Ultra-miniature, computationally efficient diffractive visual-bar-position sensor

Mehjabin Monjur, Leonidas Spinoulas, Patrick R. Gill and David G. Stork
Rambus Labs
1050 Enterprise Way, Suite 700
Sunnyvale, CA 94089 USA
dstork@rambus.com

***Abstract*—We describe the design and performance of an ultra-miniature lensless computational sensor optimized for estimating the one-dimensional position of visual bars. The sensor consists of a special-purpose wavelength-robust optical binary phase diffraction grating affixed to a CMOS photodetector array. This grating does not produce a traditional high-quality human interpretable image on the photodetectors, but instead yields visual information relevant to the bar-position estimation problem. Computationally efficient algorithms then process this sensed information to yield an accurate estimate of the position of the bar. The optical grating is very small ($120\ \mu m$ diameter), has large angle of view (140°), and extremely large depth of field ($0.5\ mm$ to infinity). The design of this sensor demonstrates the power of end-to-end optimization (optics and digital processing) for high accuracy and very low computational cost in a new class of ultra-miniature computational sensors.**

Keywords: ***Computational sensing, diffractive imager, visual-bar-position sensor, lensless smart sensor***

I. Introduction

The discipline of computational imaging involves the design of both optics and digital signal processing to achieve a desired end-to-end system performance. Because much of the overall burden of imaging or sensing can be borne by the signal processing, the constraints upon the optical components can be relaxed. [1] As such, optical systems with fewer optical components or smaller form factors can be made. True joint design requires the definition of a global or end-to-end merit function and an explicit functional relationship between this merit function and the optical and signal processing parameters. [2] Under such circumstances, the design process can rely on gradient descent in the end-to-end merit function. Approximations to this method can include taking steps alternately in the optical parameters and then the signal processing parameters and iterating until convergence.

Recently, computational imaging systems have been designed that eschew traditional refracting or reflecting optical elements (lenses or curved mirrors) and which rely instead entirely upon *diffraction*. Such devices have been demonstrated in mobile medical microscopy [3], [4] and far-field imaging. [5]–[13] Simulation studies have shown that such diffractive systems can produce digital images nearly as accurately as do ideal lensed systems, but such lensed systems are very difficult to create at the spatial scales of diffractive imagers ($\sim 100\ \mu m$). [14]

Such computational sensors require more processing than do traditional camera-based systems in order to create a digital image. If we let n denote the linear size of a square sensor array, image computation by Tikhonov regularization is an $\mathcal{O}(n^4)$ algorithm, which is rather costly for even moderately large sensors. If the optical image is shift invariant (possibly after simple image dewarping to correct for optical barrel distortion), then fast-Fourier-based deconvolution algorithm of complexity $\mathcal{O}(n^2 \ln n)$ can be used. [11] Such convolution algorithms represent a small portion of the overall computational cost of analog-to-digital conversion, data transfer, and so forth.

The above discussion centered on computational *imaging*, and while nearly all high-level computer vision and pattern recognition algorithms (object recognition, bar-code reading, face recognition, ...) operate on image data, many *sensing* and *image estimation* tasks do not require such an image. These sensing tasks are particularly attractive for low-power sensors, including

- overall brightness estimation
- color gamut estimation
- visual flow estimation
- visual orientation estimation
- axial visual flow or "looming" estimation
- image change detection
- visual mark localization and tracking

In computational sensors of the sort described here, the algorithms addressing such sensing tasks can operate on the raw sensor signals, without the need for a traditional image. For this reason, we informally consider the diffractive optical element as a *computational device*—one that performs a signal processing function in parallel, in negligible time (the time it takes light to pass through the thin grating), and with zero electrical power dissipation. Our task in the work reported here is to exploit computational sensing design methodology to simplify the optical component and reduce the computational cost for one such target application: visual-bar-position estimation. Our broad motivation is to create inexpensive, low-power, application-specific sensors for use in mobile and standalone applications in automotive, biomedical, smart architecture, smart cities, and the Internet of Things.

