

Characterization of a compressive imaging system using laboratory and natural light scenes

Stephen J. Olivas,^{1,*} Yaron Rachlin,^{2,3} Lydia Gu,^{2,4} Brian Gardiner,²
Robin Dawson,² Juha-Pekka Laine,² and Joseph E. Ford¹

¹Photonic Systems Integration Laboratory, Electrical Engineering Department, University of California at San Diego, 9500 Gilman Dr., La Jolla, California 92093, USA

²The Charles Stark Draper Laboratory, Inc., 555 Technology Square, Cambridge, Massachusetts 02139, USA

³Currently with Lincoln Laboratory, Massachusetts Institute of Technology, 244 Wood Street, Lexington, Massachusetts 02420-9108, USA

⁴Currently with Cogito Health, 7 Water Street, Suite 400, Boston, Massachusetts 02109, USA

*Corresponding author: sjolivas@ucsd.edu

Received 28 February 2013; revised 1 May 2013; accepted 19 May 2013;
posted 21 May 2013 (Doc. ID 185364); published 25 June 2013

Compressive imagers acquire images, or other optical scene information, by a series of spatially filtered intensity measurements, where the total number of measurements required depends on the desired image quality. Compressive imaging (CI) offers a versatile approach to optical sensing which can improve size, weight, and performance (SWaP) for multispectral imaging or feature-based optical sensing. Here we report the first (to our knowledge) systematic performance comparison of a CI system to a conventional focal plane imager for binary, grayscale, and natural light (visible color and infrared) scenes. We generate 1024×1024 images from a range of measurements (0.1%–100%) acquired using digital (Hadamard), grayscale (discrete cosine transform), and random (Noiselet) CI basis sets. Comparing the outcome of the compressive images to conventionally acquired images, each made using 1% of full sampling, we conclude that the Hadamard Transform offered the best performance and yielded images with comparable aesthetic quality and slightly higher spatial resolution than conventionally acquired images. © 2013 Optical Society of America

OCIS codes: (110.0110) Imaging systems; (100.0100) Image processing; (110.1758) Computational imaging.

<http://dx.doi.org/10.1364/AO.52.004515>

1. Introduction

Traditional cameras capture a full resolution image, then process and compress this information for storage, typically retaining 1% of the full (raw) image data. Even without *a priori* knowledge of the scene, perfect reconstruction can be possible with sampling below the Nyquist limit [1–3]. Compressive imaging (CI) systems acquire scene information through a series of measurements of the total energy

transmitted through image plane spatial filters (basis functions), using the minimum number of measurements required to achieve the desired image quality or, more generally, the feature-based scene information [4–7]. CI systems are of great potential interest for specific applications where focal planes with the desired spectral or spatial resolution are unavailable, or where adaptive feature-based sensing may operate with few measurements. Before undertaking these exotic applications, however, it is useful to investigate the performance of CI systems where they can be directly compared to images acquired by conventional focal planes. This paper provides the first, to our

knowledge, experimental comparison test of a CI system where we explore the impact of alternative basis functions (digital, continuous, and random) on binary, grayscale, and natural light scenes, and compare the resulting images to those acquired by a conventional focal plane imager.

The paper is organized as follows. In Section 2, we briefly summarize CI theory and discuss camera image capture in relation to sampling theory. In Section 3, we describe the experimental apparatus and methods to reconstruct the images using experimentally collected data. Section 4 explains the hardware configuration for grayscale filtering. In Section 5, we discuss system operation and the experimental setup under which the compressive images were taken. We then present results and conclude with a discussion of their significance.

2. CI and Image Sampling Theory

Any signal (or image), x , can be expressed as a weighted sum of basis functions following the equation $x = \sum_{n=1}^N [\theta(n)\psi_n]$. Here ψ_n are the basis functions, $\theta(n)$ are the weighing coefficients, and N is the total number of pixels. The number of basis functions needed to completely describe the image space is necessarily equal to the number of pixels in the image. Nyquist theory states that sampling at a rate faster than twice the highest frequency contained in the signal guarantees perfect reconstruction and proves beneficial when sampling band-limited signals. However, many images are “sparse” and can be completely represented using fewer samples, provided an appropriate basis set is used. For this reason, transformation between basis sets is useful and has become commonplace [8].

Compression algorithms decorrelate image signals into a small set of orthonormal basis coefficients leveraging the sparse representation of images. This is the premise of lossless image compression used in many cameras today, where $x = \sum_{l=1}^K [\theta(l)\psi_l]$ even though $K < N$. Karhunen–Loève expansions completely decorrelate a signal into its most sparse basis representation, maximally compacting the energy (information) contained in the signal. Reduction in dimensionality allows for more efficient use of memory and data transfer. It is possible to compress the information content further at the cost of less accurate reconstruction. For example, JPEG uses discrete cosine transforms (DCT) and JPEG2000 uses wavelet transforms as efficient basis representations for lossy or lossless compression. In most applications, however, there is no *a priori* knowledge about the image scene and therefore its bandwidth or decorrelation, which in turn limits the implementation of this rich theory to post processing.

Modern cameras compress images by first acquiring all the pixel values (basis coefficients), then transforming them into another basis set, discarding the smaller coefficients and finally encoding the remaining coefficients. This is inefficient since samples and analog to digital (A/D) conversions are

performed for coefficients that are ultimately unused. It would be ideal to sample only the values needed for perfect reconstruction. The basis set these traditional cameras use to sample the image space is a bandlimited two dimensional comb function, with one impulse per pixel. This sampling method does not statistically meet the needs of the ensemble of natural images, whose spectra follows a distribution inversely proportional to the spatial frequency [9–12]. This spectral distribution follows $S(f) = f^{-\gamma}$, where f is the spatial frequency and γ is an empirical constant, $1.8 < \gamma < 2.3$. The imager can be tailored to the statistical distribution of most images by sampling lower spatial frequencies more densely.

