Compressed Infrared Target Detection Using Stochastically Trained Least Squares

Brian Millikan, *Student Member, IEEE,* Aritra Dutta, Qiyu Sun, and Hassan Foroosh, *Senior Member, IEEE*

Abstract

Target detection of potential threats at night can be deployed on a costly infrared focal plane array with high resolution. Due to the compressibility of infrared image patches, the high resolution requirement could be reduced with target detection capability preserved. For this reason, a compressive midwave infrared imager (MWIR) with low resolution has been developed. As the most probable coefficients indices of the support set of the infrared image patches could be learnt from training data, we develop stochastically trained least squares (STLS) for MWIR image reconstruction. Using the same measurement matrix as in STLS, we construct a compressed quadratic correlation filter (CQCF) for compressed infrared target detection. We apply CQCF to the U.S. Army Night Vision and Electronic Sensors Directorate (NVESD) dataset. Numerical simulations show that the recognition performance of our algorithm matches that of the standard full reconstruction methods but at a fraction of the execution time.

Keywords

Target Detection, Compressive Sensing, Stochastically Trained Least Squares, Quadratic Correlation Filter, Linear Decoder.

I. Introduction

Capturing images of potential threats at night on many military systems requires electro-optical infrared (EO/IR) sensors, which are typically of low resolution [1], [2]. In order to improve the target detection performance, sensors of high resolution are desired but they are expensive [1]–[4]. An alternative is a low-resolution midwave infrared sensor paired with a higher resolution spatial light modulator (such as a DMD). This compressive midwave infrared imager has been used in [5]–[7] to realize a high resolution midwave infrared (MWIR) sensor, see Figure 1.

The compressive midwave infrared imager, shown in Figure 1, uses a carefully constructed sensing matrix to generate a set of measurements that is smaller than the number of samples in the original infrared image [8]. For the

The project is partially supported by National Science Foundation (DMS- 1412413).

B. Millikan is with the Department of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816, USA, and also with Lockheed Martin Missiles and Fire Control, Orlando, FL 32819, USA (e-mail: brian.millikan@knights.ucf.edu).

A. Dutta and Q. Sun are with the Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA (email: d.aritra2010@knights.ucf.edu; qiyu.sun@ucf.edu).

H. Foroosh is with the Department of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816, USA (email: foroosh@cs.ucf.edu).

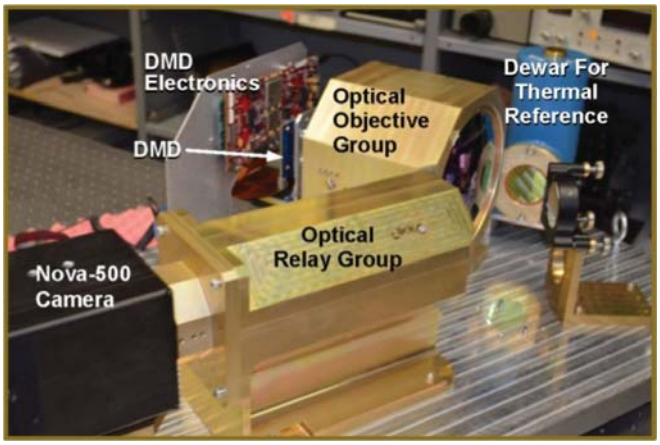


Fig. 1. Low-cost high resolution midwave infrared (MWIR) sensor testbed adapted from [7] with permission. Its design is based on a single pixel camera with multiple detectors.

case that the image is inherently sparse in some domain, it could be reconstructed with high probability [9]–[15]. On the other hand, in target detection applications with a MWIR sensor, perfect reconstruction of the infrared image is not necessary. This makes possible to use less measurements for accurate target detection.

Unlike a typical MWIR sensor, a compressive MWIR imaging sensor requires nonlinear iterative methods, such as compressive sampling matching pursuit, iterative hard thresholding, orthogonal matching pursuit and basis pursuit, to recover an infrared image with an unknown support [16]–[19]. In this paper, we use the statistical information of MWIR training data, and propose a linear decoder to reconstruct MWIR images with targets. We call this method Stochastically Trained Least Squares (STLS).

Automatic Target Recognition (ATR), which discriminates targets, is the processing component in a typical MWIR system [1], [20]. Target detection, the first step of ATR, separates potential target locations from background clutter [1], [20], [21]. One target detection filter is the quadratic correlation filter (QCF) [22]. After finding regions of interest, multi-class automatic target recognition algorithms would identify the specific target type. This step, acting as a filter for the target recognition classifier, could reduce the number of false positive detections [23], [24].

When working with a compressive midwave infrared imager, the algorithm developer must decide whether to reconstruct the image and use typical target detection techniques [6], or to modify existing target detection techniques and use the compressed data directly. In this paper, we take the second approach and modify the QCF algorithm by coupling it with the stochastically determined measurement matrix used in STLS.

The rest of the paper is organized as follows. In Section II, we introduce the STLS method for MWIR image reconstruction. In Section III, we propose Compressed Quadratic Correlation Filter (CQCF) and compare it with the QCF for infrared target detection. In Section IV, we combine STLS and CQCF for ATR applications and we demonstrate its performance on the U.S. Army Night Vision and Electronic Sensors Directorate (NVESD) dataset.

II. Stochastically Trained Least Squares

Infrared image patches, $\mathbf{x}$, in ATR usually have compressible representation,

$$\mathbf{x} = \Psi\alpha = \sum_{i=1}^{N} \alpha_i \psi_i, \tag{1}$$

in some (DCT/Wavelet/Karhunen-Loève) dictionary $\Psi = (\psi_1, \psi_2, \cdots, \psi_N)$, where $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_N)^T$ is a compressible vector in the sense that its components have sorted absolute values with fast decay. The measurement $\mathbf{y}$ of the infrared image patch $\mathbf{x}$ via a DMD is

$$\mathbf{y} = \Phi\mathbf{x} = \hat{\Phi}\alpha,$$

where $\hat{\Phi} = \Phi\Psi$ and Φ is an $m \times N$ measurement matrix of the DMD with $m < N$.

Due to compressibility of the image $\mathbf{x}$, it can be approximated by a sparse image $\Psi\alpha^*$ with small $\|\alpha^*\|_0$. One conventional selection of the sparse vector α^* is a solution of the ℓ_0 minimization problem

$$\alpha^* = \underset{\alpha}{\operatorname{argmin}} \|\alpha\|_0 \text{ subject to } \mathbf{y} = \hat{\Phi}\alpha, \tag{2}$$

which is NP-hard [8], [12], [25]–[27]. Denote the essential support of the target signal $\hat{\Phi}\alpha^*$ by S, which has cardinality s. In our ATR applications, we do not have access to the support set S. However, we can use the training data to tell us where the most probable locations T of the large coefficients are. In our simulations, we will use

$$T := \{i \in [1, N] : P[|\alpha_i| \geq \rho] \geq \tau\}, \tag{3}$$

where the probability P is learnt from the training data, ρ is the threshold of large coefficients, and τ is the threshold of the most probable locations for sparse approximation (see Section IV for details).

For compressed target detection applications, the measurement matrix Φ may depend on the statistical properties of $\mathbf{x}$. For instance

$$\Phi = \Theta_T, \tag{4}$$

where Θ is a DCT matrix and Θ_T is obtained by selecting the rows of Θ in T. This measurement matrix is what we will use to capture measurements of the infrared image patch $\mathbf{x}$.

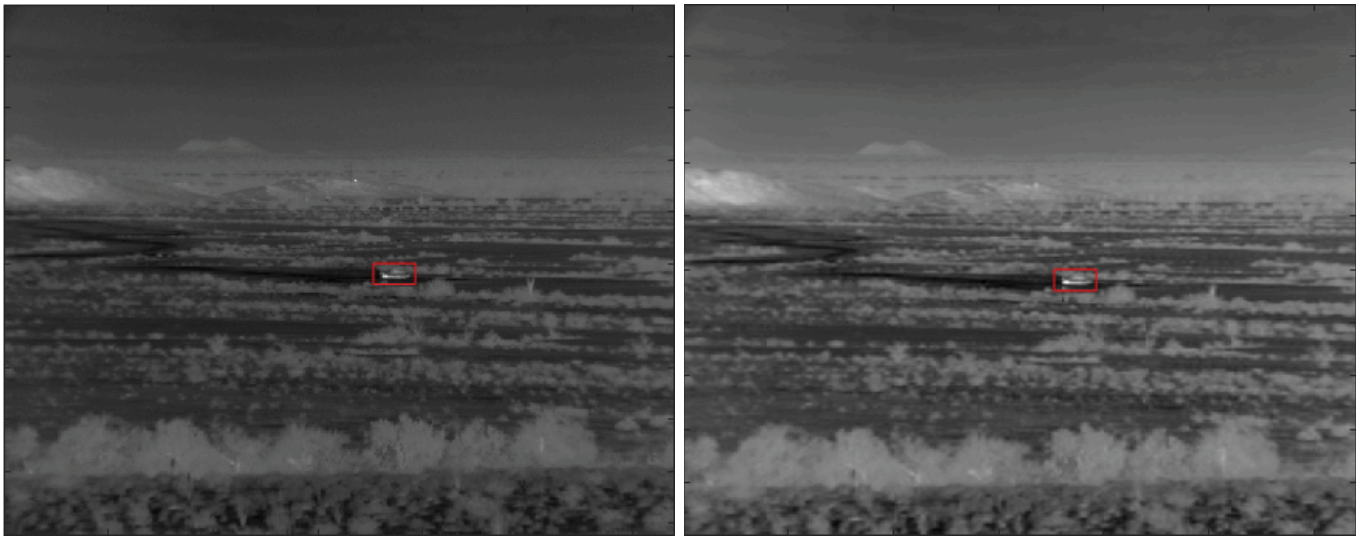
Since we have some statistical information about the support set S of the image under consideration, the probability of the set S with cardinality s to be contained in the set T with cardinality $t \geq s$ is high. Therefore we use the least squares solution α_T of the linear system

$$\mathbf{y} = \hat{\Phi}_T \alpha_T, \tag{5}$$

to approximate the compressible vector α^*. Here $\hat{\Phi}_T$ is the submatrix of $\hat{\Phi}$ obtained by extracting its columns corresponding to the indices in T. The solution to (5) is

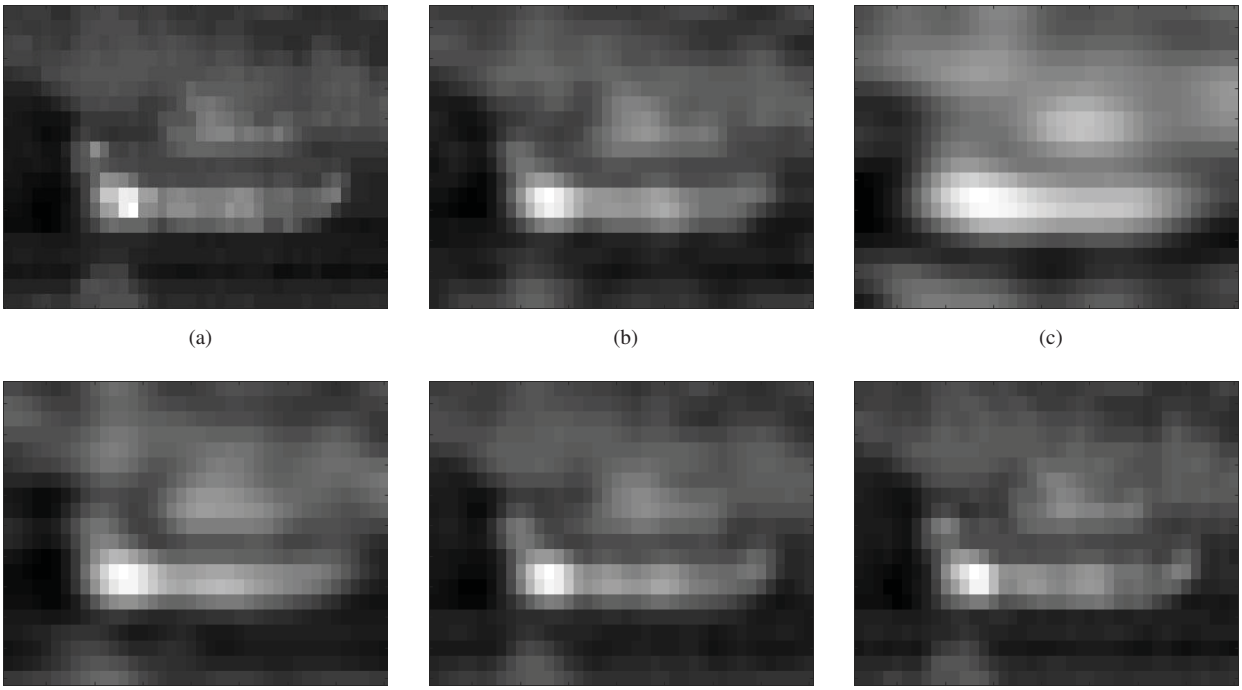
$$\alpha_T = \hat{\Phi}_T^+ \mathbf{y}, \tag{6}$$

where $\hat{\Phi}_T^+$ is the Moore-Penrose pseudoinverse of $\hat{\Phi}_T$. We call the above approach of reconstructing infrared image patches as stochastically trained least squares (STLS), see Figures 2 and 3 for the reconstruction of a midwave infrared image containing a single target.



(a) (b)

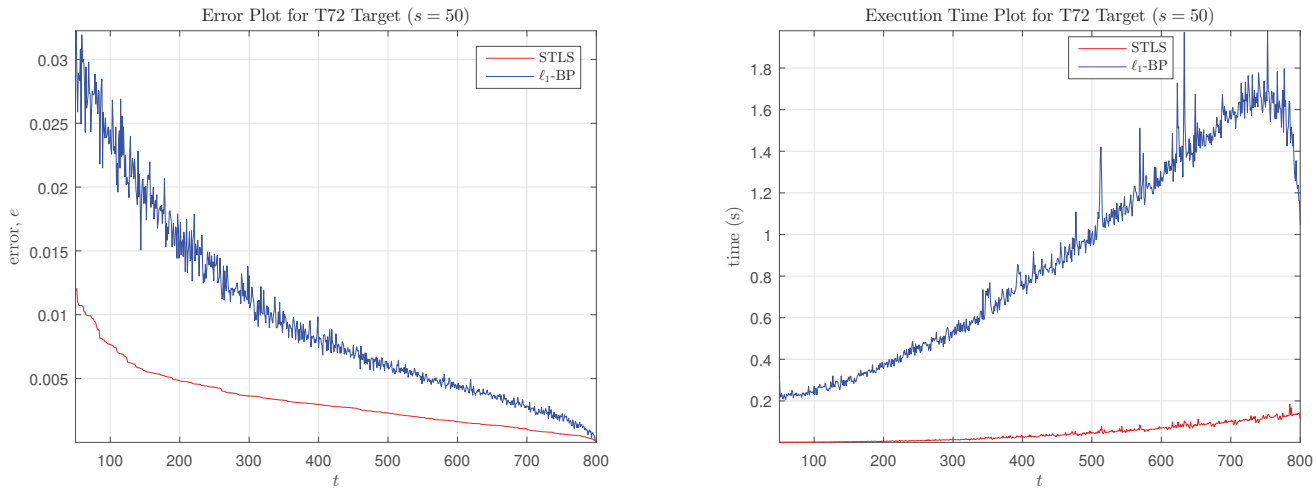
Fig. 2. (a) Original image with Main Battle Tank (T72) target. (b) Full image reconstruction using STLS from $t = 100$ measurements per 20×40 block. This reconstructed image has a PSNR of 74.2702 dB and the target image patch has a reconstruction error of 0.0078.



(a) (b) (c)

(d) (e) (f)

Fig. 3. (a) Midwave infrared 20×40 image patch containing a Main Battle Tank (T72) target. (b) Image patch reconstructed using true support S with cardinality $s = 100$ (PSNR $= 64.7373$ dB, $\rho = 0.00189$). (c) Image patch reconstructed using STLS with the set T having cardinality $t = 50$ (PSNR $= 58.1389$ dB), (d) $t = 100$ (PSNR $= 62.0743$ dB), (e) $t = 200$ (PSNR $= 65.7067$ dB) and (f) $t = 400$ (PSNR $= 70.2279$ dB). The percentage of indices of the set S contained in the set T is 43% when $t = 50$, 70% when $t = 100$, 93% when $t = 200$, and 100% when $t = 400$.



(a) (b)

Fig. 4. (a) This plot compares the reconstruction error of the ℓ_1-BP method to the STLS method for different values of t. We see that the ℓ_1-BP error is as high as 3 times that of the STLS. We also see that there is more variability in the ℓ_1-BP reconstruction error due to the random measurement process. (b) In this plot, we look at the execution times for ℓ_1-BP versus STLS for different values of t. It is observed that the execution time of ℓ_1-BP is almost 20 times as high as that of STLS for certain values of t.

Least squares is an attractive method and it has closed-form solutions. However, least squares cannot find a sparse solution without some other additional information about the infrared image patches. Compared with iterative methods, such as basis pursuit (BP), orthogonal matching pursuit (OMP) and initialized iterative reweighted least squares (IIRLS), to solve the ℓ_0 minimization problem (2) or its ℓ_1 relaxation [18], [19], [28], our STLS has better performance in terms of reconstruction error and computational time (see Figure 4).

The STLS is a linear decoder and it could be implemented in hardware. The support set T in the STLS contains the most probable coefficients indices of the support set S of the image. Comparing with the linear decoder in [29], our support set T is based on the probabilities that are defined using the statistical properties of the training data, while the support set in [29] is learnt from solving some optimization problem.

III. Quadratic Correlation Filter

The quadratic correlation filter utilizes the Fukunaga-Koontz transform (FKT) to separate target image areas from background clutter [30]. Mahalanobis et al. introduced the quadratic correlation filter (QCF) for target detection in [22]. In Section III-A, we recall the QCF algorithm. Then in Section III-B, we integrate the STLS method into the QCF algorithm to generate a compressed quadratic correlation filter (CQCF). Finally, we compare CQCF to QCF paired with iterative non-linear reconstruction methods in Section III-C.

A. Quadratic Correlation Filter For Target Detection

Denote the correlation matrices of a target image patch $\mathbf{x}_{tgt}$ and a background image patch $\mathbf{x}_{bkg}$ by

$$\mathbf{R}_{tgt} = \mathbb{E}[\mathbf{x}_{tgt}\mathbf{x}_{tgt}^T] \text{ and } \mathbf{R}_{bkg} = \mathbb{E}[\mathbf{x}_{bkg}\mathbf{x}_{bkg}^T],$$

respectively. These correlation matrices are positive semidefinite. The sum of the target and the background correlation matrices can be decomposed into the form,

$$\mathbf{R}_{tgt} + \mathbf{R}_{bkg} = \mathbf{M}\Delta\mathbf{M}^T, \tag{7}$$

where columns of the orthogonal matrix $\mathbf{M}$ are the eigenvectors of $\mathbf{R}_{tgt}+\mathbf{R}_{bkg}$ and the diagonal matrix Δ contains the corresponding eigenvalues. We then define a transform matrix

$$\mathbf{P} = \mathbf{M}\Delta^{-1/2}, \tag{8}$$

and rewrite (7) as

$$\mathbf{P}^T(\mathbf{R}_{tgt} + \mathbf{R}_{bkg})\mathbf{P} = \mathbf{I}, \tag{9}$$

where $\mathbf{I}$ is the identity matrix.

Define the target and background image patch correlation matrices

$$\widehat{\mathbf{R}}_{tgt} = \mathbf{P}^T\mathbf{R}_{tgt}\mathbf{P} \text{ and } \widehat{\mathbf{R}}_{bkg} = \mathbf{P}^T\mathbf{R}_{bkg}\mathbf{P}$$

in the new domain, which reduces the equation (9) to

$$\widehat{\mathbf{R}}_{tgt} + \widehat{\mathbf{R}}_{bkg} = \mathbf{I}. \tag{10}$$

Performing an eigendecomposition on $\widehat{\mathbf{R}}_{tgt}$ gives

$$\widehat{\mathbf{R}}_{tgt} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T, \tag{11}$$

where columns of the orthogonal matrix $\mathbf{V}$ are the eigenvectors of $\widehat{\mathbf{R}}_{tgt}$ and the diagonal matrix $\mathbf{\Lambda}$ contains the corresponding eigenvalues. Using (10) and (11), we obtain

$$\widehat{\mathbf{R}}_{bkg} = \mathbf{V}(\mathbf{I} - \mathbf{\Lambda})\mathbf{V}^T, \tag{12}$$

where the eigenvalues of $\widehat{\mathbf{R}}_{tgt}$ and $\widehat{\mathbf{R}}_{bkg}$ lie between zero and one, see Figure 5a.

Given a threshold $\epsilon \in (0, 1)$, we choose the largest $|\Omega_{tgt}|$ eigenvalues λ_i in $\mathbf{\Lambda}$ for the target class, where

$$\Omega_{tgt} := \{i \in [1, N] : \ \lambda_i \geq 1 - \epsilon\}.$$

Similarly, we choose the smallest $|\Omega_{bkg}|$ eigenvalues λ_i in $\mathbf{\Lambda}$ to represent the background class, where

$$\Omega_{bkg} := \{i \in [1, N] : \ \lambda_i \leq \epsilon\}.$$

Selecting the columns of $\mathbf{V}$ corresponding to the indices of the set Ω_{tgt}, we create the target projection submatrix $\mathbf{V}_{\Omega_{tgt}}$. Similarly we can form the background projection submatrix $\mathbf{V}_{\Omega_{bkg}}$ using the indices of the set Ω_{bkg}.

In order to classify a test image patch, $\mathbf{x}$, as target or background, we project $\mathbf{x}$ into

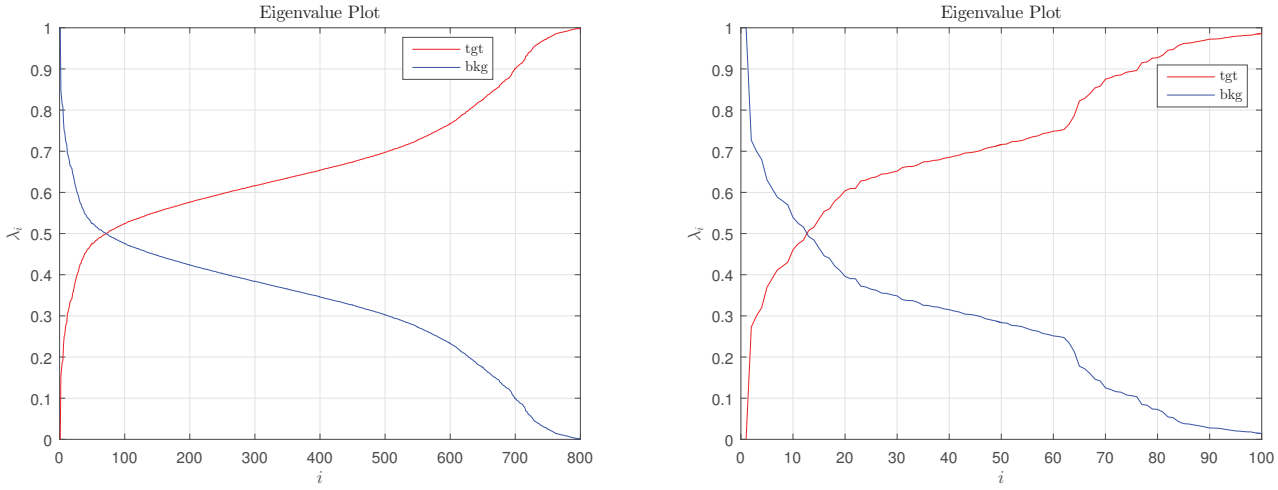
$$\mathbf{v}_{tgt} = \mathbf{V}_{\Omega_{tgt}}^T\mathbf{P}^T\mathbf{x}$$

and

$$\mathbf{v}_{bkg} = \mathbf{V}_{\Omega_{bkg}}^T\mathbf{P}^T\mathbf{x}.$$

Define a statistic

$$\varphi = \mathbf{v}_{tgt}^T\mathbf{v}_{tgt} - \mathbf{v}_{bkg}^T\mathbf{v}_{bkg}, \tag{13}$$



(a) (b)

Fig. 5. Eigenvalue plots: (a) uncompressed QCF for a 20×40 image patch and (b) compressed QCF where $t = 100$. As mentioned in [22], for the QCF and CQCF to be effective in target detection, we need to choose a proper threshold ϵ so that the eigenvalues associated with the targets and background are separated. We observe that, due to the compression, CQCF has less eigenvalues and the resulting discrimination statistic will be smaller in magnitude, and hence it could be more efficient than the QCF.

which is

$$\begin{aligned}\varphi &= \mathbf{x}^T\mathbf{P}(\mathbf{V}_{\Omega_{tgt}}\mathbf{V}_{\Omega_{tgt}}^T - \mathbf{V}_{\Omega_{bkg}}\mathbf{V}_{\Omega_{bkg}}^T)\mathbf{P}^T\mathbf{x} \\ &= \mathbf{x}^T(\mathbf{F}\mathbf{F}^T - \mathbf{G}\mathbf{G}^T)\mathbf{x}, \end{aligned} \tag{14}$$

where $\mathbf{F} = \mathbf{P}\mathbf{V}_{\Omega_{tgt}}$ and $\mathbf{G} = \mathbf{P}\mathbf{V}_{\Omega_{bkg}}$. We observe that the statistic will be large and positive for target image patches and small or negative for background image patches. We will use the statistic φ to determine whether the image patch is a target or not, see Figure 6b.

B. Compressed Quadratic Correlation Filter

In our target detection application, compressed images $\mathbf{y} = \Theta_T\mathbf{x}$ are captured using the sensing matrix Θ_T in (4). The correlation matrices for the compressed target images $\mathbf{y}_{tgt} = \Theta_T\mathbf{x}_{tgt}$ and the compressed background images $\mathbf{y}_{bkg} = \Theta_T\mathbf{x}_{bkg}$ are

$$\mathbf{R}_{tgt}^* = \mathbb{E}\left[\mathbf{y}_{tgt}\mathbf{y}_{tgt}^T\right] = \Theta_T\mathbf{R}_{tgt}(\Theta_T)^T$$

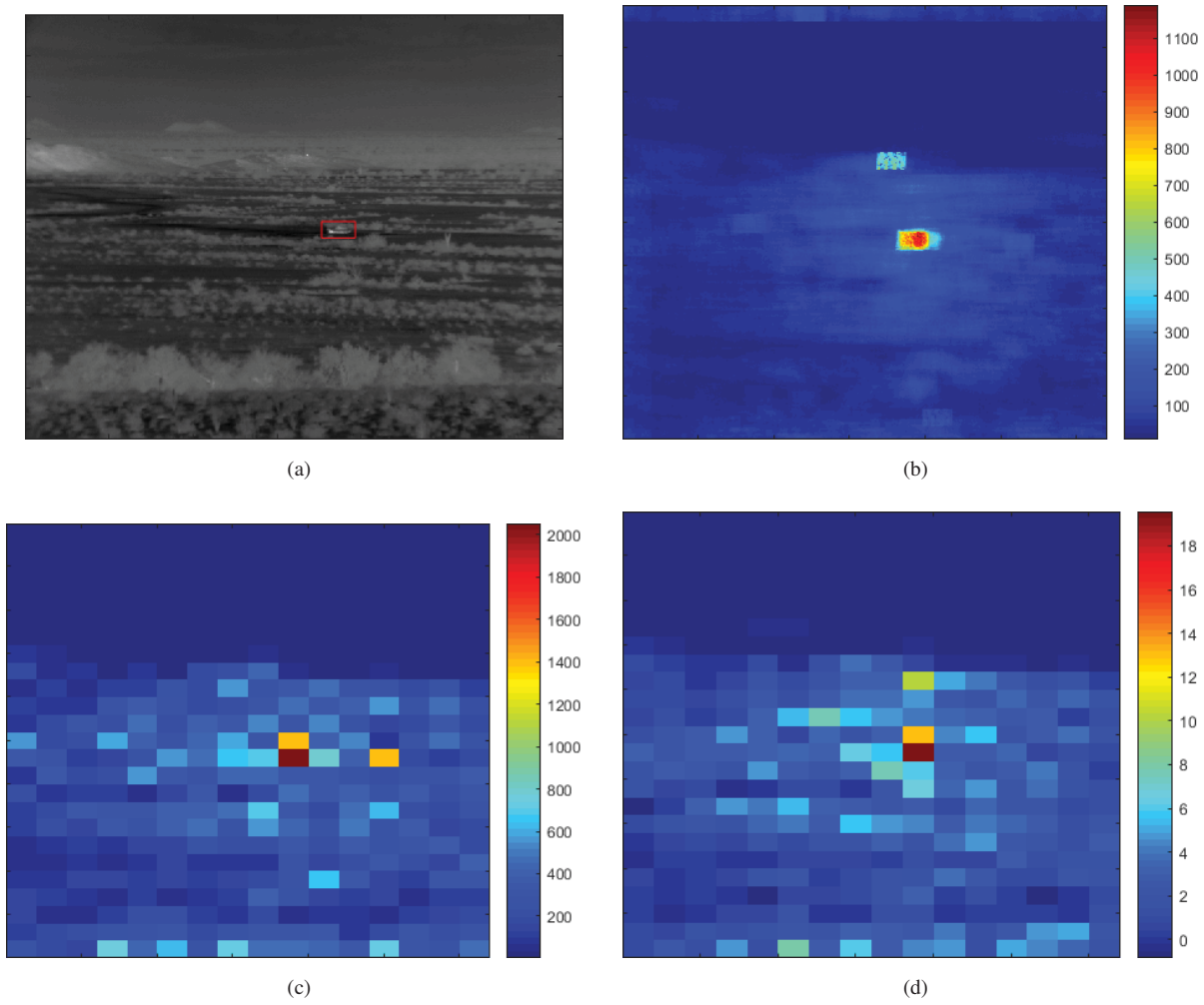
and

$$\mathbf{R}_{bkg}^* = \mathbb{E}\left[\mathbf{y}_{bkg}\mathbf{y}_{bkg}^T\right] = \Theta_T\mathbf{R}_{bkg}(\Theta_T)^T$$

respectively. Their sum can be decomposed into the form

$$\mathbf{R}_{tgt}^* + \mathbf{R}_{bkg}^* = \widehat{\mathbf{M}}\widehat{\Delta}\widehat{\mathbf{M}}^T, \tag{15}$$

where columns of the orthogonal matrix $\widehat{\mathbf{M}}$ contain the eigenvectors of $\mathbf{R}_{tgt}^* + \mathbf{R}_{bkg}^*$ and the diagonal matrix $\widehat{\Delta}$ contains corresponding eigenvalues, see Figure 5b. Similar to the QCF in Section III-A, we define the new



(a) (b)

(c) (d)

Fig. 6. (a) Original midwave infrared image containing a single Main Battle Tank (T72) target. (b) Target detection statistic φ from QCF applied to the uncompressed original image. (c) QCF applied to a 20×40 blockwise reconstruction of the image using ℓ_1-BP where $m = 100$. (d) Compressed quadratic correlation filter (CQCF) where $t = 100$. There are 542 eigenvalues for the target and 26 for the background for the discriminator in (b), 99 for the target and 1 for the background in (c), and 81 for the target and 6 for the background in (d). If we create half the range of the statistic as the threshold, we will have 1 detection in (b), and 3 detections in (c), and 2 detections in (d). We can see that φ for (b) and (c) is greater than $1,100$. However, the maximum value of $\widehat{\varphi}$ for CQCF (d) is around 20, which is due to the lower number of total eigenvalues and the disparate ratio of target to background discriminators.

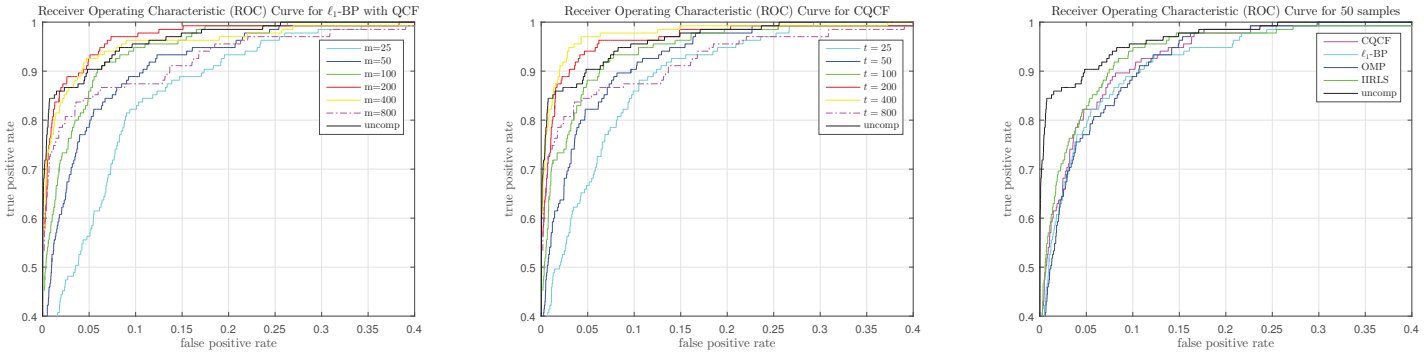
compressed target correlation matrix $\widehat{\mathbf{R}}^*_{tgt} = \widehat{\mathbf{P}}^T \mathbf{R}^*_{tgt} \widehat{\mathbf{P}}$ and the background correlation matrix $\widehat{\mathbf{R}}^*_{bkg} = \widehat{\mathbf{P}}^T \mathbf{R}^*_{bkg} \widehat{\mathbf{P}}$, where $\widehat{\mathbf{P}} = \widehat{\mathbf{M}} \widehat{\Delta}^{-1/2}$ is the transform matrix. Write

$$\widehat{\mathbf{R}}^*_{tgt} = \widehat{\mathbf{V}} \widehat{\mathbf{\Lambda}} \widehat{\mathbf{V}}^T,$$

where $\widehat{\mathbf{V}}$ is an orthogonal matrix containing the eigenvectors of $\widehat{\mathbf{R}}^*_{tgt}$ and $\widehat{\mathbf{\Lambda}}$ is the diagonal matrix containing the corresponding eigenvalues.

Given a compressed measurement $\mathbf{y}$, we define

$$\widehat{\mathbf{v}}_{tgt} = \widehat{\mathbf{V}}^T_{tgt} \widehat{\mathbf{P}}^T \mathbf{y} \text{ and } \widehat{\mathbf{v}}_{bkg} = \widehat{\mathbf{V}}^T_{bkg} \widehat{\mathbf{P}}^T \mathbf{y},$$



(a) (b) (c)

Fig. 7. (a) The ROC curves for QCF with blockwise ℓ_1-BP reconstruction for different measurements sizes m. (b) This chart shows CQCF applied to the NVESD test image dataset with similar results to (a). (c) Comparison of methods to ℓ_1-BP, OMP, and IIRLS blockwise reconstruction methods with CQCF with $t = 50$. We can see from these curves that if the compression ratio is $\frac{1}{16}$, these methods have similar target detection capability. However, CQCF requires much less execution time since it operates in the compressed domain.

and a statistic

$$\widehat{\varphi} = \widehat{\mathbf{v}}_{tgt}^T \widehat{\mathbf{v}}_{tgt} - \widehat{\mathbf{v}}_{bkg}^T \widehat{\mathbf{v}}_{bkg} \tag{16}$$

which can also be written as

$$\widehat{\varphi} = \mathbf{y}^T (\widehat{\mathbf{F}}\widehat{\mathbf{F}}^T - \widehat{\mathbf{G}}\widehat{\mathbf{G}}^T)\mathbf{y} \tag{17}$$

where $\widehat{\mathbf{F}} = \widehat{\mathbf{P}}\widehat{\mathbf{V}}_{tgt}$ and $\widehat{\mathbf{G}} = \widehat{\mathbf{P}}\widehat{\mathbf{V}}_{bkg}$.

We observe that the statistic has larger values over the target class and smaller values over the background. Therefore for compressed infrared image patch, we can use the statistic $\widehat{\varphi}$ to classify the target and the background, see Figure 6d.

C. Comparison of CQCF to QCF

The standard QCF is intended to work on an image in the entire pixel domain. For this reason, we need reconstruction techniques, such as basis pursuit, orthogonal matching pursuit and initialized iterative reweighted least squares, for target detection [18], [19], [28]. However, CQCF introduced in Section III-B operates in the compressed domain. We observe that the value of the decision statistic for CQCF is effected by the compression. Comparing with QCF, the number of distinguishing eigenvalues for target and background is significantly less, see Figure 5. This causes a decrease in the magnitude of the statistic $\widehat{\varphi}$ for CQCF versus the statistic φ for QCF. In Figure 6, we show an application of QCF and CQCF to distinguish the target main battle tank (T72) from background in a midwave infrared image.

The performance of those two methods could be measured using receiver operating characteristic (ROC) curves [31]–[33], see Figure 7. From the simulations we see that the CQCF performs similarly to the iterative recovery methods with QCF, but at a fraction of the execution time due to its closed-form linear nature.

IV. Numerical Experiments

We conducted our simulations using the SENSIAC database for ATR algorithm development [34]. This database is a mid-wave infrared (MWIR) dataset from the U.S. Army Night Vision and Electronic Sensors Directorate (NVESD). It contains 207GB of MWIR data which includes 10 vehicle target types and 2 scenarios of humans. For our experiments, we considered all 10 vehicle target types. These include a Pickup Truck (PICKUP), Sport Utility Vehicle (SUV), Armored Personnel Carriers (BTR70 and BMP2), an Infantry Scout Vehicle (BRDM2), a Main Battle Tank (T72), an Anti-Aircraft Weapon (ZSU23-4), a Self-Propelled Howitzer (2S3), an Armoured Reconnaissance Vehicle (MTLB) and a Towed Howitzer (D20). A 20×40 bounding box is formed around each target using the ground truth data to generate a target image patch, $\mathbf{x}$, for training.

A. Simulation Setting

For our simulations, coefficients α_i could be assumed to be independent with a Laplace distribution, cf. [35]. Presented in Figure 9 is the histograms of different coefficients for all ten of the vehicle target types in the NVESD dataset. Each coefficient has a corresponding population mean and variance,

$$\alpha_i \sim \text{Laplace}(\mu_i, b_i) \;\text{ for }\; i = 1, \cdots, N. \tag{18}$$

We can approximate the population mean and variance with the sample mean and variance. The i-th sample mean ($\bar{\alpha}_i$) and sample variance ($2\tilde{b}_i^2$) are evaluated by

$$\bar{\alpha}_i = \frac{1}{N_S} \sum_{j=1}^{N_S} \alpha_i^j \tag{19}$$

and

$$2\tilde{b}_i^2 = \frac{1}{N_S - 1} \sum_{j=1}^{N_S} (\alpha_i^j - \bar{\alpha}_i)^2, \;\; 1 \le i \le N, \tag{20}$$

where N_S is the total number of training target images. For our experiments,

$$N_S = 68,000 = 1700 \times 10 \times 4,$$

which consists of four different scenarios containing $1,700$ images each for all ten targets. Under the above assumptions, we can calculate the probability of having a large coefficient at i-th position as

$$P\left[|\alpha_i| \ge \rho\right] = 1 - \frac{1}{2\tilde{b}_i} \int_{-\rho}^{\rho} \mathrm{e}^{-\frac{|\alpha_i - \bar{\alpha}_i|}{\tilde{b}_i}} \mathrm{d}\alpha_i, \tag{21}$$

and define the measurement matrix $\Phi \in \mathbb{R}^{t \times N}$ as in (4) for STLS.

Shown in Figure 8 is a comparison of the large coefficient probability, τ, to the cardinality of the set T. Therefore as expected, the probabilities τ increase as the threshold ρ for a large coefficients decrease. If ρ is chosen to be too small, the probabilities will approach all 1 and will not contain any useful information about the locations of most probable coefficients. If the coefficient ρ is too large, then the probabilities will approach all 0 and again will not provide information about most probable coefficient locations. So from now on, we choose $\rho = 0.005$. The threshold τ is then chosen based on the desired set size, T. This number can be determined experimentally from the training dataset. We used set sizes of $25, 50, 100, 200, 400$ and 800 for our experiments.

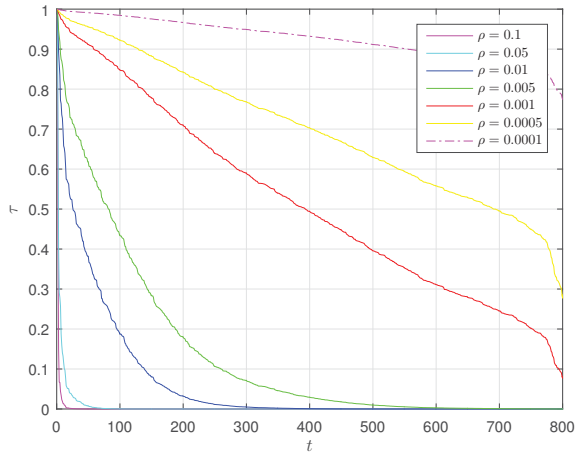


Fig. 8. Plot of the large coefficient probability, τ, versus the cardinality of the set $|T| = t$ for different values of the large coefficient threshold, ρ. If ρ is set too small or too large, the coefficient probabilities τ will not contain any useful information as they will be saturated.

The correlation matrix $\mathbf{R}_{tgt}$ for QCF can be estimated from N_S training target images. Similarly, the correlation matrix $\mathbf{R}_{bkg}$ can be also obtained using N_S random background samples from the training images.

In all simulations, we use the two-dimensional inverse DCT basis Ψ of size $N \times N$ to generate vectors α of training images extracted from the NVESD database. In CQCF, we also use a two-dimensional forward DCT basis Θ of size $|T| \times N$ for Φ.

In our experiments, the measurement matrix $\Phi \in \mathbb{R}^{m \times N}$ is a Gaussian random matrix with

$$\phi_{i,j} \sim \mathcal{N}(0, 1/m)$$

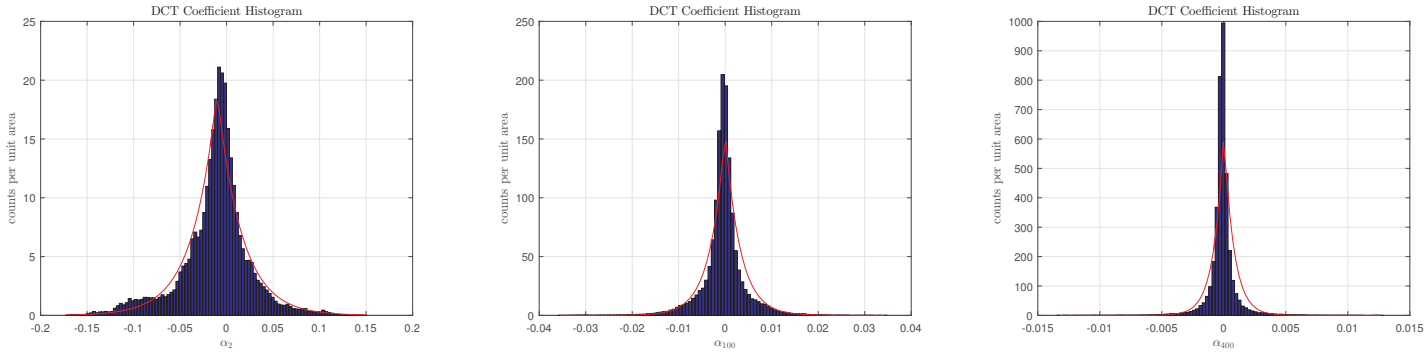
where m is the number of measurements. The above measurement matrix has been widely used to recover a sparse vector with an unknown support [8], [12], [18], [36].

B. Results

For recovery of infrared target image patches, the performance of STLS is compared against IIRLS, BP, and OMP [18], [19], [28] on the NVESD dataset, see Figure 10. Our simulation output is shown in Figure 11. These methods were compared using the normalized reconstruction metric

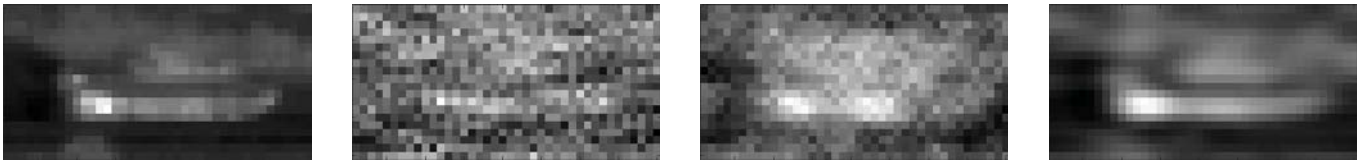
$$e = \frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_p}{\|\mathbf{x}\|_p} \tag{22}$$

with $p = 1, 2, \infty$, where $\mathbf{x}$ represents the original image patch and $\widehat{\mathbf{x}}$ represents the reconstructed image patch. We can see from Figure 11 that STLS has significantly less reconstruction error than the iterative methods, including BP, OMP, IIRLS. This is due to the fact that the measurement contains an exact copy of the most probable dominant coefficients. As long as the actual large coefficients falls into one of these "most probable" supporting bins, we will get a very good reconstruction.



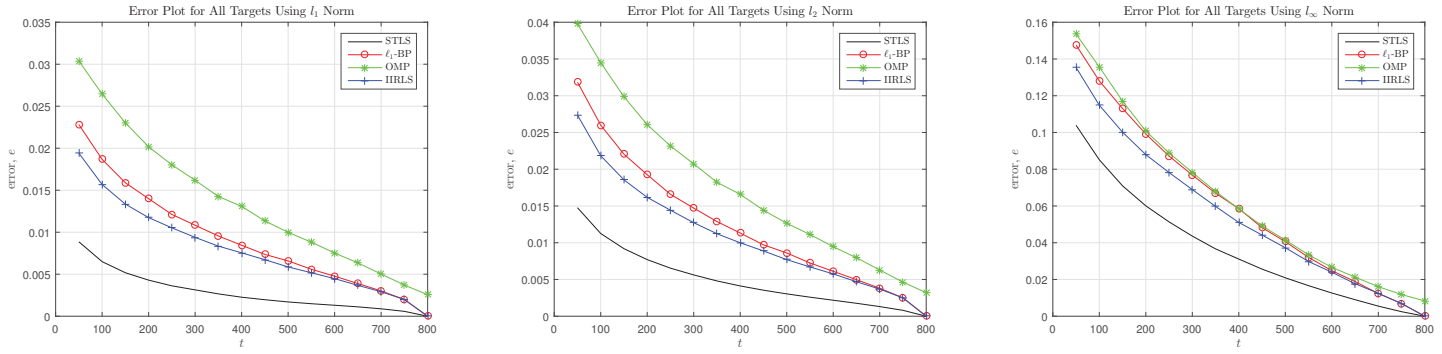
(a) (b) (c)

Fig. 9. Presented from left to right are the coefficient histograms for (a) α_2 ($\bar{\alpha}_2 = -0.0103, \tilde{b}_2 = 0.0262$), (b) α_{100} ($\bar{\alpha}_{100} = -2.3696e^{-5}, \tilde{b}_{100} = 0.0032$), and (c) α_{400} ($\bar{\alpha}_{400} = 1.5503e^{-6}, \tilde{b}_{400} = 7.9860e^{-4}$) for all ten target types.



(a) (b) (c) (d)

Fig. 10. (a) Original image (b) ℓ_1-BP reconstructed image using a Gaussian random measurement matrix (PSNR $= 51.2933$ dB, $m = 100$). (c) IIRLS (PSNR $= 54.7964$, $m = 100$) reconstructed image using a Gaussian random measurement matrix with over a $3\frac{1}{2}$ dB improvement in PSNR over ℓ_1-BP. (d) Stochastically trained least squares (PSNR $= 62.0743$) reconstructed image using $t = 100$ with over a 7 dB improvement in PSNR over IIRLS.



(a) (b) (c)

Fig. 11. Presented from left to right are ℓ_1, ℓ_2 and ℓ_∞ reconstruction error vs. measurements for six targets in the NVESD database using stochastic ℓ_2. The iterative methods, with the exception of OMP, were limited to M iterations where $M = 40$ for our simulations. OMP was limited to s iterations where s is the sparsity of the signal. The package $\ell_1 - MAGIC$ was used to implement ℓ_1 basis pursuit [37]

V. Conclusion and Future Research

Our stochastically trained least squares approach allow reconstruction of the images with a high PSNR so that subsequent processing (like target recognition) can occur without resampling. Our experimental results show that CQCF method has similar performance to more traditional approaches on accuracy, but it has minimal execution costs. We believe this algorithm could be implemented on a low-cost high resolution midwave infrared detector with negligible effects from the compression. This provides more options to design an autonomous weapon or fire control system that might employ a high resolution MWIR focal plane array.

The shift-invariance of the dictionary is an important feature for the ATR. An interesting area for future research is to find a shift-invariant dictionary that provides a better recognition accuracy.

Target detection is just the first step in an automatic target recognition system. In a typical ATR system, targets should be uniquely identified. Another interesting research area is how the compressed measurements for CQCF can be used in a target identification algorithm.

References

[1] V. C. Coffey, "Seeing in the dark: Defense applications of IR imaging," *Optics and Photonics News*, vol. 22, no. 4, pp. 26–31, 2011.

[2] X. Huang, R. Netravali, H. Man, and V. Lawrence, "Multi-sensor fusion of infrared and electro-optic signals for high resolution night images," *Sensors*, vol. 12, no. 8, pp. 10 326–10 338, 2012.

[3] G. Koretsky, J. Nicoll, and M. Taylor, "A tutorial on electro-optical/infrared (EO/IR) theory and systems," DTIC Document, Tech. Rep., 2013.

[4] E. Blasch and B. Kahler, "Multiresolution EO/IR target tracking and identification," in *2005 7th International Conference on Information Fusion (FUSION)*, vol. 1, July 2005, pp. 275–282.

[5] A. Mahalanobis and R. Muise, "Object specific image reconstruction using a compressive sensing architecture for application in surveillance systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1167–1180, 2009.

[6] A. Mahalanobis, R. Muise, and S. Roy, "Efficient target detection using an adaptive compressive imager," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 4, pp. 2528–2540, 2014.

[7] A. Mahalanobis, R. Shilling, R. Murphy, and R. Muise, "Recent results of medium wave infrared compressive sensing," *Applied Optics*, vol. 53, no. 34, pp. 8060–8070, