

Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

Adaptive three-dimensional feature specific imaging

Wm. Randall Babbitt
Krista M. Drummond
Brant M. Kaylor
Randy R. Reibel

Adaptive three-dimensional feature specific imaging

Wm. Randall Babbitt,^{a,*} Krista M. Drummond,^a Brant M. Kaylor,^b and Randy R. Reibel^b

^aMontana State University, Spectrum Lab, Box 173510, Bozeman, Montana 59717-3510, United States

^bBridger Photonics, 2310 University Way, Building 4-4, Bozeman, Montana 59715, United States

Abstract. Feature specific imaging is a computational imaging technique that minimizes the number of measurements needed to sufficiently reconstruct a scene by using *a priori* knowledge (e.g., the scene's second-order statistics) to judiciously, as well as possibly adaptively, choose the projection vectors to be measured. Here, we have developed an approach to three-dimensional adaptive feature specific imaging that takes into account the obstruction of distant objects by closer objects in the adaption of the projections and in the reconstruction algorithm. The developed system reconstructs the cross-range image of the scene at each range bin from a set of range resolved measurements from all the return from the scene at that range using only a single photodetector, while adapting to the obstruction of the scene by closer objects. Simulations and a proof-of-concept demonstration of adaptive three-dimensional feature specific imaging are presented. © 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.54.3.031104](https://doi.org/10.1117/1.OE.54.3.031104)]

Keywords: feature specific imaging; precision laser ranging; compressive sensing; computational imaging.

Paper 141038SS received Jul. 1, 2014; revised manuscript received Sep. 17, 2014; accepted for publication Sep. 18, 2014; published online Oct. 27, 2014.

1 Introduction

Computational imaging is a means of reducing the number of measurements needed to sense and reconstruct an image of a sparse scene, whether it be an image in space,^{1,2} time,³ or delay.⁴ One method for performing computational imaging is to measure the multiplication of the image of a scene with a series of projection masks. The inner product of the image and one of the projection masks constitutes a single measurement. For sparse scenes, the number of projection measurements required for highly faithful reconstruction can be significantly less than the number of pixels in the image.⁵ Typical compressive sensing⁵ approaches exploit the sparsity of a scene by making measurements on an orthogonal basis (i.e., random projection masks). Feature specific imaging (FSI)² exploits sparsity in the object features of interest by directly measuring in the sparsity basis. Both are methods to reduce the number of measurements required for reconstruction. The goal of FSI is to faithfully reconstruct the full image of the objects in a scene, or at least the object features of interest in each scene, with a number of projection measurements that is significantly less than the number of pixels in the image. In FSI, there are multiple ways to choose a set of basis vectors relevant to a set of desired object features. One method is to develop a basis from a large number of training scenes. Note that the training set typically does not include the exact scene(s) to be sensed or the exact object features of the scene to be sensed, but is, instead, an extensive set of scenes that contain similar object features of interest. For example, a training set of faces with a wide range of different eyes, noses, mouths, and chins, with or without hats and glasses can be used to faithfully reconstruct the face of a person not used in the training set.⁶ Though the image of the actual face to be reconstructed is not in the training set of faces, it is assumed that the various features in the actual

face are present in one or more of the faces used for training and, thus, all the critical features in the actual face can be reconstructed. From a mathematical standpoint, the utility of the training set is to capture the second-order statistics of the possible objects or object features of interest. The basis vectors are optimized and ordered to minimize the number of projections needed to reconstruct the objects or object features of interest in a scene. For example, supposing that the principle components of the training set were to be used for the measurement basis, one could choose the order of the principle components based on their respective eigenvalues and introduce compression by rejecting those principle components whose eigenvalues are below some threshold.

Another method is to use a predetermined basis that is known to be compressive, e.g., the Hadamard basis, discrete cosine transform (DCT) basis, or a wavelet basis. When using such a basis, it is unknown which basis vectors will provide the most information about the scene; thus, in order to reduce the number of measurements, an adaptive data-driven approach can be used wherein an initial set of measurements is used to determine the best choice of basis vectors for subsequent measurements, which are then used to best choose the next set of basis vectors, and so forth. In short, adaptive FSI⁶ (AFSI) improves the measurement efficiency by actively adapting the remaining basis vectors to be projected based on the previous measurements. Both FSI and AFSI can be developed and optimized for a variety of imaging tasks, from object reconstruction to target recognition.

Three-dimensional (3-D) FSI can be performed by projecting two-dimensional (2-D) basis vectors onto the 3-D image and measuring the integrated return as a function of range. Each 2-D range slice can then be reconstructed independently using 2-D FSI techniques. For example, high-resolution single spatial mode imaging systems⁷ could be integrated with projective or collection masks to perform 3-D FSI.⁸ However, in 3-D imaging, objects in the foreground can obscure or partially obscure objects behind them, or objects with a large range extent can be stretched

*Address all correspondence to: Wm. Randall Babbitt, E-mail: babbitt@physics.montana.edu

across multiple range bins. In either case, the preferred ordering of 2-D basis vectors will require some modification. For example, in an image of a crowd, faces in the foreground can partially obstruct faces in the background. In this example, the initial training sets for both foreground and background would be full faces. However, the obstruction of the background objects (faces) will result in a significant altering of the preferred ordering of the basis vectors needed to efficiently reconstruct the background faces, along with the unobstructed foreground objects (faces). This paper addresses this class of problems.

