EXTRACTING FOURIER DESCRIPTORS FROM COMPRESSIVE MEASUREMENTS

Puyang Wang and Vishal M. Patel

Department of Electrical and Computer Engineering
Rutgers, The State University of New Jersey
94 Brett Rd, Piscataway, NJ 08854
puyang.wang@rutgers.edu, vishal.m.patel@rutgers.edu

ABSTRACT

Fourier descriptors (FDs) are shape-based features for the recognition of two-dimensional connected shapes. We propose a method that can extract FDs of an object directly from compressive measurements without reconstructing the image. Our method entails estimating the edges via discrete horizontal and vertical image gradients from compressive measurements. Fourier descriptors are then extracted from the thresholded edges. One of the main advantages of the proposed method is that it requires fewer number of compressive measurements to estimate FDs than required to estimate the original image. Various numerical experiments on synthetic and real data demonstrate the effectiveness of the proposed method.

***Index Terms*—** Fourier descriptors, compressive sampling, compressed sensing, feature extraction.

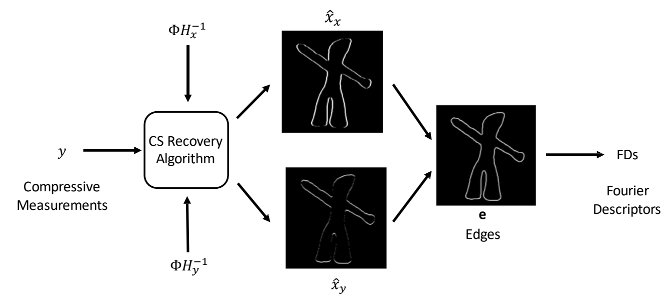


Fig. 1: An overview of the proposed method.

1. INTRODUCTION

Compressive sampling (CS) (also known as compressed sensing) is a concept in signal processing where one measures a small number of non-adaptive linear combinations of the signal. These measurements are usually much smaller than the number of samples that define the signal. From these small number of measurements, the signal is then reconstructed by a non-linear procedure [1], [2].

Since the introduction of CS in the mid 2000's, many imaging modalities have been developed that make use of CS theory for imaging. Some examples include single pixel camera [3], magnetic resonance imaging [4], synthetic aperture radar [5], ground penetrating radar [6], and millimeter wave imaging [7]. While the main objective of some of these imaging modalities is to reduce the number of samples required for imaging, in many applications, we are not interested in obtaining an exact reconstruction of the scene from compressive measurements. For instance, in the case of object recognition, one is interested in recognizing the identity of the object present in the image rather than obtaining a precise reconstruction of the object.

Many methods have been developed that explore the possibility of performing classification, detection or tracking of objects directly from the compressed measurements without reconstructing the image. In [8], a classification algorithm called *the smashed filter* was proposed that can classify objects directly from the compressive measurements. In [9], it was shown that compressive measurements can be effective for a variety of detection, classification, estimation, and filtering problems. In particular, [10] presented a technique via which background subtraction can be performed on compressive measurements of a scene. Recently [11] proposed a modification to this technique which adaptively adjusts the number of compressive measurements collected to the dynamic foreground sparsity typical to surveillance data. A more general problem regarding signal tracking using compressive observations was considered in [12]. In [13], it was shown that CS can be used to find parameterized shapes in images, by exploiting sparseness in the Hough transform domain. [14] developed a non-parametric shape estimation algorithm for inverse problems using compressed sensing.

Despite the significant advances in statistical signal processing from compressive measurements, extraction of geometric and shape-based features directly from compressed measurements has not been studied extensively in the literature. In this paper, we propose an algorithm that can extract the Fourier descriptor (FD) of an object directly from compressive measurements without reconstructing the image. The FD is a robust shape descriptor that can be used to describe the boundary of a shape in 2-dimensional space using the Fourier methods [15]. The FD is very efficient to compute and is invariant to translation, rotation and scale, which makes it a very attractive descriptor to use in many practical image retrieval and object recognition tasks [16].

