# Compressive video sensors using multichannel imagers

Mohan Shankar,[1,2] Nikos P. Pitsianis,[1,2,3] and David J. Brady[1,2,*]

[1]Fitzpatrick Institute for Photonics, Duke University, Durham, North Carolina 27708, USA

[2]Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708, USA

[3]Department of Electrical and Computer Engineering, Aristotle University, Thessaloniki 54124, Greece

*Corresponding author: dbrady@duke.edu

Received 1 October 2009; accepted 7 November 2009;
posted 7 January 2010 (Doc. ID 117940); published 3 February 2010

We explore the possibilities of obtaining compression in video through modified sampling strategies using multichannel imaging systems. The redundancies in video streams are exploited through compressive sampling schemes to achieve low power and low complexity video sensors. The sampling strategies as well as the associated reconstruction algorithms are discussed. These compressive sampling schemes could be implemented in the focal plane readout hardware resulting in drastic reduction in data bandwidth and computational complexity. © 2010 Optical Society of America

*OCIS codes:* 110.1758, 100.6640.

## 1. Introduction

The redundancies that occur in sound, image, and video content offer an opportunity to reduce the quantity of data required to represent that content, without any perceivable reduction in the quality of the representation. A single image is usually partitioned into subimages that can be adequately represented sparsely in a compressive basis such as Fourier, discrete cosine transform, and wavelets, to name a few. Thus, one would only need to handle the sparse representation of the images rather than the entire image. Moreover, the sequence of successive frames in a video stream can contain redundant information from a common background scene and objects that change only partially, if at all, and are simply translated or rotated. This redundancy could potentially lead to an even bigger reduction in the data load. This reduction of data volume is known as image or video compression. The Moving Picture Experts Group (MPEG) family of encoding and decoding standards [1] for the compression of sound, image, and video data has been used to utilize storage size and transmission bandwidth efficiently. Generally, these approaches apply the compression algorithms after the video frames have been acquired and stored.

We explore the possibilities of implementing compression in the data acquisition process. The difference between this and a conventional data compression technique is that, instead of capturing all the data and then processing it, intelligence is added to the image-capture system to reduce the data bandwidth. The captured data are then processed to obtain estimates of the full-resolution data from the compressed measurements. Compressive sensing is a technique that allows the reconstruction of a compressible signal with a small number of projections (measurements) of the object being sensed. This concept has been used to develop a single-pixel camera for both still and video capture [2], a compressive imaging sensor [3,4] as well as in compressive sampling strategies for spectroscopic applications [5–7]. Data compression could also be obtained by using well-known superresolution reconstruction techniques to obtain high-resolution images from a set of low-resolution images. The multichannel thin

0003-6935/10/1000B9-09$15.00/0

observation module by bound optics (TOMBO) camera is an example of a system with which a set of low-resolution images is obtained simultaneously and used to reconstruct a single high-resolution image [8]. Considerable research has also been carried out to obtain high-resolution images from low-resolution video sequences [9–11]. In general, these approaches are computationally intensive and have large memory requirements since they use a sliding time window to accumulate low-resolution video frames to reconstruct high-resolution image frames.

In most sensing and surveillance applications, periods of inactivity exist for the majority of the time that the system is gathering data. This constitutes a large data redundancy since there are few or no changes in the scene or object. Recognizing this redundancy, it would be highly inefficient to maintain a high data bandwidth during periods of inactivity. If the surveillance system could make use of this knowledge and reduce the bandwidth accordingly, huge power savings could be obtained. Here we explore the opportunities and challenges in the design of compressive video sensors and corresponding algorithms. In particular, we explore the possibilities of obtaining compression of video using multichannel (TOMBO) cameras that have been described previously [12–14]. A mathematical description of a multichannel imaging system is given below to help understand the process of image capture from these cameras.

### A. Multichannel Imaging Model

At each time instant, a multichannel camera generates a set of low-resolution images on the focal plane, each with a slightly different perspective of the scene due to the relative shifts of the positions of the lenslets in each subaperture [8,12] (illustrated in Fig. 1). Each lenslet implements a unique subpixel shift to create a nonredundant view of the scene. A high-resolution estimate of the scene could be obtained from these low-resolution measurements. A measurement made with a multichannel camera with $M$ subapertures at time instant $t$ can be modeled as [12]

$$x_k^{(t)} = H_k \mathbf{f}^{(t)} + n_k, \qquad k = 1, 2, \ldots, M, \qquad (1)$$

where $H_k$ is the system matrix representing the transformation from scene $\mathbf{f}^{(t)}$ to measurement $x_k^{(t)}$ in the $k$th channel and $n_k$ is the noise present in the channel. The system matrix $H_k$ applies the shift, downsampling, and blurring operation corresponding to subaperture $k$. By concatenating these $M$ measurements into a single array $\mathbf{x}^{(t)}$, the noise observations into array $\mathbf{n}$ and vertically concatenating the system matrices into a single system matrix $\mathbf{H}$, we have

$$\mathbf{x}^{(t)} = \mathbf{H}\mathbf{f}^{(t)} + \mathbf{n}.$$

With multichannel cameras, data compression could be obtained either by selecting a small block of pixels (fewer subapertures) in each time instant (spatial compression), increasing subaperture exposure time (temporal compression), or a combination of both (spatiotemporal compression [15]). Here we describe spatial compression and spatiotemporal compression strategies and demonstrate the feasibility of these schemes through the use of visible-band and infrared multichannel cameras.

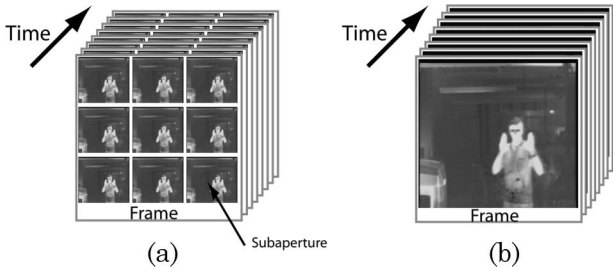



Fig. 1. Video frames from (a) a multichannel camera in which each frame consists of multiple nonredundant low-resolution representations of the scene and (b) a conventional camera.