Donoho, Candès, Romberg, and Tao [1–3] have shown that if a signal is sparse (compressible) in one basis (called the reconstruction basis), perfect reconstruction is guaranteed provided at least $\Omega > T \cdot (\log N)/C_M$ samples are measured using a basis incoherent to this reconstruction basis. Here Ω is the number of measured samples, T are the sparse basis coefficients, N is the total number of pixels, $(N - T)/N$ is the degree of compressibility, and C_M is a constant which depends on an accuracy parameter M . If the measurement basis is random, then with overwhelming probability it is incoherent with arbitrary basis sets, satisfying the conditions above. In other words, CI does not seek to measure the T sparse basis coefficients directly; instead, it seeks to measure Ω random basis coefficients. These random measurements are democratic in that all samples have an equal probability of capturing significant data [13]. Since the number of measurements does not describe the image space completely, image recovery is an ill-posed problem. Yet, if the image is sparse and the measurement basis satisfies the restricted isometry property (RIP), then an L_1 solution to the basis pursuit problem satisfies the reconstruction as well as the L_0 solution [4]. The L_1 minimization is less complex than L_0 minimization; it arrives at the solution faster and provides flexibility in numerical optimization [14,15].

Sampling with both a tailored spatial frequency distribution and compressive sensing (CS) techniques is fundamentally different than sampling in modern cameras. They all sample image projections on a basis set, but the basis sets used in CI need not be impulsive, and can extend across the image space to collect a spatial superposition of image intensity of the scene. This idea of measuring intensity superposition in the image space makes compressive imagers advantageous when samples are scarce or expensive, or when the image data can be accurately represented by a relatively small number of measurements in any basis set.

3. System Design & Configuration

CI requires that the intensity values of a scene be spatially multiplexed and summed, using specific basis patterns, to form a corresponding set of basis coefficients. The measurable coefficients are formed by

the spatial dot product of the scene intensity distribution with each basis function. Several hardware architectures for CI have been proposed [16]. The basis projections can be sampled by imaging the scene onto the spatial light modulator (SLM), then condensing the resulting energy onto a single detector (optical summing). Alternatively, a 2D array image sensor with spatially weighted sensitivity and analog summing could implement the same operation (electronic summing) [17,18]. Creating an image sensor (focal plane array) that can directly operate as a compressive imager is possible; however, fabricating one with high resolution, optical fill-factor, and sensitivity is challenging, while high resolution SLMs are commercially available. The benefits of single photodetector compressive imagers were enumerated in [4]. We chose to use this system configuration since it demonstrates the extreme case of CI, which should provide the most informative performance comparison.

To characterize single photodetector CI system performance as a function of encoding basis sets, we built the test-bed system similar to those of previous works [4–7] (shown schematically in Fig. 1). The input scene is imaged onto a 1080p Texas Instruments digital micro-mirror device (DMD) [19], so that each mirror element can direct the light from the scene into one of two direction paths, separated by $\pm 24^\circ$. The DMD specifications are given in Table 1. The positive portion of the transform is projected in the $+12^\circ$ direction and the negative portion in the -12° direction. We use a condenser lens to collect all the optical energy from the positive path onto a single-element photodetector (with noise equivalent power $NEP = 0.2 \text{ pW}/\sqrt{\text{Hz}}$). The 12° deflection angle of the DMD mirrors imposes a requirement that the lenses have an $F/\# \geq 2.573$, following the relation $F/\# = (1/2 \tan \theta) = (f/d)$, where θ is the maxi-

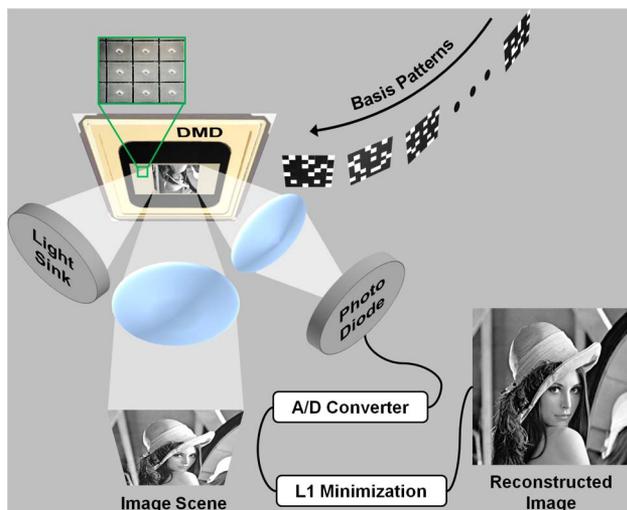


Fig. 1. Schematic of CI system. A scene is imaged onto a digital micromirror device (DMD), which serves as a SLM. A sequence of specific basis patterns are emulated by the DMD mirrors, and the filtered signals are collected by a second lens onto a single photodetector. A sink blocks undesired deflected stray light.

Table 1. Texas Instruments High Definition 0.95 1080p DMD Specifications

Mirror array size	1920 × 1080
Number of mirrors	2,073,600
Mirror pitch	10.8 μm
Fill factor	91%
Active area (mm)	20.736 × 11.664
Diagonal (mm)	23.8
Deflection angle	$\pm 12^\circ$
Settling time	12 μs
Mirror overshoot	1.5 μm
Rebound deflection	2°
Data rate	51.2 Gb/s
Binary patterns per second	24,690
Clock rate	400 MHz (2× LVDS)

mum collection angle, f is the focal length of the lens, and d is the diameter of the lens. Optomechanics must also be taken into account when satisfying this angular requirement. We chose to use a Fujinon F/9 double Gauss Copal lens with a focal length of 180 mm. The system was operated both in the lab with a back illuminated object (transparency) and outdoors with a natural daylight scene.

Since we only sample from one output path of the DMD, it is necessary to split each pattern into a sequence of two complementary patterns, one containing the negative portion and one the positive portion. For each basis pattern, two PNG image files are generated and stored to the system computer's hard drive. Photodiode measurements are taken as the DMD displays the patterns from the corresponding basis set. In order to obtain each coefficient, measurements taken from the negative portion of the basis are subtracted from measurements taken from the corresponding positive portion of the basis. In this way, the coefficients of the basis set are collected to later reconstruct the image scene. Each PNG file's resolution is 1920×1080 , corresponding to the DMD specifications, with the actual transforms occupying the center 1024×1024 mirrors for a total of 1,048,576 mirrors. This determines the maximum resolution of the system, which is about 1.04 MPix. A photo of the experimental system is shown in Fig. 2. The DMD and photodetector were configured in the system as follows. A computer running the Digital Light Innovation's Accessory Light-modulator Package (ALP) software was used to upload previously generated PNG transform basis image files to the Discovery D4100 controller board via a USB 2.0 link, which in turn controls the Texas Instruments DMD chip. A Newport Optics Corporation photodetector 918D-SL was used to measure light intensity collected from the DMD and send it to a Newport power meter 1936-C, which was sampled using a National Instruments PCIe-6363 data acquisition card in a desktop computer. We synchronized the sampling of the basis coefficients with the displaying of each basis projection by feeding the trigger from the D4100 controller board to the data acquisition card.