Figure 1 gives an overview of the proposed method. Given compressive measurements, we first extract the horizontal and vertical differences of the desired shape by reformulating the original CS reconstruction problem as the problem of reconstructing two discrete gradients. The edge of the shape is then estimated from the recovered gradients. The standard FDs are then extracted from the recovered edges. One of the main advantages of the proposed method is that it requires fewer number of compressive measurements to estimate FDs than required to estimate the original image. This is mainly due to the fact that discrete gradients are much sparser than the original image. Furthermore, unlike some of the previously proposed CS detection and recognition methods that work only when the compressive measurements correspond to partial Fourier measurements [17], our method can work with any CS measurement matrix such as

This work was supported by an ARO grant W911NF-16-1-0126.

Gaussian, Hadamard or Fourier.

The rest of the paper is organized as follows. In Section 2, we give a brief background on CS and FD. Details of the proposed compressive FD algorithm are given in Section 3. Experimental results on synthetic and real data are given in Section 4. Finally, Section 5 concludes the paper with a brief summary and discussion.

2. BACKGROUND

In this section, we give a brief background on CS and FD.

2.1. Compressive Sensing

Let $\mathbf{x} \in \mathbb{R}^N$ be lexicographically vectorized image $X \in \mathbb{R}^{N_1 \times N_2}$, where $N = N_1 N_2$. Assume that $\mathbf{x}$ is k-sparse in a basis $\mathbf{\Psi}$ so that $\mathbf{x} = \mathbf{\Psi}\mathbf{x}_0$ with $\|\mathbf{x}_0\|_0 = k \ll N$, where $\mathbf{x}_0\|_0$ is the ℓ_0-norm that counts the number of nonzero elements in $\mathbf{x}_0$. If $\mathbf{x}$ is compressible in $\mathbf{\Psi}$, then it can be well approximated by the best k-term representation. Consider a random $M \times N$ measurement matrix $\mathbf{\Phi}$ with $M < N$ and assume that M measurements, which make up a vector $\mathbf{y}$, is derived from $\mathbf{y} = \mathbf{\Phi}\mathbf{x} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{x}_0 = \mathbf{\Theta}\mathbf{x}_0$. According to CS theory, one can reconstruct $\mathbf{x}$ via $\mathbf{x}_0$ by solving the following ℓ_1-minimization problem

$$\hat{\mathbf{x}}_0 = \arg\min_{\mathbf{x}_0'} \|\mathbf{x}_0'\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{x}_0' \tag{1}$$

provided that certain conditions (i.e. Restricted Isometry Property (RIP) or incoherence property) are met [18]. Here, $\|\mathbf{x}\|_1 = \sum_i |x_i|$ is the ℓ_1-norm of $\mathbf{x}$. Various measurement matrices exist in the CS literature that have been shown to satisfy the RIP property with high probability [19]. Some of them include random Gaussian matrix, partial Fourier transform matrix and Bernoulli matrix. In the case when there are noisy observations, $\mathbf{y} = \mathbf{\Theta}\mathbf{x}_0 + \boldsymbol{\eta}$, with $\|\boldsymbol{\eta}\| \leq \epsilon$, the following optimization problem can be solved to approximate $\mathbf{x}_0$

$$\hat{\mathbf{x}}_0 = \arg\min_{\mathbf{x}_0'} \|\mathbf{x}_0'\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{\Phi}\mathbf{\Psi}\mathbf{x}_0'\| \leq \epsilon. \tag{2}$$

2.2. Fourier Descriptors

Let $a(n)$ and $b(n)$ be the coordinates of the nth pixel on the boundary of a given 2D shape containing D pixels. A complex number $z(n)$ can be formed as $z(n) = a(n) + jb(n), \quad n = 0, \cdots, D-1$. Then, the FD of this shape is defined as the discrete Fourier transform (DFT) of z

$$Z(k) = \sum_{n=0}^{D-1} z(n) e^{-j2\pi nk/D}, \quad k = 0, \cdots, D-1. \tag{3}$$

