# Communication through the Light Field: An Essay

*In the foregoing article, "The Road Ahead to the Holodeck: Light-Field Imaging and Display," James Larimer discusses the evolution of vision and the nature of light-field displays. This article looks at the physical, economic, and social factors that influence the success of information technology applications in terms that could apply to light-field systems.*

by Stephen R. Ellis

THE development of technology has greatly transformed the media used for all our communications, providing waves of new electronic information that amuse, inform, entertain, and often aggravate. Some of these media technologies, though they may initially seem solely frivolous, ultimately become so integrated that they become indistinguishable from the environment itself. A perfect example is the personal computer, initially seen as a toy of no practical use. Its key component, the micro-processor (or the modern day micro-controller), is now in myriad forms practically invisible in our watches, books, home appliances, cars, cash registers, telephones, address books, and flower pots.

The recent advances in light-field[a] displays, and in light-field capture in particular (Fig. 1), put a new twist into the process of integration because the idealized light-field display in a sense actually disappears, leaving only the light field emitted from a volume of space.[1–4] Viewers of an idealized display are thus left with a window into a world seen as a volume with all its physical features. When fully developed, this view could be indistinguishable from what they would see if they were looking through a physical window. What

*Stephen R. Ellis is with the NASA Ames Research Center. He can be reached at stephen.r.ellis@nasa.gov.*

would such a capture and display system be good for?

By analogy with previously developed technology we can be sure such displays will also amuse, inform, entertain, and aggravate us. But it seems to me hubris to claim to know what the first "killer app" for such displays will be. They may have scientific applications. They may have medical applications. They seem to be natural visual formats for home and theater entertainment, a kind of ultimate autostereoscopic display. Indeed, since they do not have specific eye points for image rendering, they are really beyond stereoscopic display. But to assert with any reasonable degree of confidence how, where, or why light-field technology will change the world seems premature, especially since there are not many established systems that include both capture and light-field display, and the constraints on their use are not well known. (Figure 1 illustrates two currently available products that are designed to capture light-field data.) However, whatever applications ultimately succeed, they will succeed because they communicate information to their users.

## Communication through the Ambient Optic Array

The physical essence of the light field is the plenoptic function, which is discussed in this issue's article, "The Road Ahead to the

Holodeck: Light-Field Imaging and Display," by James Larimer. Interestingly, the plenoptic function has a psychological/semantic aspect that Gibson[5] called the ambient optic array. This array may be thought of as the structured light, contour, shading, gradients, and shapes that the human visual system detects in that part of the light field that enters the eye. The features of the ambient optic array are primarily semantic rather than physical and relate to the array of behavior possibilities the light field opens to the viewer.

These features are in a sense the natural semantics of our environment to which we have become sensitive through the processes of evolution. The array was considered important by Gibson because it presents the viewer with information about the environment that is invariant with respect to many specific viewing parameters, *e.g.,* direction of view, motion, and egocentric position. The array thus allows viewers to determine environmental properties such as distance, slope, roughness, manual reachability of objects, or the accessibility of openings such as doors.