## 2. Spatial Compression

With this scheme we consider using only a subset of the available subapertures in each time instant as shown in Fig. 2. The number of subapertures selected at each time instant could be adaptively chosen based on user-defined metrics of the scene, for example, amount of activity in the scene. The penalty of using fewer subaperture images is loss in the spatial resolution in the reconstructed video frame. However, the system could be set up in such a way that the data rate could be low when there is little or no activity in the scene. The data rate could be progressively increased based on the amount of activity or the amount of detail that is needed to be resolved in the scene.

Mathematically, the measurements obtained using this sampling strategy can be described based on Eq. (1) by

$$x_k^{(t)} = H_k \mathbf{f}^{(t)} + n_k \qquad k = 1, 2, \ldots, P,$$

where $P$ is the number of subapertures ($P \le M$) that are considered in a particular time instant. The objective is to reconstruct the full high-resolution estimate of the scene at each time instant.

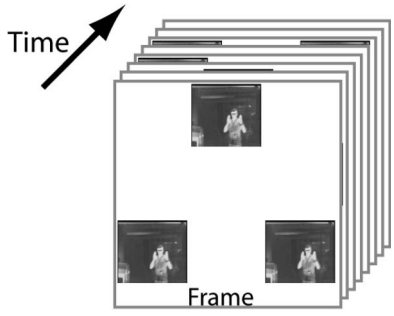



Fig. 2. Spatial-compressive measurements using a multichannel camera. A small subset of subapertures was selected in each time frame, thereby decreasing the data load.

To demonstrate spatial compression for video, a multichannel infrared camera was used (shown in Fig. 3). The quality of the reconstructed video was compared with that obtained from a conventional single-aperture infrared camera. The cameras employ microbolometer detector arrays with a 25 $\mu$m pixel pitch and operate in the far-infrared wavelength region (8–12 $\mu$m). A 3 × 3 lenslet array is positioned in front of the focal plane array to generate nine nonredundant low-resolution images of the scene. These low-resolution images are used to reconstruct a high-resolution estimate as described previously [12].

### A. Video Reconstruction

The goal of the reconstruction process is to be able to obtain full-resolution estimates of the scene at each time instant from the low-resolution measurements. The reconstruction of the full-resolution video from the subsampled video frames was approached with two algorithms: a linear least-gradient technique and an iterative Richardson–Lucy (RL) deconvolution technique. Both reconstruction techniques require accurate knowledge of the subpixel shifts in each of the subapertures to obtain a good estimate of the scene. Slight errors in the estimates of the shifts result in a poor reconstruction of the scene. The subpixel shifts are estimated during calibration experiments. If the optical system is mounted robustly, this calibration would need to be performed only once.

**Least-Gradient Reconstruction**: The least-gradient algorithm was used to reconstruct a single high-resolution frame from the nine low-resolution images from the multichannel system, as described previously [12]. The algorithm was used in reconstruction of the video sequence by applying the algorithm to each frame in the video sequence. The technique solves the image reconstruction problem for the image that varies the least in the gradient, given by

$$\mathbf{f^{(t)}}_{\mathrm{LG}} = \arg\min_{\mathbf{f}^{(t)}} \gamma(\mathbf{f}^{(t)}) = \|\nabla \mathbf{f}^{(t)}\|_2 \quad s.t. \quad \mathbf{H}\mathbf{f}^{(\mathbf{t})} = \mathbf{x}^{(\mathbf{t})} \tag{2}$$

where $\nabla$ denotes the discrete gradient operator. The solution to Eq. (2) is obtained in two steps. First, we obtain solution $\mathbf{f}_p^{(t)}$ to the linear equation $\mathbf{H}\mathbf{f}^{(\mathbf{t})} = \mathbf{x}^{(\mathbf{t})}$. The general solution to this underdetermined linear equation can then be described as $\mathbf{f}^{(t)} = \mathbf{f}_p^{(t)} - \mathbf{N}\mathbf{c}$, where $\mathbf{N}$ spans the null space of $\mathbf{H}$, and $\mathbf{c}$ is an arbitrary coefficient vector. The problem defined by Eq. (2) reduces to an unconstrained linear least-squares problem:

$$f_{\mathrm{LG}}^{(t)} = \arg\min_{\mathbf{c}} \|\nabla(\mathbf{N}\mathbf{c} - \mathbf{f}_p^{(t)})\|_2.$$

Assuming that the $\nabla\mathbf{N}$ is of full column rank, the solution to Eq. (2) can be expressed as follows:

$$\mathbf{f}_{\mathrm{LG}}^{(t)} = \mathbf{f}_p^{(t)} - \mathbf{N}(\mathbf{N}^T\nabla^T\nabla\mathbf{N})^{-1}(\nabla\mathbf{N})^T\nabla\mathbf{f}_p^{(t)}, \tag{3}$$

**Richardson–Lucy Deconvolution**: The RL deconvolution algorithm [16,17] uses the forward as well as the backward model of the system to obtain estimates of the scene iteratively based on the measurements. If $\mathbf{x}^{(\mathbf{t})}$ represents the measurement at time instant $t$, $\hat{\mathbf{f}}_k^{(t)}$ represents the current estimate; the next estimate $\hat{\mathbf{f}}_{k+1}^{(t)}$ is obtained by using the following RL deconvolution step:

$$\hat{\mathbf{f}}_{k+1}^{(t)} = \hat{\mathbf{f}}_k^{(t)} \odot \mathbf{H}^T\left(\frac{\mathbf{x}^{(\mathbf{t})}}{\mathbf{H}(\hat{\mathbf{f}}_k^{(t)})}\right),$$

where $\mathbf{H}^T$ and $\mathbf{H}$ represent the backward and forward models of the system, respectively, and the operator $\odot$ represents an elementwise multiplication. The backward model represents the process of obtaining a high-resolution estimate from the low-resolution image that includes performing image restoration (upsampling) as well as shift operation (corresponding to the subapertures chosen). The forward model represents the process of generating a measurement from the current estimate of the scene, which involves shifting, downsampling, and blurring the scene estimate depending on the subapertures being considered.

