Computational Photography: Advances and Challenges

Edmund Y. Lam

Imaging Systems Laboratory, Department of Electrical and Electronic Engineering,
University of Hong Kong, Pokfulam, Hong Kong

ABSTRACT

In the mid-1990s when digital photography began to enter the consumer market, Professor Joseph Goodman and I set out to explore how computation would impact the imaging system design. The field of study has since grown to be known as computational photography. In this paper I'll describe some of its recent advances and challenges, and discuss what the future holds.

Keywords: Digital photography, computational imaging, computational photography, image restoration, imaging systems

1. INTRODUCTION

"Telecommunication by radio shrank the world to a global village, and the satellite and computer have made imagery the language of that village," penned the late Professor Ronald Bracewell of Stanford University.[1] We have this universal language thanks to the proliferation of cameras. In the early days of photography — which spanned around a century and a half — cameras were only available to professionals and hobbyists with deep pockets, and served to record only important moments in life. However, the emergence of digital photography has drastically reduced the per-picture cost of a snapshot, and facilitated the continuous minimization of a camera's physical dimensions, so that it fits comfortably in a cellphone and virtually any other kind of mobile devices. People can now take pictures of every minute details of their lives, and not having a camera around may be something of a distant past.

In fact, not only are there more cameras, but the cameras are also more powerful. On one hand this has to do with better lens design: perfection in engineering allows for further corrections in optical aberrations, minimal weight of the lenses, and more complex designs giving us features such as a better zoom range. On the other hand, this has to do with the electronics: with digital photography, a picture becomes an array of numbers. A digital camera doubles as a computer, processing these numbers for different effects — from enhancing the visual image through techniques such as increasing the contrast, to reducing the file storage through compression. While most of these computations take place after the photographs are taken, in some cases they affect how the images are recorded. Examples include face detection, which determines the focal point of the photograph, and smile detection, which determines when the shutter is released.

The big question, however, is whether digital processing should fundamentally alter the way imaging takes place, and if so, how. The premise of computational photography is that data processing has a role that is complementary, if not of equal importance, to the lenses in the image formation process. In hindsight, this is a natural development in the history of photography.

2. DEVELOPMENT OF PHOTOGRAPHY

2.1 Pinhole photography

The Spring and Autumn Period in China, roughly from the eighth century B.C. to the fifth century B.C., was a time that spawned many influential intellectual thoughts and military developments. Confucius and Sun Tzu are perhaps most well known in the western world. Yet a lesser-known scholar, Mozi, gave us perhaps the first written account of imaging. In the Mohist canon known as *Mo Jing*, he wrote: (The original Chinese text is shown in Fig. 1.)

Further author information: (Send correspondence to Edmund Lam.)
Edmund Lam: E-mail: elam@eee.hku.hk

Figure 1. Possibly the earliest description of a pinhole camera, in the Chinese text *Mo Jing*. An excerpt is shown here.

> Inversion arises from a small hole at the intersecting point of lights; the hole projects a long image because of the hole. When light shines a man it goes straight like an arrow; a low position becomes high, and a high position becomes low. So the leg appears on top and the head appears at the bottom, and with increasing or decreasing distance of the object from the small hole, the size of the image changes.

In today's terminology, the phenomenon described above is essentially imaging with a pinhole camera—the simplest form of imaging possible. In the following century in the Western world, Aristotle, and later Euclid, also wrote about picture formation with pinholes, describing what is now known as camera obscura. We note here that in pinhole photography, there is no lens, and certainly there is no computation involved.

2.2 Lens photography and films

The nineteenth century saw a breakthrough in photography. The central role of lenses in forming images had been established over the years. Magnifying glasses, spectacles, and telescopes were all in use. The English scientist William Hyde Wollaston was credited with the invention of the first photographic camera lens known as the "Wollaston Simple Meniscus," a single-element concavo-convex lens, as it was then incorporated into camera obscuras.[2] It was later replaced by an achromatic doublet, which could correct for chromatic aberrations (which are evident with Wollaston's design) and improve the image quality, and subsequently by more and more complex designs to overcome different optical aberrations.[2]

Meanwhile, the French inventor Joseph Nicéphore Niépce took the first permanent photograph with his use of bitumen as the recording material. Bitumen hardened when exposed to light, and when the unhardened materials were washed away a positive image was formed.[3] Soon he started to use silver compounds instead, and the silver process was later refined by Louis Daguerre into what was known as the daguerreotype. Glass photographic plates were then used; after successive improvements, George Eastman developed the technology of film, leading to the bourgeoning of film cameras and film photography for over a century.