We have developed a strategy for adaptive 3-D FSI, incorporating the impacts of obscuration on down range reconstructions and including projection strategy and reconstruction techniques. Estimations of the foreground objects and their positions are used to adapt the background training sets, modifying them to include estimations of the obstructed training objects. These adapted training sets are used in subsequent choices of projection vectors with the goal of effectively reconstructing the foreground and background objects with a limited number of projections.

2 Development of 3-D Adaptive FSI Algorithm and Simulator

The adaptive 3-D scene reconstruction algorithm presented here is based on the 2-D Hadamard-based AFSI work presented in a paper by Ke et al.⁹ Hadamard-based AFSI was chosen over principal-component-based AFSI for two reasons. The first is the hardware complexity involved in projecting principal components. Hadamard projections are binary and, thus, readily produced with most spatial light modulators with a high signal-to-noise ratio and accuracy. The second reason is that the Hadamard basis is useful when a wide variety of objects are expected or when the features of interest may vary as a function of range. For example, leaves or trees in the foreground may be obstructing the view of a truck or car in the background; thus, the training set for the foreground would be various types of leaves and trees and the training set for the background would be various models of trucks and cars, each with a different and, perhaps, somewhat orthogonal set of principal components. For a large enough training set of random images, the principal components will converge on the DCT,¹⁰ and the Hadamard basis may be thought of as a binary version of the DCT.

A 3-D AFSI MATLAB[®]-based simulator was developed based on the 2-D algorithms of Ke et al.⁹ Our adaptive FSI method uses one-dimensional ranging measurements of 2-D projections to achieve 3-D imaging results. As noted in Introduction, the basic structure of the FSI, both in simulation and experiment, is to compute a set of measurements y that represent the inner product of a subset of the basis vectors, F , and the scene x .

$$y = Fx + n, \quad (1)$$

where there is additive white Gaussian noise, n , on each measurement, with variance of σ_n^2 . The goal is to reconstruct x with a set of measurements y , where the dimension of set y is much less than the set of dimension x . In general, this is possible only with prior knowledge of x . For example, x is sparse or primarily consists of features similar to the features in a training set of images.

In our 3-D AFSI algorithm, the 3-D scene x is assumed to be separable into 2-D object scenes at discrete ranges, x_k . Each scene is assumed to have size $\sqrt{N} \times \sqrt{N}$ and is represented by a vector of dimension N . The objects in any scene $x_{k'}$ can obscure objects in scene x_k ($k > k'$). For each range k , there is a training set of J_k images, $B_k^{(1)} = \{b_{k,j}^{(1)}\}$, where k is the index for range and $j = 1$ to J_k is the index for the images in the training set. Each $b_{k,j}^{(1)}$ has dimensionality N . One goal of 3-D AFSI algorithm is to develop an adapted training set at each range $A_k^{(i)} = \{a_{k,j}^{(i)}\}$ that represents the obscured training scenes that would be seen by the projector of the basis vectors and the single-pixel measurement system due to obstructions by $x_{k'}$ for all $k' < k$, where i represents the iteration index number. Initially, without having made any measurements, the training sets for each range index on the first iteration are equal: $A_k^{(1)} = B_k^{(1)}$.

On each iteration, including the first, a set of Hadamard basis vectors is chosen for projection based on the training sets and any previous measurements. There are several methods to accomplish this task. One method is to average the energies of each remaining Hadamard (those not previously projected) with respect to each scene in the current adapted training set.

$$E_n^{(i)} = \sum_k \frac{w_k^{(i)}}{J_k^{(i)}} \sum_j |H_n^{(i)} a_{k,j}^{(i)}|^2. \quad (2)$$

Range and iteration dependent weights $w_k^{(i)}$ are included and will be discussed later. A twist on this is to subtract out the average training scene before taking the average.

$$E_n^{(i)} = \sum_k \frac{w_k^{(i)}}{J_k^{(i)}} \sum_j |H_n^{(i)} (a_{k,j}^{(i)} - \langle a_{k,j}^{(i)} \rangle_j)|^2. \quad (3)$$

Another approach is to first average over the scene in the current training set and compute the energy of the average.⁹

$$E_n^{(i)} = \sum_k w_k^{(i)} \left| H_n^{(i)} \frac{1}{J_k^{(i)}} \sum_j a_{k,j}^{(i)} \right|^2. \quad (4)$$

All three of these approaches to energy measurement were tested. The results were similar for all three, though the average energy of training sets with the average removed typically gave slightly better results and all results shown were derived using this energy function [Eq. (3)]. The slight advantage of one energy method over another is problem dependent, so this preference of energy functions is not universal.

The energies associated with each remaining Hadamard basis vector is determined, and the energies are sorted in descending order. The Hadamards associated with L highest energies are chosen to make up the next set of basis projections, $f^{(i)}$. The measurements collected on iteration i , $y_k^{(i)} = f^{(i)} x_k + n$, are added to the previous measurements to form a cumulative measurement vector $Y_k^{(i)}$, along with the associated $F^{(i)}$ matrix of Hadamards used in this and previous iterations. A linear minimum mean square error (LMMSE) operator, $W_k^{(i)}$ is used to estimate the scenes in each range as

$$z_k^{(i)} = W_k^{(i)} Y_k^{(i)}, \quad (5)$$

where

$$W_k^{(i)} = R_{A,k}^{(i)} F^{(i)T} (F^{(i)} R_{A,k}^{(i)} F^{(i)T} + \sigma_n^2 I)^{-1} \quad \text{and}$$

$$R_{A,k}^{(i)} = \frac{A_k^{(i)} A_k^{(i)T}}{J_k^{(i)}}.$$

$R_{A,k}^{(i)}$ is the covariance matrix of $A_k^{(i)}$. The estimate of the observed scene at each range is, thus, estimated from the Hadamard measurements made up to that point in time and the adapted training set for each range.

