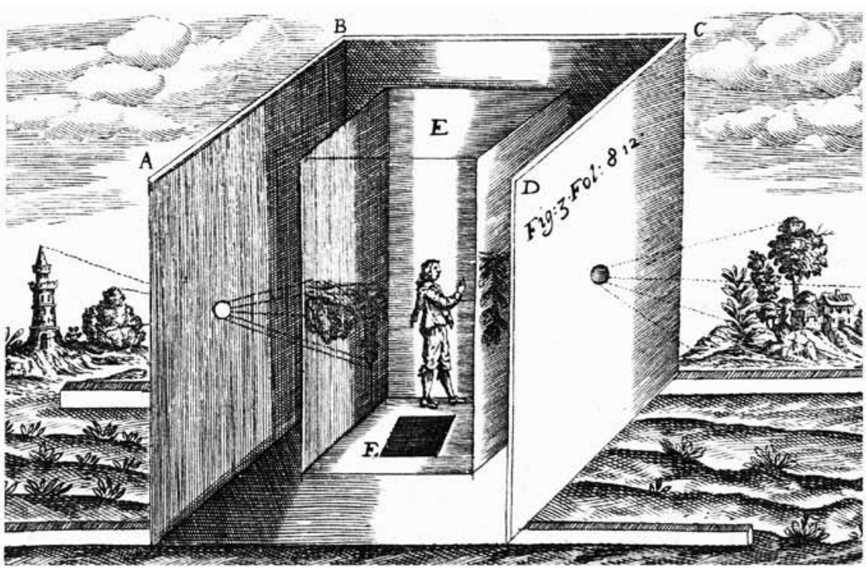
GEORGE LEGRADY

REFLECTIONS ON THE COMPUTATIONAL PHOTOGRAPH



Camera Obscura, 1671. From *Ars Magna* by Athanasius Kircher (Amsterdam, 1671).

The optically reflected image has been part of our cultural evolution since the dawn of history. One can only imagine the amazement that people felt, thousands of years ago, when they saw images of outdoor scenes projected on cave surfaces or in darkened nomadic tents. Such projections were eventually scaled down and made visible in portable, darkened boxes with light waves entering through a pinhole or lens at one end, and being projected onto a ground glass at the opposite end. The challenge of capturing and retaining the optically projected image was first resolved in 1826 by the French inventor Nicéphore Niépce (1765–1833), whose eight-hour exposure of an image of a rooftop was preserved on a sheet of pewter coated with a mixture of bitumen and lavender oil.

During the twentieth century, chemically processed photographic and cinematic images became the dominant form of visual representation, transforming Western society into an image-based culture. Our understanding of the photographic image as a chemical-based technological medium emerged and crystallized over a 170-year period, resulting in conventions of visual representation; cultural definitions, applications, and functions; and an epistemological understanding of the photographic image. Over the past two decades, with the transition to digital and increased computer functionalities, photography is in the process of being reinvented.

Eastman Kodak, which dominated the photographic supplies industry internationally throughout the twentieth century, filed for bankruptcy protection in early 2012,[1] some thirty-seven years after one of its engineers, Steven Sasson, invented the digital camera in 1975 (US Patent 4131919).[2] His prototype machine consisted of an electronic still camera that employed a basic audio-grade magnetic tape to record the data captured by a charged-couple device (CCD) that converted incoming photons into electron charges, which were then translated into pixel values. It wasn't until the early 1990s that the digital camera fully entered the market.

My first interaction with digitized images occurred in the mid-1980s, when an affordable imaging system, the AT&T Truevision Targa frame buffer placed inside an IBM AT computer, became available. The card could grab a colour image from a video stream and turn it into a bitmap image composed of a matrix of pixels at a 16-bit resolution of 32,768 colours for each pixel. Once digitized, the image could then be processed computationally or through an onscreen interactive paint program. The digital camera and related imaging technologies shifted the processing of the image

1. Kodak plans to complete its reorganization and emerge from Chapter 11 of the U.S. Bankruptcy Code in the first half of 2013. The Rochester-based company will be focusing its future business on commercial imaging. See Associated Press [New York], November 28, 2012, http://bigstory.ap.org/article/kodak-reaches-improved-financing-deal-worth-830m.