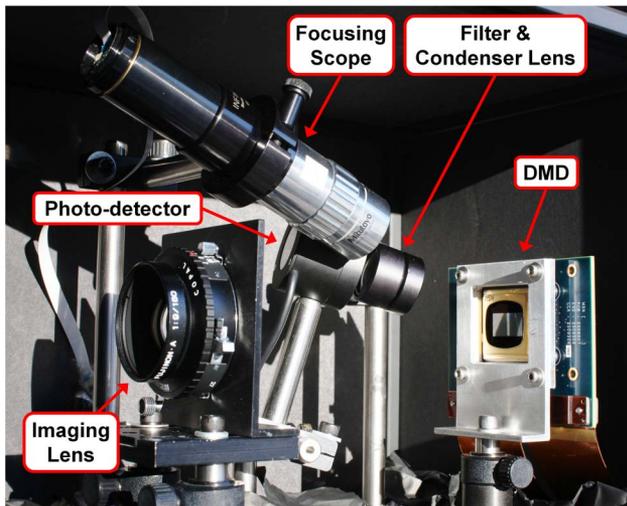


Fig. 2. Photo of experimental CI system. A scene is imaged onto the DMD which directs a portion of the imaged scene to a condenser lens, which then directs the optical energy to a single photo-detector. A focusing scope helps the user to position the imaging lens. Filters can be used to capture multispectral images. The enclosure and light sink are not shown.

Using the data obtained from the CI system, we formulate an underconstrained problem which has an infinite number of solutions, $b = Ax + z$. Here A is the sampling basis set, z is the noise term, b are the measured basis coefficients, and x is the signal (image) we reconstruct. A solution can be chosen by enforcing constraints, such as sparsity. In other words, we assume the solution has many zero or near-zero coefficients in a representation basis, which is the case for compressible signals. We also enforce smoothness in the solution, which is reasonable for most natural images. Linear programming can be used to find this solution; however, in the presence of noise, basis pursuit techniques prove to be better suited. We chose to use the NESTA software package due to its fast and accurate recovery method for solving basis pursuit problems [20]. The code converges to a solution quickly, and includes total variation smoothing parameters which can be used to deal with different amounts of noise within the data.

4. DMD Time Division Multiplexing

Displaying two-level patterns on the DMD is a far simpler process than displaying a discretized continuous signal. The DMD operates in two states, lending itself to easily display binary patterns. Having the DMD display continuous (in magnitude) patterns, however, requires time division multiplexing (TDM) over the duty cycle of the display time. See [21] for details on DMD TDM operation. When loading 8 bit PNG image files into the D4100 controller board, the onboard software automatically generates eight 1 bit files corresponding to its binary representation. The D4100 then uses these 1 bit files for pulse width modulation of the DMD. This time sequence of patterns impacts the sampling of the compressive imager and were investigated using the

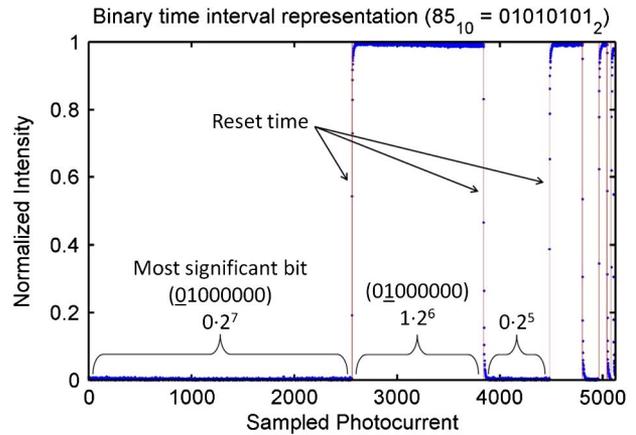


Fig. 3. Grayscale representation through the use of TDM of the DMD. An 8 bit PNG image loaded into the D4100 controller board is converted into eight 1 bit image files, each of which is displayed for a different time duration. The dataset of a single frame shows a test case where all the values of the input 8 bit PNG are set to 85_{decimal} which, when converted to binary, alternate between zeros and ones. Some samples are taken during the mirror transition reset time.

experimental data plotted in Fig. 3. To obtain this data, we input an 8 bit PNG representing the pattern where every pixel value was set to 85_{decimal} . This value was selected because its binary representation, 01010101_{binary} , requires every transition in the TDM representation. To generate this dataset, 5120 samples were taken during the display time of the PNG. This corresponds to 2560 samples for the most significant bit, and 20 samples for the least significant bit. Some samples are taken during the transition state of the mirror, when they are moving between the $+12^\circ$ and the -12° positions, as can be seen in Fig. 3.

To characterize mirror transitions as a possible source of error, we used a uniform white light input source and displayed patterns where every pixel was set to the same constant value. We input 256 PNG image files where the constant values ranged from 0 to 255, so the measured intensities should ideally also ramp from 0 to 255. We then compared two sampling modalities. For the first, we sample within each subframe time interval for the respective bits, then average and scale them by their respective place value. This sampling technique was used to selectively ignore samples taken during mirror transitions. For the second sampling modality, we sample and average all 5120 values taken during the entire TDM 8 bit frame. The comparison between these two sampling modalities for TDM of 8 bit DMD patterns is shown in Fig. 4. From the plots in Figs. 4(c) and 4(d) it is evident that there is error associated with the transitions required for TDM grayscale representation. We operated the system, taking an average over the entire TDM PNG display time, without attempting to avoid sampling mirror transitions since the resulting error was less significant. The system operation parameters the authors used for the 1 bit and 8 bit encoding are summarized in