[a]The light field, considered primarily in terms of the light rays of geometrical optics, is defined as radiance as a function of all possible positions and directions in regions of a space free of occluders. Since rays in space can be parameterized by three spatial coordinates, $x$, $y$, and $z$, and two angles, the light field, therefore, is a five-dimensional function. (See the article by Larimer in this issue.)
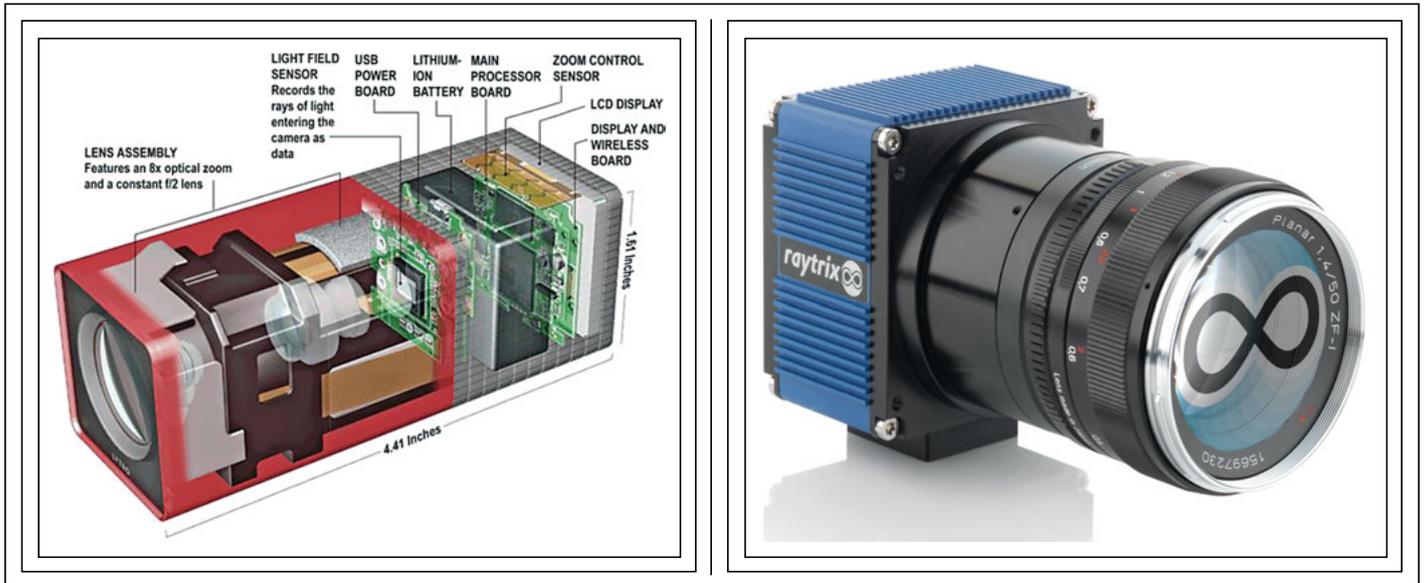
*Fig. 1:* *Two examples of light-field cameras currently offered for sale include the Lytro camera (left) and the Raytrix R11 (right). The Lytro is intended as a consumer product. It costs $400–$500 and records static light fields. Currently, the Lytro does not have a matched light-field display but does include a built-in viewer and laptop software for selecting typical photographic parameters such as focal plane, depth of focus, and view direction for creation of conventional images after the light field has been captured. The Raytrix R11 is intended as a scientific instrument to record volumetric motion as well as static light fields and can record light-field movies. It is much more expensive, costing about $20,000. Raytrix also makes an autostereoscopic viewer for use with its cameras, as well as special-purpose analytic software for tracking movement within recorded light fields.*

These are the environmental properties that guide our behavior. The last two are examples of what Gibson called affordances because they directly communicate behavioral possibilities. Such elements are, of course, the kind of visual information that displays are also intended to communicate, so it is not surprising that Gibson's work has been influential in their design.

Gibson called the informational elements of the features of the optic array "high order psychophysical variables." Examples are structures such as texture gradients in the projected view of a surface, cues to the elevation of the local horizon such as convergence points of parallel lines, and optic flow, the spatio-temporal change in the optic array due to object or viewer movement (Fig. 2). These variables may be measured as dimensionless ratios, percentages, or individualized units. They are unlike the more conventional psychophysical variables such as luminance, color, and motion, which are typically closer to the usual physical units. Optic flow, for example, can usefully be measured in terms of the viewer's personal eye height above the surface the viewer moves over; such as a measure of an optic flow rate could be eye-heights/second rather than meters/second. These features of ecological optics are not therefore themselves optical but are computationally derived from elements of the plenoptic illumination and often expressed in terms of user-specific units. These features are processed ultimately into the actors, objects, and culturally relevant elements that are the final interpretive output of our visual system.