As the algorithm iteratively processes Hadamard projections, estimates of the reconstructed images of foreground objects are used to estimate the obstruction of objects in the background scenes. In each iteration, the estimation of the obstruction function at each range is based on image estimations at prior ranges and the initial training set for that range is used to create a temporary training set for that range. The estimated scene at each range is used to create an obscuring mask for the down range training sets. For simplicity, each pixel is assumed to be two-valued: 0 and β_k . With further development, this assumption can be lifted to address scenes with multiple levels of reflectivity. For each range, an estimate of β_k can be obtained by taking the mode of $z_k^{(i)}$. The estimate $z_k^{(i)}$ is then divided by the estimate of β_k truncated between 0 and 1 to obtain the obstruction, $c_k^{(i)}$. The training scene at each range is modified to obtain an obscured training set $A_k^{(i+1)}$ given by

$$a_{k,j}^{(i+1)} = \prod_{k'=1}^{k-1} (1 - c_{k'}^{(i)}) b_{k,j}^{(i)}. \quad (6)$$

For the first range, $a_{1,j}^{(i+1)} = b_{1,j}^{(i)}$.

Three performance metrics are used to quantify the quality of the image reconstruction. The first relies on knowing the actual scenes at each range. The first metric is the root mean square error (RMSE) of each pixel in the current reconstructed image of the scene (the current scene estimations) at a given range with respect to the actual obstructed objects in the scene as would be seen from projector/detection system.

$$d_{x,k}^{(i+1)} = \sqrt{\frac{1}{N} \sum_{\text{image}} (z_k^{(i)} - x_k)^2}. \quad (7)$$

Because it requires knowledge of the exact scene, this metric cannot be used in the adaptive algorithm. The next two metrics do not depend on the actual scene(s), but on the training sets and, thus, can be used to guide the exclusion algorithm (described later). The second metric is, however, based on the prior knowledge that the images at each range were bimodal with one mode equal zero. The estimated distance is given by

$$d_{e,k}^{(i+1)} = \sqrt{\frac{1}{N} \sum_{\text{image}} (z_k^{(i)} - \beta_k c_k^{(i)})^2}. \quad (8)$$

The third metric is based on training sets at each range. The training distance is given by

$$d_{a,k,j}^{(i)} = \sqrt{\frac{1}{N} \sum_{\text{image}} (z_k^{(i)} - a_{k,j}^{(i)})^2}. \quad (9)$$

The estimation was improved and the algorithm became more efficient when on each iteration the training scenes found to be most distant from the estimated scene were removed from $A_k^{(i)}$ and $B_k^{(i)}$ to produce $A_k^{(i+1)}$ and $B_k^{(i+1)}$, respectively. At the end of each iteration, the number of training scenes for each range was reduced by an amount proportional to the number of remaining training scenes at that range, $J_k^{(i+1)} = (1 - \epsilon) J_k^{(i)}$. The proportionality constant ϵ used in the results presented here was 0.25. The minimum number of training scenes at each range was set to 100. However, it was found that the scenes at some ranges required more training scenes to be faithfully estimated. Reducing the scenes in the training set at some range below a threshold number resulted in a significant increase in the measured actual distance for that range. Fortunately, it also caused the estimated distance to rise significantly and this could be used as an indicator that the number of scenes in the training set was reduced too much. During the exclusion algorithm, if it was seen that the estimated distance at a given range increased, the training set at that range was instead increased (not decreased) as $J_k^{(i+1)} = (1 + \epsilon) J_k^{(i)}$.

The level of random Gaussian noise on the measurements was adjustable. The reflectivity of the objects was assumed to be uniform across the scene at a given range. The reflectivity was typically set to 1 for all ranges, or varied around 1; thus, the RMSE of the noise represents roughly relative RMSE with respect to unity pixel amplitudes. For a conventional image, the expected RMSE would be equal to the standard deviation of the Gaussian noise.

Another important metric for the goodness of the adaptive algorithms is the ratio of the number of pixels in the image, N , to the number of measurements or basis vectors required to achieve a measured RMSE level. We refer to this as the compression factor. Compression factors of much greater than one are desirable.

The 3-D AFSI simulator assumes that the measurements of the 3-D AFSI system are made by sequential projections of a finite set of Hadamard codes and resolving the spatially integrated returns from the observable scene at each range is distinguishable via a time of flight or chirped/frequency modulated continuous wave ranging system. (This is a critical note; although we are assuming the use of a conventional ranging system, compressive or feature specific ranging is an area that remains ripe for investigation.) Each projected Hadamard vector is applied to all ranges (though obstructed by the foreground), and thus, each Hadamard vector projected results in multiple measurements, one measurement per resolvable range bin. The number of Hadamard basis vectors projected in each adaptation iteration can be varied. The measurement at each range is of the image as seen from the imaging system's receiver (assuming a monostatic sensing system) and, thus, only the nonobscured portion of the image of the objects at each range is measured. Thus, from the actual scene at each range, an observed image at each range is derived based on the foreground obstructions due to objects ahead of that particular range. Though only 100% obscuring and 100% scattering objects were considered in this study, the simulator's obscuring function for

measurements and adaptation allowed for objects with variable transmission within any or all ranges to be processed with variable reflection/scattering characteristics for different objects.