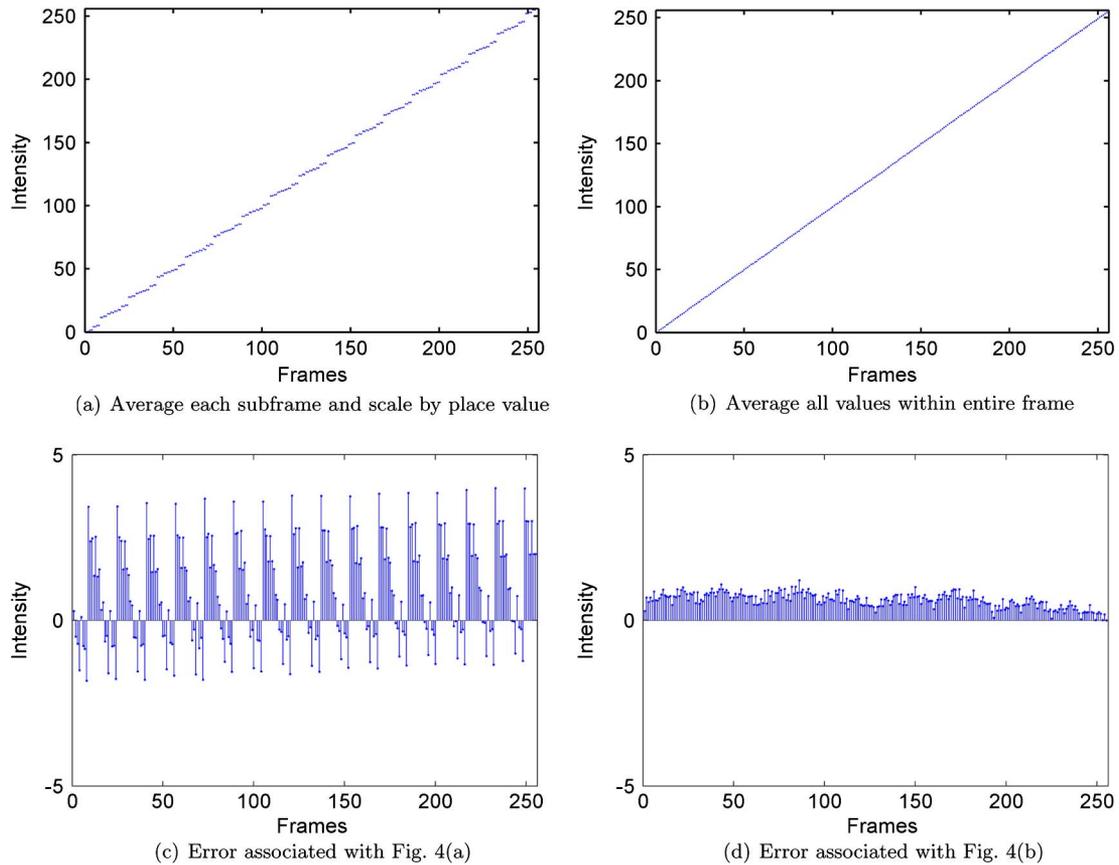


Fig. 4. Two methods of sampling the CI system patterns are compared in order to improve 8 bit system performance. To perform the comparison, 256 PNG files are run through the system, each containing uniform intensity ranging from 0 to 255. (a) Data taken by averaging samples within a subframe and scaling by their place value. (b) Data taken by averaging all values within the entire frame. Errors associated with (a) and (b) are shown in (c) and (d), respectively. Periodic fluctuations in (c) are dominated by the $2^5 - 2^4$ place value transition. Averaging over the entire frame introduces less significant error.

Table 2 for the particular case of capturing 1% of complete basis measurements. Basis load time is the time it takes to load the transform patterns onto the D4100 controller board, image acquisition is the duration of image capture, basis set memory is the memory needed to store the basis set, and measurement memory is the memory needed to store the measurements of the basis set coefficients. Many of the limitations are due to the memory size on the D4100 and the USB 2.0 data transfer rates. These are not fundamental limits; a dedicated ASIC controller would improve performance.

5. Experimental Setup

Two measurement/reconstruction approaches were utilized in the experiments. The first relies on

Table 2. Speed and Memory Requirements of Encoding Algorithms for Acquiring 1% of the Samples (10,486 Basis Patterns and Coefficients) of the Corresponding 1.04 MPix Image

Encoding (bit)	Basis Load Time (min)	Image		
		Acquisition (min)	Basis Set Memory	Measurement Memory (kb)
1	2.22	1.4	250 Mb	163
8	16.83	1.4	7.4 Gb	163

sampling techniques presented in the original formulation of CS theory [1–3], where the measurement basis is known to be incoherent with the sparse representation basis, and where sparsity pursuit reconstruction algorithms are used to reconstruct the image. The second approach uses a sparse measurement basis and relies on *a priori* knowledge (discussed in Section 2) that the statistical distribution of spatial frequencies in natural images, as an ensemble, follows a decaying exponential. Noiselet transforms were utilized as the incoherent measurement basis for the former, and Hadamard and DCT were used as the sparse basis for the latter.

The system's optical components were enclosed to prevent stray light from affecting the measurements. We took 5120 samples per basis pattern, each of which was displayed for 4 ms, with the data acquisition card operating at 200,000 samples per second. Although not the maximum DMD speed, this display time accommodated more measurements. The DMD mode of operation depended on whether the basis transform sets were two level or discrete. Our experiments tested three basis transform sets: the Noiselet, Hadamard, and the DCT, shown schematically in Fig. 5. The first two patterns are two level and the last is a discretized continuous signal. For

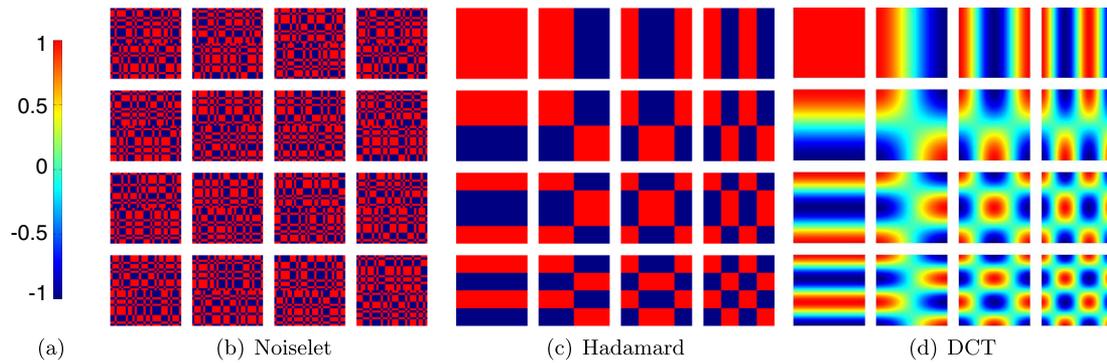


Fig. 5. Sixteen of the basis patterns depicting the (b) Noiselet, (c) Hadamard, and (d) DCT basis transform sets. The colorbar is shown in (a).