A light-field camera and light-field display system together provide a medium for recording enough of the plenoptic illumination function so as to be able to re-project it toward viewers for them to interpret the ambient optic array as if they were present at the original scene. Consequently, the ultimate success of a light-field display system will be governed not simply by the fidelity with which the light passing into the eye represents light from a real space but also by the information that that light contains and the people and things that the light makes visible.

Some of the possible characteristics of light-field displays are absolutely remarkable. Imagine one that is hand-held, one that operates solely using ambient light and does not require power, one that not only constructs the light field for unobstructed objects but also for some that ARE obstructed – letting the viewer look around corners! Such performance may
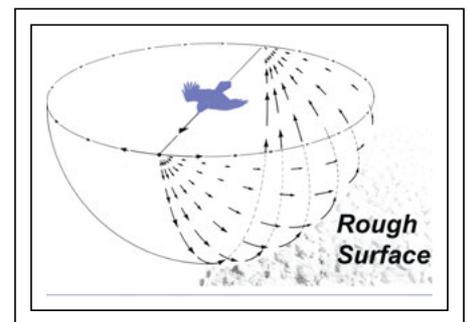


*Fig. 2:* *This is an example of optic flow, a higher order psychophysical feature that Gibson identified as specifying a direction of movement, flight in this case, over a rigid surface (Ref. 5, redrawn after Figure 7.3, p.123). Although Gibson probably would never express it so semantically, the observer learns, or appreciates from preexisting knowledge, the connection between the parameters of the flow field and their own velocity vector.*

actually currently be in the works[6,7] but nothing is exactly off the shelf and it's hard to know what the first commercial system with long-term viability will look like, how interactive it will be, if it will capture motion, or what visual resolution it will support.

Consequently, it seems to me better to consider now the list of challenges that the inventors of such displays will need to overcome in order to develop products with widespread appeal than to guess what the "killer app" will be. I will not be considering the detailed technical challenges so much as the overall performance issues. These issues are somewhat generic, but I believe they do apply to the technology; indeed, they apply to any new technology. They may be captured in a short list (Table 1).

It seems to me that the factors constraining widespread adoption of light-field systems may be broken into three categories: physical, economic, and social. Each of these has several elements that may be considered in somewhat arbitrary order.

## Physical Challenges
Time counts! The benefits of amazing technology can be greatly, even completely, eclipsed if the users are forced to wait endlessly for them. At its inception, the World Wide Web was rightly lampooned as the World Wide Wait and would never have become as pervasive as it is if its original latency problems had not been solved.

Time can influence performance in many different ways. Insufficiently fast update rates can disturb image quality,[8] reading rate,[9] and

**Table 1:** Would-be inventors of light-field systems face challenges in physical, ecoomic, and social terms.

| Physical | Economic | Social |
|---|---|---|
| Time counts | Costs shock | Novelty wanes |
| Size matters | Costs hide | Message mediates |
| Resolution clarifies | | Content rules |
| Power regulates | | Story conquers |
| | | Empowerment enables |
| | | Art matters (too) |

subjective sense of speed.[10] Even displays accurately presenting smooth motion through high dynamic frame rates can be problematic if visual shifts of their content give rise to vection, a subjective sense of self-motion. A possibly apocryphal example of this problem is the supposed effect of the first camera pan movement on a movie audience that, reputedly unprepared for it, promptly fell over. A related more likely true example is the reported reaction to the Lumière brothers' first showing of a movie clip of a directly approaching steam locomotive. The audience, strongly affected by the sight, seems to have scattered in fear (*The New York Times*, 1948).[11]

Size matters! Great ideas in awkward, heavy packaging do not make it. One major problem with the early tablet displays was the difficulty sharing them among a small group the way a modern tablet such as an iPad may be easily passed around like a sheet of cardboard. A reason volumetric displays, which to some extent already present light fields in that they actually create visible, dynamically deformable, physical structures (Fig. 3), have not caught on as consumer products is that they are moderately large, heavy (up to 60 pounds.) desktop products, and therefore are more transportable than really portable.