3. PROPOSED METHOD

In order to extract the FDs from a given image, one has to first estimate the boundary of the object. This can be done by detecting the edges in the image. Hence, given the compressed measurements $\mathbf{y}$, the first step of our FD extraction algorithm involves recovering the horizontal and vertical discrete gradients of $\mathbf{x}$ from $\mathbf{y}$. Note that the horizontal and vertical gradients of X can be implemented by 2D convolutions as

$$X_x = [-0.5, 0, 0.5] * X, \quad \text{and} \quad X_y = [-0.5, 0, 0.5]^T * X, \tag{4}$$

where $*$ denotes 2D convolution and X_x and X_y denote the horizontal and vertical gradients of X, respectively. One can reformulate the above convolutions in matrix forms as

$$\mathbf{x}_x = \mathbf{H}_x \mathbf{x}, \quad \mathbf{x}_y = \mathbf{H}_y \mathbf{x}, \tag{5}$$

where $\mathbf{H}_x$ and $\mathbf{H}_y$ are the corresponding convolution matrices and $\mathbf{x}_x$ and $\mathbf{x}_x$ are lexicographically vectorized gradients X_x and X_y, respectively. Hence, one can rewrite the CS observation model in terms of the discrete gradients as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{H}_x^{-1}\mathbf{x}_x = \mathbf{\Phi}\mathbf{H}_y^{-1}\mathbf{x}_y. \tag{6}$$

As a result, we can recover the discrete gradients from $\mathbf{y}$ by solving the following two optimization problems

$$\hat{\mathbf{x}}_x = \arg\min_{\mathbf{x}_x} \|\mathbf{x}_x\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{\Phi}\mathbf{H}_x^{-1}\mathbf{x}_x \tag{7}$$

$$\hat{\mathbf{x}}_y = \arg\min_{\mathbf{x}_y} \|\mathbf{x}_y\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{\Phi}\mathbf{H}_y^{-1}\mathbf{x}_y. \tag{8}$$

Note that reconstruction of gradients from compressive measurements will depend on the conditioning of $\mathbf{\Phi}\mathbf{H}_x^{-1}$ and $\mathbf{\Phi}\mathbf{H}_y^{-1}$. It is very difficult to prove any general claim that these resulting sensing matrices satisfy a RIP or a mutual incoherence property. This remains an open problem. We will further investigate this in our future work.

Once the discrete gradients are estimated, the next step is to estimate edges from the recovered gradients. Gradient-based edge detection has been the foundation of image processing for decades and there are various edge detection operators existing in the literature such as Canny, Sobel and Prewitt [20]. One can easily adapt any of these gradient-based edge detectors to our proposed method by accordingly modifying $\mathbf{H}_x$ and $\mathbf{H}_y$. While sophisticated edge detection methods can be applied to general grayscale images for better edge detection, in the case of binary images finding edges is rather simple since the gradient magnitude is also binary. Hence, for binary images we obtain the boundary (i.e. edges) of $\mathbf{x}$ as the thresholded gradient magnitude

$$\mathbf{e} = threshold(\sqrt{\hat{\mathbf{x}}_x^2 + \hat{\mathbf{x}}_y^2}). \tag{9}$$

Gradient magnitude thresholding results in binary image gradients which essentially are the boundary of an object in the image. Let $a(n)$ and $b(n)$ be the coordinates of the nth non-zero pixel on the boundary $\mathbf{e}$ containing D pixels. Then the FD of the boundary $\mathbf{e}$ can be found as

$$Z(k) = \sum_{n=0}^{D-1} z(n) e^{-j2\pi nk/D}, \quad k = 0, \cdots, D-1, \tag{10}$$

where $z(n) = a(n) + jb(n)$. The proposed compressive FD (CSFD) extraction algorithm is summarized in Algorithm 1.

4. EXPERIMENTAL RESULTS

In this section, we present results of our proposed CSFD method on synthetic and real data. We compare the performance of our method with that of a naive method (see Figure 2) in which the image is first reconstructed by solving the following L1TV algorithm [21]

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}_0} w_1 \|\mathbf{x}_0\|_{TV} + w_2 \|\mathbf{x}_0\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{\Phi}\mathbf{x}_0, \tag{11}$$

Algorithm 1: The CSFD algorithm.