In contrast to pinhole photography, lens photography relies on lenses in the imaging system, but no electronics or computations are required.

2.3 Digital photography and post-processing

Advents in semiconductor technology since the mid-twentieth century allowed for the emergence of two dominant types of digital sensors: the charge-coupled device (CCD), and the complementary metal-oxide semiconductor (CMOS) sensor. Both contain photoactive regions divided into a regular array of photodetectors, with each photodetector resulting in a pixel of the digital image. Consequently, there is no need to "develop" the picture

	physical systems (lenses)	processing systems (computations)
pinhole photography	✘	✘
lens photography	✔	✘
digital photography	✔	✔
computed photography	✘	✔
computational photography	≪ boundary removed ≫	

Table 1. A classification of various types of photography, in terms of their use of lenses and computations.

from the film, nor is it necessary to purchase new recording materials (films) when one likes to take more pictures. With an ever-declining computer memory cost, saving images is inexpensive, and taking a snapshot is effectively free.

Digital photography began to enter the consumer market in the 1990s. During the first decade or so, the design philosophy was essentially to replace the film with a digital sensor, with everything else unchanged. This was particularly the case for single-lens-reflex (SLR) cameras, where to lure high-end users to adopt digital photography, the camera manufacturers made sure that other investments of the photographers, particularly their lens collections, would be compatible with their new digital camera bodies. The major technical advancement in this era was the increase of spatial resolution of the digital cameras. The earlier ones delivered far lower resolution (which was somewhat simplistically equated with pixel counts) than was the case for film cameras, making the former more like a toy than a tool for serious photography; but sensor resolution increased rapidly, thanks to Moore's Law in driving the semiconductor manufacturing process.

Very early on, people realize the potential and the necessity of putting computations in digital photography.[4] Image compression is virtually a must, given the cost of on-camera storage and the size of the pictures; in addition, many image enhancement techniques already exist, and it does not take a lot of efforts to migrate them to the camera's processing. Most of these, however, are post-processing methods, such as converting a color photograph to sepia tone, and some image restorations.[5,6] Digital zooming can also be considered post-processing, as the camera simply interpolates the pixels to give a false sense of zooming, without affecting the actual optical system. Nevertheless, we consider digital photography embracing both lenses in the physical system and computations in the processing.

2.4 Computed photography and graphics

It is instructive to summarize the above three stages in a table, as shown in Table 1, to see the increasing sophistication of photography involving lenses and computations. In the interest of symmetry, one could ask whether there is a form of photography that does not involve lenses but requires computations? In a way, yes — and here we call it computed photography*. A more common name, though, is computer graphics. Photographs are not "taken," but simulated through computations. To enhance photorealism, graphics researchers studied the process of image formation and light modeling. This eventually led them to computational photography, where there is a physical camera.

2.5 Computational photography

In a sense, computational photography is digital photography — putting both the lenses and computations to work in arriving the final imagery. This idea is not entirely new in the optics community, which has been described as integrated computational imaging systems.[7] In astronomy and remote sensing, for example, adaptive optics has a bit of this flavor whereby a wavefront sensor identifies the atmospheric distortion and the computation informs the mirrors to deform and correct for the distortions.[8] In microscopy at the other end of the spectrum of physical dimensions, computational holography is used for imaging volumetric objects. The digital hologram

---

*The term computed photography is sometimes used to refer to what we will describe next as computational photography, but here we distinguish the two.

contains the three-dimensional information, but is not immediately recognizable; computation is needed to recover, for example, a sectional image of the underlying 3D object.[9, 10]

Yet computational photography is also more than digital photography in the way that the camera designer can now shift some of the imaging work to the digital processing, and with that one can significantly broaden the nature of image capture. One no longer aims at an aesthetically pleasing (or even recognizable!) image at the sensor plane, but that the physical system (lenses) is seen as modulating the signal (object), leaving the job of recovering the information, i.e. image formation, to the processing system by means of computations. Several new architectures have been reported, which lead to new imaging capabilities. As they become practical for consumer photography, they will certainly challenge the *status quo* of the current design philosophy of digital photography.