### B. Experiments and Results

The performance of the compression scheme was observed by considering one, four, and all nine subaperture images from the multichannel camera at every time instant and comparing the quality of the reconstructions. The video from the camera was captured in its native form (full frames), the sampling strategy and the reconstruction were implemented on each frame. The reconstruction algorithms were applied to every subaperture and then combined to obtain the overall scene estimate.

Video was recorded of a person rolling a chair from a distance toward the camera. On the chair, a set of Kapton heater strips was mounted on a support

Fig. 3. Multichannel infrared camera (left) used to demonstrate compressive measurements for video. It contains a 3 × 3 lenslet array to create nine nonredundant low-resolution images of the scene. Video from a conventional single-aperture infrared camera (right) is used to compare the quality of the video reconstructions.

structure. Single frames from the reconstructed video when considering one, four, and all nine subaperture images in every time instant are shown in Figs. 4(a)–4(c) (Media 1), respectively. The corresponding frame from the conventional infrared camera is shown in Fig. 4(d) (Media 1). Figure 5 shows a zoom into that part of the image that contains the Kapton heater from all four images. Figure 6 shows plots across a cross section of the zoomed-in images. There is an improvement in the resolution by considering four and nine subaperture images rather than considering just one. The regions in which the reconstructed image fails to resolve details in the scene are indicated by circles in Fig. 6.

Reconstructed frames from the video sequence using the RL deconvolution algorithm while considering one, four, and all nine subapertures are shown in Figs. 7(a)–7(c), (Media 2), respectively. The corresponding frame obtained from the conventional camera is shown in Fig. 7(d) (Media 2). Figure 8 shows the corresponding zoomed-in versions of the images and Fig. 9 shows plots across one row of the zoomed-in images. The regions in which the reconstructed image is unable to resolve details in the scene are marked by circles.

The performance of the sampling strategy was also observed by use of an infrared bar target. By use of an infrared collimator (Santa Barbara Infrared Model STC-630) and target projector (Model 14001), a bar target at a certain temperature was projected. The camera was mounted on a rotation stage at the aperture of the collimator and rotated at a certain fixed rate, capturing frames of data as the camera rotated. The results of applying the sampling strategy using the least-gradient and the RL reconstruction techniques are shown in Figs. 10 and 12, respectively. The corresponding plots across one row in the images are shown in Figs. 11 and 13, respectively.

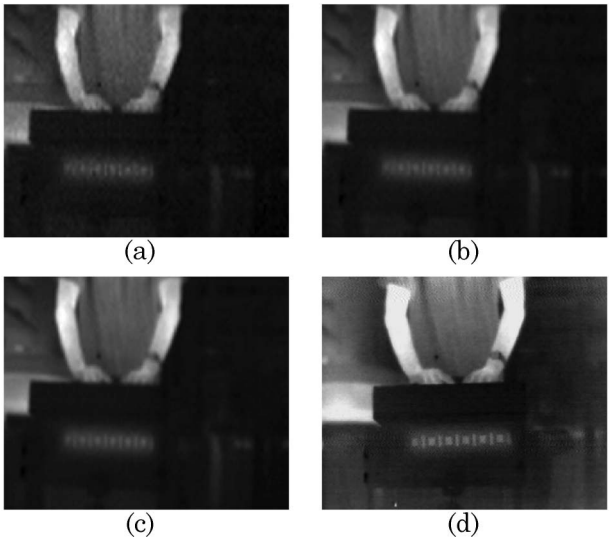



Fig. 4. (Media 1) Comparison of corresponding frames from the least-gradient video reconstruction at a certain time instant when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction, and (d) the corresponding frame obtained with a conventional infrared camera.

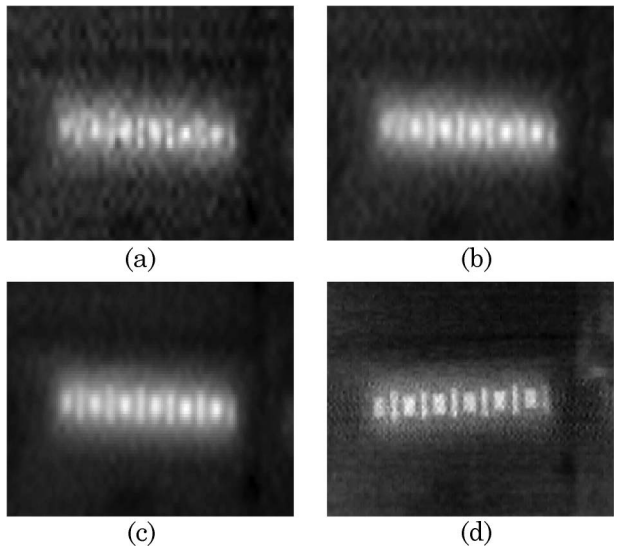



Fig. 5. Zoom into regions of the least-gradient reconstructed images shown in Fig. 4 when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction, and (d) the conventional camera image.