Input: $\mathbf{y}, \mathbf{H}_x^{-1}, \mathbf{H}_y^{-1}, \mathbf{\Phi}$
Estimate $\mathbf{x}_x$ and $\mathbf{x}_y$ by solving (7) and (8), respectively.
Obtain the boundary:
1) binary image: $\mathbf{e} = threshold(\sqrt{\hat{\mathbf{x}}_x^2 + \hat{\mathbf{x}}_y^2})$.
2) grayscale image: $\mathbf{e} = edge(\hat{\mathbf{x}}_x, \hat{\mathbf{x}}_y)$.
Let $z(n) = a(n) + jb(n)$, where $a(n)$ and $b(n)$ are the coordinates of the nth pixel on $\mathbf{e}$.
$Z(k) = DFT(z(n))$
Output: Fourier descriptors: $Z(k), k = 0, \cdots, D-1$.

where w_1 and w_2 are the regularization parameters. Then the FDs are extracted from the estimated image. Once the FDs are estimated, relative error is used to measure the performance of different methods. The relative error is defined as

$$\text{Relative Error} = \frac{\|Z - \hat{Z}\|_2}{\|Z\|_2}, \tag{12}$$

where Z are the magnitudes of true FDs and $\hat{Z}$ are the magnitudes of estimated FDs. It is possible to use any CS reconstruction algorithm in our approaches, however, in our experiments with the synthetic data, we employed a highly efficient algorithm that is suitable for large-scale applications known as the spectral projected gradient (SPGL1) algorithm [22][23] for solving (7) and (8). Note that due to space limitations, we only present the results comparing the quality of the extracted FDs in this paper. Object recognition using the estimated FDs will be presented elsewhere.

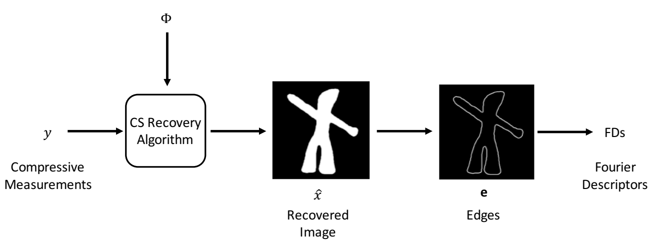
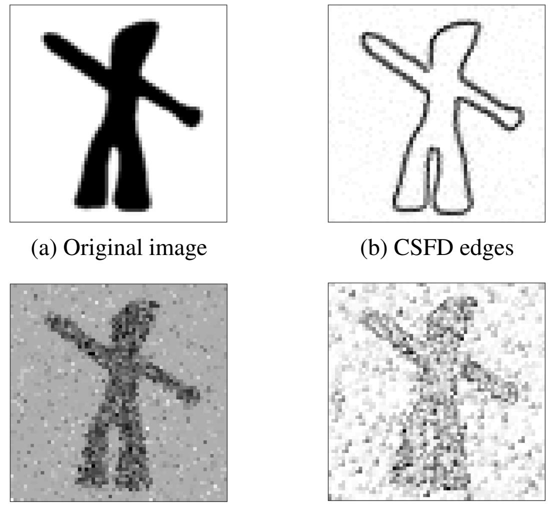


Fig. 2: A naive method for estimating FDs from compressive measurements.

4.1. Synthetic Data

In the first set of experiments, we reconstructed a 64×64 grayscale image shown in Figure 3 (a) from only 40% of its Gaussian measurements. In this case, the CS matrix $\mathbf{\Phi}$ corresponds to an $M \times N$ Gaussian matrix. Figure 3 (b) shows the edges reconstructed by our method. One can clearly see that our method is able to fully recover the boundary of the image from only 40% of the measurements. Figure 3 (c) shows the reconstructed image using the L1TV method and the corresponding edges are shown in Figure 3 (d). In contrast to our method, the L1TV method fails in extracting the edges. This makes sense because the edges are much sparser than the image. To illustrate this further, in Figure 4, we plot the sorted absolute values of X_x, X_y and X corresponding to the original image. One can see from this figure that the discrete gradients decay much faster than the original pixel intensities. This means that our method can take advantage of this and be able to estimate the edges with far fewer measurements than required by the L1TV method. This in turn will result in better estimation of FDs when our method is used compared to L1TV.



