the lens will be well focused. Objects at other distances—not lying on a "plane of best focus"—will be imaged as a blur, sometimes called the *circle of confusion*. If the object lies far enough from this plane that the circle of confusion is larger than some nominal diameter (typically a pixel), we say that the object lies outside the camera's depth of field.

As photographers know, introducing an aperture stop (diaphragm) into such an optical system and partially closing it reduces the effective diameter of the lens. This shrinks the circle of confusion for objects off the plane of best focus, hence increasing the camera's depth of field. Conversely, if you open up the diaphragm, you expand the circle of confusion, thereby decreasing its depth of field.

If the aperture is made extremely large, let's say as wide as the distance to the plane of best focus as shown in Figure 5c, the depth of field becomes so shallow that only objects lying on that plane are sharp. Interestingly, if an object lying outside the depth of field is small enough that for every point on the plane of best focus, at least some of its rays still reach the lens, the object no longer obscures the camera's view of these points.

Five hundred years ago, Leonardo da Vinci observed that if you hold a needle in front of your eye, since the needle is narrower than the pupil of the eye, it adds a haze to your view of the world, but it does not completely obscure any part of it.

An obvious application of this principle is to "see through" objects consisting of many small parts, like trees or crowds of people. It's inconvenient to build a camera with a lens that is larger than a tree leaf, not to mention a person, but we can simulate such a camera by capturing and resampling a light field. For example, if we have an array of $N \times N$ cameras pointing at a scene,
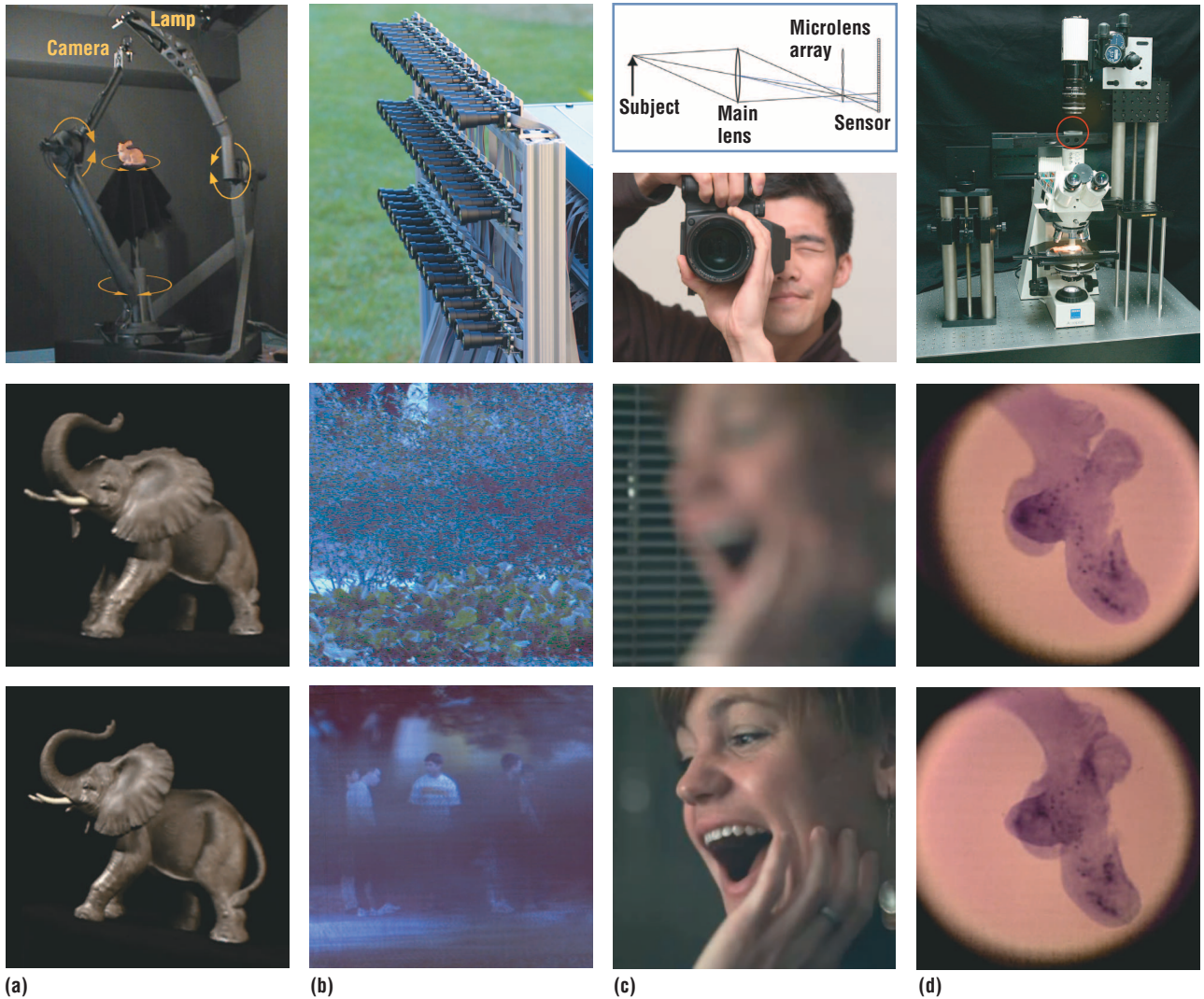
(a)   (b)   (c)   (d)