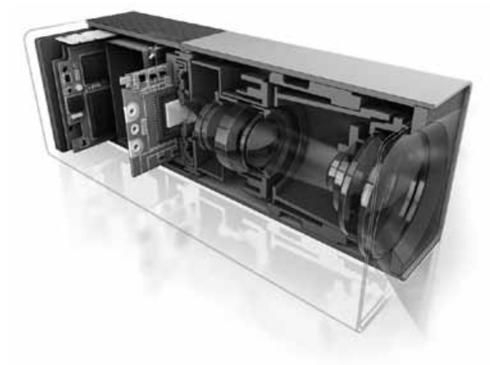
2. http://www.google.com/patents/US4131919?printsec=abstract#v=onepage&q&f=false.

The world's first camera, known as the "chambre de la découverte" (the chamber of discovery), made by Nicéphore Niépce before 1826. Sliding box camera. 30.5 x 31.5 x 25 cm.

from development and reproduction with chemicals to a mathematically based algorithmic operation, in which focus, contrast, tonal range, sharpness, and other parameters could be corrected by adjusting pixels' numerical red, green, and blue colour values.

Digital images can be processed in any number of ways and encoded with additional data not visible to the viewer. Every digital image today has encoded exchangeable image file format (EXIF) information about its origins. EXIF is used by most digital cameras to add descriptive data to the image file such as creation date and time and camera settings; some cameras also include Global Positioning System (GPS) coordinates. The field of steganography, which involves concealing data so that only sender and intended recipient are aware of them, has evolved to a significant extent with techniques to encode messages in unused bytes in the data files of digital images, sounds, and texts. Such techniques are commonly used to encode documents for copyright purposes. The practice of encoding copyright emerged from that of watermarking computerized documents, which consists of creating a watermark image

Steven Sasson's digital camera prototype.



The commercially developed Lytro camera that emerged from the lightfield research.

that is printed "under" the text; the watermark may convey relevant information related to the document, such as its source, production date, and the printing house. Steganography is closely related to cryptography, a practice that produces a form of secure communication in which ordinary messages are converted into unrecognizable data that are then decoded by the intended receivers. With its origin in the Greek words *steganos* (covered or protected) and *graphein* (writing), steganography essentially means "to hide in plain sight," and suggests the presence of information that is not visible until it is revealed or unmasked. In sound samples, sonic information can be embedded into other sounds by reducing it to a low-level signal. Images can also be hidden within other images without the viewer noticing their presence.

Digital imaging technologies have generated a wide range of interdisciplinary practices, disciplines, and industries that engage with the creation, duplication, analysis, treatment, storage, classification, and authentication of images. Much of the handling of images today is based on metadata – additional information embedded in the file that gives pertinent details such as origin, production date, and other properties. There are many categories of metadata, but it can be generally defined as "data that describe data"; for instance, a person's name, gender, height, age, citizenship, and social security number are all forms of metadata.

In a paper that I wrote in the 1990s, "Image, Language, Belief and Synthesis,"[3] I explored the complex situation of the authoritativeness of the photograph in its transition from analog to digital. To this day, photographs function as authoritative evidential documents, still convincing in their transparent visual descriptions of the world even though much of our image culture and many media representations today are obvious technological enhancements or simulations. Images created through algorithmic simulations were of great interest at the time that I wrote the paper, as they replicated physical or imagined phenomena, generating discussion about the relationship of data to image. If the data were true, then the resultant image had to be believed to some degree. Photographs, once digitized, challenge notions of truth in reporting, as post-capture manipulations through digital processing increase the difficulties in identifying data transformation. Although digital photographs may be processed, manipulated, and reconstructed in post-production, the visual information in the original is limited to what was recorded at the moment of capture. For instance, a still photograph of a sports event cannot provide information about what occurred prior to or after the millisecond when the image was recorded. There is no return to the subject source, if one is limited to a standard digital camera.