(a) Original image (b) CSFD edges

(c) L1TV reconstruction (d) L1TV edges

Fig. 3: Reconstruction results with 40% of Gaussian measurements.

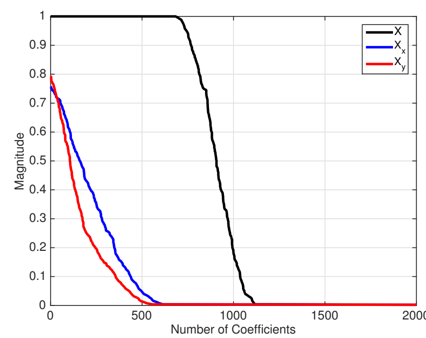
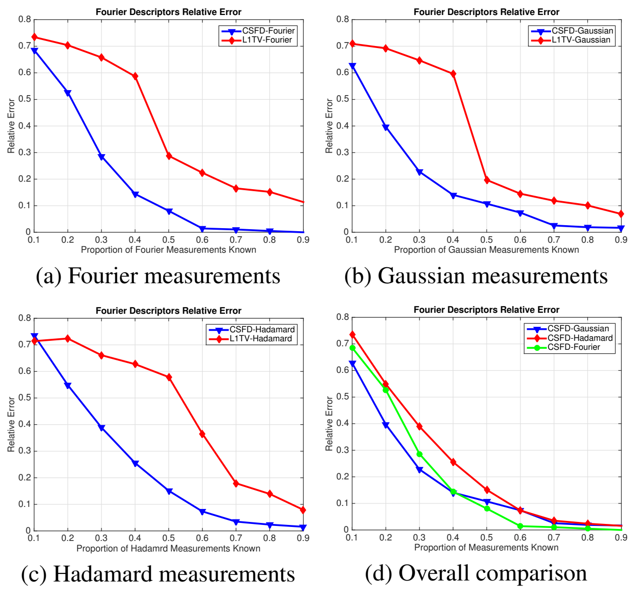


Fig. 4: Magnitude of image X, horizontal gradient X_x, and vertical gradient X_y coefficients in decreasing order for the 64×64 *gumby* image.

In the second set of tests, we study the performance of different methods in extracting FDs as we vary the number of compressive measurements. We use partial Fourier, Gaussian and partial Hadamard matrices to conduct experiments. These three classes of measurement matrices are extensively used in many CS applications [24]. The proportion of measurements known was varied from 0.1 to 0.9 in the increments of 0.1 and for each proportion of measurements the algorithms were applied 5 times. Figure 5 shows the average relative error corresponding to different methods and different sensing matrices. In particular, Figure 5 (a) - (c) show the results corresponding to Fourier, Gaussian and Hadamard measurements , respectively. As can be seen from these figures, the proposed CSFD method is able to extract FDs better than the naive method. In particular, our method achieves very low relative errors when the number of measurements are around 50%. In contrast, the L1TV method does not produce good relative errors since it requires more number of measurements to reconstruct the image. Furthermore, the performance of our method with respect to different sensing matrices in Figure 5 (d) shows that our method performs comparably in all three cases. This experiment clearly shows that one can extract FDs from compressive measurements directly without reconstructing the image.