The results show a clear improvement in the spatial resolution when more subapertures are considered. This confirms intuition since the quality of the reconstruction must improve as the number of measurements increases. The improvement by taking all nine subaperture images over taking four is more evident in the reduction of noise in the reconstructed image. For objects that are at close range to the camera, data from fewer subapertures could be

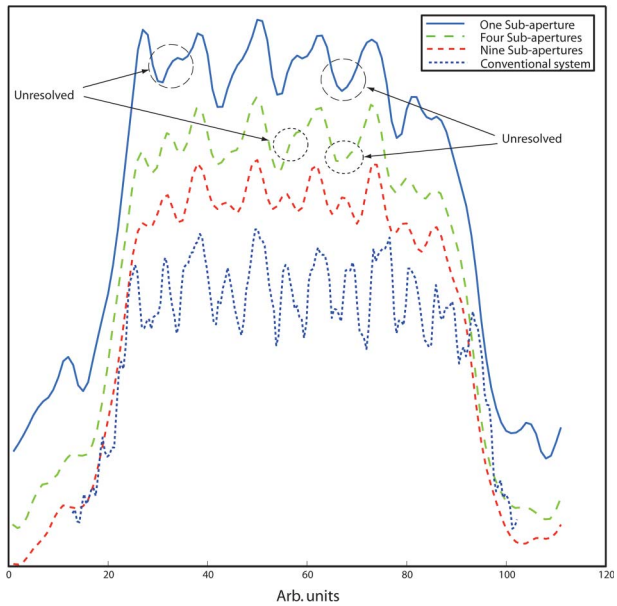



Fig. 6. (Color online) Plots across the cross section of a row in the image from the least-gradient reconstructed images shown in Fig. 4. The circles indicate regions in which the image reconstructed by using fewer subapertures is unable to resolve details that are resolved when more subapertures are considered.

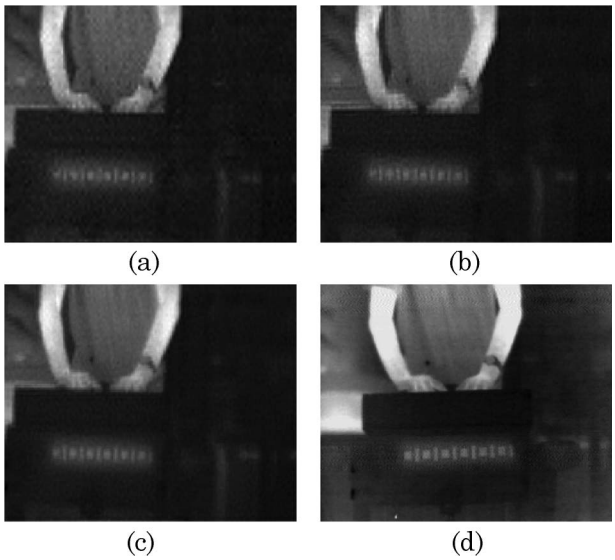



Fig. 7. (Media 2) Comparison of corresponding frames from the RL video reconstruction at a certain time instant when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction, and (d) the corresponding frame obtained with a conventional infrared camera.

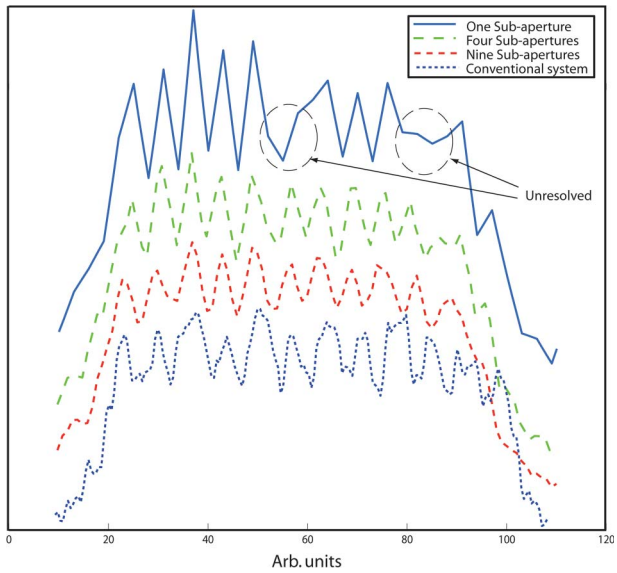



Fig. 9. (Color online) Plots across the cross section of a row in the image from the RL reconstructed images shown in Fig. 7. The circles indicate regions in which the image reconstructed by using fewer subapertures is unable to resolve details that are resolved when more subapertures are considered.

sufficient to be able to obtain a reasonable representation of the scene. However, at much larger distances, the improvement in resolution by considering more (or all) subaperture images could be critical to tasks such as object recognition or identification. A significant reduction in data bandwidth (and an increase in frame rate) could be obtained by the use of fewer subapertures instead of reading out information from the entire focal plane. This, however, would result in a lower quality reconstruction that must be tolerated by the specific application using this camera.

## 3. Spatiotemporal Compression

In this scheme, the availability of multiple subapertures to perform temporal aliasing at different time instants is exploited. This can be implemented by changing the exposure time and the time of readout corresponding to the different subapertures uniquely. A particular sampling scheme that was considered is illustrated for a four-channel ($2 \times 2$) multichannel system in Fig. 14. As indicated in Fig. 14, the modified sampling scheme involves increasing the exposure time from each subaperture

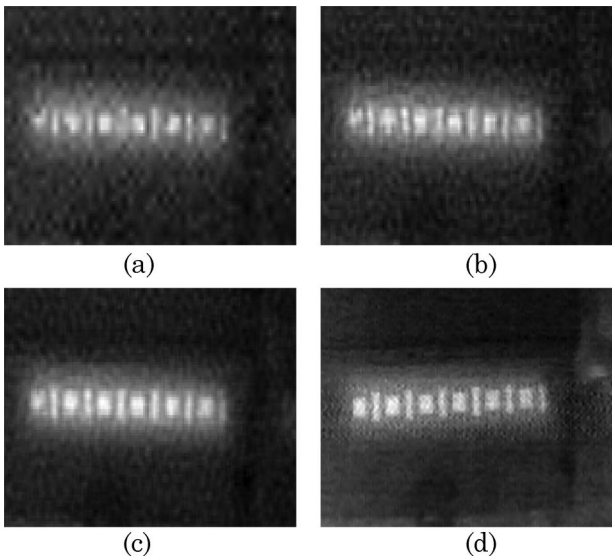



Fig. 8. Zoom into regions of the RL reconstructed images shown in Fig. 7 when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction, and (d) a conventional camera image.