3. ADVANCES IN COMPUTATIONAL PHOTOGRAPHY

Here we highlight some of the representative work in computational photography. This is by no means exhaustive, and in fact, some of the developments predate the emergence of this term. Nevertheless, they serve as good examples of what the resulting "new" designs can be, and many current computational cameras build upon these early developments.

3.1 Wavefront coding

In imaging a scene with varying depth, in principle only one plane can be in focus, while the others would suffer from different degrees of defocus. In practice, one defines a circle of confusion, and if the resulting point spread function (PSF) is within the acceptable boundary, one considers those planes to be in focus as well. The range of distances from the camera is called the depth of field.

We can, however, restore a defocus image back to focus provided we know the PSF.[4, 5] However, for most scenes that contain multiple depths, this often requires a segmentation process before one can apply the appropriate sharpening to the different regions with similar PSF. Wavefront coding, pioneered by Cathey and Dowski, tackles this with a different philosophy: is there a way one can deliberately introduce a known distortion to the image such that the resulting PSF stays mostly invariant for objects at a range of depths? If so, we can undo this known distortion subsequently in a digital restoration process. They argue, with mathematical reasons and physical demonstrations, that this is possible with a cubic phase mask.[11, 12] The phase mask is put along the path of the main lens, resulting in a blurry image even for object that would have been in focus without the phase mask. In essence, the system trades off the ability to form sharp image at focus with the ability to maintain a reasonable signal-to-noise ratio (SNR) at the other out-of-focus planes for the subsequent image restoration. Because the PSF is known, a standard image restoration scheme such as the Wiener filter can be used.

3.2 Lenticular array

In nature, compound eye is common among insects. With apposition compound eyes, the animal collects a number of images simultaneously, and combines them in the brain to interpret the surrounding scene.[13] This inspires similar optical system designs.[14, 15] Among them, the TOMBO system is particularly noteworthy in this endeavor, primarily because its optical design allows for a very thin imaging system. Short for "thin observation module by bound optics" (the acronym also means dragonfly in Japanese), it divides a planar photodetector array into groups of pixels with a separation layer on top to avoid cross-talk, and above the separation layer is a micro-lens array.[16, 17] Therefore, every time an array of low resolution (LR) images are captured simultaneously, instead of a single high resolution (HR) image. Image reconstruction is needed to bring this set of LR images to form a HR image, such as with the "pixel rearrange" method.[18]

More sophisticated algorithms from the sub-field of superresolution imaging can also be used to deliver possibly better image reconstruction results. In addition, a theoretical study combining wavefront coding and multi-lens system has also been carried out.[19]

3.3 Single pixel camera

In a way, taking a picture is a "parallel process" — we are simultaneously recording the intensity data at different spatial locations. Conceivably, we can turn that into a sequential process, where at any single time we are recording intenity information at a particular location. We then perform a raster-scan of the object to obtain the full picture. This way of imaging is used, for example, in confocal microscopy. An advantage of this is that the sensor design can be extremely simple, where a "bucket detector" or a single-pixel sensor would suffice.

Of course, a major disadvantage is the amount of time needed to capture the entire picture. In the last decade, the mathematical theory known as compressed sensing promises to reduce substantially the time taken for the image capture. For an $N \times N$ image, a pixel-by-pixel raster scan takes N^2 captures; but if instead every time we record a linear projection of the scene, under appropriate conditions — essentially that the object is "sparse," which is the cornerstone of compressed sensing[20] — the number of projections needed is substantially smaller than N^2. This is the working principle behind a single-pixel camera for digital camera[21] and for terahertz imaging.[22]

3.4 Plenoptic camera

The plenoptic camera was developed by arguing another "deficiency" in conventional photography: the light distribution is lost in the recording process. Light rays from different angles arriving at the same photodetector are simply aggregated from the detector's perspective, yet if they can be separated, we would have the means to reconstruct the image that would be captured when the sensor is placed at other positions.