(a) Fourier measurements (b) Gaussian measurements

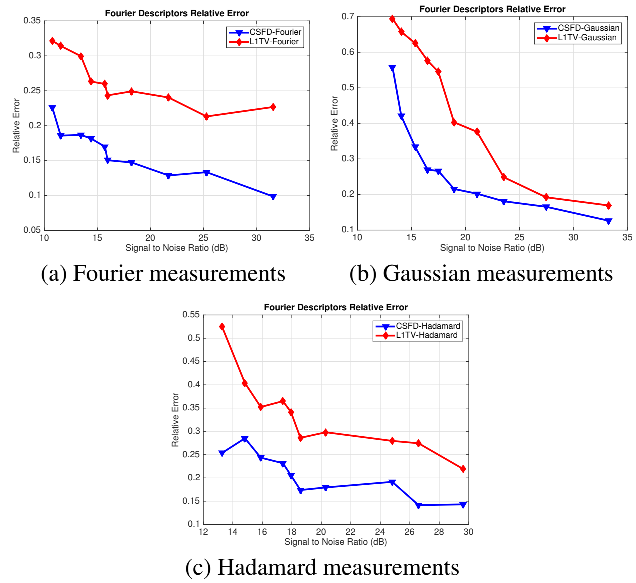
(c) Hadamard measurements (d) Overall comparison

Fig. 5: Fourier descriptor reconstruction relative errors of CSFD and L1TV when Fourier, Gaussian and Hadamard measurements are used.

The next experiment is very similar to the last experiment except that, this time, a significant amount of additive white Gaussian noise was added to the compressive measurements. This time, the proportion of measurements was fixed to 60%. The noise standard variation starts from $\sigma = 0.01$ and increases to $\sigma = 0.1$ in increments of 0.01. The resulting average FD relative errors are shown in Figure 6. As can be seen from these figures that our method is able to extract FDs from noisy compressive measurements better than the L1TV method. This experiment shows the robustness of our method in the presence of additive noise.

4.2. Real Data

In the final set of tests, we conduct experiments using the real data collected by the single pixel camera (SPC) [3]. We used the 64×64 image of a black-and-white ‘R’ as the test image (See Figure 7 (a)). According to the specific implementation details of the SPC described in [3], the digital micromirror device (DMD) is equivalent to a Bernoulli measurement matrix $\mathbf{\Phi}$ that takes 0/1 values with equal probability. In this experiment, rather than using the SPGL1 algorithm, we use the CVX optimization toolbox [25][26] to reconstruct the gradients from compressive measurements. Results are shown in Figure 7 (b). As can be seen from this figure, our method performs much better than the L1TV method and even with only 30% of measurements, CSFD produces an average relative error as low as 0.0183 compared to the average relative error of 0.4546 corresponding to the L1TV method. This experiment shows that our method can be used in practical CS applications to extract FDs directly from compressive measurements without reconstructing the image. As a result, our approach saves computational time and resources by

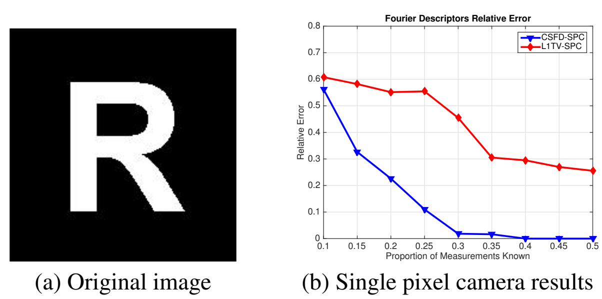


(a) Fourier measurements (b) Gaussian measurements

(c) Hadamard measurements

Fig. 6: Fourier descriptor reconstruction relative errors of CSFD and L1TV when Fourier, Gaussian and Hadamard measurements are used.

avoiding the operations related to reconstruction. This has significant implications in robotics and computer vision applications when CS methods are used for imaging.



(a) Original image (b) Single pixel camera results

Fig. 7: Experiments with the real data. (a) Original image. (b) Fourier descriptor reconstruction relative errors of CSFD and L1TV on real SPC data.

5. CONCLUSION

We presented a compressive feature extraction method based on FDs. Our method entails extracting FDs from the estimated edges via discrete horizontal and vertical image gradients. Various experiments have shown a great improvement in quality of reconstruction as well as a robustness to noise over a naive method in which FDs are extracted after reconstructing the image from compressive measurements. In the future, we will investigate the object recognition performance of the proposed CSFD method. We will also study the conditioning of the sensing matrices in our method.