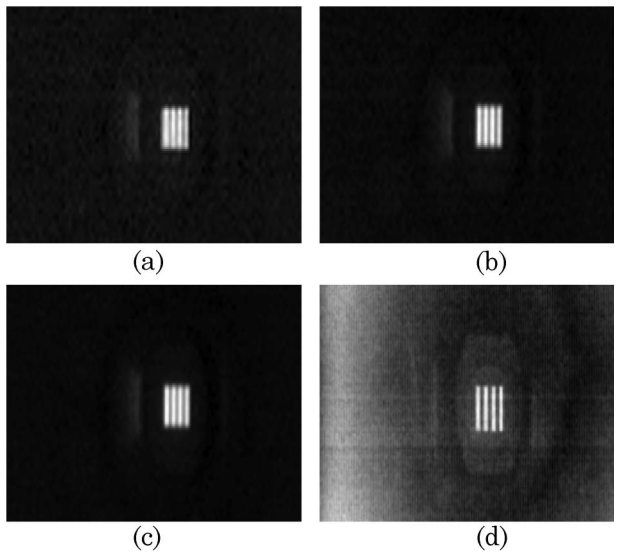



Fig. 10. Comparison of corresponding frames from the least-gradient video reconstruction at a certain time instant using a bar target when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction, and (d) a conventional camera image.

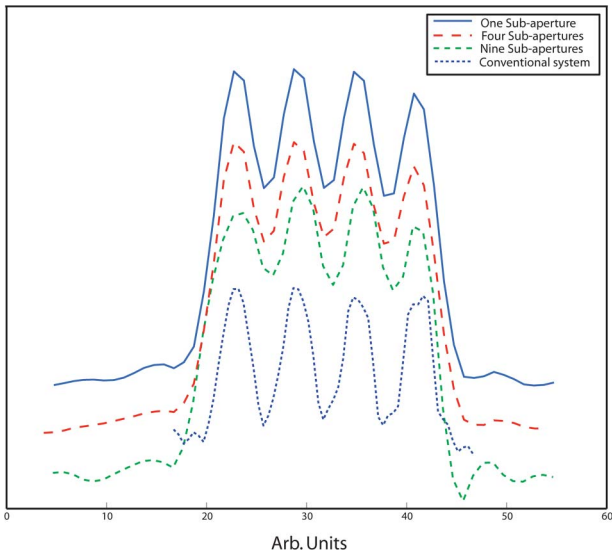



Fig. 11. (Color online) Plots across the cross section of a row in the least-gradient reconstructed bar-target images shown in Fig. 10.

(by the same amount) and initiating the sampling for each at different time instants. At time instant $t_1$, the frame from subaperture 1 is captured and the exposure time is set to four video frames (four subsequent frames are captured). At time instant $t_2$, the frame from subaperture 2 is captured, with the exposure time again set to four video frames, at time instant $t_3$, collect from subaperture 3, and so on. In doing so, the image from subaperture 1 at time instant $t_4$, the one from subaperture 2 at time instant $t_5$ and so on are obtained. At time instant $t_7$, all four images (which represent sums over different time windows from different subapertures) are recorded. The objective now is to estimate the high-resolution frames for the seven time instants from the four low-resolution time-averaged sums. For the subsequent sums, the time window is shifted by one time frame (start from time instant $t_2$ for the second set of sums, and so on) and the process is repeated.

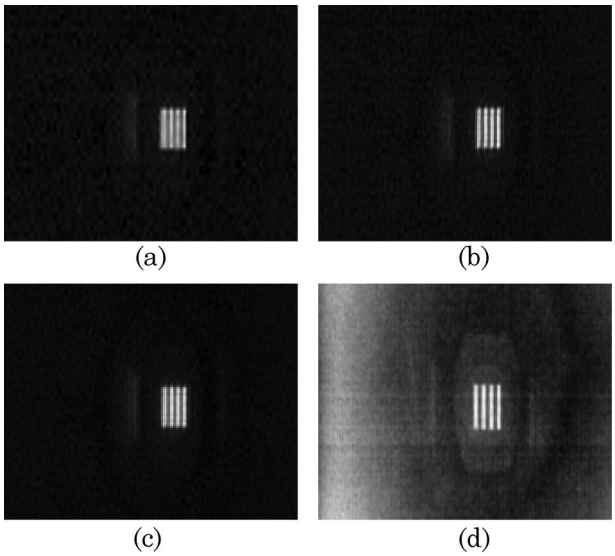



Fig. 12. Comparison of corresponding frames from the RL video reconstruction at a certain time instant using a bar target when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.

