Invited Paper

Image formation challenges in the MOSAIC platform

M.E. Gehm[a,b], D.R. Golish[a], E.M. Vera[a], J. Hahn[c], D.L. Marks[c], and D.J. Brady[c]

[a]Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ;
[b]College of Optical Sciences, University of Arizona, Tucson, AZ;
[c]Department of Electrical and Computer Engineering, Duke University, Durham, NC

ABSTRACT

The DARPA MOSAIC program applies multiscale optical design (shared objective lens and parallel array of microcameras) to the acquisition of high pixel count images. Interestingly, these images present as many challenges as opportunities. The imagery is acquired over many slightly overlapping fields with diverse focal, exposure and temporal parameters. Estimation of a consensus image, display of imagery at human-comprehensible resolutions, automated anomaly detection to guide viewer attention, and power management in a distributed electronic environment are just a few of the novel challenges that arise. This talk describes some of these challenges and presents progress to date.

Keywords: multiscale imaging, image stitching, image formation

1. INTRODUCTION

High-resolution, wide Field-of-View (wFoV) imagery has great utility in a variety of defense, security, and medical applications. Acquisition of such images, however, has generally proved problematic. The limit on the resolution arises in a fundamental sense from the size of the coherent aperture (the diffraction limit) and in a practical sense from how close to the diffraction limit the system can be designed and fabricated. It is relatively simple to design and manufacture a diffraction limited system for a small aperture and small field; however, as the aperture and field size grow, so does the required complexity of the optical system. As described in a recent publication,[1] the solution to this challenge is break the longstanding tradition of single-axis optical design. Just as parallel architectures enable digital processor scaling, breaking lens design into hierarchical scales enables camera capacity to increase linearly in objective aperture size.

This new strategy, which is discussed briefly below and in much greater detail in the aforementioned publication,[1] provides a clear solution to the optical design and fabrication challenges. It does this recognizing that modern imaging systems are hybrid optical/electronic/processing systems and that each component technology has its individual strengths and weaknesses. By considering a novel architecture, the high-resolution, wFoV challenge is moved from the optical domain into the electronic and processing domains where, we argue, the scaling is more favorable.

Nonetheless, the electronic and processing solutions (as well as the optical design strategies that will complement the new techniques) are not yet fully identified and understood. The new architecture introduces a number of new challenges in this domain that we are now beginning to address. This manuscript identifies and discusses a number of the key challenges and discusses our recent progress in the development of a multiscale optical system from within the context of the DARPA MOSAIC project—an effort where multiscale optical design will be used to create a 50 gigapixel imager with a 120° FoV.

Corresponding address: gehm@ece.arizona.edu

Visual Information Processing XX, edited by Zia-ur Rahman, Stephen E. Reichenbach, Mark A. Neifeld,
Proc. of SPIE Vol. 8056, 80560T · © 2011 SPIE · CCC code: 0277-786X/11/$18 · doi: 10.1117/12.883695

2. MULTISCALE OPTICAL DESIGN

Multiscale optical design draws its inspiration from modern high-performance computing architectures that replace the monolithic processor of earlier designs with a large collection of relatively simple processors or processing cores. This transition is driven by the fact that, as computational requirements increase, a parallel or hierarchical attack on the problem becomes more cost effective than continual development of ever more powerful sequential processors.

The optical version of this insight is the realization that, as optical performance requirements increase (e.g. to giga- or terapixel performance over extremely large fields), continual development of ever more complicated *sequential optical designs* is cost ineffective relative to the development of parallel or hierarchical optical systems utilizing relatively simple designs.

A multiscale design, then, starts with a single, large, shared objective (i.e a common coherent aperture). This large objective has sufficient aperture to achieve the desired resolution, but is *not* fully corrected. Following the objective is a parallel array of smaller optical systems (termed *microcameras*) that take a portion of the intermediate image formed by the objective, correct for the residual *local* aberrations, and form an image of their individual FoVs (typically on a per-camera focal plane). A schematic of a multiscale design is shown in Fig. 1.

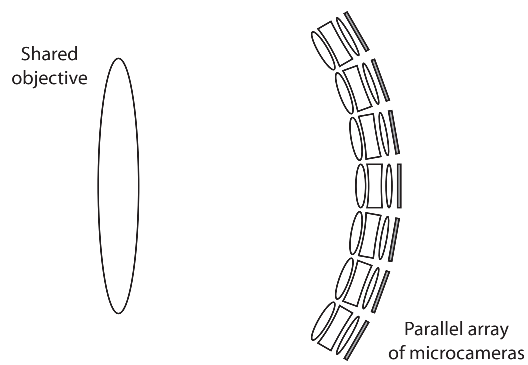

Figure 1. Schematic showing the basic design of a multiscale optical system.