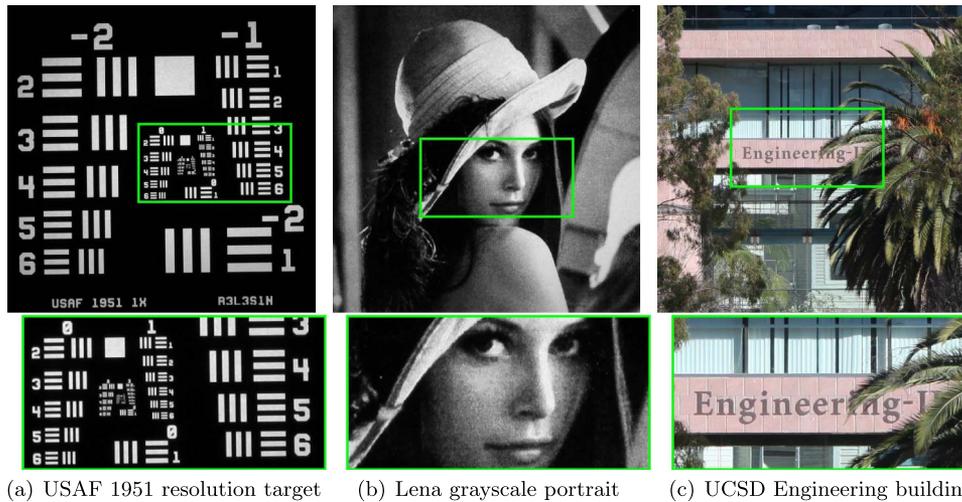


Fig. 6. Ground truth images taken with a Canon 5D Mark II DSLR camera. The scenes include (a) and (b), back-illuminated chrome on glass transparencies in a lab setting, and (c) a daylight illuminated building.

reasons discussed in Section 4, the samples for 1 and 8 bit transforms were averaged over the entire pattern display time. The experimental conditions are discussed in the following.

Two chrome on glass transparencies were back illuminated using a 10 W white light four-element

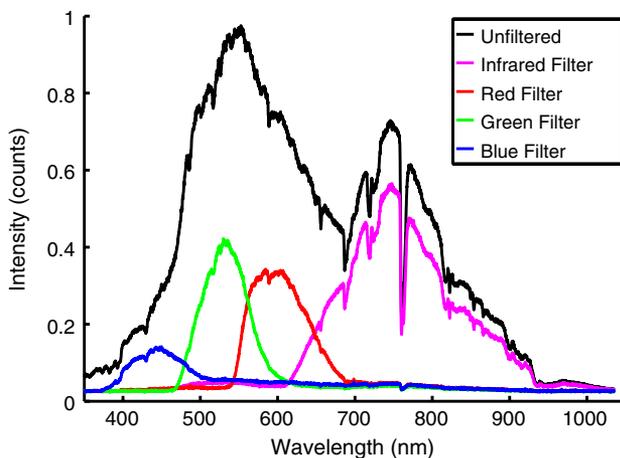


Fig. 7. Spectral data of the outdoor image scene [Fig. 6(c)] taken using an Ocean Optics Spectrometer with and without color filters in place.

(Seoul Semiconductor model number P7) LED and diffusive screen in a lab setting. A USAF 1951 resolution target, whose ground truth image is shown in Fig. 6(a), served as a binary transparency containing sharp edges. A portrait of Lena, a standard test image whose ground truth image is shown in Fig. 6(b), served as a grayscale image with more texture and features. A camera with a sensor pixel pitch and resolution comparable to the mirror pitch of the DMD were unavailable, so in order to form a ground truth image we used a commercial digital SLR camera with resolution superior to the CI system. Both of these indoor lab images were taken using a Canon DSLR 5D Mark II using a Sigma 50 mm focal length lens set to $f/14$ with an exposure of $1/125$ and $1/500$ s, respectively. The sensor in the Canon Mark II has 5616×3744 pixels with a pitch of $6.4 \mu\text{m}$, of which we used a 1024×1024 pixel region to image the scene. For the lab images, there was 18.3 lux of illuminance and 11.91 mW of power at the camera lens. No optical filters were used. During these indoor tests, stray light was well controlled and its effects were negligible. Contrast was limited by the DMD.

For an outdoor field test of the system in a daylight setting, the input scene was of a University of

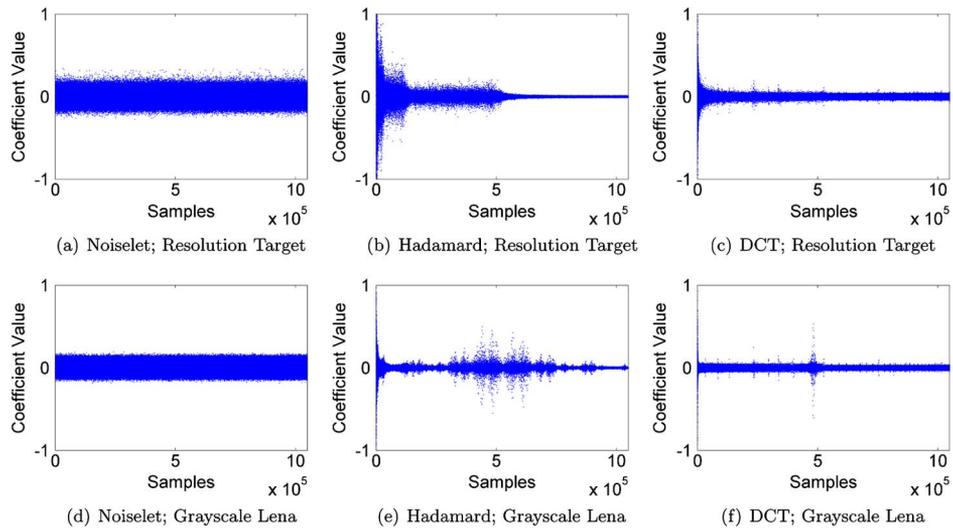


Fig. 8. Experimentally acquired (a), (d) Noiselet, (b), (e) Hadamard, and (c), (f) DCT complete datasets from lab-imaged, back-illuminated chrome on glass transparencies of a 1951 USAF resolution target (a)–(c) and a grayscale Lena portrait (d)–(f). Each sample corresponds to a basis pattern and produces a basis coefficient. The ground truths of the corresponding images are shown in Figs. 6(a) and 6(b). As opposed to Noiselet transform, highly spatially structured patterns, such as Hadamard and DCT, interact with the image, making certain coefficients stand out.