Resolution clarifies! The physical world is high-res! But the limiting factor actually is the eye rather than the world. The number of pixels needed to fill out the eye's resolution in a full 4pi steradian view is on the order of 600 Mpixels.[12] Restricting this analysis to a fixed head position but allowing normal eye movements within a field of regard cuts the number down to about 120 Mpixels, which is still large.

One of the compelling features of ordinary photographs is that they still can win in the resolution game, with chemically processed slow-speed film having ~5000 lines/in. (~20 lines/mm). Only recently are widely used electronic displays beginning to claim to approach the human-visual resolution, with Apple's so-called Retina displays. By the time light-field capture-display systems become commercially available and widely used, the purveyors of these displays will need to note that the public will likely have a well-established expectation of very high visual resolution for electronic imagery, probably exceeding that of current HDTV, thus placing a premium on anti-aliasing and other techniques to manage artifacts due to relatively lower pixel resolution.

Power regulates! One of the virtues of the Palm PDA was the relatively long run time supported by its battery technology and low power drain. I had one and loved to brag to my friends using iPhones that I could often wait days and days between charges. Power is, of course, not a major issue if you can run a system hooked up to the grid, but to the extent any part of the system is mobile, power can be critical. A system for which power is likely to be an issue is, for example, the recently announced Google head-mounted display. Its very small head-portable form factor coupled with its wireless connection, video capture capability, and more or less continuous all-day use, possibly outdoors, is likely to strain its battery power, especially if the only battery and computing system used is going to be incorporated within its spectacle-frame mount.

## Economic Challenges
Costs shock! A $1500 personal display of uncertain application will not generally find an immediate mass market unless it can interact with available critical content. Totally novel and amazing display capabilities are not insensitive to price. This fact is not news. But the costs involved are not always obvious.

Costs hide! It is hard to put a cost on hassle but it can be high. Indeed, the persistent attempts to find better autostereoscopic displays attest to the high cost of conventional stereoscopic displays requiring viewing spectacles. Gestural interfaces have been touted as intuitive, powerful, and the next great thing in UI but often their cost in fatigue is not appreciated. They have only really become widely used after they were implemented on horizontal touch surfaces that could support the weight of the users' hands. In Gibson's terms, one of the affordances of the horizontal touch screen is support for the weight of a hand. This need for support is often unappreciated and; in fact, hand support is one of the virtues of the mouse and joystick. Hand gestures in the air will wear out even the dedicated gamer, as was discovered by the users of the Mattel Power Glove.[13]

## Social Challenges
Perhaps the most compelling constraints on acceptance and dissemination of new technology are social. For example: Novelty wanes! In the mid 1980s, Ivan Sutherland's idea[14] for an "ultimate display" as a personal simulator, oxymoronically a.k.a. virtual reality, was reinvented at a much cheaper price point,

$100,000s per system vs. $1,000,000s for his system. One of these less expensive simulators was even adopted by Matsushita Electric Works, Ltd., as a medium for customer involvement in the design and sale of customized home kitchens. The novelty of giving anyone interested in a custom kitchen access to such an unusual system initially filled the company's show room in Tokyo and made international news. But, in time, partly because of poor dynamic performance,[13] the show room returned mainly to the use of conventional interactive computer graphics and architectural visualization.

Actually, the personal head-mounted display that Matsushita used disappeared from the show room for a variety of reasons. For one thing, there was no way to conveniently share the design view of the customized kitchen with others since the system was basically one of a kind and could not serve as a medium for communication. Message mediates! But also there was the problem of developing content for the visualization. A major effort was required to prepare the existing CAD data for visualization with the HMD. There was essentially very little pre-existing content that could be easily imported into the virtual environments that the HMD could be used to view. This situation contrasts with the rapid spread of the World Wide Web. Both the Web and virtual reality (VR) were both similarly initially impeded by poor dynamic performance, but the large amount of interesting and useful pre-existing content that existed on the Internet gave the Web a boost; users were willing to wait because there was much to wait for. Content rules!