One can understand intuitively how local aberration correction can be achieved using far fewer optical surfaces when compared with full-field correction. This intuition has been formally verified by showing that expansion of the aberration polynomials about a local origin reduces terms dependent on high orders of the field angle, moving that wavefront error into lower-order terms that are more easily corrected.[1]

The full multiscale concept therefore allows for local correction of spatially-varying aberrations introduced by the objective. An early realization in the MOSAIC project was that by using a *monocentric* objective (i.e. one in which all of the optical surfaces share a common center of curvature), the aberrations can be made independent of field angle. Thus, provided that the microcameras are placed with spherical symmetry relative to the same center of curvature, the aberrations seen by the individual microcameras are *identical*. This requires only a single microcamera prescription, greatly reducing the design and fabrication challenges.[2,3] The specifics of the MOSAIC optical design are presented in other manuscripts associated with this conference.

We turn now to the image formation challenges that arise in this parallel architecture. In the standard sequential architecture, image formation is primarily a task of the optics, which create a nominally well-formed image that is then sampled by the focal plane (and possibly corrected via post-processing). In the parallel approach, each microcamera produces its own well-formed image of a *sub*-FoV. Therefore, image formation in the multiscale case necessarily consists of not only the action of the objective and individual microcameras but also the electronics and processing components that consolidate and combine the sub-images to form the overall image of the FoV. In this, multiscale imaging has much in common with earlier imaging array concepts (which were parallel arrays of microcameras, with no common objective). However, the presence of the objective, while greatly increasing system performance, introduces a number of constraints that make image formation in multiscale systems fundamentally different from earlier image arrays. In what follows we assume the current MOSAIC approach, which is based on a monocentric objective for the reasons given above.

In this context, the main image formation challenges that we have identified so far include: optimizing the sub-FoV overlap, determining sub-FoV registration, and compositing the sub-FoVs into the final system FoV. The following section describes these challenges in greater detail and presents our progress to date in addressing them.

3. KEY IMAGE FORMATION CHALLENGES AND PROGRESS TO DATE

3.1 SUB-FOV OVERLAP

The first challenge that arises is ensuring that there is sufficient sub-FoV overlap to avoid gaps in the FoV and facilitate any necessary registration between neighboring sub-FoVs. In traditional imaging arrays, sub-FoV overlap is solely a function of microcamera prescription and placement, and they remain key factors in the multiscale version as well. However, the presence of the objective couples its prescription with that of the microcameras, and monocentric design requires microcamera placement to occur on a spherical shell. These additional factors complicate the design.

The simplest way to visualize the problem is to consider the FoVs of the individual microcameras *on the spherical intermediate image surface of the objective*. The sub-FoV of an individual, cylindrically-symmetric microcamera appears as a small circular region on this spherical surface. Arranging the microcameras so that these circles cover the spherical surface as efficiently as possible (defined as minimizing the number of microcameras required to cover the desired system FoV while still providing the required amount of sub-FoV overlap) turns out to be an open, research-grade mathematical problem. The problem is further complicated by the fact that the microcamera focal planes are *emphatically not* cylindrically-symmetric, resulting in actual sub-FoVs that do not have a circular shape on the intermediate image surface. This introduces the *clocking*, or rotation, of the microcamera about its individual optical axis as an additional free parameter during the design of the sub-FoV overlap. Thus, given a particular microcamera prescription (thereby fixing the magnification and the effective size of the sub-FoVs on the intermediate image surface), the positioning of each microcamera can be described by three angular parameters: the polar angle (θ), the azimuthal angle (ϕ), and the clocking angle (γ).