This motivates the design of a plenoptic camera. Originating primarily in the fields of computer vision and graphics, and limiting to geometric optics assumptions, the plenoptic function or a light field (which is not a field in the sense of waves) records not only the position where a ray emanates, but also the angle, and possibly the wavelength and time.[23] In the context of photography, where we can assume light traveling in one axial direction, the light field can be reduced to 4D.[24] The plenoptic camera makes use of a conventional main lens, but in addition, places a lenticular array in the original sensor position, thus allowing the light rays from different angles to be separated.[25] With the light field as the raw data, one can, for example, perform post-capture refocus, which generates a lot of interest. In lieu of the lens array, it has also been shown that one can achieve similar results with masks placed appropriately along the optical path.[26]

4. CHALLENGES OF COMPUTATIONAL PHOTOGRAPHY

So, what lies ahead? Despite the many advances in computational photography, researchers continue to work on the following aspects:

- The physics of the imaging process

  As geometric optics can be viewed as a simplified treatment of wave optics, one can also ask how light field is related to some other mathematical entities in the latter. These entities in fact lie in phase-space optics. In recent years, it has been shown that light field relates to a form of Wigner distribution.[27] Earlier on, the analysis of wavefront coding has also been performed using ambiguity function.[11] In recent years, there are also growing interest in techniques such as fractional Fourier transform and linear canonical transform, further enriching the arsenal of tools to analyze imaging systems.

- The mathematics of image reconstruction

  In different forms of computational photography, a reconstruction process to recover the image from the raw data is an integral part of the imaging system. Most of the computations often boil down to solving an ill-posed, linear inverse problem.[28, 29] Ongoing research focuses on three main aspects: finding ways to incorporate suitable prior knowledge (whether it's sparsity, edge sharpness, spectrum smoothness, or compact support, to name a few), improving image resolution, and executing fast numerical schemes to solve the large-scale problem efficiently.[30] The latter often makes use of recent advances in iterative schemes for convex optimization.

- Applications

In a way, computational photography is like opening the Pandora's box of imaging, where people are now exploring all kinds of ways to unleash the computational power for different applications. We have mentioned extended depth-of-field and refocus; in addition, people look at image forensics,[31] re-photography,[?] flash photography,[32] and high dynamic range imaging,[33] and many more.

5. CONCLUSIONS

We have looked into the historical developments of photography and how various technologies came to be known collectively as "computational photography" developed over the past 10 to 15 years has ushered into a new era of taking pictures. Looking into the future, imagery as the language of the global village will only promise to be more colorful and rich with artistic expressions!

ACKNOWLEDGMENTS

REFERENCES

[1] Bracewell, R. N., [*Two-dimensional Imaging*], Prentice Hall, Englewood Cliffs, New Jersey (1995).

[2] Kingslake, R., [*A History of the Photographic Lens*], Academic Press (1989).

[3] "History of photography." http://en.wikipedia.org/wiki/History_of_photography. Retrieved on July 27, 2011.

[4] Lam, E. Y., "Image restoration in digital photography," *IEEE Transactions on Consumer Electronics* **49**, 269–274 (May 2003).

[5] Lam, E. Y. and Goodman, J. W., "Discrete cosine transform domain restoration of defocused images," *Applied Optics* **37**, 6213–6218 (September 1998).

[6] Lam, E. Y., "Digital restoration of defocused images in the wavelet domain," *Applied Optics* **41**, 4806–4811 (August 2002).

[7] Mait, J. N., Athale, R., and van der Gracht, J., "Evolutionary paths in imaging and recent trends," *Optics Express* **11**, 2093–2101 (September 2003).

[8] Tyson, R. K., [*Introduction to Adaptive Optics*], SPIE (2000).

[9] Zhang, X., Lam, E. Y., and Poon, T.-C., "Reconstruction of sectional images in holography using inverse imaging," *Optics Express* **16**, 17215–17226 (October 2008).

[10] Lam, E. Y., Zhang, X., Vo, H., Poon, T.-C., and Indebetouw, G., "Three-dimensional microscopy and sectional image reconstruction using optical scanning holography," *Applied Optics* **48**, H113–H119 (December 2009).

[11] Dowski, E. R. and Cathey, W. T., "Extended depth of field through wave-front coding," *Applied Optics* **34**, 1859–1866 (April 1995).