The decision on the set of Hadamard projections for the next iteration can be based on equal weighting of the energies at all the ranges, or different weightings for different ranges, which can vary with iteration. The weighting of the energies could be used to better balance the simultaneous needs of projecting basis vectors that represent the background scenes and of reconstructing faithful images of the foreground scenes that are obstructing background scenes and thwarting simple FSI of far scenes. One weighting scheme that was explored was one that weighted the near range with a decay exponential (as a function of iteration), weighted the mid-range with an exponential that would peak half way through the desired number of iterations, and weighted the far range with an increasing exponential. For example,

$$E_n^{(i)} = \sum_k w_k^{(i)} e^{-\text{abs}[i-(k+1)N_i/(N_k-1)]/N_i} \left| H_n^{(i)} \frac{1}{J_k^{(i)}} \sum_j a_{k,j}^{(i)} \right|^2, \quad (10)$$

where N_i is the number of desired iterations and N_k is the number of ranges.

The thought was to estimate the foreground scenes early, determine their obstruction of the farther ranges, and then concentrate on the farther ranges progressively. For the scenes tested, no advantage was seen in this approach over constant uniform weighting. In other scenarios, the weighting of the energies at different ranges could be adjusted to meet particular imaging mission objectives or to adjust to different degrees of expected obscuration at each range. Further research is needed.

3 Simulation Studies of 3-D AFSI Algorithms and System

As described above, the 3-D AFSI algorithm uses adaptive measurements in two ways. The first is to iteratively use estimates of foreground scenes to adaptively improve measurements of more distant scenes via an adaptive obscuring mask. The second is to adaptively eliminate scenes from the training set for each range based on the measurements made up to that point. This is referred to as adaptive exclusion.

For the 3-D AFSI simulations (and the experiments described later), the objects explored were polygons of various shapes, sizes, rotations, and displacements. A software object generator was developed in MATLAB[®] that places a set of different polygons within a specified scene. The type (number of sides), radius, rotation, and displacement of each polygon could be specified individually. As polygons are added to the scene at a given range, they could either be ORed or XORed into the scene. Training sets for each range were developed that included a large number of objects with two or more polygon types per scene. Different sets of polygons could be trained on for each range. The training set included polygons with features similar to those in the actual scene to be reconstructed at each range, but the actual scenes were not included in the training set. The training set(s) could also contain a set of diverse objects that represented features

that had nothing in common with the actual objects in the scene.

Tests of the 3-D AFSI simulator found that the LMMSE operator was a powerful tool for reconstructing with trained features. It is important to note that if the actual scene at each range contains features that are not trained, the LMMSE operator will tend to substitute trained features for the untrained features. Thus, it is imperative to have an extensive training set of possible object features included in the training set scenes.

3.1 3-D AFSI with Adaptive Obscuring Function and Low Noise Without Adaptive Exclusion

The simulation tests of the 3-D AFSI algorithm presented here used 32×32 pixel scenes at three ranges. The actual scenes and the images as seen through obstructions are shown in Fig. 1. There are two object shapes at each range, consisting of polygons of different radii that have undergone a displacement and rotation. For example, range 1 has two merged (ORed) polygons. Range 2 has a pentagon on the left and a hexagon (slightly distorted due to pixelization) on the right. Range 3 has a rotated square inside (XORed with) a rotated pentagon. The bottom row shows the images of the same scenes, as they would be seen from the viewpoint of the 3-D imaging system due to obstruction of the scene at each range by the objects in front of them. Range 1 is not obstructed. Range 2 is obstructed by the objects in range 1. Range 3 is obstructed by both the objects in ranges 1 and 2. For each range, a scene was chosen randomly and removed from a set of 2560 created scenes to produce an initial training set of 2559 scenes for each range. Because the actual scenes were removed, the training set has scenes with similar object features, but none of the training set scenes match the actual scenes. A random sampling of scenes in the training sets at the three ranges is shown in Fig. 2.

For the initial simulations, the objects in the scenes at all ranges had the same reflectivity. The black and white levels were 0 and 1, respectively. The Hadamard basis vectors were also 32×32 . A Gaussian noise with variance σ_n^2 was added to each measurement, which resulted in an RMSE for the measured pixels equal to σ_n for a scene produced using



Fig. 1 Top row: 32×32 actual scenes from the closest range on left to farthest range on right. The white and black areas have reflectivity 1 and 0, respectively. Bottom row: 32×32 obstructed images of the scenes as viewed by three-dimensional (3-D) adaptive feature-specific imaging (AFSI) imager after each scene is obstructed by the front objects.

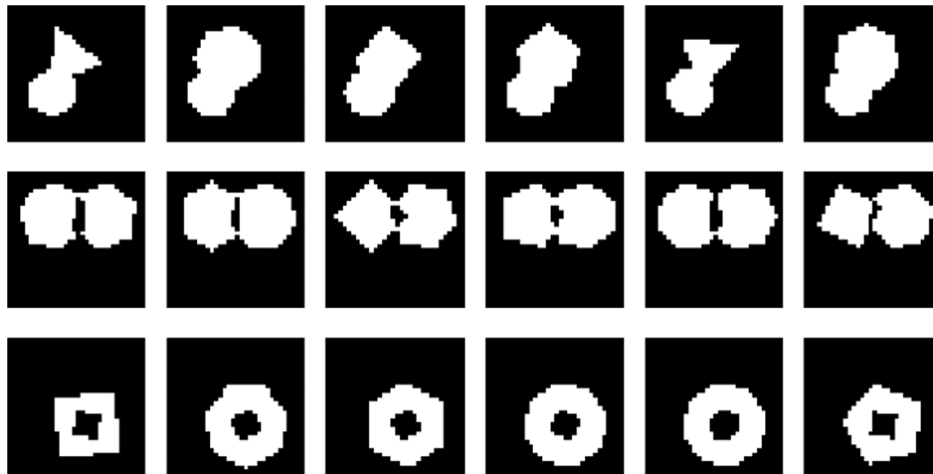


Fig. 2 Six randomly chosen examples of the 32×32 scenes that make up each of the training sets for the closest range (top row), mid-range (middle row), and far range (bottom row).