The pointing angles (θ and ϕ) of the microcameras were determined early in the project by placing the individual microcamera optical axes on the nodes of a modified godesic dome structure. While not a provably-optimal covering strategy for the sphere (it results in a variation in the inter-camera angles between microcameras), it has the benefit of being relatively straightforward to calculate and produces an apparently sensible design. The remaining free parameters are the microcamera magnification and the individual clocking angles γ. The goal was to *maximize* the magnification, thereby minimizing sub-FoV overlap and the corresponding division of light between multiple microcameras, while still ensuring no gaps in the composite FoV and providing overlap regions suitable for any registration tasks that might be required. In the end, three different clocking schemes were considered, termed *original*, *alternating*, and *symmetric*. Their general structure, as applied to the early MOSAIC 'MC0' testbed is shown in Fig. 2. In each case, display of the central, on-axis microcamera has been suppressed for clarity. Each subplot shows how the individual microcamera sub-FoVs tile the spherical intermediate image surface. Here X and Y refer to $\theta_x = \theta \cos\phi$ and $\theta_y = \theta \sin\phi$, respectively.

There are several key things to note in the plot. First, that the geodesic method of determining microcamera pointing results in a five-fold rotational symmetry, with a 'spoke' of microcameras oriented along the $+X$ axis and four others arranged symmetrically around the center. Second, that the rectangular focal place array crops the sub-FoV from a circle to a lozenge shape with two parallel edges connected by circular arcs. The centers of these parallel edges are the locations of *minimal extent* of the individual sub-FoVs. If we can ensure that these points always overlap another sub-FoV, even in the presence of a reasonable microcamera misalignment, then we simultaneously ensure that there can be no gaps in the overall FoV.

Looking closely at the three proposed schemes, we see that the original and symmetric configurations produce locations (primarily along the spokes) where the flat edge of one microcamera sub-FoV overlaps the flat edge of its neighbor's sub-FoV. This is the most fragile of all possible situations, requiring the minimal amount of pointing perturbation to produce a gap. In contrast, the alternating design never produces this situation and

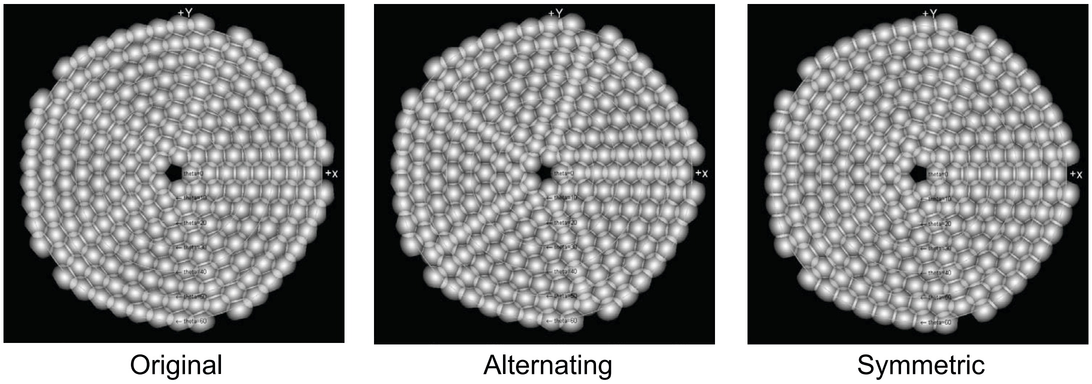

Figure 2. Microcamera clocking schemes considered in the MOSAIC program.

is correspondingly far more robust. Further, the overlap regions produced in the alternating design are more uniform in shape, allowing a fixed magnification to produce overlap regions of useful size across the entire FoV.

3.2 REGISTRATION

Just as the sub-FoV overlap problem in multiscale imagers differs in a fundamental way from earlier imaging arrays, so to does the nature of sub-image registration differ from the image registration problem in computer vision and photogrammetry applications. In those areas, the registration is typically *unconstrained*—that is, there is little (if any) prior knowledge regarding the magnification and relative pointing of the different imagers that formed the individual sub-images. In the multiscale case, we encounter the opposite extreme—there is almost perfect knowledge regarding these parameters; the only deviation from design/calibration is the result of drift within and between microcameras. Barring mechanical impact, the most likely sources of this drift are thermal expansion/contraction and gravitational and inertial deformation. With proper optomechanical design, one can expect these effects to be small, but not nonexistent—in particular, thermal issues will be challenging to avoid, especially if dynamic power management strategies produce thermal loading with a spatio-temporal variation across the array.

The scale of the registration problem also differs from earlier systems. From looking at Fig. 2, we see that, in the regions between the spokes, the microcamera arrangement approximates a hexagonal lattice structure. Therefore, as the number of microcameras increases, the average number of neighbors for each microcamera approaches 6. Thus, for N microcameras in an array, there are approximately $K = 6N/2 = 3N$ unique double overlap regions. The early 'MC0' testbed will eventually incorporate $N \approx 200$ microcameras, yielding $K \approx 600$ overlap regions (and corresponding registration computations). For the more ambitious 'MC1' and 'MC2' designs that will be explored later in the program, we anticipate $N > 3500$, resulting in $K > 10^4$ overlap regions and corresponding computations.