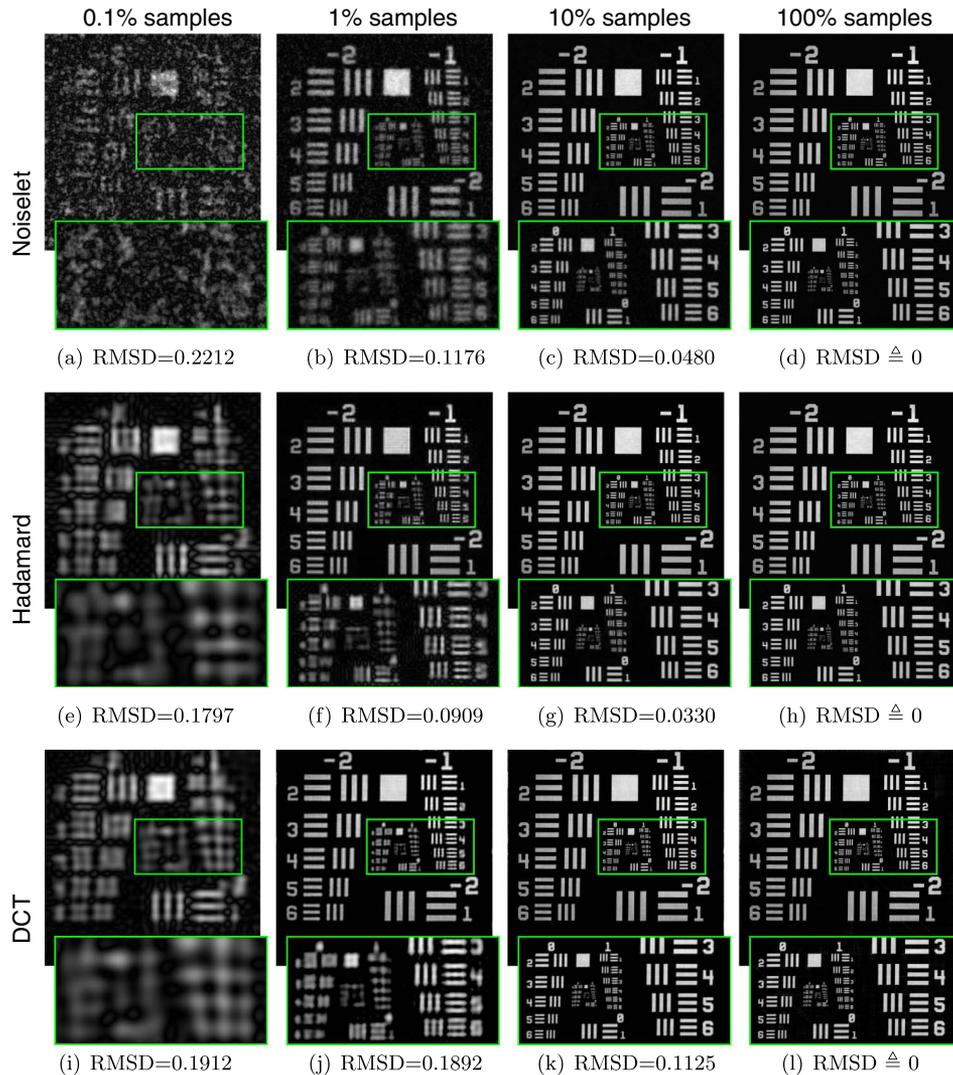


Fig. 9. (a)–(d) Noiselet, (e)–(h) Hadamard, and (i)–(l) DCT reconstructed images of a 1951 USAF resolution target using 0.1, 1, 10, 100% of samples recorded in a lab setting. Error is calculated with reference to the respective 100% reconstruction.



Fig. 10. (a)–(c) Noiselet, (d)–(f) Hadamard, and (g)–(i) DCT reconstructions of a grayscale image using 1%, 10%, 100% of samples recorded in a lab setting. Error is calculated with reference to the respective 100% reconstruction.

California San Diego engineering building. A ground truth image taken using a Canon DSLR 5D Mark II with a Canon 70–300 mm focal length zoom lens set to $f = 100$ mm, $f/9$ and with an exposure of $1/200$ s is shown in Fig. 6(c). In addition to taking natural environmental experimental data, we sought to form a multispectral image using optical filters at the photodetector to create RGB color and infrared images. The RGB color filters had center wavelengths of 450, 550, and 650 nm, each with a FWHM bandwidth of 70 nm. The infrared filter passed wavelengths longer than 700 nm. In this outdoor setting, there was 665 lux incident on the system, with 22, 227, 76, and 11 lux passing through the four respective filters. The outdoor power reading measured using the Newport power meter was 8.9 W and with the respective filters present was 484, 623, 969, and 5.19 W. A plot of the outdoor spectra (shown in Fig. 7) taken during the data acquisition was measured using an Ocean Optics Spectrometer USB

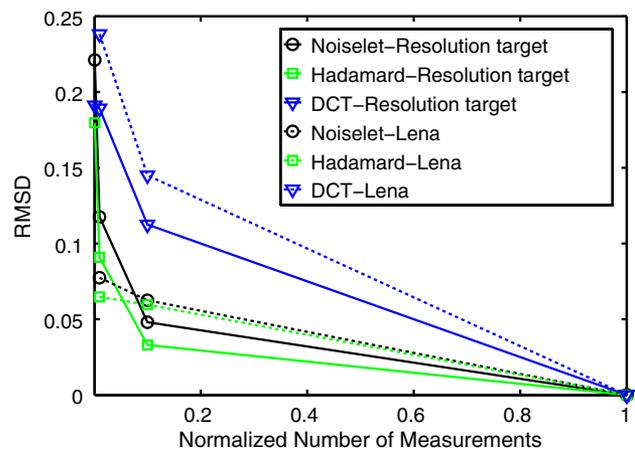


Fig. 11. Plot of the RMSD error versus the number of measurements used in the reconstruction of the two lab scenes using the three basis sets. Measurements are taken from Fig. 9 (solid line) and Fig. 10 (dashed line) and are normalized to the number of functions in the complete basis set. The Hadamard transform has the least error followed closely by the Noiselet and last the DCT.