simple Hadamard reconstruction when all 1024 Hadamard basis vectors are projected and measured and weighted equally in the image reconstruction. This was tested on both low ($\sigma_n = 0.01$) and high ($\sigma_n = 0.4$) noise cases. The complete Hadamard reconstruction of the scenes for the high- ($\sigma_n = 0.4$) noise case is shown in Fig. 3. The measured RMSE distance of each pixel is relative to the reflectivity of the white pixels at each range. Since the white pixels are set to 1 for the first example, the measured RMSE is also the relative RMSE for this example. In a later example, the reflectivity of each scene will be varied.

The goal of FSI is to reach a faithful reconstruction of the scenes with M measurements, where M is much less than N . The results of this simulation are shown in Figs. 4 and 5. For all the cases, in each iteration of the algorithm, eight Hadamard basis projections are randomly selected out of the set of unused basis functions, projected, and measured before the next image reconstruction is performed. The training sets and choices of Hadamard projections were governed by the algorithm described above, using Eq. (4) for the energy calculation. The green (lower) curves in Fig. 4 are the low-noise case, which show a steady improvement up to 56 projections and then levels off and slowly approaches the low RMSE equal to σ_n . The blue (upper) curves in Fig. 4 represent the high-noise case. Here, the adaptive choice of the Hadamard projections and the reconstruction power of the LMMSE operator (given prior knowledge of the scene) resulted in RMSEs at all three ranges that are significantly less than $\sigma_n = 0.4$. Thus, after only 96 Hadamard

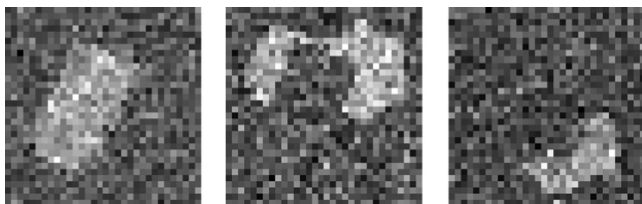


Fig. 3 Reconstruction of the scenes at the three ranges (from left to right are the near, mid-, and far ranges) when the full set of 1024 Hadamards (no compression) are projected and the measurements are used for simple Hadamard reconstruction of the scene. The noise level was set to $\sigma_n = 0.4$.

measurements (a compression factor of $1024/96 = 11$), the LMMSE operator has picked out and reconstructed the essential object features in the scenes at all three ranges.

3.2 Effectiveness of Methods Used in 3-D AFSI Algorithm

The next study explores the effects of varying some of the adaptation methods used in our 3-D AFSI algorithm. While the results (shown in Fig. 6) vary slightly from one run of the simulator to the next, the results presented here are representative of the typical RMSE achieved by the simulator under the parameter settings tested. It is important to compare the performances of the different algorithm settings within each range. The figure shows the degradation of the performance of the 3-D AFSI algorithm with respect to the full algorithm

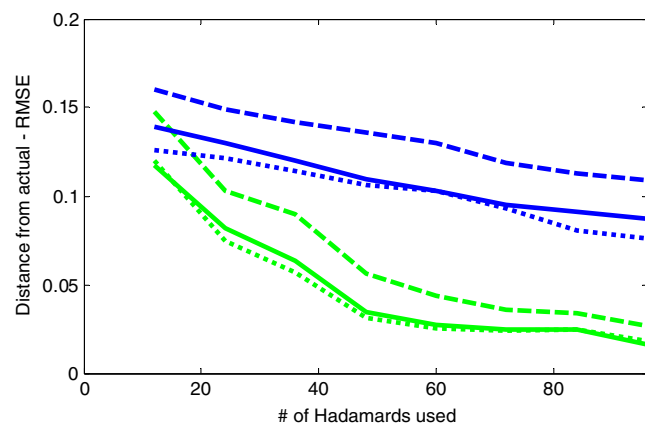


Fig. 4 The actual distance [root mean square error (RMSE) pixel error with respect to actual scene] versus number of Hadamard basis projections for 3-D AFSI simulation for 32×32 scenes at three ranges with $\sigma_n = 0.01$ (lower green curves) and $\sigma_n = 0.4$ (upper blue curves). The solid, dashed, and dotted lines show the actual distances for ranges 1, 2, and 3 (close, mid-, and far ranges), respectively. In each iteration of the algorithm, eight Hadamard basis projections are selected as per the algorithm described in the text. The total number of Hadamards projected after eight iterations is 96. Faithful estimation of the scenes was achieved with much less than $N = 1024$, resulting in a compression factor of ~ 11 for achieving a relative RMSE of 0.02 to 0.03.