Given the fact that we come at the problem with significant prior knowledge regarding pointing and magnification, we have rejected the traditional unconstrained registration strategy and instead view the state of the array as a *perturbation* from the design/calibration state. We maintain a parameterized model of the array based on the optical and optomechanical design (and eventually system calibration). The registration task then will focus on finding perturbations to this parameterized model that best explain the observed information from the overlap regions. A side-benefit of this approach is that we can use this parameterized model as a *forward model*, allowing us to simulate the focal plane measurements across the entire array, given a suitable input scene. Using this approach on the design prescription of the 'MC0' testbed, we have analyzed the performance our our parameterized system model and find that produces errors of only sub-pixel magnitude. Figure 3 shows a synthetic input scene and several simulated 'MC0' focal plane measurements that were generated via the forward model.

Finally, given that we expect small perturbations dominated by the slow timescales inherent with thermal drift, we consider an approach where the registration task is performed only *periodically*, amortizing the relatively

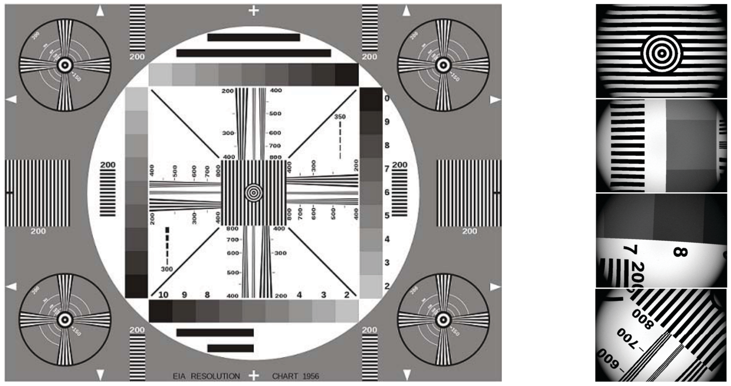

Figure 3. (Left) Synthetic test chart used as input to the parameterized system model. (Right) Several simulated 14 megapixel (reduced for publication) focal plane measurements for different microcameras. The strong vignetting of the microcameras is clearly seen. The rotation of the object within each image depends on the clocking of the particular microcamera.

large computational load across many frames. How this impacts the image compositing task is discussed in the next subsection.

3.3 IMAGE COMPOSITING

As mentioned above, we consider a registration approach that produces periodic updates to a parameterized model. This means that our methodology for image composition will effectively be *blind*. The focal plane measurements are mapped to object space using the parameterized model with the current parameter values and without regard to the data (the only exception being when new parameters are estimated from the measurements of the overlap regions). There are two central challenges to the blind composition approach. The first is determining how to structure the compositing algorithm so that it facilitates rapid construction of the final image. The second is deciding upon a methodology for combining measurements in the overlap regions.

To maximize the compositing speed, we have decided to structure the algorithm so that it is as parallelizable as possible. In this way, all available computing resources can be brought to bear. To this end, we have settled on a compositing approach that is compatible with a MapReduce framework.[4]

MapReduce is a modern approach to distributed data processing that can facilitate massively parallel implementations in some cases. Application of a MapReduce approach is contingent on the processing task having several properties. First, the input data must be representable as a list of {key, value} pairs. Second, the processing task itself must be separable into two subtasks with particular properties. The first subtask (called the *map* step), must operate on individual {key, value} pairs, converting them to another {key, value} pair (changing either the key, the value, or both). The second subtask (called the *reduce* step), must operate independently on any set of new (post map-step) {key, value} pairs with a common key value, producing a single {key, value} pair for that particular key value. The output of the processing step is the final list of these resulting {key, value} pairs.