But content in isolation has limited impact. For display content to be really useful, it needs to be woven into a story. Indeed, one way to think of the design of an interactive, symbolic information display is to imagine it as a backdrop for a story being told to a user. An air traffic display, for example, has a stage, characters, action, conflicts, rules, outcomes – all the elements of real-time drama. In fact, training in drama is not a bad background for an air-traffic controller and I know of at least one tower manager who actually has a degree in theater from Carnegie Mellon University.

Another example of the key role of content and story is the introduction of the first cell phone. Though it weighed about 2 pounds, cost on the order of $4000 (~$9000 in current dollars), and had only limited talk time, there was significant initial demand, even if the phone clearly did not have an immediate mass market. Motorola had the foresight to make sure at least some of the necessary cellular infrastructure was in place before its first public demonstration in 1973. Users could talk to each other and to others anywhere in the world who were on the phone network.[14] Story conquers!

Probably the single most important element in the widespread deployment of new information technology is the provision of a sense of personal empowerment. Part of the amazing success of the microcomputer revolution of the 1970s and early '80s was that the microcomputer enabled a single programmer, working mainly alone, to create useful software products such as word processors that previously had typically been built by a group working on a mainframe system. The word-
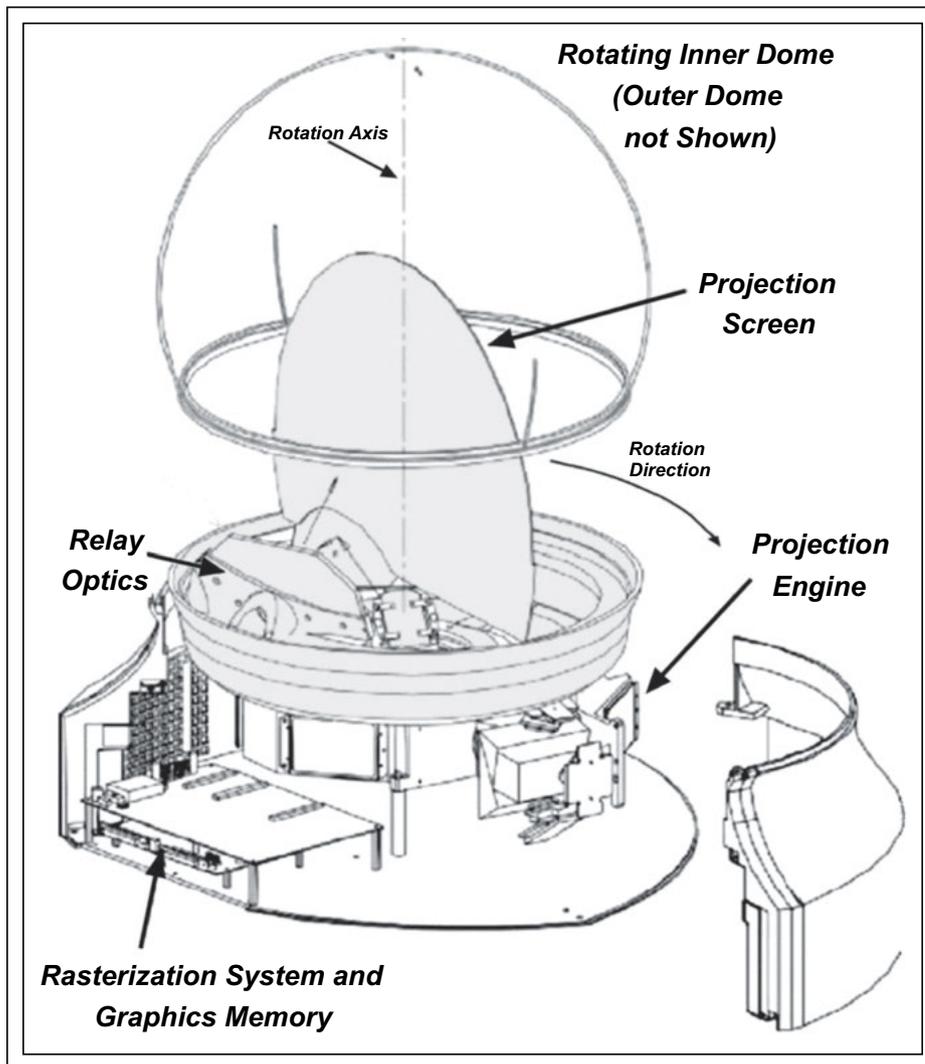