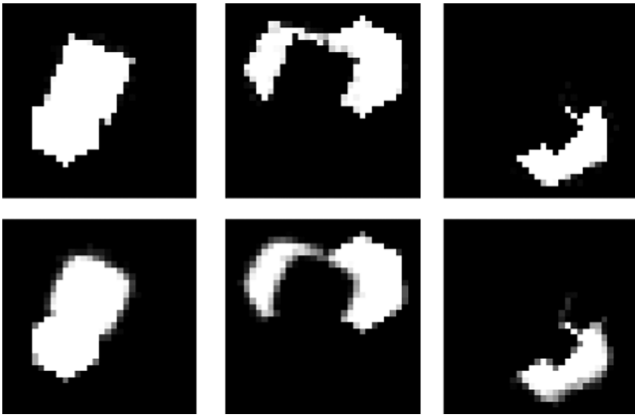


Fig. 5 The estimated images for the three ranges (left to right) for $\sigma_n = 0.01$ (top row) and $\sigma_n = 0.4$ (bottom row). The number of Hadamards projected was 96, for a compression factor of 11.

described above when (1) exclusion is turned off (constant J_k), (2) the energy function used is given by the average of the energies (versus the energies of the averages), and (3) unequal exponential weightings are used in the energy function (described earlier in text), versus using constant weighting.

The effects of not employing adaptation for obstruction (AFO) are demonstrated. The main effects of not employing AFO are to distort the reconstruction of the scenes at far ranges. Figure 7 compares two reconstruction methods both in the high- and low-noise regimes. Scenes reconstructed using the 3-D AFSI algorithm with AFO are compared to

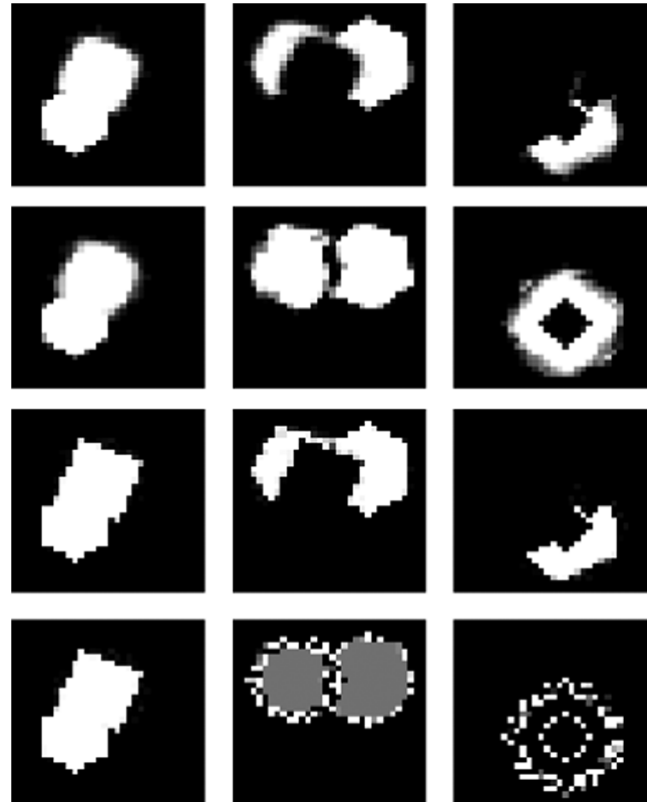


Fig. 7 Simulated algorithm with and without adaptation for obstruction (AFO) in the low- ($\sigma_n = 0.01$, top two rows) and high- ($\sigma_n = 0.4$, bottom two rows) noise regimes. The estimated scenes for the three ranges are shown left to right. The estimated scenes in the top row and third row down were produced with the algorithm with AFO. The estimated scenes in the second from top row and bottom row were produced with the algorithm without AFO. Ninety-six Hadamards were projected for a compression factor of 11.

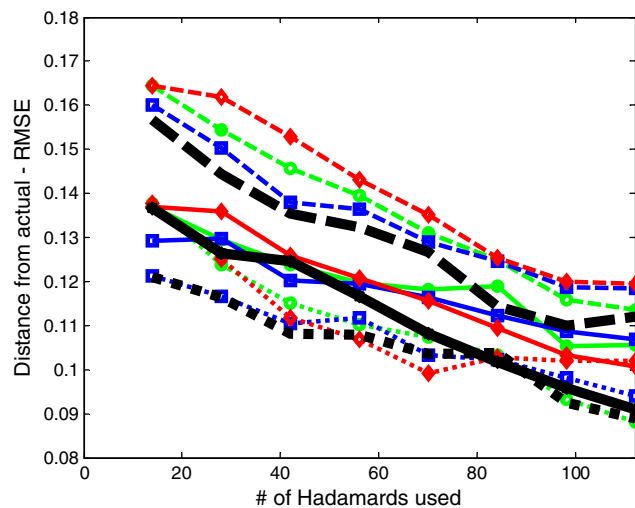


Fig. 6 The progression of the actual distance measurements versus number of Hadamards projected for different simulator settings. The results for each range (solid is close-range, dashes is mid-range, and dotted is far range) are shown for each setting (color) of the simulator. The thick black curves are the full simulator as used throughout the paper. The red curves with diamond markers are the full simulator with exclusion turned off (constant J_k). The blue curves with square markers are the full simulator, but using the energy function given by the average of the energies (versus the energies of the averages). The green curves with circle markers are the full simulator, but instead of constant weighting, the exponential weightings described earlier is used. The added Gaussian noise level was $\sigma_n = 0.4$ for all cases. The total number of Hadamard projections after 14 iterations was 112, for a compression factor of 9 to achieve the final RMSE levels.

scenes using an energy function and LMMSE that uses the unobstructed (raw) training sets for adaptation of the Hadamards and estimation of the scenes. In the low-noise cases, the algorithm without AFO estimates fuller images (showing unobstructed portions) and roughly estimates the unobstructed shapes with slightly less detail of the edges of the actual objects compared to the algorithm with AFO. In the high-noise limit, the algorithm without AFO fails significantly, while the algorithm with AFO faithfully reproduces the obstructed images with good estimation of the objects' edges. If the goal was to estimate the unseen portions of the objects, using the algorithm without AFO might be useful, but it is likely that the results of the algorithm with AFO (weights of obstructed training set) could be used to estimate the continuation of the obstructed objects with high fidelity. This is a topic for further research.