Blind composition of multiple microcamera images can be cast in this form. The initial data set is a list of pixel values from the focal planes, where the {key, value} pairs take the explicit form {(camera #, pixel #), pixel value}. Here the key is itself a tuple expressing the microcamera number and the pixel number within the camera. The parametric system model can then be applied to these {key, value} pairs independently, and is thus perfectly parallelizable. This is the map step. The resulting data is a new list of {key, value} pairs with the form $\{(\theta,\ \phi)$, (pixel value, relative illumination)}. Here both the key and value in the {key, value} pair are tuples. The key describes the angular coordinates of the point in object space that illuminated the original camera pixel, while the value describes that pixel value, as well as the expected relative illumination of the microcamera for that pixel (describing the vignetting). The irradiance of that particular (θ, ϕ) coordinate in object space can then be estimated from the pixel value and relative illumination of every measurement corresponding to that particular key value. This can also be done independently of all other sets of key values, and is thus perfectly parallelizable. This is the reduce step.

Determining the specifics of the reduce step is the second compositing challenge alluded to above. Given a set of pixel measurements (and associated relative illumination values), how do we estimate the irradiance of the corresponding point in object space? We consider an individual pixel measurement m as being related to the source irradiance i via

$$m = ri + n, \tag{1}$$

with r the relative illumination corresponding to that particular camera/pixel location and n a random AWGN component. In this case we can easily derive the maximum-likelihood estimator (MLE) for the irradiance i

$$i_{\mathrm{MLE}} = \frac{\sum_k r_k m_k}{\sum_k r_k^2}, \tag{2}$$

where the sum is over the set of all pixel measurements corresponding to the object space location. We are currently using this MLE approach in our reduce step.

We have investigated the performance of the compositing algorithm using the simulated measurements of our system forward model. In Fig. 4 we see the resulting composited FoV using simulated measurements from approximately 200 microcameras in a nearly-full 'MC0' configuration, as well as a closer view of how three sub-images are composited—the latter demonstrating the effectiveness of the MLE approach in overcoming the vignetting of the microcameras. The input scene was an approximately 800 megapixel image of the Minneapolis government plaza area.[5] These results were generated using the MapReduce-based algorithm, but with fully *sequential* processing (using a single processor). We are currently creating a fully parallel implementation.

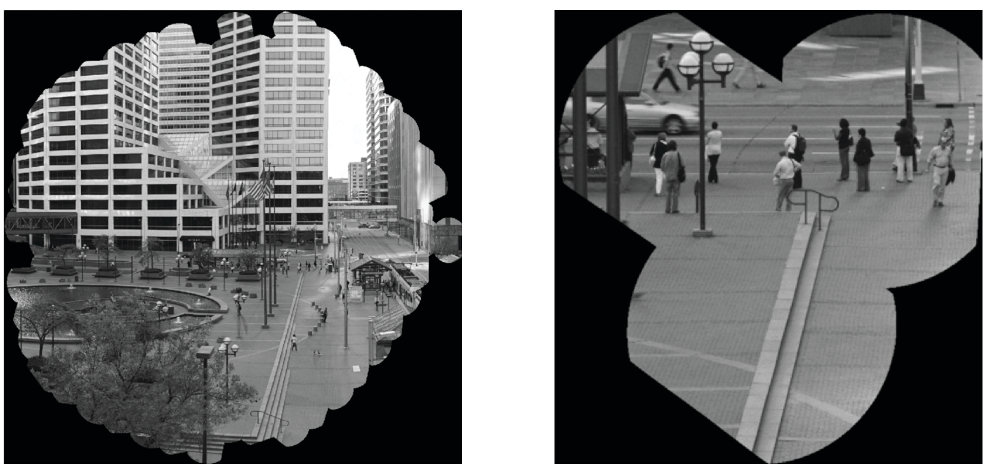

Figure 4. (Left) Composited FoV. (Right) Close-up of how three sub-images are composited.

4. INITIAL RESULTS

While waiting for the fabrication and assembly of the 'MC0' hardware, we have also constructed a small compositing testbed consisting of a triad of microcameras and a shared objective. This testbed was later expanded to include a pair of microcamera triads. At the time, the final 14 megapixel focal planes and control electronics were not yet available, so we utilized a set of existing 5 megapixel detectors and associated electronics. As they were not designed for the 'MC0' objective and microcameras, the focal plane boards were too large to allow tight packing of the microcameras (i.e. there is a FoV gap between the three cameras that is not present in the real system). Nonetheless, this testbed provided an interesting opportunity for exploring our compositing approach. Figure 5 shows the testbed and the 'first light' image. The testbed is shown in its expanded configuration with a pair of microcamera triads. The first light image shows a portion of the main engineering building at Duke University.