*Fig. 3: This example of an innovative contemporary, swept-screen, volumetric, multiplanar display was made by Actuality Systems. The gray central assembly along with the attached inner dome rotates at high speed while beams of light are scanned by deflecting mirrors over the surface of a 10-in. spinning disk.[17] Precise timing, placement, and coloring of the scanning light dot allow this system to produce a computer-graphics-generated, true-three-dimensional, full-parallax, colored image that can be used for information display. (For more on this display, and Actuality Systems, see "The Actuality Story" in the May/June 2010 issue of* Information Display.)

processing software in turn empowered writers with many functions previously the province of publishers. Thus, those with sufficient training and intellectual ability really could use the microcomputers as personal intelligence amplifiers in a symbiotic way as conjectured by Licklider.[15] This kind of personal empowerment has now become even more varied and widespread with the availability of powerful search and communication applications like those in Google and Twitter. Empowerment enables!

So we have now the challenge to the purveyors of light-field display systems: Can enough quickly processed, high-resolution, naturalistically colored light fields be captured for the display of storied content to provide useful visual information to personally empower the display's users?

There does seem already to be acknowledgement of some of the components of this question. The developers of the Lytro camera are tacitly acknowledging the preceding need for content by working first on the capture system, deferring the display for later. But additionally they would be wise to find ways to import existing three-dimensional data into light-field display formats so as to be able to benefit from all the existing volumetric or stereoscopic imagery available on the internet and elsewhere.

But they also will need to acknowledge that in addition to the natural information in the light field, such as full-motion parallax, which supports the natural semantics of our environment, synthetic light-field displays will also need to allow the introduction of artificial semantics. These semantic elements can take the form of geometric, dynamic, or symbolic enhancements of the display. Geometry can be warped, movement can be modified, and symbols can be introduced, all in the interest of communication of specific information. Such enhancements can turn a pretty picture into a useful spatial instrument in the way cartographers do when they design a map. Geometry of the underlying spatial metric can be warped as in cartograms (Fig. 4). Control order can be reduced through systems using inverse dynamics. Symbolic elements can be resized to reflect their importance.[16] There can be truth through distortion! Consequently, the naturally enhanced realism of the coming light-field systems may be only the beginning of the design of the next "ultimate display." Art matters too!

## References

[1]S. Grobart, "A Review of the Lytro Camera," Gadgetwise column, *The New York Times* (5 pm, February 29, 2012).

[2]C. Perwass and L. Wietzke, "Single Lens 3D-Camera," *Proc. SPIE* **8291**, *Human Vision and Electronic Imaging XVII*, 829108 (February 9, 2012); doi:10.1117/12.909882; http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1283425

[3]F. O'Connell, "Inside the Lytro," *The New York Times* (5 pm, February 29, 2012).

[4]M. Harris, "Light-Field Photography Revolutionizes Imaging," *IEEE Spectrum* (May 2012).

[5]J. J.Gibson, *The Ecological Approach to Visual Perception* (Earlbaum, Hillsdale, NJ, 1979/1986).

[6]R. Raskar, MIT Media Lab Camera Culture Group (2013); http://web.media.mit.edu/ ~raskar/

[7]D. Fattal, Z. Peng, T. Tran, S. Vo, M. Fiorentino, J. Brug, and R. G. Beausoleil, "A multi-directional backlight for a wide-angle, glasses-free three-dimensional display, *Nature* **495**, 348–351 (2013).