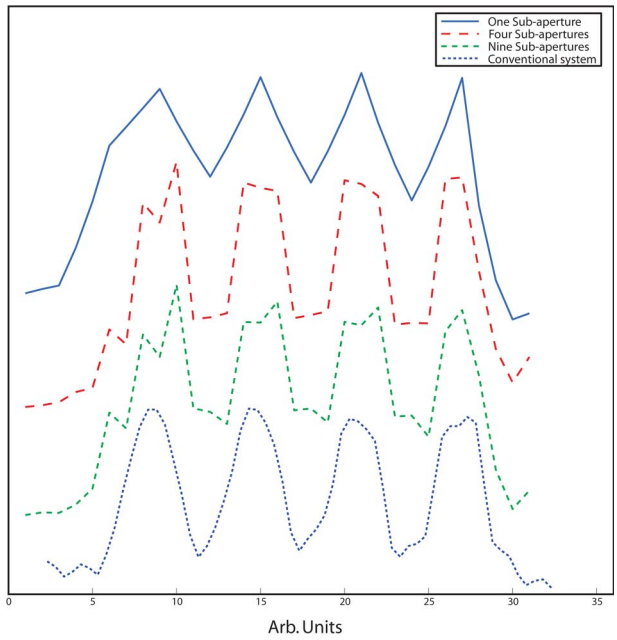



Fig. 13. (Color online) Plots across the cross section of a row in the RL reconstructed bar-target images shown in Fig. 12.

### A. Video Reconstruction

The system model corresponding to this sampling strategy can be described based on the notation used in Eq. (1) and Fig. 15 as

$$S_{1q} = \sum_{p=q-6}^{q-3} x_{1p}, \tag{4}$$

$$S_{2q} = \sum_{p=q-5}^{q-2} x_{2p}, \tag{5}$$

$$S_{3q} = \sum_{p=q-4}^{q-1} x_{3p}, \tag{6}$$

$$S_{4q} = \sum_{p=q-3}^{q} x_{4p}, \tag{7}$$

which is valid for $q \geq 7$, where $q$ represents the frame number (time instant) in the time-compressed data stream; $p$ represents the frame number (time instant) in the original data stream; $x_{1p}$, $x_{2p}$, $x_{3p}$, and $x_{4p}$ represent the four subaperture images in frame $p$ of the original data stream; and $S_{1q}$, $S_{2q}$, $S_{3q}$, and

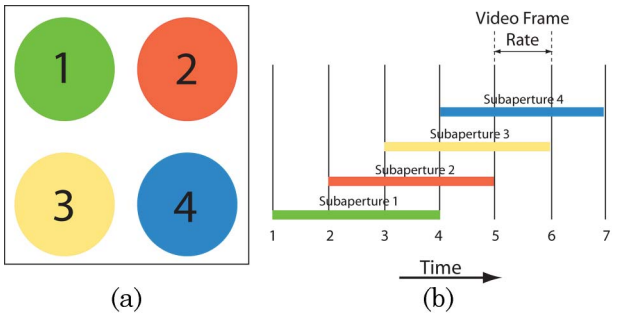



Fig. 14. (Color online) Timing diagram illustrating the spatiotemporal sampling scheme using a multichannel camera. The circles in (a) and their associated colors indicate different subapertures. In (b), the vertical lines indicate sampling instants for conventional video and the color strips represent the sampling instants for the various subapertures with the modified sampling scheme.

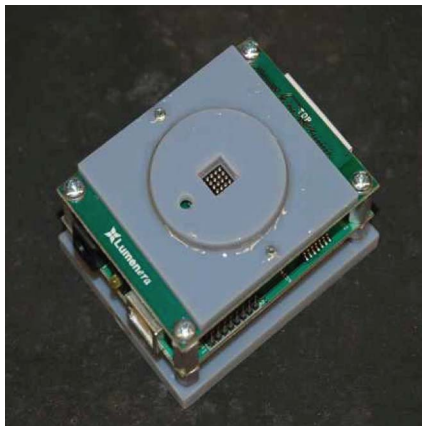


Fig. 16. (Color online) Visible-band multiaperture camera used to demonstrate spatiotemporal compressive sampling for video. The 4 × 4 lenslet array replaces the conventional single-aperture optics.

$S_{4q}$ represent the four subaperture images in frame $q$ of the time-compressed data stream.

As an example, in frame 7 ($q = 7$), $S_{17}$ contains the temporal sum of frames from subaperture 1 ($x_{1p}$) from time instants $p = 1, 2, 3, 4$; $S_{27}$ contains the temporal sum of frames from subaperture 2 ($x_{2p}$) from time instants $p = 2, 3, 4, 5$; $S_{37}$ contains the temporal sum from subaperture 3 ($x_{3p}$) from time instants $p = 3, 4, 5, 6$; and $S_{47}$ contains the sum from subaperture 4 ($x_{4p}$) from time instants $p = 4, 5, 6, 7$. Similarly, $S_{18}$ contains the sums from time instants $p = 2, 3, 4, 5$; $S_{28}$ contains the sums from time instants $p = 3, 4, 5, 6$ and so on.

The process of obtaining the time-averaged measurements from an original high-resolution video sequence $\mathbf{F}$ involves operating on all three dimensions (two spatial dimensions and time) of a video sequence. This can be represented by