Over the past six years, frenzied engineering research activities have been paving the way for a revolutionary direction in digital photography called computational photography, creating a major paradigm shift in how we are to understand the photographic image by pushing the boundaries of what a "true" photographic image may be. The evolution of this paradigm is an outcome of the confluence and synthesis of computation embedded into the camera itself so that much of the post-production work traditionally done on a computer after the image has been created is now produced at the moment, or prior to the moment, of image capture. Many of these new computation techniques enhance the image-recording process by incorporating contextual data encoded into machine vision, resulting in high-dimensional images whose properties, such as focus and spatial composition, may be adjusted in the viewing process after the image has been created. Such images expand the boundaries

3. George Legrady, "Image, Language, Belief and Synthesis," in G. Legrady, Jean Gagnon, and Pierre Dessureault, *George Legrady: From Analogue to Digital: Photography and Interactive Media = de l'analogie au codage numérique : la photographie et l'interactivité*, CD-ROM (Ottawa: National Gallery of Canada, 1998), http://www.mat.ucsb.edu/g.legrady/glWeb/publications/p/image.html. The article was previously published in *Critical Issues in Electronic Media*, ed. Simon Penny (Albany: State University of New York Press, 1995).

of the photographic image. One of the memorable scenes in the culturally influential science fiction film *Blade Runner* (1982)[4] involves the interaction by police officer Deckard (played by Harrison Ford) with a photo-analysis machine through which Deckard manages to navigate around corners and behind walls to find the information he is looking for.

This innovative "moving around" within the visually captured space inside an existing two-dimensional photograph reveals a radical rethinking of what the photograph could be if enough additional visual information were recorded at the moment of capture. The scene in *Blade Runner* proposes the ability to discover additional information within the photograph after the fact of its original capture by recalculating the vantage point inside the flat space of the photograph.[5] What this suggests is a conflict between the still image and how it represents the world on which it is based. The nineteenth-century French sculptor Auguste Rodin (1840–1917) framed this scenario in an insightful way when he described how the artist constructs a painting in comparison to taking a photograph: "It is art that tells the truth and photography that lies. For *in reality, time does not stand still*. . . . The artist [, Rodin continues,] condenses several successive moments into a single image"[6] that then synthesizes the event more precisely than does a still frame frozen in time.

Research developments in camera optics, computational enhancement at the moment of image capture, and expansion of photographs' presence in the data cloud are some of the ways that the photograph is being redefined. Shree Nayar,[7] who directs the Computer Vision Laboratory[8] at Columbia University, one of the leading labs in the field, defines four research directions for his team, starting with digital photography, through which image processing is applied to existing images to enhance them. "Computational photography" involves computational processing during exposure to create multiple views, multiple foci through light field applications, and high-dynamic range exposure. "Computational imaging/camera" expands on the hardware through various new optics, such as multiple lenses, fish-eye panoramic imaging, and a radial catadioptric[9] system that captures a scene from a large number of viewpoints using mirrors to recover 3D information. At the receiving end of the light photons, Nayar's lab is experimenting with new types of detector sensors to push the boundaries of how pixels can be further enhanced. Finally, the "corneal imaging system" is an unusual experiment in which information within a scene is gathered on the basis of visual

4. *Blade Runner* (1982), directed by Ridley Scott, Warner Bros., http://bladerunnerthemovie.warnerbros.com.
5. Elissa Marder, "*Blade Runner*'s Moving Still," *Camera Obscura* 9, no. 27 (September 1991), 88–107, doi:10.1215/02705346-9-3_27-88.
6. Quoted in Paul Virilio, *The Vision Machine* (Bloomington: Indiana University Press, 1994), 2.
7. http://www.cs.columbia.edu/~nayar.
8. http://www1.cs.columbia.edu/CAVE.
9. "A catadioptric optical system is one where refraction and reflection are combined in an optical system, usually via lenses (dioptrics) and curved mirrors (catoptrics)." See http://en.wikipedia.org/wiki/Catadioptric_system.

The Frankencamera from Marc Levoy's Computer Graphics Lab at Stanford University, California.

data reflected in a subject's eye. This information can then be computationally reformulated to accurately represent what the subject may be seeing while looking at the photographer.

Images recorded through a set of multiple lenses positioned in a two-dimensional matrix result in enhanced views as each lens returns a slightly different vantage point, which can then be correlated to explore variances in depth of field and 3D sampling of a scene. This research, initiated in the mid-1990s at the Stanford Computer Graphics Lab directed by Marc Levoy,[10] has led to computer graphics experiments in the light-field camera, a system that uses a microlens array to capture 4D light field information in a scene resulting in an image with variable focus that can be adjusted during the viewing process.