Figure 6. Devices built in the Stanford Computer Graphics Laboratory for capturing light fields. (a) Spherical gantry with four motorized motions (orange arrows). The inner arm typically holds a detector or camera, the outer arm holds a light source or video projector, and the object sits on the central platform. Below are two frames from a light field captured using the gantry (http://graphics.stanford.edu/projects/gantry). (b) Multicamera array, consisting of 128 VGA-resolution cameras with telephoto lenses (48 were used here). Below is the view from one camera, and a synthetic aperture photograph created by summing the views from all cameras, allowing us to see through foliage. (c) Plenoptic camera, in which a microlens array has been inserted between the main lens and digital sensor of a Mamiya medium-format SLR. The optical design is shown at top (see text for details). Below are two synthetic refocusings of a snapshot taken by the camera. (d) Light field microscope (LFM), in which a microlens array (red circle) has been placed at the intermediate image plane of a standard microscope. Below are two perspective views of an embryo mouse lung, computed from one snapshot. Specimen from Hernan Espinoza.

we can simulate the focusing effect of a lens as large as the array in the following way.

Consider a single pixel in the output image. Using geometrical optics, calculate the point's location on the plane of best focus that would be imaged onto this pixel by the giant lens. Now select the image sample from the view recorded by each camera, possibly with interpolation from neighboring samples, whose line of sight passes through that point. Add together these $N \times N$ samples, as shown in Figure 5d. Repeat this procedure for each of the $P \times P$ pixels in the output image.

Thus, after work proportional to $N^2 \times P^2$, we have constructed a perspective view of the scene, but using a synthetic camera having a large aperture and therefore a shallow depth of field. Aaron Isaksen and colleagues[9] describes this process as "reparameterizing the light field"; I prefer to call it synthetic aperture photography or "digital refocusing." Figure 6 shows some images computed in this way.

DEVICES FOR CAPTURING LIGHT FIELDS

Having surveyed some computational techniques

researchers have applied to light fields, let's shift gears and talk about devices that have been proposed for capturing them. (Light fields can also be created by rendering images from 3D models, but our focus here is on photography.) In most cases, the density with which we can sample a light field depends on the device we employ. This sampling density, and the physical scale of the device (room-size versus microscopic), limits or enables specific applications of the computational techniques we have been surveying.

Moving cameras

Let's start by assuming the range of viewpoints to be captured spans a long baseline (from feet to miles). For static scenes, we can capture a light field by moving a single camera through the scene. Examples in which the camera translates across a plane include our original work on light field rendering[3] and the Digital Michelangelo Project.[14] Examples in which the camera moves across the surface of a cylinder or sphere include the inward-looking camera Apple built to construct QuickTime VR data sets,[7] a similar, but more precise, gantry we built in our lab (Figure 6a), and an outward-looking system Microsoft Research/China developed to construct "concentric mosaics."

Light fields can also be constructed using a handheld camera, assuming that the camera's pose (position and direction of view) can be estimated.[15] In the Stanford CityBlock Project,[12] we used optical flow algorithms from the computer vision literature or, alternatively, sensors fixed to the camera, for this task.