$$\mathbf{S} = \mathbf{THF}, \tag{8}$$

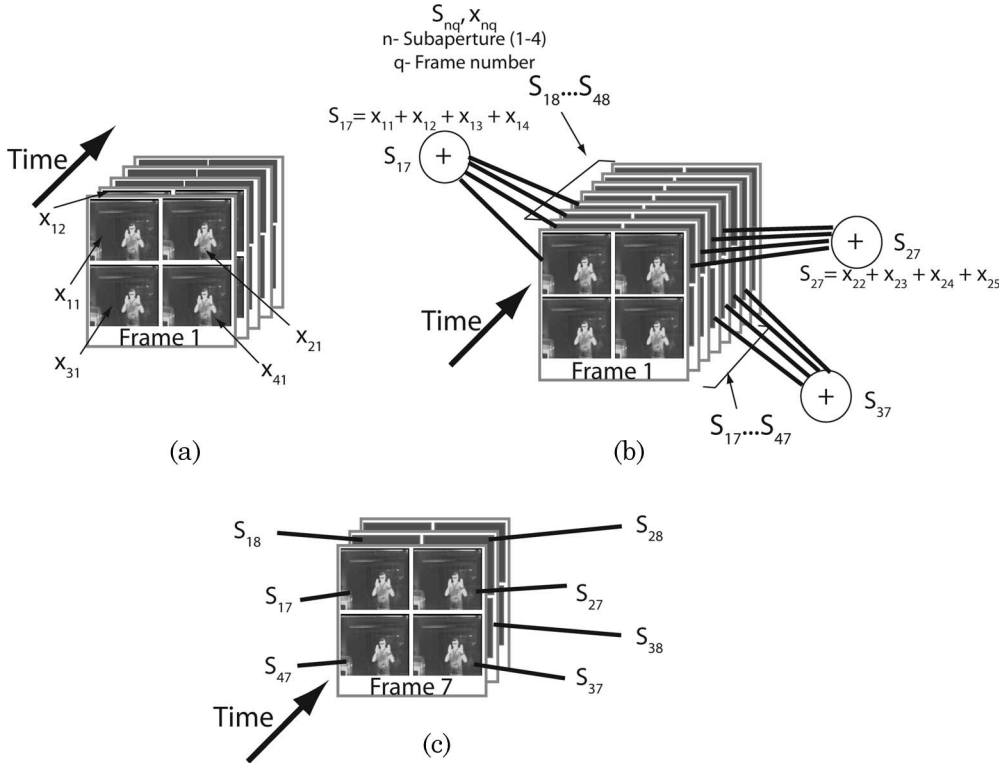
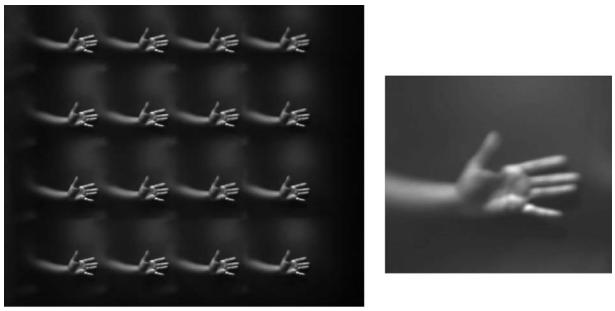



Fig. 15. Description of the sampling strategy involving spatiotemporal compression, shown for a 2 × 2 multichannel system. (a) The original video stream; (b) compression performed by considering different sums of subapertures [$S_{1q}$, $S_{2q}$, $S_{3q}$, and $S_{4q}$ as defined by Eqs. (4)–(7)] at different time instants; (c) the four sums are shown for different time instants; these are used to estimate the high-resolution frames.

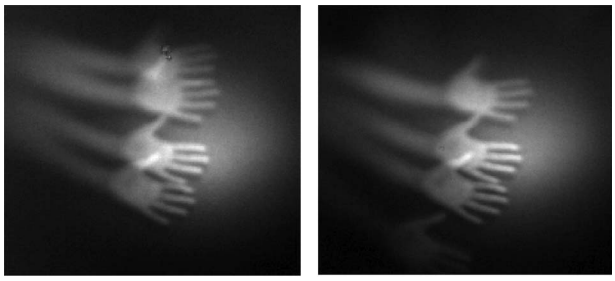


(a) (b)

Fig. 17. Images (a) obtained with the visible 4 × 4 multichannel camera and (b) reconstructed from the 16 low-resolution images in (a).

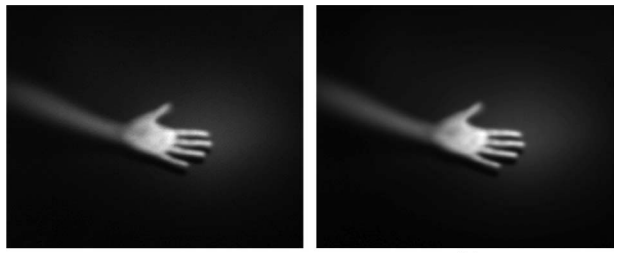
where **H** represents the system matrix corresponding to the process of shifting and downsampling high-resolution frames to obtain multichannel frames. These shift and downsampling operations are performed on the first two dimensions (rows and columns) of the video sequence to obtain the subaperture images. Operator **T** operates on the time dimension to produce the time aliased measurements at any time instant. The system matrix **H** could be large depending on the size of the images. The problem could be made simpler if the operators in **H** could be expressed separably as a Kronecker product of the operator on the rows and the operator on the columns.

The reconstruction process involves solving for the high-resolution image sequence **F** given the measurements **S**. This solution is obtained in two steps. First, the least-gradient algorithm is used to obtain the solution along the temporal dimension (modeled by **T**). This results in a nonaliased image sequence estimate that is still spatially compressed. Next, the least-gradient algorithm is applied again (using the system matrix **H**) to obtain the estimates of the high-resolution image sequence for each frame from the corresponding low resolution image sequence. Equations (4)–(7) show that the four sums are calculated on a time window that is shifted by one time frame. Therefore Eq. (8) would need to be solved along the time dimension **T** only for the first set of frames. Subsequent nonaliased frames are estimated by considering only the temporally aliased subaperture 4 [given by Eq. (7)] and then subtracting out estimates of frames that have already been estimated in previous time instants.



(a) (b)

Fig. 18. (Media 3) A temporally aliased frame at a certain time instant in the video (averaged from four previous time frames) obtained from (a) subaperture 1 and (b) subaperture 2.



(a) (b)

Fig. 19. (Media 4) (a) Reconstructed frame from the spatiotemporal compressed video and (b) frame from the multichannel camera with no temporal aliasing.