We begin in Section II with a description of the image estimation task at hand, then turn in Section III to our overall design methodology. Next, we discuss in Section IV both our hardware design, in particular our special optical phase grating and the optical signals it captures, as well as the digital signal

 ISBN: 978-1-61208-425-1

processing. We present the overall sensor performance and computational cost in Section VI, and conclude in Section VII with a brief summary and suggestions for future research.

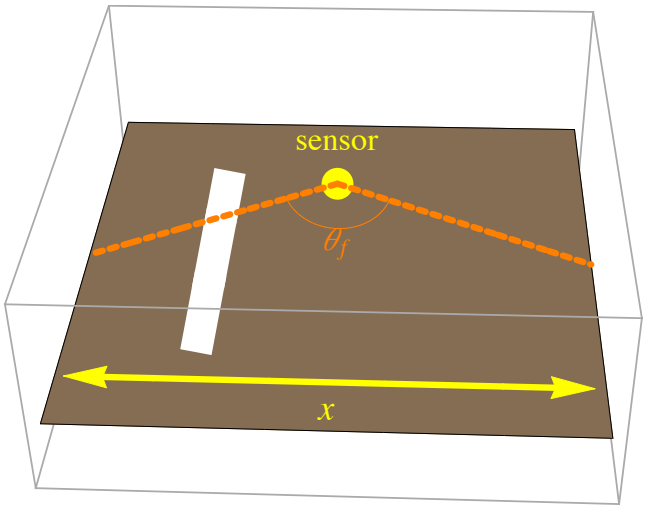

Fig. 1. The image sensing task is to estimate the left-right position, x, of a long vertical visual bar on a dark background in the far field (distance from sensor greater than 2 mm). The full-field angle-of-view of the sensor is $\theta_f = 140°$.

II. Image Estimation Task

Figure 1 illustrates our image estimation task. The target is a vertical bar on a dark background of unknown horizontal position x within a visual field of $\pm 70°$ or $\theta_f = 140°$ full field. The bar can be of arbitrary spectral composition in the visible range $400\ nm < \lambda < 700\ nm$. Throughout, the field luminance is greater than roughly $500\ lux$, but performance does not depend significantly upon luminance as long as it is within a few orders of magnitude of this value. Such an image estimation task appears in numerous practical applications from machine inspection (e.g., alignment of moving parts *in situ*), horizon tracking in unmanned aerial vehicles, lane-tracking in autonomous vehicles, estimation of the height of liquids in medical equipment such as test tubes, and others. We report here end-to-end simulation design and testing results, including optical diffraction and digital signal processing.

III. Sensor Design Methodology

As mentioned above, computational sensing and imaging relies upon the joint design of both the optics and the signal processing for a desired end-to-end performance of the digital imaging system. The ideal, true joint method is to form a global criterion or merit function, such as predicted image mean-squared error, then determine how the optical and the signal processing parameters affect this merit function, and then iteratively adjust all parameters simultaneously to optimize this global merit function. [2] For complicated optical systems, the dependency of the merit function upon a system parameter, such as the shape parameter governing a diffraction grating, cannot be determined analytically. In such cases, a number of related or approximate design methodologies can be used instead: one can adjust the optical parameters then the signal processing parameters, then the optical parameters and so on until global convergence of the merit function is reached. Alternatively, one can design a fixed optical element that has desirable application-specific properties and then optimize the signal processing. In our current work, we employed the latter technique.

We model the sensor as a linear system, that is,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{x}$ is a vector of inputs from the scene, $\mathbf{y}$ is the vector of photodetector pixel responses, $\mathbf{A}$ the system matrix describing the linear transformation performed by the two-dimensional optical grating, and $\mathbf{n}$ the additive noise, which describes photodetector noise, Poisson photon statistics, quantization noise, etc. (Other models, such as simple multiplicative noise, could also be assumed.) Here $\mathbf{x}$ is n-dimensional and $\mathbf{y}$ and $\mathbf{n}$ are m-dimensional, and hence $\mathbf{A}$ has dimensions $n \times m$. During design, we computed the system matrix $\mathbf{A}$ once for each candidate diffraction grating, then used it with bars at different positions $\mathbf{x}$ to simulate the signals on the photodetector array. We then implement the signal processing algorithms to estimate the bar position (cf., Sect. V). All simulation steps were implemented in *Matlab*.

IV. Sensor and Phase Grating

We are not aware of any general theory for deriving analytically the relationship between grating parameters of candidate phase gratings and a final merit function reflecting the accuracy of the final bar position estimate. The physical processes of diffraction, the constraints of grating manufacturability, and so forth, make such an analytical relationship complicated indeed. For that reason, we explored a number of grating designs "by hand," guided by knowledge of manufacturing and physical constraints. The active portion of the gratings we explored were $120\ \mu m$ in diameter, commensurate with sensor hardware described elsewhere. [8] Because the sensing problem appears one dimensional, we first tested linear (vertical) gratings. These yielded poor results in part because of geometrical effects due to sources at different depths: the projected images were curved toward the periphery, and that curvature depended upon the optical wavelength of light. Because the intermediate images varied significantly and in highly nonlinear ways for source bars of different spectral composition and spatial depth we could not design simple signal processing methods that reliably estimated the bar position across all such variations.

We then explored several classes of two-dimensional gratings. A traditional Fresnel diffraction grating gave informative images for bars emitting at its single design wavelength, but not for bars emitting elsewhere throughout the visible range. For this reason, basic Fresnel gratings and gratings closely related to it, were unacceptable. A particularly intriguing class of gratings were based on fractals, which were designed

 ISBN: 978-1-61208-425-1

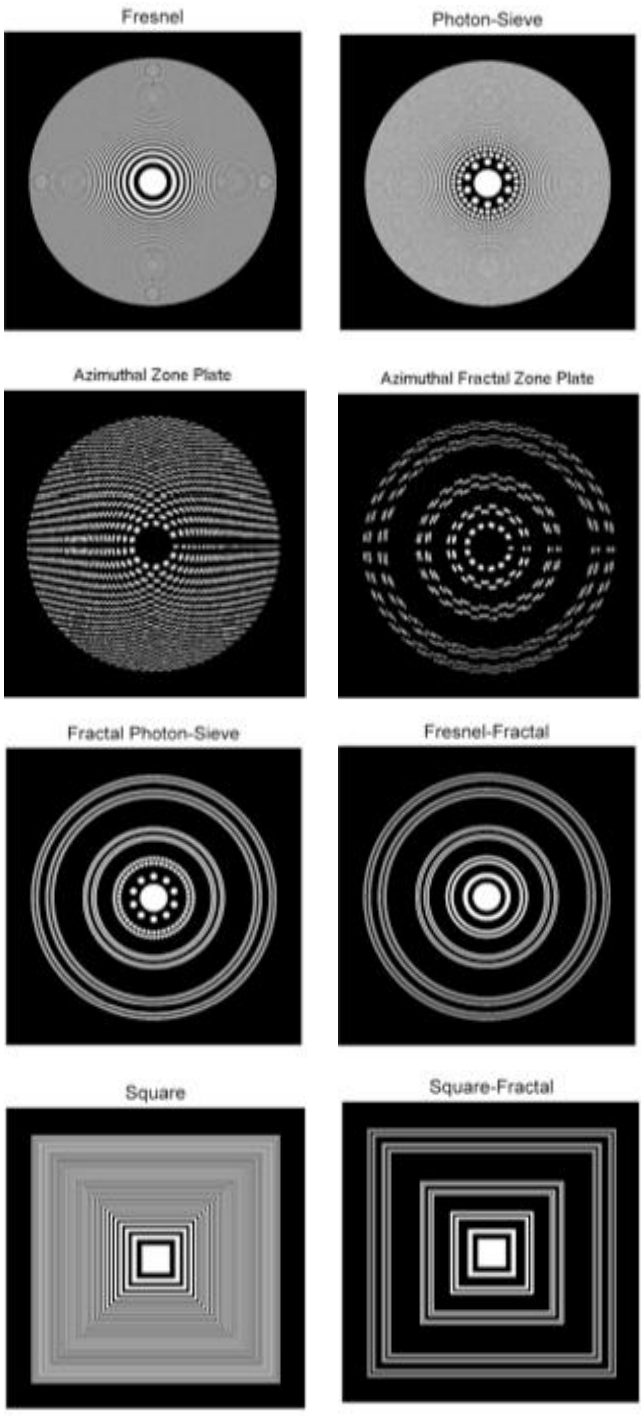

Fig. 2. Representative candidate binary phase gratings tested for one-dimensional bar position estimation.

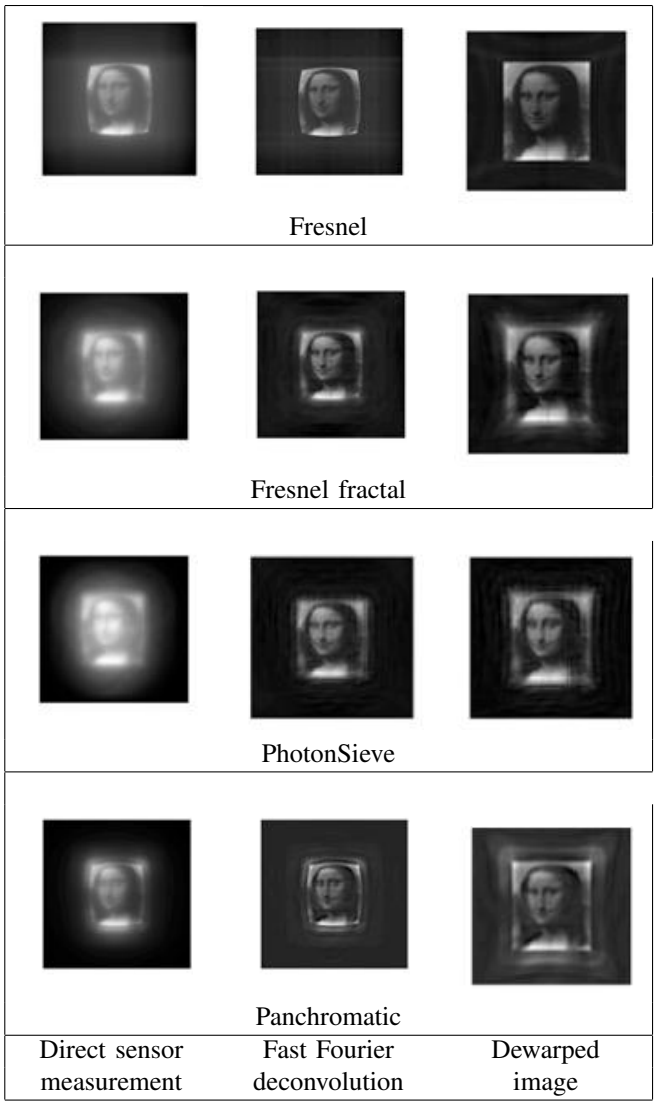

Fig. 3. Images produced by several of the candidate binary phase gratings shown in Fig. 2 for input scene tendered in blue light. These images were computed from the raw sensor signals by fast Fourier deconvolution, described elsewhere. [11] The images in the right column are radially dewarped versions of those in the middle column.

to be relatively insensitive to wavelength. [15] While these yielded intermediate images that were indeed robust to spectral variations, the spatial shape of the resulting images were somewhat complicated so no simple signal processing algorithms could accurately estimate the bar location. Figure 2 shows representative candidate gratings we explored.

Figure 3 shows the imaging performance of several grating designs (including the panchromatic Fresnel zone plate, see below), designed at the intermediate wavelength $\lambda = 550\ nm$ but rendered at $\lambda = 470\ nm$. The image of Leonardo's *Mona Lisa* was computed using fast Fourier deconvolution. [11] In traditional computer vision methodology, the ideal intermediate optical image would be sharp lines independent of wavelength, but such images cannot be achieved using binary phase gratings. Our goal, then, is the create optical images such that signal processing can estimate the visual bar location reliably, and at low computational cost, regardless of wavelength and distance of visual bar.

The grating that provided an easily processed image despite variations in bar spectral composition was the *panchromatic binary Fresnel zone plate*, governed by Eq. 2:

 ISBN: 978-1-61208-425-1

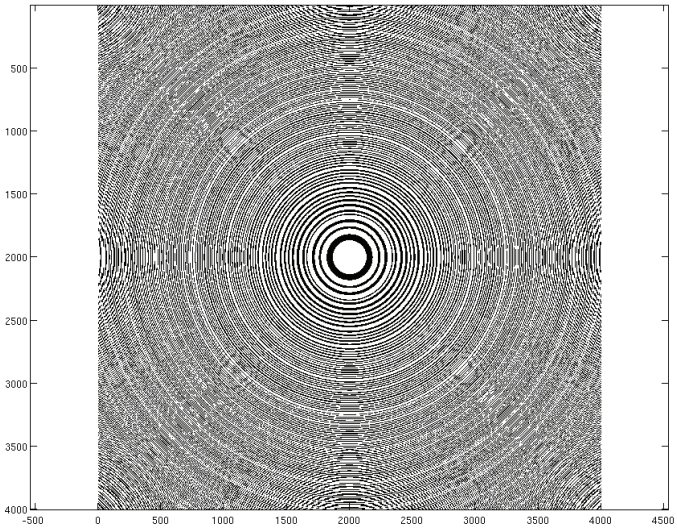

Fig. 4. Panchromatic binary phase Fresnel zone plate generated according to Eq. 2. This circularly symmetric grating yield optical images on the photodetector matrix that are robust to variations in source spectral content, as shown in Fig. 5.

$$PFZP(x,y) = \Theta\left[\sum_{i=1}^{3}\left|a + \frac{jk_i}{2\pi z_0}e^{-jk_i(x^2+y^2)}\right|^2\right], \quad (2)$$

where $\Theta[\cdot]$ is a Heaviside step or threshold function, a a bias constant, z_0 a spatial distance corresponding to the distance of a virtual point source that interferes with normal plane waves, the k_i the wave numbers of the optical wavelengths chosen from the range desired and $j = \sqrt{-1}$ is the unit imaginary number. This grating can be considered the mixture of separate Fresnel zone plates each designed with different wavelengths. Figure 4 shows a typical panchromatic Fresnel zone plate created with three wavelengths $\lambda_i = 470, 550$ and $700\ nm$, corresponding to blue, green and red ranges of the optical spectrum. This grating yielded intermediate images of visual bars such as shown in Fig. 5.

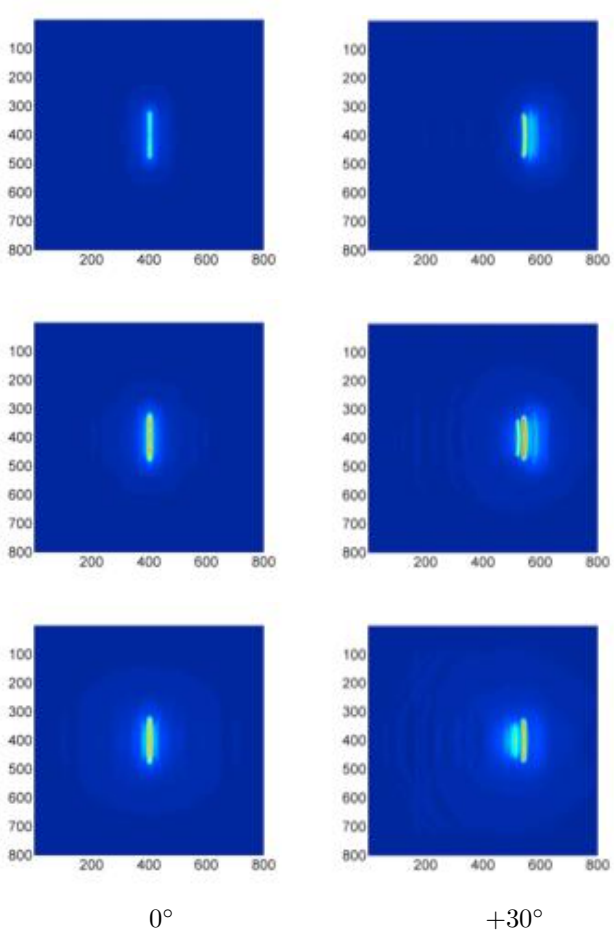

Fig. 5. Raw signals produced by the panchromatic Fresnel zone plate of Fig. 4 for a visual bar at 0° and 30° for $\lambda_i = 470, 550, 700\ nm$, top to bottom. Notice that on the axis (0°) the intermediate image is sharp and centered in the sensor matrix while off axis (30°) the image is structured, has multiple lobes, and is wavelength dependent. (Through simple signal processing lobes due to different discrete incident wavelengths can be isolated.)

V. Signal processing

Every grating we tested produced wavelength-robust, narrow intermediate optical images of visual bars at the center of the visual field ($\theta = 0°$) but somewhat complex multimodal images of bars at large field angles (e.g., $\theta = 40°$). It is likely that a method of spatially varying dictionary learning [16] or complex dewarping followed by Bayesian or other pattern classification method [17] could be used to estimate the bar location from such optical images but such methods are computationally costly. Our goal was to find an algorithm with low computational cost, and thus we sought algorithms that were spatially independent, that is, applied the same algorithmic steps throughout the sensor domain.

The estimation problem is one-dimensional (the left-right position of the visual bar) and our first algorithmic step was to project the sensor signals onto a one-dimensional horizontal line. This step integrates signals over vertical lines and hence reduces the effects of image noise and variations due to optics. The optical signals were reliable and accurate throughout the full field of view. The next computational step is to estimate some measure of the center or central tendency of the one-dimensional projected signal. We tried several computationally efficient methods to this end:

- peak or mode
- peak or mode after Gaussian smoothing
- mean of a Gaussian fit to the signal
- gradient pre-processing,

three of which are illustrated in Fig. 6.

VI. Sensor performance

Figure 7 shows the results of five different signal processing algorithms for the most robust diffraction grating, the panchromatic binary Fresnel zone plate. Each subfigure shows three curves representing the estimated bar position (ordinate) versus the actual bar position (abscissa) in long, medium and short

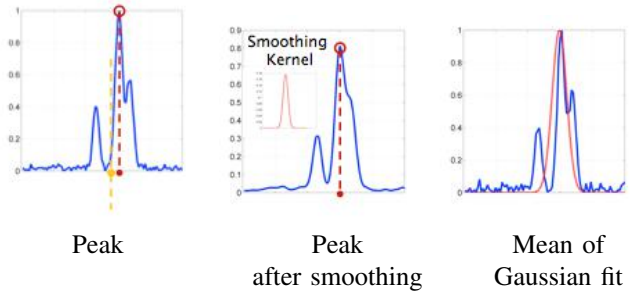

Fig. 6. The blue curves represent the projection of the full two-dimensional sensor signal along the vertical direction (i.e., parallel to the visual bar). The methods of central tendency estimation (for source angle) that were most accurate, reliable and robust to source spectral variation were the peak and the peak after a Gaussian smoothing (see below). This latter estimate did not depend upon the variance of the smoothing Gaussian, throughout a wide range of such variances.

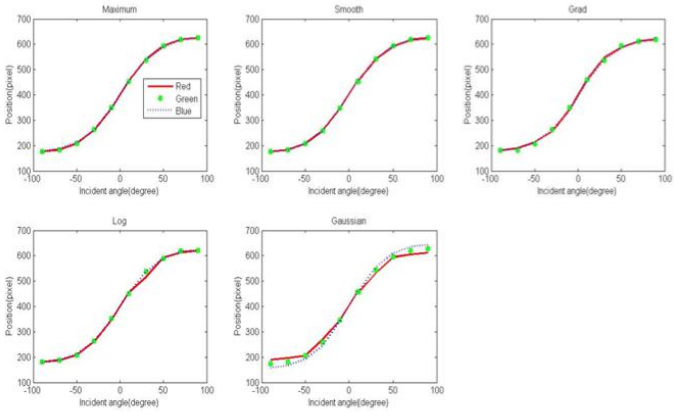

Fig. 7. The estimated bar center (in pixels on the the sensor) versus bar position in the field of view for a panchromatic Fresnel zone plate. Each subfigure contains three curves (for $\lambda = 470, 550$ and 700 nm light), but in several subfigures the curves overlap significantly and hence appear as one curve. The five signal processing algorithms are (left-to-right, top-to-bottom) are: maximum or peak; peak after Gaussian smoothing; gradient-based estimation; peak after log processing; mean of a Gaussian fit.

wavelengths—red, green and blue. (In most subfigures, the three component curves overlap significantly and hence appear as one curve.) Each curve has a general sigmoidal or ogive shape due to geometric effects and Snell's law of light passing through the grating at large incident angles (a version of barrel distortion).

The desiderata for the sensor are:

- all three color curves overlap perfectly throughout the entire angle range, indicating that the location estimate is independent of spectral composition of the visual bar
- all three curves are monotonic (a bijection), so that the unique angle estimate can be computed from the image center estimate by an inverse function or lookup table
- all three curves extend through a large angle of view
- antisymmetry of response with respect to central axis, indicating geometric consistency.

Both first two subfigures—peak and peak after Gaussian smoothing–show excellent performance on all four desiderata and accuracy of roughly $0.2°$ throughout the field of view, raising to $0.4°$ at the extremes of the field of view. The difference between the computational costs of these two methods are negligible compared to the full computational costs. Either of these methods, then, would be acceptable in a fielded sensor application.

The overall space computational cost of the estimation algorithm was *1.0 kB* without lookup table for inverting the sigmoidal curves (as in Fig. 7), and time cost *6.0 Mflop/sec* at video rates, and hence easily implemented in embedded processors or special CMOS. The signal projection step of the algorithm could be parallelized in a SIMD microarchitecture, but such speedups are not needed in practical applications.

VII. Conclusions and Future Directions

We have designed and tested through extensive simulations an ultra-miniature lensless sensor for estimating the one-dimensional position of a visual bar throughout a large field of view and regardless of the spectral composition of the bar. Our end-to-end design approach led to an optical element (panchromatic binary Fresnel zone plate) that while somewhat complicated in design, is simple to manufacture and mount on a CMOS image sensor. The signal processing operates on the raw sensor signal (rather than a reconstructed or computed image) and is very computationally efficient. There are a number of directions for future work in end-to-end optimization based on these results, such as extending these methods to other image sensing functions. Finally, the deep challenge of developing a general-purpose theory for global end-to-end optimization of diffractive electro-optical systems incorporating physical constraints—analogous to the theory developed for traditional reflection and refraction based imagers [2]—remains elusive.

Acknowledgments

We thank Thomas Vogelsang for helpful comments. M. Monjur and L. Spinoulas of Northwestern University performed this work as 2014 summer interns in Rambus Labs.

References

[1] W. T. Cathey and E. R. Dowski, Jr., "A new paradigm for imaging systems," *Applied Optics*, vol. 42, no. 29, pp. 6080–6092, 2002.

[2] D. G. Stork and M. D. Robinson, "Theoretical foundations of joint design of electro-optical imaging systems," *Applied Optics*, vol. 47, no. 10, pp. B64–75, 2008.

[3] D. Tseng, O. Mudanyali, C. Oztoprak, S. O. Isikman, I. Sencan, O. Yaglidere, and A. Ozcan, "Lensfree microscopy on a cellphone," *Lab on a chip*, vol. 14, pp. 1787–1792, 2010.

[4] A. E. Cetin, A. F. Coskun, B. C. Galarreta, M. Huang, D. Herman, A. Ozcan, and H. Altug, "Handheld high-throughput plasmonic biosensor using computational on-chip imaging," *Light: Science and Applications*, vol. 3, pp. e122–, 2014.

[5] P. R. Gill, "Odd-symmetry phase gratings produce optical nulls uniquely insensitive to wavelength and depth," *Optics Letters*, vol. 38, no. 12, pp. 2074–2076, 2013.

[6] P. R. Gill and D. G. Stork, "Lensless ultra-miniature imagers using odd-symmetry phase gratings," in *Proceedings of Computational Optical Sensing and Imaging (COSI)*, Alexandria, VA, 2013.

 ISBN: 978-1-61208-425-1

[7] ——, "Digital camera with odd-symmetry spiral phase gratings supports full-resolution computational refocusing," in *Advanced Photonics (Optical Society of America Sensors Congress)*, 2013.

[8] ——, "Hardware verification of an ultra-miniature computational diffractive imager," in *Proceedings of Computational Optical Sensing and Imaging (COSI)*, Kohala Coast, HI, 2014.

[9] D. G. Stork and P. R. Gill, "Lensless ultra-miniature computational sensors and imagers," in *SensorComm 2013*, Barcelona, Spain, 2013.

[10] ——, "Optical, mathematical and computational foundations of lensless ultra-miniature diffractive imagers and sensors," *International Journal on Advances in Systems and Measurements*, vol. 7, no. 3-4, pp. 201–208, 2014.

[11] P. R. Gill and D. G. Stork, "Computationally efficient Fourier-based image reconstruction in a lensless diffractive imager," in *Computational Optical Sensing and Imaging (COSI)*, Arlington, VA, 2015 (submitted).

[12] P. R. Gill, M. Kellam, J. Tringali, T. Vogelsang, E. Erickson, and D. G. Stork, "Computational diffractive imager with low-power image change detection," in *Proceedings of Computational Optical Sensing and Imaging (COSI)*, Alexandria, VA, 2015 (submitted).

[13] D. G. Stork, "Computational diffractive sensing and imaging: Using optics for computing and computing for optics," in *Society for Information Display*, San Jose, CA, 2015, p. (in press).

[14] L. Spinoulas, O. Cossairt, P. R. Gill, D. G. Stork, and A. K. Katsaggelos, "Performance comparison of ultra-miniature diffraction gratings with lenses and zone plates," in *Computational Optical Sensing and Imaging (COSI)*, Alexandria, VA, 2015 (submitted).

[15] W. D. Furlan, G. Saavendra, and J. A. Monsoriu, "White-light imaging with fractal zone plates," *Optics Letters*, vol. 32, no. 5, pp. 2109–2111, 2007.

[16] J. Mairal, J. Ponce, G. Sapiro, A. Zisserman, and F. R. Bach, "Supervised dictionary learning," in *Advances in Neural Information Processing Systems 21*, D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, Eds., 2009, pp. 1033–1040.

[17] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York, NY: Wiley, 2001.

 ISBN: 978-1-61208-425-1