Arrays of cameras

You need multiple cameras to capture long-baseline light fields of a dynamic scene. These can be film cameras, digital still cameras, or video cameras. The latter are better for capturing critical moments in a fast-moving event, since the array can free-run until the critical moment occurs. If the cameras are arranged along a 1D path, then displaying the views in rapid succession gives an impression of orbiting around a scene that has been frozen in time.

Pioneered by Dayton Taylor, this technique was made famous by the 1999 movie *The Matrix*. To our knowledge, imagery from these systems has never been fed into a light field viewer, where pixels from different images could be combined to generate new views. Doing so would let the virtual observer move toward the objects being imaged, rather than only along the camera's path, but the new views would exhibit horizontal parallax only.

If the cameras are arranged in a 2D array, then a full light field is captured. As Figure 6b shows, we have built such an array, which in addition to capturing video light fields can also capture ultrahigh speed video by staggering the cameras' triggering times, high-dynamic-range video by varying their exposure times, or high-resolution panoramas by splaying their direction of view.[16]

Other systems we know of are the 3D Room at Carnegie Mellon University, a 16-camera array built by the University of Tokyo, and a 64-camera array built by the Massachusetts Institute of Technology's Computer Graphics Group.

Arrays of lenses

If the range of viewpoints spans a short baseline (from inches to microns), then we can replace the multiple cameras with a single camera and an array of lenses. The use of lens arrays to capture light fields has its roots in Gabriel Lippman's 1908 invention of integral photography. The operating principle behind these arrays is simple. If you place a sensor behind an array of small lenses (lenslets), each lenslet records a perspective view of the scene observed from that position on the array. This constitutes a light field, whose uv resolution, as Figure 2c shows, depends on the number of lenslets, and whose st resolution depends on the number of pixels behind each lenslet.

Placing a "field" lens on the object side of the lenslet array, and positioning this lens so that the scene is focused on the array as shown in the Figure 6c diagram, transposes the light field; now its st resolution depends on the number of lenslets and its uv resolution on the number of pixels behind each lenslet. The first arrangement has the advantage of being physically thin. However, the resolution of views computed from it will be low, so if system thickness is not an issue, the latter arrangement is preferred.

In this arrangement, only the field lens needs to be corrected for aberrations, not each lenslet. The superiority of this arrangement, combined with recent improvements in technology for manufacturing microlenses smaller than 1 mm, has led researchers to propose inserting microlens arrays between the sensor and main lens of a photographic camera, thereby creating a plenoptic camera.[17] We have built such a camera by modifying a Mamiya medium-format SLR body, as Figure 6c shows.[18]

Starting from the light fields recorded by a plenoptic camera, you can create perspective flybys and multi-perspective panoramas, although the range of available viewpoints is limited by the diameter of the camera's aperture. (It works best in macrophotography, where the scene is close to the camera and therefore large relative to the camera's aperture.)

> **Researchers have proposed inserting microlens arrays between the sensor and main lens of a photographic camera, thereby creating a plenoptic camera.**

More interestingly, you can perform synthetic aperture photography. This essentially allows the photographer to refocus a snapshot after it has been captured, as the figure shows. The tradeoff is a loss in spatial resolution.

Specifically, for a microlens array having $P \times P$ microlenses and $N \times N$ pixels beneath each microlens, we can compute views having $P \times P$ pixels, and if the camera's main lens has a relative aperture (F-number) of f/A, we can refocus these views anywhere within the range of depths that would be in focus if the camera's lens were stopped down to $f/(A \times N)$. For example, our prototype plenoptic camera has a 16-megapixel sensor, an $f/4$ main lens, and a microlens array with 300×300 microlenses. Thus, $P = 300$, $N = 14$, and we can refocus anywhere within the depth of field of an $f/56$ camera.

Unfortunately, the computed images are only 300×300 pixels, barely enough to be useful. However, the number of pixels in modern digital cameras continues to increase. If the sensor in a full-frame 35-mm digital camera is reengineered to have pixels as small as a point-and-shoot camera (about 2 microns) and employs an array with microlenses 20 microns on a side—that is, $N = 10$—it would be possible to compute images having $1,800 \times 1,200$ pixels and the refocusing range of an $f/40$ camera. Such a camera would undoubtedly find adherents.