Finally, though the reflectivity of the scenes used in the above examples was fixed to a value of 1.0, the current algorithm can handle having a different reflectivity at each range. In Fig. 8, the estimated scenes are shown for reflectivities of 0.9, 1.6, and 1.1 for the near, mid-, and far ranges. The algorithm did well at estimating the reflectivities of the first two ranges and was off by 10% in estimating the reflectivity of the far range. Note that the different reflectivities did not affect the ability of the algorithm to adapt to the obstructions.



Fig. 8 Estimation of scenes and scene reflectivities for scene reflectivities of 0.9, 1.6, and 1.1 for the near, medium, and far ranges, respectively. The high-noise regime ($\sigma_n = 0.4$) was used here to compare with the experimental results, which have similar noise levels. The reflectivity levels measured for the estimated scenes at each range were constant over much of the objects seen in each scene. Ninety-six Hadamards were projected for a compression factor of 11.

4 Experimental Demonstration of 3-D AFSI Algorithm

To provide a proof-of-concept demonstration of the 3-D AFSI algorithm, the 2-D FSI imaging setup shown in Fig. 9 was built and used. The 2-D FSI system consisted of a video projector (Texas Instruments' Pico Projector, Dallas, Texas) to create Hadamard projections with a variable number of pixels and widths. An e-ink reader (Amazon Kindle, Seattle, Washington) was used for the scenes, which allowed for accurate and quick variability of the scenes, as well as the e-ink technology, which provided good contrast, high resolution, and front-scattering scenes that are polarization and angle insensitive. The projector was mounted on an x - y - z stage that allowed for the projections to be centered in the scenes on the Kindle. The Kindle was mounted on a rotation stage whose axis was aligned through the center of the scenes on the Kindle. The rotation stage was used to align the axes of the Hadamard projections and scenes. The z -axis stage allowed for slight scale adjustments of the projections, so that the size of the pixels of the projections matched the size of the pixels of the scenes. Here, pixel refers to the pixilation of our scenes and projection basis, not the actual pixels of the projector or Kindle. A red color filter is used in front of the photodetector to allow detection of only the red cycle of the projector's video projection. The single-pixel detector is mounted next to the projector, up and off to one side to avoid backreflections off the Kindle. Even with this measure in place, a slight glare can be seen in the reconstructed images.

The measurements were made with a 48 kilosamples per second analog to digital converter synchronized to the video projection. This allowed digital gating of the measurement to integrate only when red light was projected from the projector (versus blue and green projections). The integrated

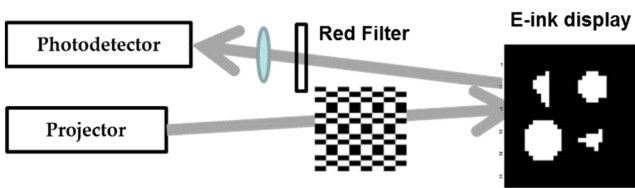


Fig. 9 The experimental two-dimensional FSI setup with projector for projecting Hadamard vectors onto e-ink display. The scattered light from the display passes through a red filter and the spatially integrated light is collected on a single large-area photodetector. The sampling of the photodetector is synchronized with the red period in the video projection cycle.

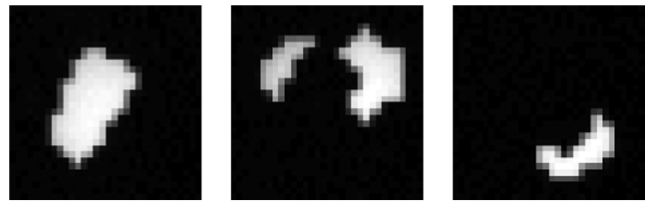


Fig. 10 Experimental reconstruction of the obstructed 32×32 scenes at near, mid-, and far (left to right) ranges, based on measurement of all the 1024 Hadamard projections, with no knowledge of the training set.

measurements of the red light were averaged over several video cycles. The measurement for a totally black scene (not necessarily zero light projection) was subtracted from the measurement of the scenes for each Hadamard projection.

The first demonstration used the same set of 32×32 scenes as in the simulation results in the previous section, except that they are now displayed on the e-ink display. The experimental demonstration was performed by first making measurements of all the Hadamard basis projections for the unobscured scenes at closest range and then repeating the full set of Hadamard projections for the obscured scenes at the mid-range and for the doubly obscured scenes at the far range. The obstructed scenes are created in MATLAB® and loaded onto the Kindle. The full set of Hadamard projection measurements was uploaded into the simulator, but only used selectively in the simulator as the algorithm calls for particular Hadamards to be measured. The estimated scenes from a simple Hadamard reconstruction using all 1024 Hadamard measurements are shown in Fig. 10.