The Lytro camera[11] is a commercial product developed by Stanford graduate student Ren Ng in Levoy's lab. The Frankencamera, also from Levoy's lab, was designed as an experimental platform for computational photography research. It

10. http://graphics.stanford.edu/~levoy.

11. https://www.lytro.com/camera.

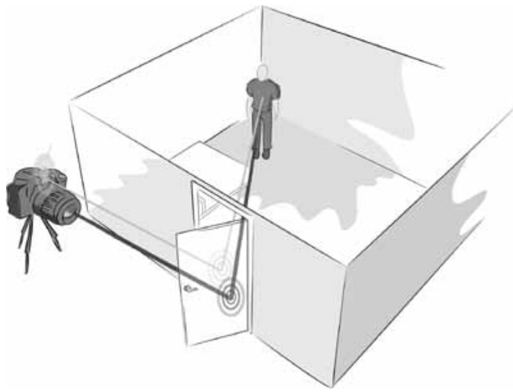
has a somewhat cumbersome look, reminiscent of Sasson's first digital camera, as its design is intended for programmable multifunctionality. The lab produced an application-programming interface software package (FCam)[12] to provide capture and post-processing functionalities for the camera on the Nokia N900 Smartphone.

Daniel Vaquero, a computer science doctoral researcher at the University of California, Santa Barbara, who was fortunate to collaborate with labs at Stanford, Nokia, and IBM Watson, worked with the Frankencamera and Nokia software to develop a research project called "Composition Context Photography," for which he developed software for the Nokia that would record preliminary compositions that an image maker might choose to make prior to capturing the image. Recording of and feedback on preliminary compositions may enhance the ability to create images.

As artists know, assessment of aesthetic value is a complex and challenging question, and the idea of quantifying it in photographic images to be translated into machine vision seems unrealistic. Nonetheless, researchers Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang[13] at Pennsylvania State University have taken on the challenge of formulating measurable, quantifiable ways to evaluate what is a good photograph, with the intention of developing software that will be embedded into cameras to guide image makers to create better compositions. To perform their research, Datta and his colleagues have identified various image primitives, such as colour relationships, hue, the rule of thirds, familiarity measure (defined as learning to rate images from the experience gathered by seeing other images), texture, size and aspect ratio, region composition, depth of field, and other image components that they have computationally analyzed from large online photo collections such as Flickr. Large datasets may be a noisy statistical resource if compared to a sampling of what experts in the field may agree on, but it does provide a basis by which to begin to understand whether large communities use universal aesthetic standards when they assess images. Once these and other, similarly ephemeral, properties of images are quantifiably defined, our cameras will be able to tell us what to photograph, thus raising the questions of whether aesthetic analysis of the image is still within the domain of the arts and whether science can quantitatively measure what may be an aesthetic solution and implement this as a camera function. Artists find the proposition questionable, as they know that aesthetic value is a continuously changing process, based on the culture's continuous absorption and rejection of norms. Nonetheless, the project is intriguing.

12. https://graphics.stanford.edu/papers/fcam/html.

13. http://infolab.stanford.edu/~wangz/project/imsearch/Aesthetics/ECCV06/datta.pdf.



“Looking Around Corners using Femto-Photography.” Research at Ramesh Raskar’s Camera Culture Group of the MIT Media Lab, Cambridge, Massachusetts.

Photosynth,[14] a software application from Microsoft, originally developed at the University of Washington, provides the most immediate experience of connecting one’s images with others in the Internet data cloud in such a way that the image becomes an information node in relation to the multitude of other images that can be correlated to it. Its visual and informational properties are sufficiently extended in detail to be viewed from many different angles, positioned in a three-dimensional volumetric space with other photographs taken at that scene at different times. With more than a billion people now using networked mobile cameras, there is tremendous real-time synchronization of any image taken with all others that were produced in the same geospatial location. Photosynth provides an expanded platform in which to contextually explore one’s image in relation to others, reminiscent of *Blade Runner*’s Esper machine, as we now can navigate around our image in its visual space, surrounded by other images accessed in real-time from the cloud.