Microscopes

Moving further down the scale of scenes we might image, if we place a microlens array at a microscope's intermediate image plane, we can capture light fields of microscopic specimens in a single photograph.[19] As in Lippman's original proposal, this *light field microscope* sacrifices spatial resolution to obtain angular resolution. Unlike the plenoptic camera, diffraction places an upper limit on the product of spatial and angular resolution in a microscope light field. The exact limit depends on the numerical aperture of the microscope objective lens. For readers familiar with photography, numerical aperture NA can be converted to F-number A using the approximate formula $A = 1/(2\ NA)$.

Despite this limit, we can produce useful light fields with this arrangement. From these, we can employ light field rendering to generate perspective flyarounds, at least up to the angular limit of rays we have captured. Since microscopes incorporate a special "telecentric" aperture stop that causes them to produce orthographic views, perspective views represent a new way for microscopists to look at their specimens.

Similarly, we can use synthetic aperture photography to produce a *focal stack*—a biologist's terminology for a sequence of images each focused at a different depth. Focal

stacks aren't new, but manual techniques for capturing them by moving the microscope stage vertically and capturing an image at each position are time-consuming and hence not applicable to moving (live) or light-sensitive specimens. Figure 6d shows our prototype and examples of the perspective views we can compute using it.

THE FUTURE OF LIGHT FIELDS

What opportunities exist for new research at the intersection of light fields, photography, and computational imaging?

First, every technique I've described in this article could benefit from better instrumentation. We especially need better ways to capture large collections (thousands) of viewpoints. We should also explore collecting light fields at very large scales (terrestrial) as well as very small scales (electron microscope). At both extremes, the trend has been away from optomechanical solutions and toward optoelectronic solutions. Improvements in our ability to run these systems at high speeds, or to trigger them in controlled ways, suggest that temporal multiplexing will become an increasingly useful strategy.

> The 3D structure of microscope light fields can be analyzed using algorithms for reconstruction from projections.

Slower progress has been made in the display of light fields. Although researchers have built autostereoscopic displays for light fields, even end-to-end 3D television systems (Wojciech Matusik and colleagues at Mitsubishi Electric Research Labs, and Bahram Javidi and colleagues at the University of Connecticut), this is a fundamentally harder problem than capturing a light field because interpolation between sparse samples can't be performed digitally, only optically. Nevertheless, slow but steady increases in display resolution, coupled with novel fabrication technologies, could lead to breakthroughs here as well.

Second, light fields can be used to reconstruct a 3D shape using computer vision algorithms. For example, *shape-from-stereo* operates by finding corresponding features in two or more views of a scene taken from different observer positions. For each correspondence, we can triangulate to determine the feature's 3D location. Alternatively, *shape-from-focus* examines a collection of views taken from one position but with varying focus. The depth associated with each pixel can then be determined by deciding which focus setting made that pixel appear sharpest. Unfortunately, occlusions make it hard to find corresponding features or to decide when an object is sharply focused.

However, having more images allows us to peek around occlusions, as Figure 6b shows. Therefore, it's easy to imagine that, given a light field instead of a small collection of images, we should be able to improve the
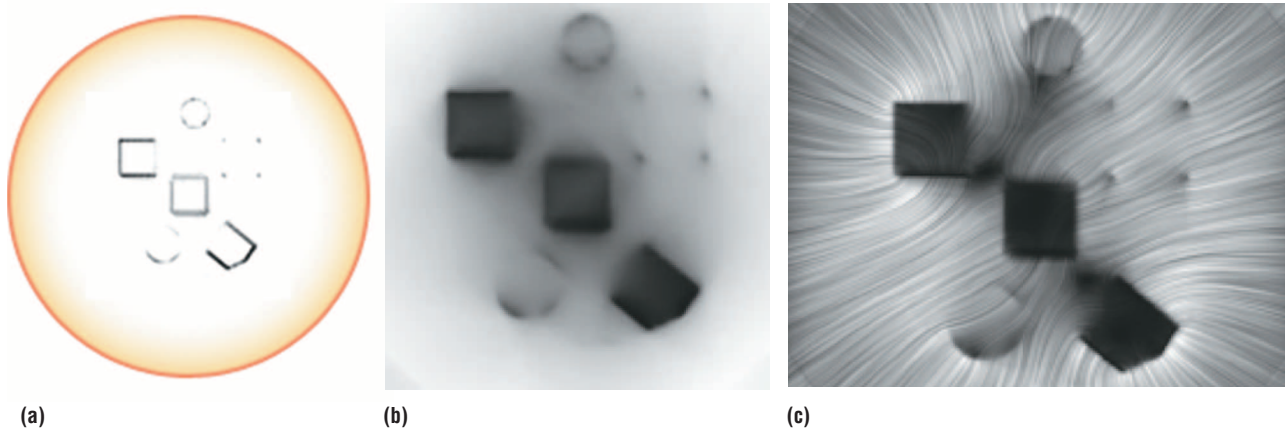
Figure 7. Visualization of a vector irradiance field in flatland. (a) The scene is a collection of points, lines, and arcs of varying opacities (darker means more opaque), some of which have been arranged to form closed figures. Illumination (yellow haze) impinges on the scene uniformly from all directions (red circle). (b) Total irradiance arriving at each point in the scene, taking into account occlusion by the lines and arcs. (c) The irradiance vector direction at each point, visualized using Brian Cabral's line integral convolution (LIC) technique. In 3D, this vector direction can be interpreted as the orientation for facing a flat surface placed at different points in a scene to most brightly illuminate it. Note the saddle between the two axis-aligned squares; a surface placed here and oriented in either of two opposite directions would receive equal illumination.