By only using the measurements as needed, the simulator bases its estimates on a significantly compressed number of experimental Hadamard projection measurements in order to estimate the scene. The noise level of the experiment was estimated to be 0.4 by measuring the RMSE of the Hadamard image reconstructions of the scenes with all 1024 Hadamard projections (with no knowledge of the training set). These noise levels are relative to the measured reflectivity amplitudes of the objects in the Hadamard-base reconstruction. The amplitudes for the near, mid-, and far ranges are 2.0, 1.6, and 1.9, respectively. The noise level of 0.4 was incorporated into the postprocessing algorithm and used by the LMMSE operator in reconstructing scenes from a limited set of the experimental data. No additional noise is added.

The estimated scenes after only 48 Hadamard projections are shown in Fig. 11, representing a compression factor of 21. The measured amplitudes (1.6, 1.2, and 1.3) roughly



Fig. 11 Experimental demonstration of 3-D AFSI for 32×32 scenes. The reconstructed scenes for the near, mid-, and far ranges (from left to right) with full AFSI adaptation. Images shown are with 48 projections, for a compression factor of 21. The measured amplitudes of the objects are 1.6, 1.2, and 1.3.

followed those measured with full Hadamard projection. The estimated distances for the three ranges are 0.2, 0.2, and 0.13, which are relative to measured amplitudes of 1.6, 1.2, and 1.3, respectively.

5 Summary

In conclusion, a 3-D AFSI algorithm was developed and tested via simulation and experiment. The adaptation of obstruction algorithm setting proved effective in forming a scene estimation where only part of an object is visible. The periodic ordering of the Hadamards, based on the Hadamard energies for all the training scenes summed over all ranges, worked best. The adaptation of the training sets by exclusion of distant scenes provided a benefit; however, this result may vary with different training sets and different types and sizes of objects in the scenes. The results show that 3-D AFSI is an effective means of reconstructing object features with a significantly reduced measurement set for scenes at multiple ranges and in environments where scenes at differing ranges can be obscured by foreground objects.

Acknowledgments

The authors are thankful to Dr. Mark Neifeld, Dr. Zeb Barber, and Steven Crouch for helpful discussions. The authors gratefully acknowledge the support of the Air Force Research Laboratory (AFRL) under Small Business Innovative Research Phase II contract number FA8650-10-C-1722. The opinions expressed are those of the authors and not necessarily those of AFRL.

References

1. R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Process. Mag.* **24**(4), 118–121 (2007).
2. M. A. Neifeld and P. Shankar, "Feature-specific imaging," *Appl. Opt.* **42**(17), 3379–3389 (2003).
3. M. Mishali and Y. Eldar, "From theory to practice: sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Sel. Topics Signal Process.* **4**(2), 375–391 (2010).
4. W. R. Babbitt, Z. W. Barber, and C. Renner, "Compressive laser ranging," *Opt. Lett.* **36**(24), 4794–4796 (2011).
5. E. J. Candes, "Compressive sampling," in *Proc. of the Int. Congress of Mathematicians*, Vol. 3, pp. 1433–1452, European Mathematical Society, Madrid, Spain (2006).
6. P. K. Baheti and M. A. Neifeld, "Recognition using information-optimal adaptive feature-specific imaging," *J. Opt. Soc. Am. A* **26**(4), 1055–1070 (2009).
7. Z. W. Barber et al., "Accuracy of active chirp linearization for broadband frequency modulated continuous wave lidar," *Appl. Opt.* **49**(2), 213–219 (2010).
8. B. M. Kaylor et al., "Face recognition via a projective compressive sensing system," *Proc. SPIE* **8254**, 82540G (2012).
9. J. Ke, A. Ashok, and M. A. Neifeld, "Object reconstruction from adaptive compressive measurements in feature-specific imaging," *Appl. Opt.* **49**(34), H27-39 (2010).
10. A. Ashok and M. A. Neifeld, "Compressive imaging: hybrid measurement basis design," *JOSA A* **28**(6), 1041–1050 (2011).

Wm. Randall Babbitt is a professor of physics at Montana State University. He received his BS degree in physics from Stanford University in 1982 and his PhD degree in physics from Harvard University in 1987. He is the author of more than 75 journal papers and has 14 issued patents. His current research interests include optical signal processing in the time and spatial domain, compressive sensing, spectral holography, and remote education to rural students.

Krista M. Drummond received her BS degree in electrical engineering from Montana State University in 2012. She is currently a master's student in the College of Optical Sciences at the University of Arizona, working in the OSC Polarization Lab under Dr. Russell Chipman. Her research is focused on using polarization properties to determine a detailed characterization of ice layers.

Brant M. Kaylor received his BS in optics from the University of Rochester in 2002 and his MS in optical sciences from the University of Arizona in 2004. He joined Bridger Photonics Inc. as an optical scientist in 2008 and has extensive technical expertise in optical engineering, LADAR imaging, computational imaging, optical signal processing, and biomedical imaging. He has played a critical role in the development of the company's precision LADAR, active three-dimensional (3-D) imaging, and computational imaging systems.

Randy R. Reibel graduated with a dual major in physics and computer science from Western Washington University. After receiving his PhD from Montana State University in physics, he worked as a laser physicist under a cooperative research agreement on-site at the Air Force Academy for Directed Energy Solutions. In 2007, he cofounded Bridger Photonics Inc., where he is currently the chief operations officer, concentrating on high-resolution coherent lidar systems and 3-D imagers.