[8]D. M. Hoffman, V. I. Karasev, and M. S. Banks, "Temporal Presentation Protocols in Stereoscopic Displays: Flicker Visibility, Perceived Motion, and Perceived Depth," *J. Soc. Info. Display* **19**, No. 3, 271–297 (2011).

[9]M. J. Montegut, B. Bridgeman, and J. Sykes, "High Refresh Rate and Oculomotor Adaptation Facilitate Reading from Video Displays," *Spatial Vision* **10**, No. 4, 305–322 (1997).

[10]S. R. Ellis, N. Fürstenau, and M. Mittendorf, "Visual Discrimination of Landing Aircraft Deceleration by Tower Controllers: Implications for Update Rate Requirements for Virtual or Remote Towers," *Proceedings of the Human Factors and Ergonomics Society,* 71–75 (2011).

[11]"Louis Lumière, 83, a screen pioneer," credited in France with the invention of motion pictures, *The New York Times*, Book Section, (June, 7, 19480), p. 19.

[12]A. B. Watson, Personal communication (2013).

[13]S. R. Ellis, "What Are Virtual Environments?, *IEEE Computer Graphics and Applications* **14**, No. 1, 17–22.(1994).

[14]I. E. Sutherland, "The Ultimate Display," *Proceedings of the IFIP Congress*, 506–508 (1965).

[15]M. Bellis, "Selling the Cell Phone, Part 1: History of Cellular Phones," accessed (March 24, 2013); http://inventors.about.com/library/weekly/aa070899.htm

[16]J. C. R. Licklider, "Man-Computer Symbiosis," *IRE Transactions on Human Factors in Electronics* **HFE-1**, 4–11 (March 1960).

[17]S. R. Ellis, "Pictorial Communication" in *Pictorial Communication in Virtual and Real Environments*, 2nd ed., S. R. Ellis, M. K. Kaiser, and A. C. Grunwald (eds.) (Taylor and Francis, London, 1993), pp. 22–40.

[18]G. E. Favalora, "Volumetric 3D Displays and Application Infrastructure," *IEEE Computer,* 37–43 (2005).

[19]D. A. Keim and S. C. North, "CatroDraw: A Fast Algorithm for Generating Contiguous Cartograms," *IEEE Transactions on Visualization and Computer Graphics* **10**, No. 1, 95–110. ■
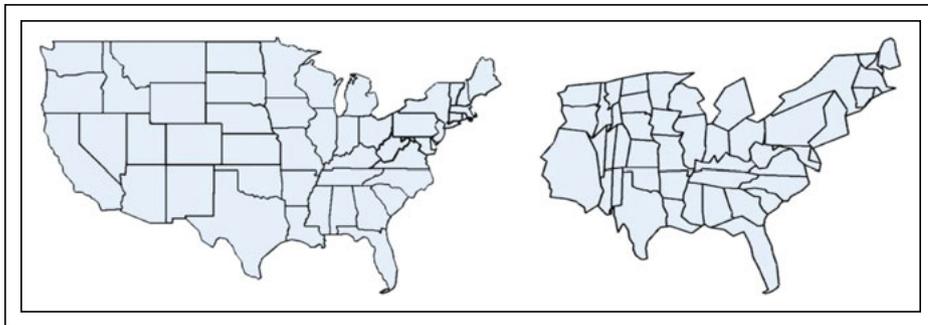
*Fig. 4: Cartograms are topological transformations of ordinary cartographic space based on geographically indexed statistical data. For example, in this classic cartogram based on 1998 U.S. data discussed by Keim, North, and Panse,[18] the areas of U. S. states in a conventional map (left) can be made proportional to their populations in a cartogram (right). Their paper develops new transformation techniques that can improve preservation of some geometric properties such as adjacency and shape of the state borders while making others such as area proportional to arbitrary statistical indices.*