performance of these algorithms. We are actively working on this problem in our laboratory.[20]

Most objects in macroscopic scenes are opaque, forcing us to use vision algorithms to analyze them. In microscopic scenes, objects are often thin enough to make them partially transparent. This means that the 3D structure of microscope light fields can be analyzed using algorithms for reconstruction from projections, such as tomography and 3D deconvolution. These algorithms are fundamentally more robust than computer vision algorithms. This robustness allows us to transform microscope light fields into volumetric data sets with relative ease.[19] We can then use volume rendering techniques to visualize these data sets.

Although I've focused here on capturing 4D light fields, many close relatives to light fields bear examination. In this article, I've used the 4D light field to characterize the appearance of objects under unchanging illumination. If we relax this assumption, two 4D light fields are of interest—one characterizing the light incident on the object and another characterizing the light leaving the object.

If the object is geometrically complex—it contains more than one flat surface—the incoming light along any ray can affect the outgoing light along any other ray due to multiple reflections, refractions, and other optical effects. We can capture this dependency by defining a proportionality function that relates the outgoing radiance along each ray to the incoming radiance along each other ray. This function is commonly called the *reflectance field*, or light transport matrix.

Reflectance fields are an active research area in applied physics and computer graphics. Indeed, many of the devices I've described can be modified to measure these fields. Unfortunately, the full reflectance field is 8D, making it enormous, and to date nobody has ever measured one. However, researchers have measured subsets and lower-dimensional slices of this field. For example, if the viewpoint is fixed and only the illumination is allowed to vary, the result is a 4D reflectance field.

Gershun's paper on the light field considered the light passing through a point as a sum of vectors, one per direction impinging on the point, with lengths proportional to their radiance. Integrating these vectors over the sphere of incoming directions produces a scalar value—the total irradiance at that point, and a resultant direction. In computer graphics, this has been called the *vector irradiance field*, but aside from a 1994 paper by James Arvo,[21] it hasn't been systematically studied.

Figure 7 shows a visualization of the magnitude and direction components of this vector field for a simulated scene. Interestingly, as I've defined the illumination for this particular scene, the scalar irradiance at each point is equivalent to the fraction of the surrounding circle that can be seen from that point, and the irradiance vector points in the average direction of unoccluded points on the circle. Geometers call this function the *ambient occlusion map* or *visibility map*. The safest escape route for a robot placed in this scene would be along the field lines in Figure 7c. Hmmm, the field lines of a light field—isn't that what Faraday was talking about in 1846?

I would be remiss if I didn't end this article with one or two unfounded speculations that people can point to with derision in 20 years. When light fields were first introduced to computer graphics 10 years ago, we pro-

posed only one application—creation of new perspective views—and it seemed impractical to capture enough imagery to make this application useful. As a result, light fields were considered mainly of theoretical interest. In the intervening decade, computer speeds, memory, and bandwidth have doubled more than six times, the resolution of high-end digital cameras has increased a hundredfold, and low-end digital cameras have become tiny, cheap, and ubiquitous.