### B. Experiments and Results

To demonstrate the spatiotemporal compression technique, a visible band multichannel camera was used. The detector is a Lumenera CCD with 5.6 $\mu$m pixels. A 4 × 4 microlens array was used to obtain sixteen low-resolution images in each frame [14]. The camera, an image obtained from this camera, and the reconstructed image are shown in Figs. 16 and 17, respectively. Of the sixteen available subapertures, only four (2 × 2) were considered to demonstrate the proof of concept. Video from this camera was captured at its native (full) resolution. The sampling strategy and the reconstruction were applied on this captured video. The reconstructed video sequence was compared with that obtained from the multichannel camera with no temporal aliasing.

The experiment consisted of moving a hand in front of the camera up and down across the field of view of the camera. Two temporally aliased frames (corresponding to the subapertures 1 and 2) at a certain time instant are shown in Fig. 18 (Media 3). From these aliased video frames, the estimates for the original video sequence (at native frame rate) were obtained. A representative frame from the reconstructed video (Media 4) is shown in Fig. 19(a) and the corresponding reconstructed frame from the camera without any temporal aliasing is shown in Fig. 19(b).

This compressive sampling scheme uses temporal averaging to reduce the data bandwidth from the focal plane. The motion blur that is associated with the increased exposure times has been removed by using the reconstruction algorithm. The quality of the reconstruction depends on the quality of the estimates of the first sequence of frames since subsequent frames are obtained from previous estimates. Any artifacts that appear while estimating the first set of frames would propagate through the video sequence. Therefore, it might be necessary to apply the reconstruction periodically while acquiring the video to minimize the appearance of artifacts. The amount of compression that can be achieved depends on the computational load that the system can handle and the level of aliasing artifacts that can be

tolerated with the increase in the number of time-averaged frames.

## 4. Conclusion

We have explored the use of multichannel cameras to demonstrate compressive measurements for video. Using these cameras, two types of compressive sampling scheme, spatial and spatiotemporal compressions, were illustrated. The associated reconstruction algorithms necessary to obtain high-resolution estimates from low-resolution measurements have also been described. Implementation of these schemes require a modification to the readout mechanism of a conventional focal plane array, whether it is to select a subset of subapertures at each time instant or to increase the exposure times of subaperture frames. The results from both experiments show that significant savings could be obtained with compressive sampling. There is an associated loss in resolution that is due to subsampling and whether this could be tolerated depends on the specific application. The applications that could most benefit from using these strategies are those in which cameras would need to be powered for extended periods of time, with little or no activity in the scene, especially in video-based surveillance systems. The system could also be setup in such a way as to increase the data bandwidth when activity is detected or if high-resolution details are desired in an area being monitored. The techniques described above obtain full-resolution video stream (spatial and temporal) after reconstruction. Greater savings could be obtained if either of the resolution requirements were relaxed. By distributing the data processing load between the focal plane readout and the postreadout, the sampling strategies we described make the system more efficient and result in low-power video sensors.

This research is supported by the Microsystems Technology Office of the Defense Advanced Research Projects Agency (DARPA), contract HR0011-04-C-0111.

## References

1. Moving Picture Experts Group, "http://www.mpeg.org.."
2. D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. K. Kelly, and R. G. Baraniuk, "A new camera architecture based on optical-domain compression," Proc. SPIE **6065**, 606509 (2006).
3. N. P. Pitsianis, D. J. Brady, and X. Sun, "Sensor-layer image compression based on the quantized cosine transform," Proc. SPIE **5817**, 250–257 (2005).
4. N. P. Pitsianis, D. J. Brady, A. Portnoy, X. Sun, T. Suleski, M. A. Fiddy, M. R. Feldman, and R. D. TeKolste, "Compressive imaging sensors," Proc. SPIE **6232**, 62320A (2006).
5. D. J. Brady, N. P. Pitsianis, X. Sun, and P. Potuluri, "Compressive sampling and signal inference," U.S. patent 7,432,843 (7 October 2008).
6. D. J. Brady, N. P. Pitsianis, X. Sun, and P. Potuluri, "Compressive sampling and signal inference," U.S. patent 7,463,174 (9 December 2008).
7. D. J. Brady, N. P. Pitsianis, X. Sun, and P. Potuluri, "Compressive sampling and signal inference," U.S. patent 7,463,179 (9 December 2008).
8. J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Marimoto, N. Kondou, D. Miyazaki, and Y. Ichioka, "Thin observation module by bound optics (TOMBO): concept and experimental verification," Appl. Opt. **40**, 1806–1813(2001).
9. L. Hong, "Superresolution video reconstruction," Proc. SPIE **5022**, 631–642 (2003).
10. R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," IEEE Trans. Image Process. **5**, 996–1011 (1996).
11. A. Tekalp, M. Ozkan, and M. Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," IEEE Trans. Acoust. Speech Signal Process. **3**, 169–172 (1992).
12. M. Shankar, R. Willett, N. Pitsianis, T. Schulz, R. Gibbons, R. T. Kolste, J. Carriere, C. Chen, D. Prather, and D. Brady, "Thin infrared imaging systems through multichannel sampling," Appl. Opt. **47**, B1–B10 (2008).
13. A. Portnoy, N. Pitsianis, X. Sun, D. Brady, R. Gibbons, A. Silver, R. Te Kolste, C. Chen, T. Dillon, and D. Prather, "Design and characterization of thin multiple aperture infrared cameras," Appl. Opt. **48**, 2115–2126 (2009).
14. A. D. Portnoy, N. P. Pitsianis, X. Sun, and D. J. Brady, "Multichannel sampling schemes for optical imaging systems," Appl. Opt. **47**, B76–B85 (2008).
15. M. Shankar, N. P. Pitsianis, and D. J. Brady, "Spatio-temporal sampling for video," Proc. SPIE **7076**, 707604 (2008).
16. W. H. Richardson, "Bayesian-based iterative method of image restoration," J. Opt. Soc. Am. **62**, 55–59 (1972).
17. L. B. Lucy, "An iterative technique for the rectification of observed distributions," Astron. J. **79**, 745–754 (1974).