6. REFERENCES

[1] David L Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] Emmanuel J Candès, Justin Romberg, and Terence Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on information theory*, vol. 52, no. 2, pp. 489–509, 2006.

[3] M. F. Duarte, M. A. Davenport, D. Takbar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, March 2008.

[4] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing mri," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72–82, March 2008.

[5] V. M. Patel, G. R. Easley, D. M. Healy Jr., and R. Chellappa, "Compressed synthetic aperture radar," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 244–254, April 2010.

[6] A. C. Gurbuz, J. H. McClellan, and W. R. Scott, "A compressive sensing data acquisition and imaging method for stepped frequency gprs," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2640–2650, July 2009.

[7] V. M. Patel and J. N. Mait, "Compressive passive millimeter-wave imaging with extended depth of field," *Optical Engineering*, vol. 51, no. 9, pp. 091610–1–091610–7, Sept 2012.

[8] Mark A. Davenport, Marco F. Duarte, Michael B. Wakin, Jason N. Laska, Dharmpal Takhar, Kevin F. Kelly, and Richard G. Baraniuk, "The smashed filter for compressive classification and target recognition," in *Proc. SPIE*, 2007, pp. 64980H–64980H–12.

[9] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal processing with compressive measurements," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 445–460, April 2010.

[10] Volkan Cevher, Aswin Sankaranarayanan, Marco F. Duarte, Dikpal Reddy, Richard G. Baraniuk, and Rama Chellappa, *Compressive Sensing for Background Subtraction*, pp. 155–168, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[11] G. Warnell, D. Reddy, and R. Chellappa, "Adaptive rate compressive sensing for background subtraction," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2012, pp. 1477–1480.

[12] N. Vaswani, "Kalman filtered compressed sensing," in *2008 15th IEEE International Conference on Image Processing*, Oct 2008, pp. 893–896.

[13] A. C. Gurbuz, J. H. McClellan, J. Romberg, and W. R. Scott, "Compressive sensing of parameterized shapes in images," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2008, pp. 1949–1952.

[14] J. C. Ye, "Compressed sensing shape estimation of star-shaped objects in fourier imaging," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 750–753, Oct 2007.

[15] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Transactions on Computers*, vol. C-21, no. 3, pp. 269–281, March 1972.

[16] Dengsheng Zhang and Guojun Lu, "Review of shape representation and description techniques," *Pattern recognition*, vol. 37, no. 1, pp. 1–19, 2004.

[17] Vishal M Patel, Ray Maleh, Anna C Gilbert, and Rama Chellappa, "Gradient-based image recovery methods from incomplete fourier measurements," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 94–105, 2012.

[18] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, July 2007.

[19] Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[20] Djemel Ziou, Salvatore Tabbone, et al., "Edge detection techniques-an overview," *Pattern Recognition and Image Analysis C/C of Raspoznavaniye Obrazov I Analiz Izobrazhenii*, vol. 8, pp. 537–559, 1998.

[21] Michael Lustig, David Donoho, and John M Pauly, "Sparse mri: The application of compressed sensing for rapid mr imaging," *Magnetic resonance in medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[22] Ewout Van Den Berg and Michael P Friedlander, "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008.

[23] E Van Den Berg and MP Friedlander, "Spgl1: A solver for large-scale sparse reconstruction," 2007.

[24] Yaakov Tsaig and David L Donoho, "Extensions of compressed sensing," *Signal processing*, vol. 86, no. 3, pp. 549–571, 2006.

[25] Michael Grant and Stephen Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, V. Blondel, S. Boyd, and H. Kimura, Eds., Lecture Notes in Control and Information Sciences, pp. 95–110. Springer-Verlag Limited, 2008.

[26] Michael Grant, Stephen Boyd, and Yinyu Ye, "Cvx: Matlab software for disciplined convex programming," 2008.