Fig. 12. Reconstructed images with the red, green, and blue bandpass filters were combined to form outdoor color images. Reconstructions were made using 1% of the (a) Noiselet, (b) Hadamard, and (c) DCT transforms. Daylight changed during the image acquisition, resulting in variations in color and illumination.

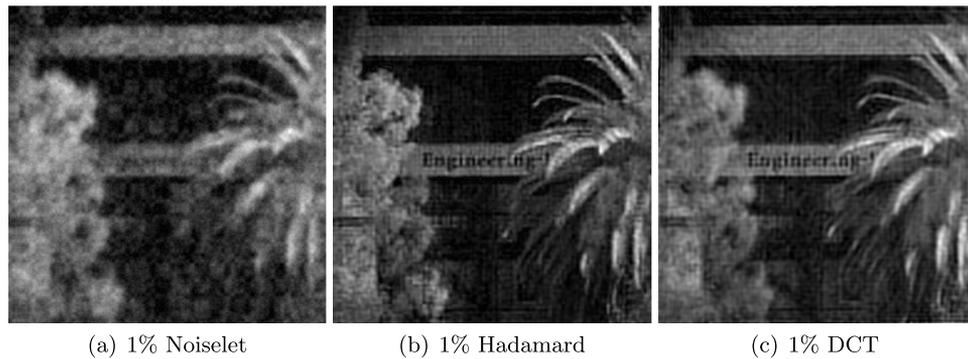


Fig. 13. Infrared reconstructed images using 1% of samples recorded outdoors using an IR bandpass filter.

4000-VIS-NIR-ES in the absence and presence of these filters. This plot shows the relative energy levels the CI system performed under during image acquisition. In this setting, stray light increased noise to nearly twice the noise floor of the photodetector. For example, with the blue filter in place, the noise floor was 4 nW while the stray light increased this value to 6.8 nW; the average value of the signal present during imaging was approximately 460 nW.

6. Experimental Results

The six complete 1,048,576 sample datasets captured for the two laboratory images are shown in Fig. 8. The data shown in Figs. 8(a) and 8(d) supports the idea that the Noiselet sampling technique is “democratic” (see Section 2), since the coefficient values are nearly uniform for all samples. When using a spatially structured pattern, such as the Hadamard [Figs. 8(b) and 8(e)] or DCT [Figs. 8(c) and 8(f)] transforms, the measured coefficient values stand out for certain spatial frequencies. This is due to interactions between the structure of the basis patterns and the image. In general, magnitudes of the coefficients are inversely proportional to spatial frequency. This is supported by both datasets but more so by the data from the resolution target than from the Lena portrait. This can be expected since the statistical model cited in Section 2 acts upon the ensemble of natural images and not the few

specific images with which we tested the functionality of the compressive imager. The image reconstruction computation took less than 1 min in every case. The laboratory reconstructions of the resolution target and the Lena image are shown in Figs. 9 and 10, respectively. The root mean squared difference (RMSD) values were calculated between reconstructions made using some and all samples. The plot in Fig. 11 summarizes the CI system error for the two images using the three measurement transforms.

Figures 12 and 13 show reconstructions of an outdoor scene when only 1% of the samples are taken, namely 10,486 samples. It is important to note that the Noiselet and Hadamard images took about 7 min per channel, while the DCT images took about 20 min per channel due to the larger memory needed to store and transfer the 8 bit DCT patterns. The D4100 controller board has limited memory, making it necessary to load and run the patterns several times, imposing a limit on the operation of the compressive imager. Since the 1 and 8 bit RGB color images shown took about 21 and 60 min to be captured, respectively, changes in the daylight scene became significant. Additionally, since different color filters were used at different times, there are color artifacts in these regions of changing illumination.

Figure 14 shows the comparison of the compressive images made using 1% of the total samples (10,486



Fig. 14. Comparison of the compressive imager with a conventional camera. (a) Images taken by a Canon 5D Mark II where the pixel data has been binned to 1% of its original 1.04 MPix size, namely 10,404 pixels. (b) Shows the images in (a) upsampled to the same resolution (1.04 MPix) as the compressive imager. (c) Shows the CI result using 1% (10,404 samples) of the lowest spatial frequency Hadamard data.

samples) with Canon 5D images where the pixel data has been binned to 1% of the pixel count (10,486 pixels). Images in Fig. 14(a) are compressed to 1% of their original size by binning the pixels. Images in Fig. 14(b) are upsampled versions of those in Fig. 14(a) to show a less pixelated result for comparison. Images in Fig. 14(c) show the resulting compressive images made using the 10,486 lowest spatial frequency samples measured with the Hadamard transform.

7. Discussion and Conclusion

In general, the CI system images acquired with different basis set representations conformed to expectations. Images taken with the Hadamard

and DCT transforms increase in resolution as more samples acquired using higher spatial frequency patterns are included. Images acquired with the Noiselet basis set, which targets all spatial frequencies equally, exhibit a different behavior. With the Noiselet basis, higher spatial frequency information is present but more samples are needed to converge to a visually appealing solution. This is apparent from the 1951 USAF resolution target (Fig. 9), as well as from the grayscale “Lena” portrait (Fig. 10). The DCT reconstruction deals with contour edges less effectively than the Hadamard reconstruction due in part [see Fig. 10(g)], to the DMD’s TDM encoding. The system performance is reflected in the RMSD error; however, the error values do not reveal