With these trends in mind, it is probably safe to predict that some light field applications will become commercially practical within the next five years. In fact, I predict that in 25 years, most consumer photographic cameras will be light field cameras. Whether they use this extra information to improve focus, to refocus, to extend the depth of field, or to change the viewpoint, I won't venture to guess. I also predict that photograph albums won't be filled with holograms, autostereoscopically displayed light fields, or Harry Potter talking movies. Most personal albums, whether paper or electronic, will still consist of ordinary images. ■

Acknowledgments

References

1. J. Kajiya, "The Rendering Equation," *Proc. ACM Siggraph*, ACM Press, 1986, pp. 143-150.
2. A. Gershun, "The Light Field," Moscow, 1936, trans. by P. Moon and G. Timoshenko, *J. Math. and Physics*, vol. 18, 1939, pp. 51-151.
3. M. Levoy and P. Hanrahan, "Light Field Rendering," *Proc. ACM Siggraph*, ACM Press, 1996, pp. 31-42.
4. S.J. Gortler et al., "The Lumigraph," *Proc. ACM Siggraph*, ACM Press, 1996, pp. 43-54.
5. E.H. Adelson and J.R. Bergen, "The Plenoptic Function and the Elements of Early Vision," *Computation Models of Visual Processing*, M. Landy and J.A. Movshon, eds., MIT Press, 1991, pp. 3-20.
6. P. Moon and D.E. Spencer, *The Photic Field*, MIT Press, 1981.
7. S.E. Chen, "QuickTime VR—An Image-Based Approach to Virtual Environment Navigation," *Proc. ACM Siggraph*, ACM Press, 1995, pp. 29-38.
8. J-X. Chai et al., "Plenoptic Sampling," *Proc. ACM Siggraph*, ACM Press, 2000, pp. 307-318.
9. A. Isaksen, L. McMillan, and S.J. Gortler, "Dynamically Reparameterized Light Fields," *Proc. ACM Siggraph*, 2000, pp. 297-306.
10. C.L. Zitnick et al., "High-Quality Video View Interpolation Using a Layered Representation," *ACM Trans. Graphics*, vol. 23, no. 3, 2004, pp. 600-608.
11. A. Zomet et al., "Mosaicing New Views: The Crossed-Slits Projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 741-754.
12. A. Román and P.A. Lensch, "Automatic Multiperspective Images," to be published in *Proc. Eurographics Symp. Rendering*, 2006.
13. A. Agarwala et al., "Photographing Long Scenes with Multiviewpoint Panoramas," to be published in A*CM Trans. Graphics*, vol. 25, no. 3, 2006.
14. Levoy et al., "The Digital Michelangelo Project," *Proc. ACM Siggraph*, ACM Press, 2000, pp.131-144.
15. C. Buehler et al., "Unstructured Lumigraph Rendering," *Proc. ACM Siggraph*, 2001, pp. 425-432.
16. B. Wilburn et al., "High-Performance Imaging Using Large Camera Arrays," *ACM Trans. Graphics*, vol. 24, no. 3, 2005, pp. 765-776.
17. T. Adelson and J.Y.A. Wang, "Single Lens Stereo with a Plenoptic Camera," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, 1992, pp. 99-106.
18. R. Ng et al., *Light Field Photography with a Hand-Held Plenoptic Camera*, tech. report CTSR 2005-02, Stanford Univ., 2005.
19. M. Levoy et al., "Light Field Microscopy," to be published in *ACM Trans. Graphics* , vol. 25, no. 3, 2006.
20. V. Vaish et al., "Reconstructing Occluded Surfaces Using Synthetic Apertures: Stereo, Focus, and Robust Measures," to be published in *Proc. Conf. Computer Vision and Pattern Recognition*, 2006.
21. J. Arvo, "The Irradiance Jacobian for Partially Occluded Polyhedral Sources," *Proc. ACM Siggraph*, ACM Press, 1994, pp. 335-342.

Marc Levoy is a professor of computer science and electrical engineering at Stanford University. His research interests include volume rendering; 3D scanning; light field sensing and display; computational imaging; digital photography; and applications of computer graphics in art history, preservation, restoration, and archaeology. Levoy received a PhD in computer science from the University of North Carolina. He is a member of the IEEE and the ACM. Contact him at levoy@stanford.edu.