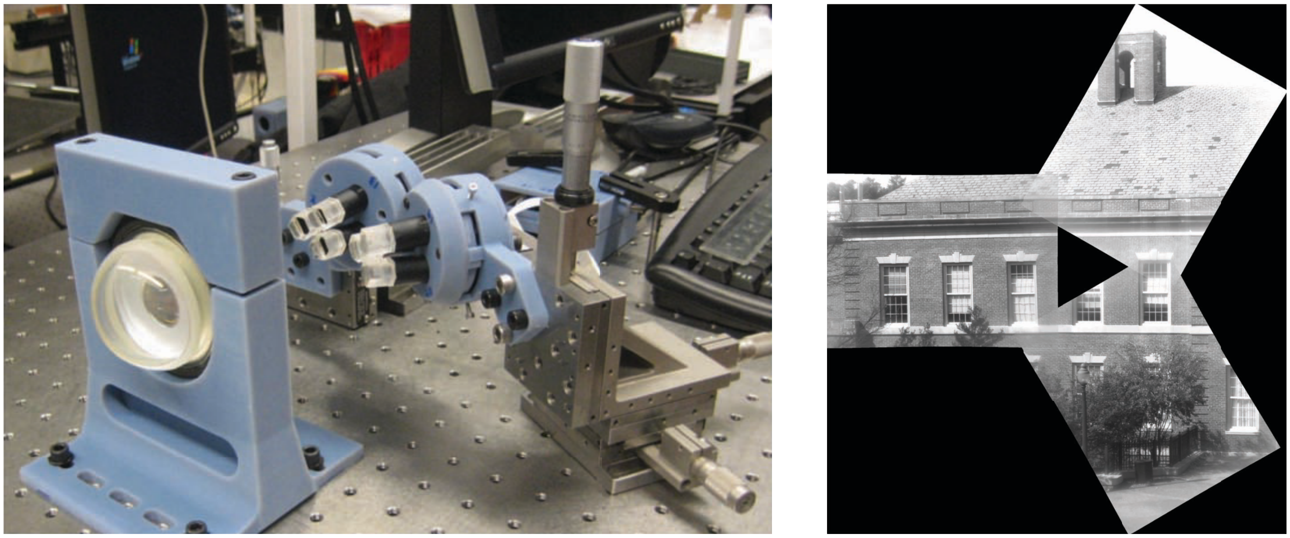

Figure 5. (Left) Compositing testbed. (Right) First light image from the testbed generated by our compositing approach and parameterized system model.

The first light image, while not perfect, is nonetheless impressive given the limitations of the testbed system—the optomechanics are rapid-prototyped polymer of limited precision; the parameterized model used in the compositing algorithm is that of the ideal optical design and not the as-built microcamera optics; and the exposure and gain of the individual microcameras were only roughly hand-tuned. We expect greatly improved performance given the fully calibrated 'MC0' system.

5. CONCLUSION AND FUTURE WORK

We have identified several key image formation challenges that arise in the context of multiscale optical design for high-pixel count imaging, such as is found in the DARPA MOSAIC project. The most notable of these challenges are the design of the microcamera configuration to ensure proper sub-FoV overlap, registration of the sub-FoVs in the presence of inter- and intra-microcamera drift, and rapid compositing of the sub-FoVs into the overall system FoV.

Completion of the 'MC0' testbed is imminent. We are certain that, with access to this high-performance multiscale testbed, we will be able to validate the performance of our planned approaches as well as identify additional image formation challenges that arise in this context. We will report on the current status of the system.

ACKNOWLEDGEMENTS

We gratefully acknowledge the Defense Advanced Research Projects Agency (DARPA) for financial support via award HR-0011-10-C-0073.

REFERENCES

[1] Brady, D. J. and Hagen, N., "Multiscale lens design," *Opt. Express* **17**, 10659–10674 (Jun 2009).
[2] Ford, J. E. and Tremblay, E., "Extreme form factor imagers," in [*Imaging Systems*], *Imaging Systems* , IMC2, Optical Society of America (2010).
[3] Marks, D. L. and Brady, D. J., "Gigagon: A monocentric lens design imaging 40 gigapixels," in [*Imaging Systems*], *Imaging Systems* , ITuC2, Optical Society of America (2010).
[4] Dean, J. and Ghemawat, S., "Mapreduce: simplified data processing on large clusters," *Commun. ACM* **51**, 107–113 (January 2008).
[5] Nelson, T., "Minneapolis government center." Posted to gigapan.org and provided via personal communication (2010). http://www.gigapan.org/gigapans/46074/.