subtle differences in the reconstruction quality. The outdoor color images (Fig. 12) and outdoor infrared images (Fig. 13) are similar in that the Hadamard transforms perform the best overall, the Noiselet reconstructions are data starved for low sample counts, and the DCT reconstructions exhibit artifacts at the image edges and color artifacts due to the longer exposure times. For this reason, we concluded that the Hadamard transform was the best representation for general-purpose CI, and used it in our comparison to the images acquired by a conventional (but equal resolution) focal plane imager (Fig. 14). The comparison shows clearly that, as the theory predicts, the CI images with 1% of full sampling did obtain substantially identical resolution when compared to a focal plane with 1% of the original 1.04 MPix image resolution. Given the potential SLM and detector hardware limitations, this was not a foregone conclusion. The 20 dB dynamic range of the SLM clearly limited the contrast for the outdoor images [comparing outdoor images in Figs. 14(b) and 14(c)]. In general, however, the results confirm that a general-purpose CI system can use a standard basis representation to acquire images with a wide range of scenes and lighting conditions. As expected, the CI imager's exposure requirements make it impractical when compared to a conventional imager; however, here we used it as a test bed to quantify and validate performance of this emerging technology.

There are several hardware configurations that could improve performance of the CI system. The most obvious is to sample both paths of light reflected from the DMD by replacing the light sink with a second photodetector (see Fig. 1). We successfully reconstructed images using this configuration in lab experiments but crosstalk between the transform and the surrounding mirrors of the DMD worsened performance since the DMD resolution did not match that of our transforms. This implementation doubles the acquisition speed by capturing the positive and negative portions of each pattern simultaneously. Color artifacts in the color images could be mitigated using a rapidly switching color filter to acquire spectral data for each pattern sequentially, rather than taking three separate images. Increasing DRAM memory in the D4100 controller board, such that all the patterns could be loaded prior to image capture, would significantly speed up data acquisition to a 1% sampling runtime of 1.4 min or less. Even faster acquisition could be achieved, and dynamic range improved, by multiplexing parts of the spatially filtered signals onto multiple parallel detectors, as proposed by Ke *et al.* [22], providing a continuum between conventional focal planes and compressive imagers. In the longer term, ASIC image sensors with high pixel counts and programmable on-chip signal aggregation will be able to integrate the pattern encoding directly into the image sensor itself, eliminating the need for an external SLM and enabling CI systems to function with high-performance image formation optics [17,18]. Finally, the most significant performance

improvement in CI will be enabled by making use of the intrinsic architectural flexibility for feature-specific imaging [23,24], including face recognition [25,26], to dramatically decrease the basis set size required to acquire a conclusive measurement.

This work was supported by the Charles Stark Draper Laboratory, University R & D program contract SC001-00000628A.

References

1. D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).
2. E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory* **52**, 489–509 (2006).
3. E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Trans. Inf. Theory* **52**, 5406–5425 (2006).
4. D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, "A new compressive imaging camera architecture using optical-domain compression," *Proc. SPIE* **6065**, 606509 (2006).
5. M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "An architecture for compressive imaging," in *2006 IEEE International Conference on Image Processing (IEEE, 2006)* pp. 1273–1276.
6. M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.* **25**(2), 83–91 (2008).
7. L. McMackin, M. A. Herman, B. Chatterjee, and M. Weldon, "A high-resolution SWIR camera via compressed sensing," *Proc. SPIE* **8353**, 835303 (2012).
8. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. (Prentice-Hall, 2006).
9. D. L. Ruderman, "Origins of scaling in natural images," *Vis. Res.* **37**, 3385–3398 (1997).
10. R. P. Millane, S. Alzaidi, and W. H. Hsiao, "Scaling and power spectra of natural images," in *Proceedings of Image and Vision Computing New Zealand*, D. G. Bailey, ed. (Massey University, 2003), pp. 148–153.
11. W. H. Hsiao and R. P. Millane, "Effects of occlusion, edges, and scaling on the power spectra of natural images," *J. Opt. Soc. Am. A* **22**, 1789–1797 (2005).
12. G. Frenkel, E. Katzav, M. Schwartz, and N. Sochen, "Distribution of anomalous exponents of natural images," *Phys. Rev. Lett.* **97**, 103902 (2006).
13. R. G. Baraniuk, "Compressive sensing [Lecture Notes]," *IEEE Signal Process. Mag.* **24**, 118–121 (2007).
14. S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University, 2004).
15. E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.* **25**, 21–30 (2008).
16. M. A. Neifeld and J. Ke, "Optical architectures for compressive imaging," *Appl. Opt.* **46**, 5293–5303 (2007).
17. R. Robucci, L. K. Chiu, J. Gray, J. Romberg, P. Hasler, and D. Anderson, "Compressive sensing on a CMOS separable transform image sensor," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2008 (IEEE, 2008)*, pp. 5125–5128.
18. Y. Oike and A. E. Gamal, "A 256 × 256 CMOS image sensor with $\Delta\Sigma$ -based single-shot compressed sensing," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC) (IEEE, 2012)*, pp. 386–388.
19. L. J. Hornbeck, "Digital light processing: a new MEMS-based display technology," in *Technical Digest of the Institute of Electrical Engineers of Japan 14th Sensor Symposium (IEEJ, Kawasaki, Japan, 1996)*, pp. 297–304.
20. S. Becker, J. Bobin, and E. Candès, "NESTA: a fast and accurate first-order method for sparse recovery," *SIAM J. Imaging Sci.* **4**, 1–39 (2011).

21. D. Doherty and G. Hewlett, "10.4: phased reset timing for improved digital micromirror device (DMD) brightness," Soc. Inf. Disp. Symp. Dig. Tech. Pap. **29**, 125–128 (1998).
22. J. Ke, P. Shankar, and M. A. Neifeld, "Distributed imaging using an array of compressive cameras," Opt. Commun. **282**, 185–197 (2009).
23. M. A. Neifeld and P. Shankar, "Feature-specific imaging," Appl. Opt. **42**, 3379–3389 (2003).
24. P. K. Baheti and M. A. Neifeld, "Feature-specific structured imaging," Appl. Opt. **45**, 7382–7391 (2006).
25. H. S. Pal, D. Ganotra, and M. A. Neifeld, "Face recognition by using feature-specific imaging," Appl. Opt. **44**, 3784–3794 (2005).
26. P. K. Baheti and M. A. Neifeld, "Recognition using information-optimal adaptive feature-specific imaging," J. Opt. Soc. Am. A **26**, 1055–1070 (2009).