[12] Cathey, W. T. and Dowski, E. R., "New paradigm for imaging systems," *Applied Optics* **41**, 6080–6092 (October 2002).

[13] Land, M., "The optics of animal eyes," *Contemporary Physics* **29**, 435–455 (September/October 1988).

[14] Sanders, J. S. and Halford, C. E., "Design and analysis of apposition compound eye optical sensors," *Optical Engineering* **34**, 222–235 (January 1995).

[15] Hamanaka, K. and Koshi, H., "An artificial compound eye using a microlens array and its application to scale-invariant processing," *Optical Review* **3**(4), 264–268 (1996).

[16] Tanida, J., Kumagai, T., Yamada, K., Miyatake, S., Ishida, K., Morimoto, T., Kondou, N., Miyazaki, D., and Ichioka, Y., "Thin observation module by bound optics (TOMBO): Concept and experimental verification," *Applied Optics* **40**, 1806–1813 (April 2001).

[17] Tanida, J., Shogenji, R., Kitamura, Y., Yamada, K., Miyamoto, M., and Miyatake, S., "Color imaging with an integrated compound imaging system," *Optics Express* **11**, 2109–2117 (September 2003).

[18] Kitamura, Y., Shogenji, R., Yamada, K., Miyatake, S., Miyamoto, M., Morimoto, T., Masaki, Y., Kondou, N., Miyazaki, D., Tanida, J., and Ichioka, Y., "Reconstruction of a high-resolution image on a compound-eye image-capturing system," *Applied Optics* **43**, 1719–1727 (March 2004).

[19] Chan, W.-S., Lam, E. Y., Ng, M. K., and Mak, G. Y., "Super-resolution reconstruction in a computational compound-eye imaging system," *Multidimensional Systems and Signal Processing* **18**, 83–101 (September 2007).

[20] Candès, E. J. and Wakin, M. B., "An introduction to compressive sampling," *IEEE Signal Processing Magazine* **25**, 21–30 (March 2008).

[21] Duarte, M. F., Davenport, M. A., Takhar, D., Laska, J. N., Sun, T., Kelly, K. F., and Baraniuk, R. G., "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine* **25**, 83–91 (March 2008).

[22] Xu, Z. and Lam, E. Y., "Image reconstruction using spectroscopic and hyperspectral information for compressive terahertz imaging," *Journal of the Optical Society of America A* **27**, 1638–1646 (July 2010).

[23] Adelson, E. H. and Bergen, J. R., "The plenoptic function and the elements of early vision," in [*Computational Models of Visual Processing*], 3–20, MIT Press (1991).

[24] Levoy, M. and Hanrahan, P., "Light field rendering," in [*ACM SIGGRAPH*], 31–42 (August 1996).

[25] Ng, R., Levoy, M., Brédif, M., Duval, G., Horowitz, M., and Hanrahan, P., "Light field photography with a hand-held plenoptic camera," Tech. Rep. CSTR 2005-02, Stanford University Computer Science Department (2005).

[26] Veeraraghavan, A., Raskar, R., Agrawal, A., Mohan, A., and Tumblin, J., "Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing," in [*ACM SIGGRAPH*], (August 2007).

[27] Zhang, Z. and Levoy, M., "Wigner distributions and how they relate to the light field," in [*IEEE International Conference on Computational Photography*], (April 2009).

[28] Bertero, M. and Boccacci, P., [*Introduction to Inverse Problems in Imaging*], Taylor & Francis (1998).

[29] Lam, E. Y. and Goodman, J. W., "Iterative statistical approach to blind image deconvolution," *Journal of the Optical Society of America A* **17**, 1177–1184 (July 2000).

[30] Vogel, C. R., [*Computational Methods for Inverse Problems*], SIAM (2002).

[31] Choi, K. S., Lam, E. Y., and Wong, K. K., "Automatic source camera identification using the intrinsic lens radial distortion," *Optics Express* **14**, 11551–11565 (November 2006).

[32] Eisemann, E. and Durand, F., "Flash photography enhancement via intrinsic relighting," *ACM Transactions on Graphics* **23** (August 2004).

[33] Debevec, P. E. and Malik, J., "Recovering high dynamic range radiance maps from photographs," in [*ACM SIGGRAPH*], (August 2008).