To penetrate the flat dimensions of a photograph on its own is not a realistic possibility but, as *Blade Runner*’s Esper Machine implies, a photograph may potentially represent a broader range of its spatial environment if the capture technology can record it. The technique of ray-tracing was introduced a few decades ago to enhance realism

14. http://photosynth.net.

in computer graphics. The technique consists of tracing the path of light between virtual objects in a virtual space, and simulating on the surface of these objects how the light rays bounce around and reflect off each other. Some rays are absorbed, others are reflected to varying degrees, and still others are refracted due to surface variability and environmental conditions (think of how water in a glass will bend visual information perceived through it). The process of calculating each point in a 3D simulated scene is computationally expensive, but, given the continuous increase in computational performance, a next step in computational photography is to further explore how to harvest visual information beyond the limitation of incidental light rays received by a directional lens. A paper titled "Dual Photography"[15] by Pradeep Sen,[16] a researcher in image synthesis – a field that explores simulating realistic, photographic-looking images that integrate complex renderings of lighting, texture, shading, and occlusion – exploits the "reciprocity principle" phenomena discovered in the nineteenth century by the German physicist Hermann von Helmholtz, known for his work in physics and visual perception. The reciprocity principle of "If you can see me I can see you"[17] is realized in the "Dual Photography" research by pointing a camera at a scene, and then lining up a projector that provides structured illumination at a different angle facing a scene so that two vantage points can be computationally calculated – one from the camera's point of view and the other from the projector's – resulting in two variable representations of a scene.

Another project that explores the seemingly impossible task of recording what is beyond the line of sight has been proposed by Ramesh Raskar's[18] team at the Camera Culture Group[19] at the Massachusetts Institute of Technology (MIT). The group calls its research "Femto-Photography," referring to a superfast camera that can record at a rate of one trillion exposures per second. The research involves the release of laser pulses into a multidimensional space beyond the direct sight of the camera lens. As the pulses bounce around the space, reflecting off objects near openings such as slightly ajar doors, and return to the camera lens, the camera records the incoming information and translates the sum of pulses into a visual-spatial dimension, not unlike the mechanism for sonar detection of sunken ships undersea, to sketch out the hidden geometry of surrounding spaces beyond the sight of the camera lens. With this project and the "Dual Photography" project, the barrier to representation beyond the camera's vantage point has been broken, and we can expect images to have further dimensionality as research advances.

15. P. Sen et al., "ACM Transactions on Graphics (TOG)," *Proceedings of ACM SIGGRAPH 2005* 24, no. 3, (July 2005), 745–55, http://graphics.stanford.edu/papers/dual_photography.
16. http://www.ece.ucsb.edu/~psen.
17. http://graphics.stanford.edu/papers/dual_photography.
18. http://www.media.mit.edu/people/raskar.
19. http://web.media.mit.edu/~raskar.

My goal in this essay has been to give a brief overview of the breadth of activity in current engineering research that focuses on next-step expansions and reinventions of the camera and photography. I am particularly interested in the promise of multi-dimensional photography as suggested some three decades ago by the science fiction movie *Blade Runner*. Computational photography involves advances in optics, sensors, and computer vision applications within the camera so that multiple layers of information with changing scene parameters can be recorded and then fused to enhance the representation. In contrast to the rapid pace of technological innovations, Western culture's understanding of the photographic image's transformation into a greater "something else" is evolving at a much slower pace, widening the gap between the cultural meaning of images and the technological instruments that create them.

\ **George Legrady** is a digital media arts installation artist whose current work engages with data visualization, swarm robotic camera interaction, and creation of multi-image narratives in static lenticular displays. In the early 1990s, he began to explore the potential of computer processing for narrative-based digital interactive installations, which have been exhibited internationally. He is director of the Experimental Visualization Lab and a professor in the Media Arts & Technology PhD program at the University of California, Santa Barbara (UCSB). He is also a member of the ExpVisLab artists' collective participating in Le Mois de la Photo à Montréal 2013. His work is currently funded by two National Science Foundation grants and the Robert W. Deutsch Foundation. \ www.georgelegrady.com