

Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

Direct-detection synthetic-aperture coherent imaging by phase retrieval

James R. Fienup

Direct-detection synthetic-aperture coherent imaging by phase retrieval

James R. Fienup*

University of Rochester, Institute of Optics, Rochester, New York, United States

Abstract. This paper describes a way to synthesize a larger coherent aperture from smaller apertures combined with motion, when only intensities are measured. It relies on collecting intensity patterns in two planes for each aperture, for example, the aperture plane and an image plane, and using a phase-retrieval algorithm to reconstruct the optical field in the aperture plane. As the sensor moves forward, a larger two-dimensional aperture is synthesized, allowing a much finer resolution image to be reconstructed. An algorithm to correct for the relative pointing (tip and tilt phases) and piston errors between different apertures and at different times is needed to phase up the synthetic aperture. Results of simulations, including the effects of speckle, are shown, and practical considerations are evaluated. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.OE.56.11.113111](https://doi.org/10.1117/1.OE.56.11.113111)]

Keywords: synthetic aperture; imaging; phase retrieval; phase-error correction; image reconstruction; heterodyne.

Paper 171359 received Aug. 28, 2017; accepted for publication Nov. 1, 2017; published online Nov. 21, 2017.

1 Introduction

To achieve fine resolution imagery at a given long distance and in a given wavelength band, one needs a collection system (telescopes) having a large enough effective aperture. As an alternative to building and deploying larger single-aperture systems (which become increasingly bulky, heavy, and costly), one can perform aperture synthesis. This can be done either passively (using reflected sunlight) as in Michelson stellar interferometry or actively as in coherent laser illumination with phase-sensitive detection. Michelson interferometry requires two or more simultaneous apertures having substantial motion of one aperture relative to another and, for a variety of reasons, is poorly suited to looking downward at the earth. One could use laser synthetic-aperture radar (SAR) in which temporal or chirped frequency sensing provides range information and forward motion provides along-track resolution; temporal heterodyne sensing over large temporal bandwidths is required for fine range resolution. Digital holography, also known as spatial heterodyne, can achieve fine resolution in angle-angle space by a string of apertures in the cross-track direction combined with aperture synthesis in the along-track direction; it can employ narrow laser bandwidths but must still interfere the return field from the object with a local oscillator (LO), requiring stable LO distribution from a master laser to all the telescopes. One could employ multiple small apertures underneath the wings of an aircraft, or on a group of small satellites, or on a moving ground vehicle, to synthesize a large two-dimensional (2-D) aperture with fine resolution. Images from laser illumination systems exhibit speckle that degrades the effective resolution unless one gathers multiple speckled images, with different speckle realizations, and averages together their intensities.

Because of a desire to avoid the complications of LOs, including distribution, stability, relative Doppler, and timing,

this paper considers an alternative imaging architecture similar to digital holography but employing direct detection and phase retrieval instead of spatial or temporal heterodyne. The technique requires, for each telescope, the simultaneous detection of the received beam in at least two planes, where typically one would select an image plane, where one obtains a low-resolution intensity image of the object and a pupil plane (which is a reimaged aperture plane). Then one would use a phase-retrieval algorithm to reconstruct the field in the plane of the aperture. This is done for each laser pulse and for each telescope within an array of telescopes. An example of a system configuration is shown in Sec. 2. Now having the fields within each aperture position, one can synthesize a larger coherent aperture. Since the relative phases between the different apertures will be unknown, these relative phases must be reconstructed from the measured and processed data. Uncertainties in the relative pointing of the different telescopes result in relative linear phase errors, and uncertainties in the relative distances from the target center and each telescope result in piston errors between telescopes, and these are assumed to change for different laser pulses and for different telescopes. These phase errors must be sensed and corrected from the measured and processed data. Section 3 describes the algorithms developed for these purposes and shows that successful image reconstruction and interaperture phasing can (in simulation) be accomplished even for very low signal-to-noise ratios (SNRs), as low as four photons per speckle within the detection planes, which is equivalent to an SNR of 2. It also shows the effect on image quality of having different numbers of speckle realizations, each of which requires an additional synthetic aperture of data to be collected. It is recommended that approximately 10 speckle realizations of each image be collected for high-contrast targets and possibly more for low-contrast targets.

Section 4 examines system requirements, including pulse repetition frequency (PRF) and laser coherence length, and derives the relationship among laser power, wavelength, area of the scene illuminated, and other parameters. The area

*Address all correspondence to: James R. Fienup, E-mail: fienup@optics.rochester.edu

coverage rate, collecting multiple images, is shown to be proportional to the laser power available and inversely proportional to the number of speckle realizations averaged to get one image. Speckle “boiling” (the “memory effect”) is shown to be negligible for long-range imaging. For low SNRs, the direct-detection approach is shown to be somewhat noisier than heterodyne sensing, but for higher SNRs it yields images of quality comparable to heterodyne sensing. Finally, a comparison with Fourier ptychography is made.

2 System Concept

For the purpose of performing systems analysis and simulations, a particular reference approach was chosen. Many variations on this theme are possible. Figure 1 at the top shows an example of a sparse array of apertures, one small aperture (i.e., pupil) per small telescope, moving together. Their horizontal separations are necessitated by the sizes of the telescope support structures. Employing several short laser pulses, they advance downward in the vertical direction and synthesize the larger aperture shown at the bottom. The aperture transmitting the laser illumination beam can be one of the receive apertures or can be another telescope which is either moving along with the receive array (approximately monostatic case, which we will assume here) or can be somewhere entirely different (bistatic case). Furthermore, the single transmitter can be replaced with multiple spatially separated transmitters, which can be advantageous from the point of view of rapidly synthesizing a large aperture from a smaller number of telescopes. There are many possibilities for the number and arrangement of transmitters and receivers to achieve a desired synthetic aperture.

Aperture synthesis with coherent electromagnetic radiation (laser light) requires that the complex-valued optical fields be sensed. This is commonly done by heterodyne techniques, interfering the receive field with an LO or, in holographic terms, a reference beam. Due to an assortment of difficulties in dealing with the LO for an array of receivers,

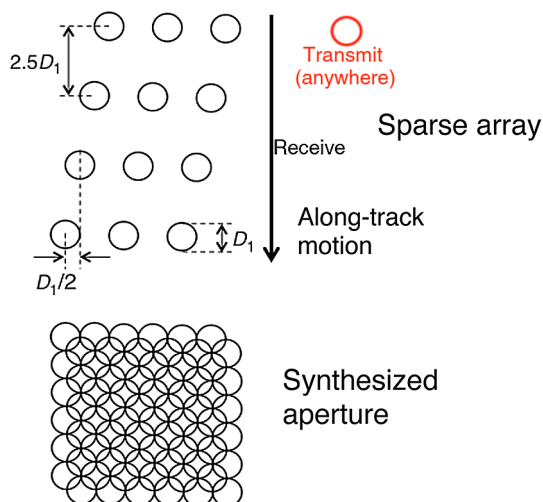


Fig. 1 Face-on view of an array of telescopes. Top: Sparse array of telescope apertures moving together; D_1 is the diameter of one of the telescope apertures. Bottom: roughly square aperture synthesized from telescope pupils using multiple laser pulses (some of the synthesized aperture outside the square area is not shown).

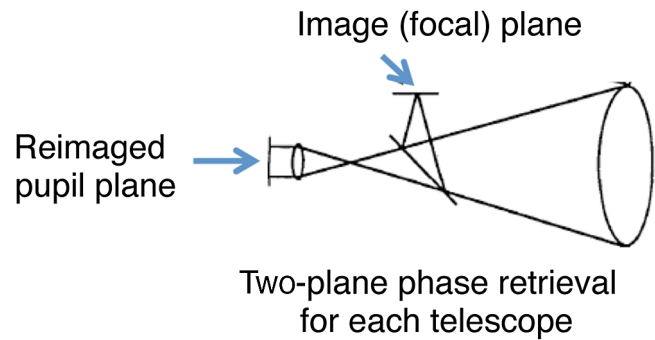


Fig. 2 Side view of a single telescope, with the object to the right, using two-plane detection to determine the optical fields with phase retrieval (a subset of a system shown in Ref. 1).

we wish to obtain complex-valued fields using only direct detection of intensity. This can be accomplished with a collection approach shown in Fig. 2.¹

If one collects a 2-D array of speckle intensity in each of two planes, then one can recover a diffraction-limited field across the entrance aperture (the large lens on the right) using a phase-retrieval algorithm such as the Gerchberg–Saxton algorithm.^{2,3} Each of the telescopes shown in the top of Fig. 1 would employ a pair of detection planes. The detection planes would likely be in the pupil plane of the telescope and in an image plane, as illustrated here, but other planes (and numbers of planes) are possible as well.

The fields within each of those pupils would have a random piston phase relative to one another. They would also likely have a random tip/tilt phase as well, corresponding to a translation of the image (due to nonidentical pointing of the different telescopes) relative to all the other pupils. The exact locations where the reconstructed pupil fields should be located within the synthetic aperture might also be known with insufficient accuracy if the locations of the telescopes with respect to one another are imperfectly known. Solving for these unknown phase and location terms would presumably be possible since they represent at most five additional parameters per aperture, as compared with thousands of phase values being solved by the phase-retrieval algorithm. The overlap of the different neighboring apertures making up the synthetic aperture, seen in Fig. 1, is very important for robust estimation of these five additional parameters. Note that if one were to use annular apertures, as shown in Fig. 3, then one still retains the overlap regions essential for solving for those additional parameters. While annular apertures do not perform quite as well as filled apertures (since some of the field is missing), we will show later that high-quality images can be reconstructed with annular apertures.

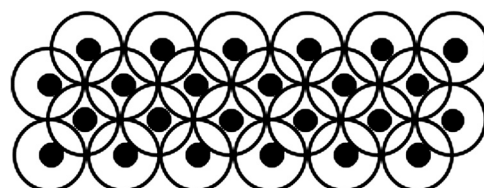


Fig. 3 Portion of a synthetic aperture for annular-aperture telescopes.

It is also physically possible to reconstruct fields from a single intensity measurement,⁴ but that approach places stringent demands on the illumination beam and lacks the high degree of robustness that we seek. For the sake of a more robust phase-retrieval algorithm, we will consider a two-plane direct-detection approach for which each aperture has two planes of intensity measurements and no LO.

An alternative to the architecture described above is Fourier ptychography.^{5,6} It is a method for synthesizing a coherent aperture from image-plane intensities, mostly used for microscopy. In the remainder of this paper we will concentrate on the two-plane phase-retrieval approach, but the Fourier ptychography approach (also employing phase retrieval), while very different in some respects, is expected to have performance that is in the same ballpark as two-plane phase retrieval, as discussed in Sec. 4.6.

3 Image Reconstruction Algorithms and Simulation Experiments

As mentioned earlier, the first step is to use a Gerchberg–Saxton type of iterative algorithm to reconstruct the complex-valued field within each aperture. Step 2 is to assemble the reconstructed fields into a synthetic aperture; while doing so, it is necessary to correct piston tip and tilt terms from the different telescopes. This was done using a fast subpixel registration algorithm,⁷ employing the overlap regions of the neighboring apertures within the synthetic aperture. For most of the simulations reported here, it was assumed that the transverse translations of the telescopes were known; hence we are solving for three additional unknowns per telescope.

Past experience shows that it is the number of photons per speckle that dictates image quality.⁸ For digital experiments, if the object field fills the array of numbers, then there is one speckle per sample in the aperture field obtained by computing a fast Fourier transform of the object array. If the object fills a width of $1/(q_{ap})$ times the width of the array, then the aperture field will have one speckle per q_{ap} pixels in each dimension, where q_{ap} is the sampling ratio in the aperture, relative to Nyquist sampling. $q_{ap} = 1$ is Nyquist sampled for the fields and $q_{ap} = 2$ is Nyquist sampled for the intensities. For our experiments, we added only shot (photon) noise to the measurements, that being the most fundamental source of noise. Further realism would be had by adding detector read noise, dark current, background noise, quantization noise, etc.

3.1 Six-Aperture Simulations

A large number of data sets were simulated and images were reconstructed, for the purpose of speed, with the six-aperture synthetic aperture shown in Fig. 4. When showing these synthetic apertures here, the sums of the circular apertures are shown, but the actual synthetic aperture is averaged in the areas of overlap rather than summed.

For reference, an ideal incoherent image through a single aperture is shown in Fig. 5. In this rendition of the 1951 US Air Force bar target, the finest pair of three bars that can easily be distinguished is indicated by the arrow, which is group 1, element 6, or (1,6) for short. Each group is worth a factor of two in resolution, which is the same as a delta-NIIRS = 1, where NIIRS is the National Image Interpretability Rating Scale.⁹ Each element within

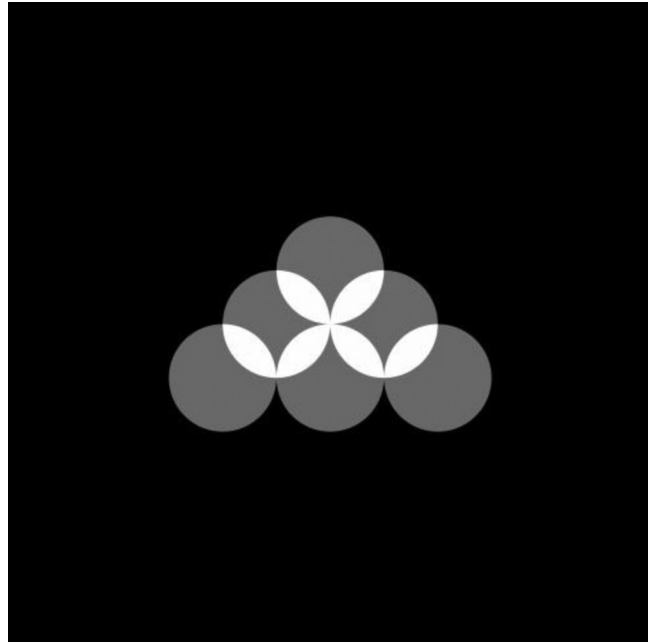


Fig. 4 Six-aperture synthetic aperture.

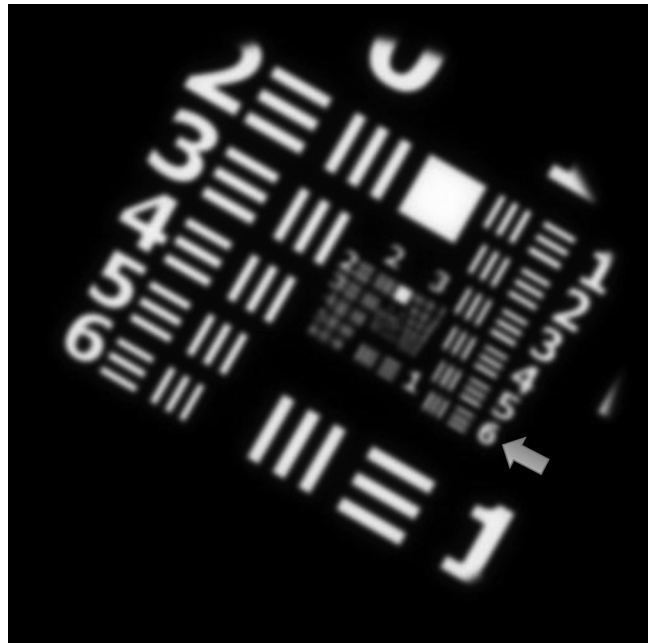


Fig. 5 Ideal incoherent image through a single aperture. In all images of bar targets, the arrow points to the finest pair of three bars that can easily be distinguished.

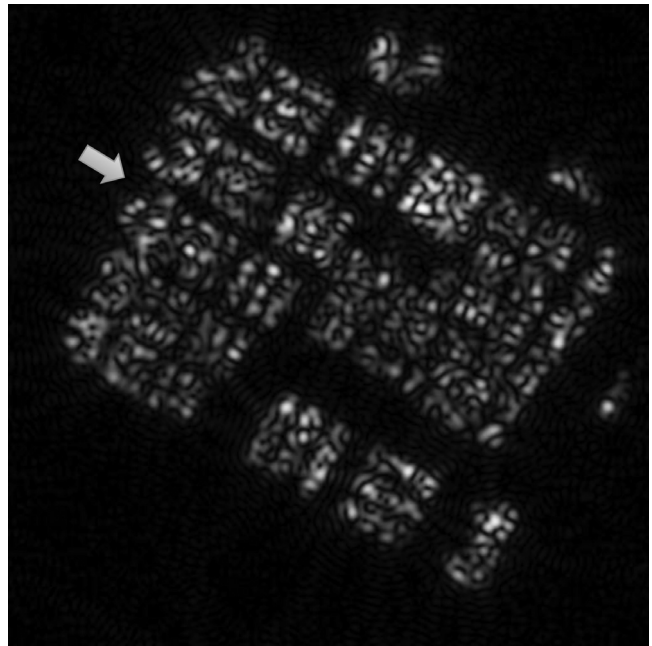
a group is worth a factor of $2^{1/6} = 1.12246$, or 12.2% additional resolution, equivalent to a delta-NIIRS of $1/6 = 0.167$. Each delta-NIIRS of 0.1 is worth a factor of $2^{0.1} = 1.0718$ in resolution. A delta-NIIRS of 0.1 is considered to be just perceptible.¹⁰ Hence a delta-element, being worth 0.167 delta-NIIRS, would be significantly greater than (1.67 times) just discernible. Table 1 shows the relationship among delta-groups, delta-elements, delta-NIIRS, and ratios of effective resolution.

Table 1 Relationship between delta-groups, delta elements, delta-NIIRS (loss in NIIRS), and resolution ratio.

Delta groups	Delta elements	Delta NIIRS	Resolution ratio
0	0	0	1
0	0.5	0.083	1.059
0	1	0.167	1.122
0	2	0.333	1.260
0	3	0.500	1.414
0	4	0.667	1.587
0	5	0.833	1.782
0	6	1.000	2.000
1	0	1.000	2.000
1	1	1.167	2.245
1	2	1.333	2.520
1	3	1.500	2.828
1	4	1.667	3.175
1	5	1.833	3.564
1	6	2.000	4.000
2	0	2.000	4.000
2	1	2.167	4.490
2	2	2.333	5.040
2	3	2.500	5.657
2	4	2.667	6.350
2	5	2.833	7.127
2	6	3.000	8.000
3	0	3.000	8.000

For comparison with the incoherent image, in which we can discern (1,6), Fig. 6 shows an ideal coherent, speckled image through the same single aperture, in which we can discern (0,3) or (0,4), or 1 group plus 2 or 3 elements worse resolution, equivalent (according to Table 1) to a loss of 1.3 to 1.5 in NIIRS and a factor of 2.5 to 2.8 in resolution. When viewing such poor images, one should “zoom out” or demagnify the image, putting the finest detail within a favorable part of the contrast sensitivity curve of the human visual system, allowing one to discern the greatest detail. Note that for much larger synthetic apertures, as will be shown later, one must zoom in or magnify the images to see the finest detail.

From Figs. 5 and 6, we see that the effect of speckle in coherent imaging (of any type) of optically rough objects plays a major role in effective resolution (a factor of 2.5 to 3) and image interpretability. Speckle is much more dominant for optical and near-infrared than it is for microwave SAR; in the microwave wavelength regime, the world is


Fig. 6 Ideal coherent, speckled image from a single aperture.

much smoother and targets often contain multiple glints that can aid in target recognition.

By gathering multiple (N_s) images with different transmitter/receiver locations relative to the object, each image having a different realization of the speckle pattern, one can average the speckled intensities together to yield a reduced-speckle image, which approaches an incoherent image as the N_s approaches infinity. The speckle contrast, initially unity, is reduced by the factor $1/\sqrt{N_s}$.¹¹ Figure 7 shows an ideal speckle-reduced image with $N_s = 100$ from a single aperture. It has a noisier appearance than the ideal


Fig. 7 Ideal single-aperture speckle reduced image, the average of $N_s = 100$ image intensities.



Fig. 8 Ideal image from six-aperture synthetic aperture with $N_s = 100$.

incoherent image shown in Fig. 5, but it has almost the same resolution. Figure 8 shows an ideal speckle-reduced image with $N_s = 100$ from the six-aperture synthetic aperture. With (2,6) discernible, it has a full factor of two better resolution than the single-aperture image, on account of the synthetic aperture having approximately twice the effective width as the single aperture.

Figure 9 shows the result of simulating data with $N_{pps} = 100$ photons per speckle, performing the image reconstruction, including two-plane phase retrieval, correcting the relative piston, tips, and tilt (PTT) phase for each aperture, and



Fig. 9 Reconstructed image for six-aperture synthesis with $N_{pps} = 100$, $N_s = 100$.

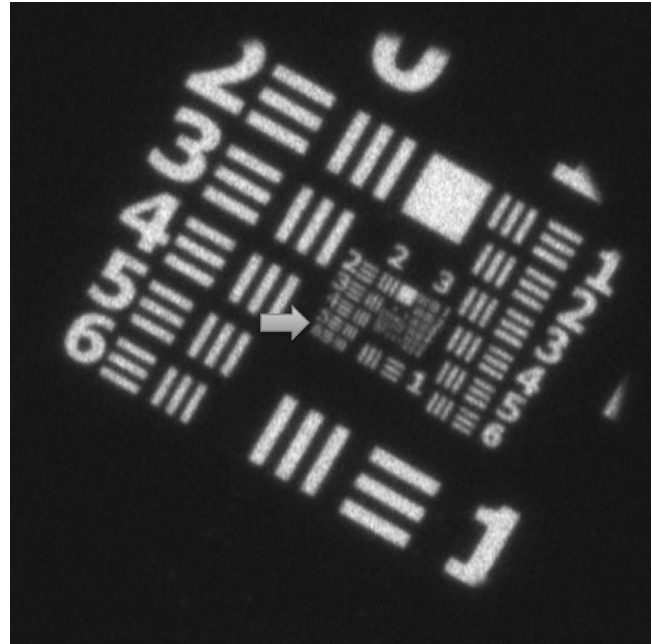


Fig. 10 Reconstructed image for six-aperture synthesis with $N_{pps} = 4$, $N_s = 100$.

averaging over $N_s = 100$ speckled images. Its resolution and quality is comparable to that of the ideal (noise-free) image shown in Fig. 8, demonstrating that our two-plane pupil field reconstruction algorithm and aperture-phasing algorithm are working well, and that $N_{pps} = 100$ is more than enough signal for success. Figure 10 shows the same thing but for the low light level $N_{pps} = 4$, which still has almost the same resolution as for the higher SNR for this high-contrast target.

Table 2 shows results from a number of reconstructions for a variety of signal levels and speckle realizations averaged for this high-contrast target. It shows the importance of having at least several speckle frames over which to average. Averaging over 10 speckle frames improved the resolution roughly by a factor of 2. It also shows that resolution does not improve much above $N_{pps} = 2$ or 4 (very low light levels), for this high-contrast target; the reconstruction algorithms worked very well even for these very low signal levels.

In Ref. 12, we showed that the same algorithm worked for the large 72-aperture synthetic aperture shown in the bottom of Fig. 1, for the higher SNRs ($N_{pps} = 100$), but the pupil-phasing algorithm (after fields were successfully reconstructed over individual apertures) was not adequate for low SNRs. To work well for low SNRs and large synthetic apertures, a least-squares reconstruction algorithm is needed, but time did not permit its implementation. In addition, reconstruction was possible with annular apertures, despite missing areas in the synthesized aperture, although the images produced had faint halos surrounding the bright points of the image. Those halos were successfully removed from a speckle-averaged image ($N_s = 10$) by a Wiener-Helstrom filter designed for incoherent images.¹² In what follows, we explore the combination of the more difficult annular apertures with the more difficult case of an object of realistic contrast.

Table 2 Results of image reconstruction for the six-aperture synthesis for varying number of photons per speckle (N_{pps}) and number of speckle realizations averaged (N_s). Upper line: group, element just discernible; lower line: delta-NIIRS relative to ideal incoherent image.

N_{pps}	Infinity	N_s			
		100	10	1	
Infinity w/o reconstruct	3,1	2,6-3,1	2,4-2,5	1,5	Group, El delta-NIIRS
	0	0.083	0.416	1.333	
100		2,6-3,1	2,4-2,5	1,5	
		0.083	0.416	1.333	
4		2,5-2,6	2,2-2,3	1,3-1,4	
		0.250	0.583	1.583	
2		2,4-2,5	2,1-2,2	1,3-1,4	
		0.416	0.916	1.583	
1		2,2-2,3	1,3	No bars	
		0.750	1.667	>3	
0.5		1,1	No bars	No bars	
		2.000	>3	>3	

3.2 Nine-Annular-Aperture Simulations with Realistic Contrast

All the imagery shown so far was for a high-contrast US Air Force bar target. One would expect that a low-contrast scene would require a larger SNR than a high-contrast scene. In addition, it can be expected that the aperture-phasing algorithm will have decreasing accuracy with decreasing scene

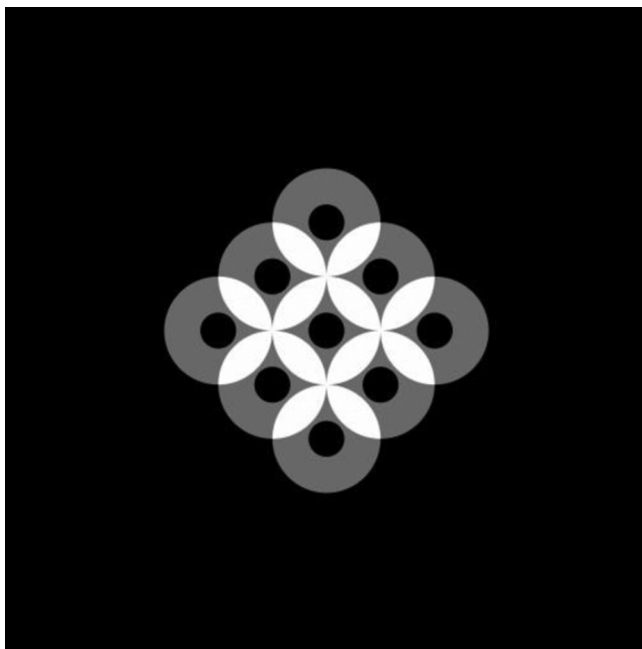


Fig. 11 Nine-annular-aperture synthetic aperture.



Fig. 12 Ideal (noise-free) incoherent image for nine-annular aperture synthesis.

contrast, since it relies on a cross-correlation algorithm to match areas of the images from the different individual apertures. To illustrate that effect, a series of simulation experiments was performed on an image with a more realistic contrast. For this study, the nine-annular-aperture synthetic aperture shown in Fig. 11 was used.

For reference, Fig. 12 shows a simulated ideal incoherent image through the nine-annular-aperture synthetic aperture (without Wiener filtering). The original photograph was taken with a Nikon D-90 DSLR camera, from a roof-top of a four-story office building, of a parking lot on a sunny day and includes (left to right) a front loader, a man walking, a water truck, and a pick-up truck, with trees in the background. Figures 13–16 show ideal images for $N_s = 1, 10, 100,$ and 1000 speckle frames averaged, respectively. For $N_s = 1$ (no speckle averaging), one can easily discern that there is an object in the location of the water truck, but not the two other vehicles. For $N_s = 10$, one can detect all three vehicles and discern their sizes and shapes. For $N_s = 100$, one can also see the walking man and easily see features on the vehicles such as tires. The resolution for $N_s = 100$ is comparable to that of the ideal incoherent image, but it has a noisier appearance. For $N_s = 1000$, one gets something approaching the ideal incoherent images. From these results we see that for more realistic, lower-contrast (than the bar targets) scenes, one needs a greater number than $N_s = 10$ speckle realizations to be able to extract all the information from the images. Further simulations would be required to quantify that number.

The images of the realistic-contrast scenes aforementioned were for the noise-free ideal images with averaging over speckle realizations. Next, image reconstruction experiments were performed, varying the SNR. For one to appreciate the kind of data that the reconstruction algorithms work on, Fig. 17 shows the noise-free image from a single annular aperture (perhaps the water truck is detectable).

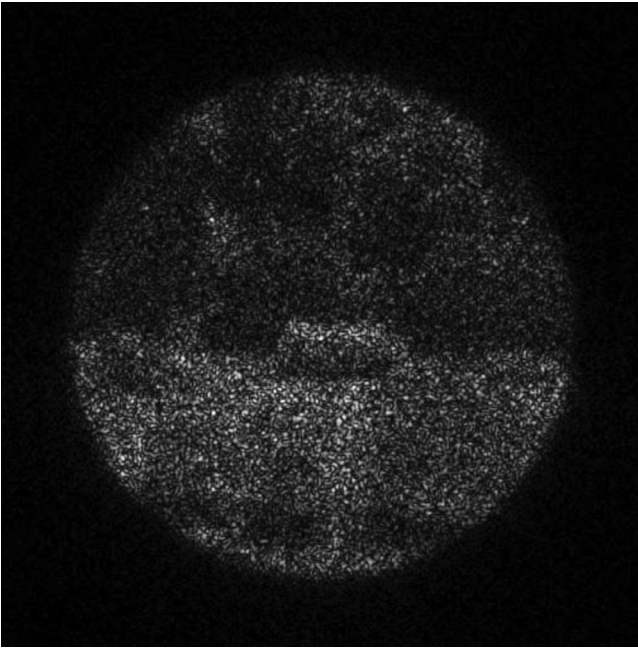


Fig. 13 Ideal (noise-free) coherent image for nine-annular aperture synthesis, $N_s = 1$.



Fig. 15 Ideal (noise-free) coherent image for nine-annular aperture synthesis, $N_s = 100$.

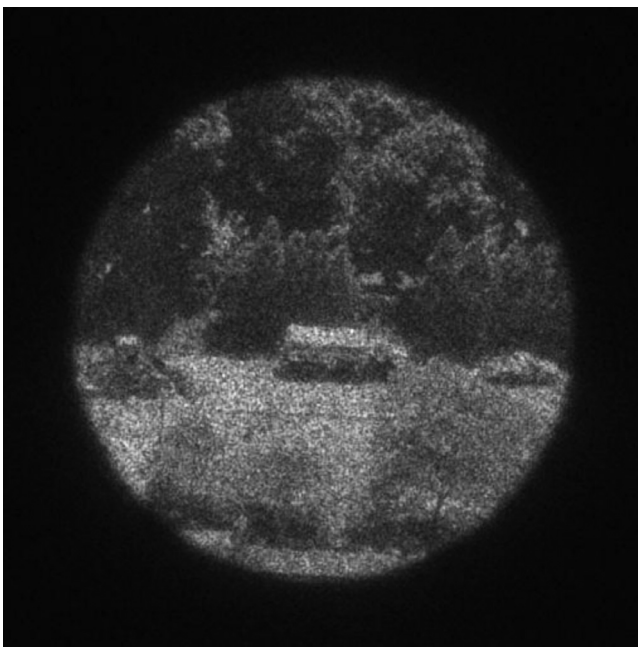


Fig. 14 Ideal (noise-free) coherent image for nine-annular aperture synthesis, $N_s = 10$.



Fig. 16 Ideal (noise-free) coherent image for nine-annular aperture synthesis, $N_s = 1000$.

Figure 18 shows a noisy image from a single annular aperture, with $N_{\text{pps}} = 4$ (shot noise only; no detector noise). This is the image frame of data going into the Gerchberg–Saxton-like algorithm for estimating the fields in the aperture plane. One can see the individual photon events. Because of the manner in which the simulations were performed, it is sampled at the final resolution of the synthetic aperture, so it is oversampled by about a factor of three in each dimension before the photon noise was applied.

Figure 19 shows the image reconstructed for the nine-annular-aperture synthesis with $N_{\text{pps}} = 4$, $N_s = 10$. While that amount of noise allowed the aperture phasing to be adequate for the high-contrast bar target, it was substantially degraded for this realistic-contrast target: the image is not just noisier, it is also blurred compared with the ideal noise-free image shown in Fig. 14, because of reduced-quality aperture phasing as well as the phase retrieval for individual apertures in the presence of this large amount of noise.

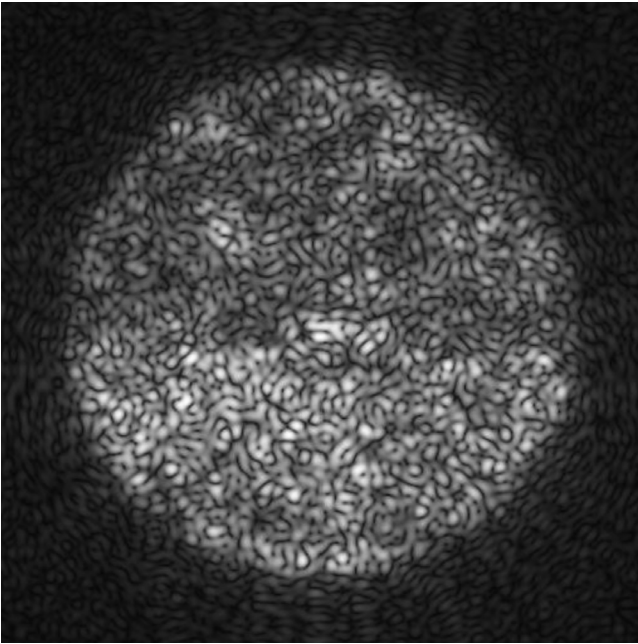


Fig. 17 Ideal (noise-free) coherent image for a single annular aperture, $N_{\text{pps}} = \text{infinity}$, $N_s = 1$.

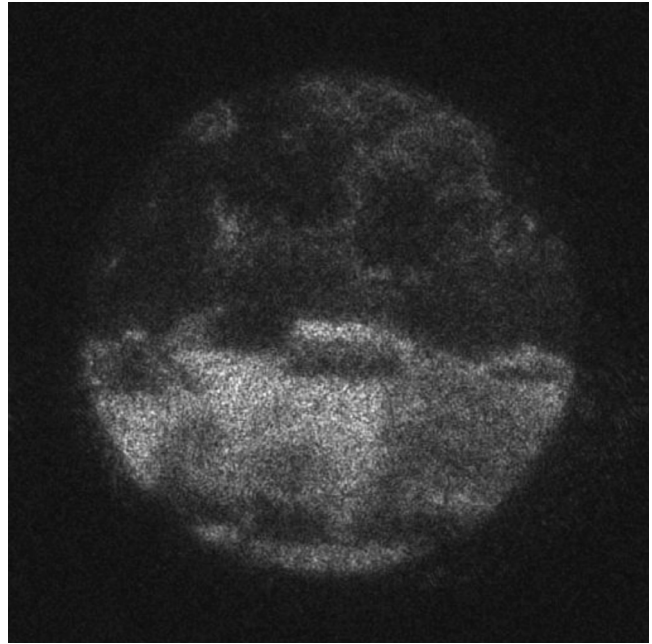


Fig. 19 Reconstructed image for nine-annular-aperture synthesis with $N_{\text{pps}} = 4$, $N_s = 10$.

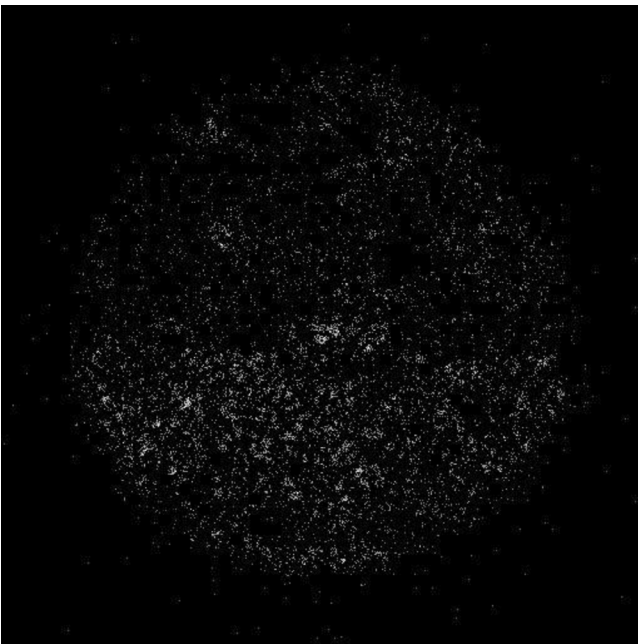


Fig. 18 Noisy coherent image for a single annular aperture, $N_{\text{pps}} = 4$, $N_s = 1$.

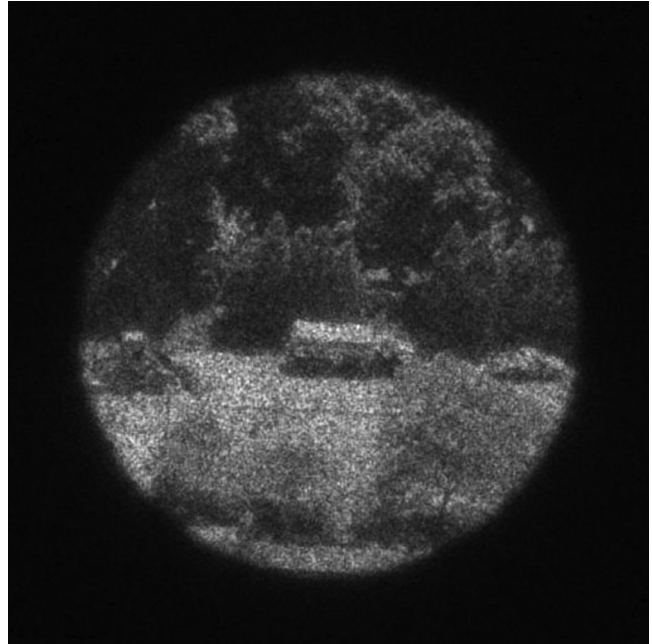


Fig. 20 Reconstructed image for nine-annular-aperture synthesis with $N_{\text{pps}} = 100$, $N_s = 10$.

For the case of higher SNR, shown in Fig. 20, with $N_{\text{pps}} = 100$ (SNR = 10), the aperture-phasing algorithm worked very well, as can be seen by comparing this image with the ideal noise-free image shown in Fig. 14, where they match even at the level of individual speckles. Hence, one needs N_{pps} to be something greater than 4 but less than 100 for the case of a target with more realistic contrast. Further simulations will be required to narrow down the number of photons needed.

3.3 Aperture Array Phasing Accuracy

The relative phasing of the different apertures within the synthetic aperture, including relative PTT phases where the tip and tilt phases are equivalent to telescope relative pointing errors, were corrected with a subpixel accuracy registration algorithm. The requirements on the accuracy can be thought of as follows. For images, the effects of diffraction due to the finite aperture, and that due to motion blur (similar to summing images suffering from misregistration), both can be

described by multiplicative transfer functions in the Fourier domain and by convolutions in the image domain, similar to the result in [Ref. 13, Eq. (8.5-4)] for the case of diffraction and aberrations. Furthermore, if two Gaussians are convolved together (multiplied in the Fourier domain), the result is a Gaussian having a variance equal to the sum of the variances of the two original Gaussians. Similarly, we assume here that the variances of the convolution of some other spread functions approximately add, making the width of the convolution approximately the square root of the sum of the squares of the individual widths. Then for images, if one has a root-mean-squared (rms) pointing error of s diffraction-limited resolution elements, then the net resolution element (here, the final resolution of the synthesized aperture) will have width $\sim\sqrt{1+s^2}$ times the diffraction-limited resolution. In that case, a misregistration by 0.25 resolution elements rms will degrade the resolution by a factor of 1.03, a misregistration by 0.5 resolution elements rms will degrade the resolution by a factor of 1.12 (one element in the bar target), and a misregistration by 1 resolution element rms will degrade the resolution by a factor of 1.41 (three elements in the bar target). Based on those numbers, one might require a residual telescope pointing error, after correction, to be around 0.5 resolution elements or less.

In the 72-aperture simulations done earlier¹² for the high-contrast bar targets for $N_{\text{pps}} = 100$, the actual pointing errors in the estimates from the noisy data were 0.12 resolution elements rms, which are negligible. For $N_{\text{pps}} = 4$, the actual pointing errors in the estimates from the noisy data were 0.77 resolution elements rms, slightly more than the goal of 0.5 resolution elements, yielding a nonnegligible error. This error would probably be lower than 0.5 if the data were fed into a least-squares reconstruction algorithm, which we did not have a chance to implement. This becomes a bigger problem with the realistic-contrast scene than with the bar targets, but the size of the effect has not yet been quantified. For that realistic-contrast target, it appeared that for $N_{\text{pps}} = 4$ the phasing algorithm yielded significant residual errors whereas for $N_{\text{pps}} = 100$ the phasing algorithm worked very well.

4 Practical Considerations

A first-principles study of major trade-off issues was performed to determine the feasibility of the direct-detection, multiaperture, synthetic-aperture, active, coherent imaging system concept. The system architecture shown in Fig. 1 was assumed, although the analysis could be readily applied to other cases.

4.1 PRF, Speckle Velocity, and Pulse Length

The PRF of the laser should be sufficient to allow no large gaps in the synthetic aperture. The PRF must be fast enough so that, for the example of four rows of apertures shown in the top of Fig. 1, the first row of telescopes moves forward by half its diameter between pulses (the speckles move across the aperture at twice the speed that the aperture moves forward, since both the apertures and the transmitter are moving forward at the same speed). In the bottom of that figure, the first telescope pupil in the synthetic aperture appears again one diameter from its previous position. (The manner in which we are speaking of these things is accurate for imaging

in a direction perpendicular to the direction of motion of the synthetic aperture; various geometrical effects must be taken into account when pointing forward or to the side.) Accounting for the fact that the speckles in the pupil plane move backward at the same speed, v_p , as the forward motion of the array, a speckle speed relative to the aperture will be $v_s = 2v_p$. For a telescope of diameter D_1 , the synthetic aperture will move by one aperture width in time $D_1/(2v_p)$, which suggests a time between pulses of $D_1/(2v_p)$ and a PRF $= 2v_p/D_1$. For example, for three scenarios, a ground vehicle with $D_1 = 10$ cm and $v_p = 15$ m/s, an unmanned aerial vehicle with $D_1 = 10$ cm and $v_p = 150$ m/s, and a low-earth-orbiting satellite with $D_1 = 20$ cm and $v_p = 7.8$ km/s, the PRFs would be 300 Hz, 3 kHz, and 78 kHz, respectively, requiring very fast detector arrays.

As mentioned in the image reconstruction section, in order to achieve adequate image quality, multiple synthetic-aperture images, each with an independent speckle pattern, are usually needed. Collecting multiple speckle patterns would seem to be less expensive than building a much larger array of telescopes to achieve the desired resolution.

The length of an individual pulse should be short enough to freeze the speckles at the detector, or one must have an optical compensation for translating speckles. The speckles will be moving at twice the speed of the imaging platform, assuming that the laser illuminator is near the receiver traveling at approximately the same velocity. The diameter of one of the speckles in the pupil plane will be

$$d_s = \frac{\lambda R}{w_o}, \quad (1)$$

where w_o is the width of the illuminated area of the ground (projected perpendicular to the line of sight), λ is the wavelength, and R is the range to the object. The number of speckles across a telescope aperture would be $D_1/d_s = (D_1 w_o)/(\lambda R)$. To Nyquist sample the intensity, one would like to have two samples per speckle or $2(D_1 w_o)/(\lambda R)$ samples across the aperture with a sample spacing of $d_s/2 = \lambda R/(2w_o)$. Since the pupil-plane detection is designed to be performed on a demagnified image of the pupil, as shown in Fig. 2, the sample spacing of the detector array will be reduced by the magnification factor, but the number of speckles remains the same. The number of speckles within the pupil is approximately the same as the number of resolution elements across the image, the total number of spatial modes being invariant.

The speckles move their own width in time $d_s/(2v_p) = \lambda R/(2v_p w_o)$. To keep the speckle from moving by 1/4 its diameter (less than that would be desirable), the laser pulse duration would have to be less than 1/4 of that, or $\lambda R/(8v_p w_o)$. This is equivalent to a coherence length of $c\lambda R/(8v_p w_o)$, where c is the speed of light. This coherence length allows only objects/scenes of depth half the coherence length (due to the round-trip distance of the reflected light), or of depth $c\lambda R/(16v_p w_o)$, to be imaged coherently. Note that reducing the width, w_o , of the illuminated object increases the speckle width proportionally and increases the allowable uncompensated speckle motion proportionally. For wide-field imaging (large w_o) with a fast-moving platform, the allowed object depth may be unacceptably small, in which case it would be necessary to have, as part of the

receiver (or transmitter) optics that scans along with the speckle motion to freeze it over a longer pulse length. Since the pupil will be reimaged to a smaller scale, this can be done with a fast-steering mirror or with an acousto-optic modulator on the reduced-diameter beam.

Independent of the illumination diameter w_o , for a width D of the synthesized aperture, the synthesized aperture subtends an angle of D/R , but the angular motion is half that, $D/(2R)$. If 10 such synthetic apertures were gathered sequentially without pause in-between, for 10 speckle realizations, the total angular subtense of the synthetic aperture would be $5D/R$. For most scenarios of interest, for long-range imaging, one can collect many synthetic apertures without the object appearing to be different due to different illumination and viewing angles.

4.2 Doppler Shift and Timing

While heterodyne systems must compensate for the Doppler shift due to the radial velocity of the imaging platform relative to the object, in order for the return beam to properly interfere with the LO, this direct-detection approach is unbothered by Doppler shifts.

For heterodyne detection, great care must be taken so the return pulse arrives at the same time as an LO pulse at the detector array. In the case of our direct-detection system, it is only necessary that the shutter be open when the return pulse arrives and that the shutters of the two detector arrays per telescope are open at the same time.

4.3 Link Budget

If the product of the two-way atmospheric transmittance, transmitter and receiver optical transmittance and quantum efficiency is η , and the mean object intensity reflectivity is τ_o , then a laser pulse with energy E_p joules (per pulse) at range R , reflected from the target area and falling onto a single collecting element of length and width d_d results in a mean number of photons at a pupil-plane detector element of

$$N_{pd} = \frac{E_p \tau_o \eta \lambda d_d^2}{2\pi R^2 hc}, \quad (2)$$

where h is Planck's constant, c is the speed of light, and a Lambertian reflecting surface is assumed. A factor of 2 in the denominator comes from the fact that half of the light goes to a pupil-plane detector array and the other half goes to an image-plane detector array. For simplicity we assume the detector pitch is also d_d (unity duty cycle). Since it is photons per speckle that determines image quality, employing Eq. (1) we see that the number of photons per speckle is

$$N_{pps} = \frac{E_p \tau_o \eta \lambda d_s^2}{2\pi R^2 hc} = \frac{E_p \tau_o \eta \lambda^3}{2\pi hc w_o^2}. \quad (3)$$

For a given wavelength and pulse energy, we see that there is a direct trade-off between the number of photons per speckle and the area, w_o^2 , of the illuminated object or scene. Solving for the illuminated width that gives a desired N_{pps} (~ 4 for a high-contrast target, according to our image reconstruction simulations) we get

$$w_o = \sqrt{\frac{E_p \tau_o \eta \lambda^3}{2\pi hc N_{pps}}}. \quad (4)$$

It is interesting to note that Eqs. (3) and (4) are independent of range (aside from absorption losses through the atmosphere) and independent of resolution! For a given resolution, if one doubles the range, then the aperture must double in diameter to preserve the resolution, which results in gathering the same total number of photons. For a given range, to make the resolution twice as fine, one must double the aperture diameter, gathering four times the number of photons, but those photons are spread over four times as many resolution elements, so the number of collected photons per resolution element (i.e., per speckle) stays the same.

Suppose that one has a laser with the high average power $E_L = 200$ W and that we need $N_{pps} = 4$ photons/speckle (adequate for a high-contrast target), and that $\tau_o = 0.2$ and $\eta = 0.5$ and $\lambda = 1 \mu\text{m}$. Then, Eq. (4) predicts the following scenarios:

Case 1a. If we have a continuous PRF of 300 Hz (as for the ground vehicle platform), which would give us 0.67 J/pulse, then $w_o = 290$ m. However, one could collect several of these images in 1 s and mosaic them together into a larger image.

Case 1b. If we have a continuous PRF of 78 kHz (as for the satellite platform), which would give us 2.6 mJ/pulse, then $w_o = 18$ m, a very tiny instantaneous field of view. However, one could collect many small images and mosaic them together into a larger image.

Case 2. If during 1 s only 10 laser pulses are transmitted, allowing for a single synthetic aperture, which would allow for 20 J/pulse, then $w_o = 1600$ m.

Case 3. If during 1 s the laser transmits 70 pulses making up 10 synthetic apertures ($N_s = 10$), which would allow for 2.9 J/pulse, then $w_o = 600$ m.

Note that cases 2 and 3 are independent of the speed of and distance to the sensor platform. The area coverage rate is proportional to E_L , which may be produced by either one or multiple lasers.

Case 3 is probably the scenario one would choose to collect, taking advantage of the 10 speckle realizations in order to achieve the desired image quality.

These calculations are for the case of a high-contrast target such as the USAF bar targets. For targets of more natural contrast, larger SNRs (larger values of N_{pps}) and larger numbers of speckle realizations, N_s , are needed.

4.4 Speckle Boiling

Not only do speckles translate as the illumination angle relative to the object changes but they also "boil,"¹¹ a phenomenon sometimes referred to as the "memory effect." One might wonder about how this affects image quality. Fortunately the effect is small for remote sensing scenarios. According to Ref. 9, an intensity correlation of e^{-2} occurs for an angular change of

$$\Delta\theta_i = \frac{1}{\pi \sin \theta_i} \frac{\lambda}{\sigma_h}, \quad (5)$$

where θ_i is the angle of incidence (with respect to the surface normal), $\Delta\theta_i$ is the change in angle allowed, and σ_h is the standard deviation of the height differences or surface roughness (found within a resolution element). This was derived using a single-scattering model. On a nominally flat surface tilted at an angle of 45 deg with respect to the line of sight, $\sin \theta_i = 0.414$, and for a surface height (roughness) standard deviation of 0.1 mm, this gives $\Delta\theta_i = 4.4$ mrad, as compared with the angular extent of, say, 0.6 mrad for a 0.6-m synthetic aperture at a distance of 1 km, making it well within the “memory effect.” The effect is proportionally smaller for longer distances. Hence, this effect should be quite small for collections of flat surfaces viewed from long distances. The effect on image quality for large apertures at small distances, for which speckle boiling is significant, is an interesting topic.

4.5 Heterodyne Versus Direct Detection

Direct detection with phase retrieval was chosen over heterodyne sensing for this study in order to overcome the difficulties in heterodyne sensing, including the logistics of distributing the LO to multiple, possibly disjoint telescopes, subwavelength stability of the LO relative to the illumination beam, temporal delay needed to interfere the LO with the return beam, correcting the relative Doppler shift of the LO relative to the return beam, etc. Since the telescopes in the group are relatively close to one another, however, reducing some of the difficult logistics, it is worth comparing heterodyne and direct detection.

First, with heterodyne detection, nearly all the light reflected from the object and captured by the aperture of a telescope can go to the single heterodyne channel. This avoiding of splitting the light into two approximately equal channels would appear to imply a $\sqrt{2}$ increase in SNR, but that is not the case. For direct detection, still all the photons are collected, but by two detector arrays rather than just one. Furthermore, introduction of the LO, if performed with spatial heterodyne, requires more pixels in the detector array than for direct detection, but direct detection requires two detector arrays rather than just one. From this perspective, heterodyne sensing may or may not have an advantage over direct detection.

Heterodyne definitely benefits from the “heterodyne advantage,” namely, that read noise and dark current essentially go away if the LO is much brighter than the light from the object, which almost always will be the case. This benefit goes away if one employs photon-limited detectors with direct detection.

The computation of the field from heterodyne data can be a purely linear process, whereas phase retrieval reconstruction is inherently nonlinear. One would expect heterodyne to have an image SNR advantage for this reason.

Once the fields in the individual apertures are reconstructed for each telescope, then the remainder of the process—synthetic-aperture assembly, aperture phasing, pupil registration and high-resolution image formation, and averaging over speckle realizations—would be the same whether one performs heterodyne sensing or direct detection with phase retrieval.

For the purpose of comparison with direct detection employing phase retrieval, images were computed to have the same noise statistics as for heterodyne detection (since

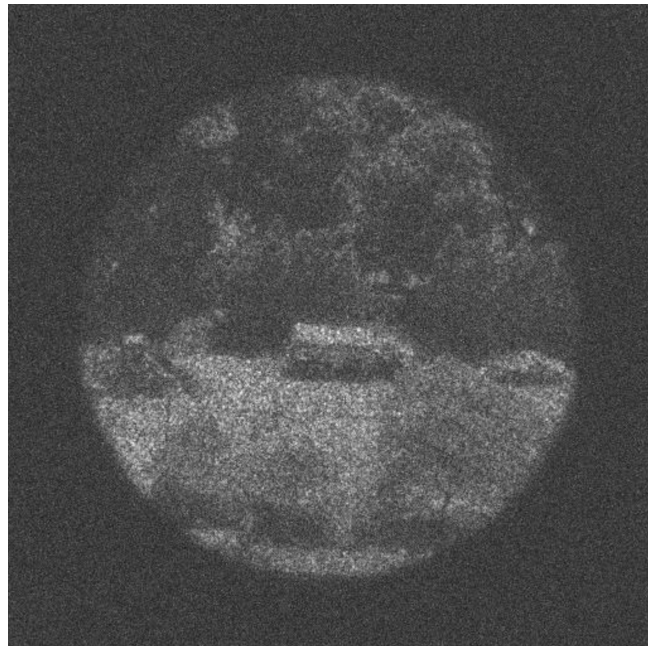


Fig. 21 Heterodyne image with noise appropriate to eight photons per speckle in the pupil plane, $N_s = 10$. Perfect aperture phasing was assumed.

simulating the entire heterodyne process was beyond the scope of this effort). To compare with Fig. 19 (for direct detection with phase retrieval), zero-mean complex Gaussian random noise was added to an ideal nine-annular aperture synthetic aperture field (having no phase errors) to give it the noise that would be expected with eight photons per speckle, double the number per detector array as for the direct-detection approach. It was assumed that there were zero registration errors or aperture-phasing errors. Figure 21 shows the resulting image. It is sharper than the

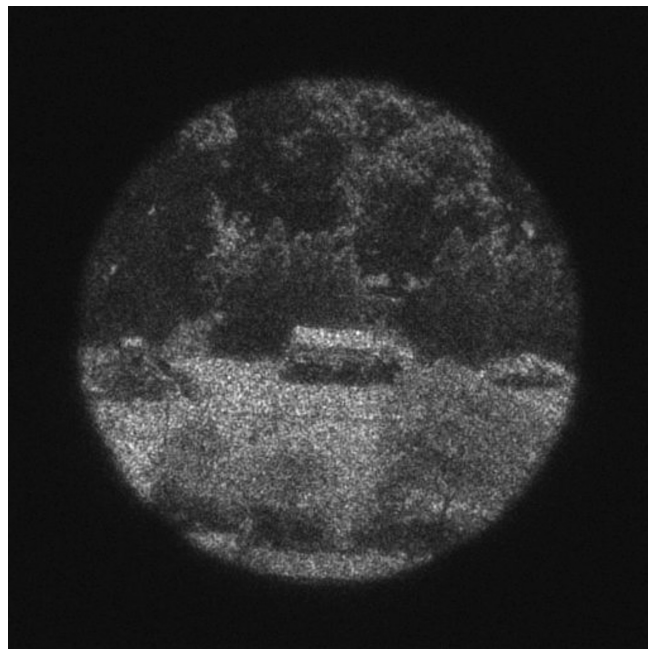


Fig. 22 Heterodyne image with noise appropriate to 200 photons per speckle in the pupil plane, $N_s = 10$. Perfect aperture phasing was assumed.

direct-detection image shown in Fig. 19, with a better definition of the rectangular shape of the water truck in the center of the image. This comparison is unfair in that the heterodyne image was not subjected to the aperture-phasing errors of the direct-detection result. The noise characteristics of the two images are different. With detection only in the pupil plane, the heterodyne image has noise spread over the entire computational window of the image, whereas the phase-retrieval image, constrained by the focal-plane noisy image, has most of its noise energy confined to the illuminated region of the object.

Figure 22 shows the heterodyne image with noise appropriate for 200 photons per speckle for comparison with the direct-detection result shown in Fig. 20. In this case, the results are virtually indistinguishable. This is probably because the residual speckle noise dominates over the photon noise. Hence, the noise advantage of heterodyne over direct detection is only a factor at low light levels.

4.6 Fourier Ptychography Versus 2-Plane Phase Retrieval

Another direct-detection alternative to the architecture analyzed in this paper is Fourier ptychography.^{5,6} It is a method for synthesizing a coherent aperture from image-plane intensities, mostly used for microscopy. It involves coherently illuminating the sample from multiple different angles to synthesize a larger aperture. For the long-range imaging application of interest here, however, similar ideas and algorithms can be used to perform aperture synthesis with moving apertures. It is similar to the aperture synthesis described for the two-plane intensities already mentioned, but with the following differences. First, only measurements of the low-resolution focal-plane images are employed. Second, the overlaps of the individual aperture locations within the synthetic aperture are much denser than the overlaps shown in Fig. 1. Hence there is much more redundancy within the synthetic aperture, and the synthetic aperture will be smaller than, and the resolution will be poorer than, for the two-plane approach for a given number of telescopes and laser pulses. In microscopic imaging applications, Fourier ptychography has been shown to be robust with the dense sampling. Requiring only a single detector array in each telescope rather than a beamsplitter and two detector arrays makes the individual telescopes simpler than the two-plane approach. The image reconstruction is different, too. Rather than reconstructing a field across each pupil, synthesizing an aperture from those fields, and then phasing the pupils within the aperture, in Fourier ptychography the phase retrieval is performed directly on the synthesized pupil-plane array. A single synthesized complex array is found that is consistent with all of the measured low-resolution image intensities. Each measured low-resolution image intensity must agree with the images obtained by computing the Fourier transforms of a circularly (or whatever the shape of the individual telescopes) windowed portion of the estimated synthetic-aperture field (and taking the squared magnitude). Note that no measurements are made in the pupil plane in this case. The Fourier ptychography approach (also employing phase retrieval) is expected to have performance that is in the same ballpark as two-plane direct-detection phase retrieval.

5 Conclusions

In this paper, we showed a new way of performing aperture synthesis using coherent light but with direct detection (no LO or reference beam) and phase retrieval. This allows for imaging with much finer resolution than with a single fixed aperture. Simulations show its feasibility and systems analysis shows its practicality.

Three sets of algorithms are needed: (1) reconstructing the phase of each pupil from the pupil and image plane intensities, (2) assembly of the pupils into a synthetic aperture and correcting relative phase errors (PTT) based on overlapping portions of the synthesized aperture (the tip and tilt correction corresponding to correcting the relative pointing errors between the telescopes); and possibly (3) pupil registration correction might be needed as well.

From the image reconstruction studies described in Sec. 3, we showed the following in regard to the two-plane direct-detection phase-retrieval approach to sparse-aperture synthetic-aperture imaging:

- Algorithm (1), individual pupil field reconstruction, worked well even for low SNRs ($\text{SNR} = 2$, equivalent to four photons per speckle in each plane of detection for a high-contrast target).
- Algorithm (2), aperture phasing, worked well for the same low SNRs with modest-sized (nine apertures) synthetic apertures, but greater SNRs were needed for large (72-aperture) synthetic apertures. To work well for low SNRs and large synthetic apertures, a least-squares reconstruction algorithm is needed, but time did not permit its implementation.
- Annular-aperture telescopes as well as filled-aperture telescopes can be made to work, despite missing areas within the synthetic aperture. Wiener–Helstrom filtering was effecting in cleaning up halos in the reconstructed images from a speckle-averaged image from annular-aperture telescopes.¹²
- For high-contrast objects, we recommend a laser power such that one can achieve a minimum of about four photons per speckle. The requirements for low-contrast objects are for a greater light level but were not quantified.
- For high-contrast objects, averaging 10 speckle realizations yields about a factor of about 2 improvement in resolution over a single speckle realization. For low-contrast objects, a greater number of speckle realizations are needed.

In Sec. 4, on practical considerations, we found that

- The area coverage rate is proportional to the laser power, the transmittances of the atmosphere and optics, the reflectivity of the object, the quantum efficiency of the detectors, and the cube of the wavelength; and it is inversely proportional to the number of photons per speckle needed to achieve the desired image quality. One example: assuming a 200-W average power laser, needing 10 speckle realizations, and needing four photons per speckle per plane, using 70 pulses at 2.9 J/pulse, one could image one 600 m × 600 m area or 1/3 km²/s.

- The PRF needed to avoid gaps in the synthetic aperture is proportional to the speed of the platform and requires a fast detector array.
- Speckle “boiling” (the “memory effect”) is negligible for most long-range imaging.
- Heterodyne detection is superior from an SNR perspective for low light levels, but the advantage is negligible at higher light levels when speckle noise is the dominant source of noise.

Finally, we note that there are many different geometrical configurations of telescope apertures and laser transmitters that can be employed with differing numbers of each, allowing for a flexible, extensible imaging architecture.

Acknowledgments

Portions of this paper were presented in Ref. 12. Thanks go to Dr. Thomas Karr, DARPA/STO, for suggesting this problem. This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA). The views, opinions, and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

References

1. J. R. Fienup and A. M. Kowalczyk, “Phase retrieval for a complex-valued object by using a low-resolution image,” *J. Opt. Soc. Am. A* **7**, 450–458 (1990).
2. R. W. Gerchberg and W. O. Saxton, “A practical algorithm for the determination of phase from image and diffraction plane pictures,” *Optik* **35**, 237–246 (1972).
3. J. R. Fienup, “Phase retrieval algorithms: a comparison,” *Appl. Opt.* **21**, 2758–2769 (1982).
4. J. R. Fienup, “Lensless coherent imaging by phase retrieval with an illumination pattern constraint,” *Opt. Express* **14**, 498–508 (2006).
5. G. Zheng, C. Kolner, and C. Yang, “Microscopy refocusing and dark-field imaging by using a simple LED array,” *Opt. Lett.* **36**, 3987–3989 (2011).
6. G. Zheng, R. Horstmeyer, and C. Yang, “Wide-field, high-resolution Fourier ptychographic microscopy,” *Nat. Photonics* **7**, 739–745 (2013).
7. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, “Efficient subpixel image registration algorithms,” *Opt. Lett.* **33**, 156–158 (2008).
8. P. S. Idell and A. Webster, “Resolution limits for coherent optical imaging: signal-to-noise analysis in the spatial-frequency domain,” *J. Opt. Soc. Am. A* **9**, 43–56 (1992).
9. J. C. Leachtenauer et al., “General image-quality equation: GIQE,” *Appl. Opt.* **36**, 8322–8328 (1997).
10. R. D. Fiete and T. Tantaló, “Image quality of increased along-scan sampling for remote sensing systems,” *Opt. Eng.* **38**, 815–820 (1999).
11. J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications*, Roberts and Company Publishers, Greenwood Village, Colorado (2007).
12. J. R. Fienup, “Synthetic-aperture direct-detection coherent imaging,” *Proc. SPIE* **10410**, 104100B (2017).
13. J. W. Goodman, *Statistical Optics*, 2nd ed., Wiley, Hoboken, New Jersey (2015).

James R. Fienup received his AB from Holy Cross College, and MS and PhD degrees in applied physics from Stanford University, where he was a National Science Foundation graduate fellow. After performing research at ERIM, he became the Robert E. Hopkins Professor of Optics at the University of Rochester. He is a fellow of SPIE and OSA, a member of the National Academy of Engineering, and a recipient of SPIE’s Rudolf Kingslake Medal and Prize.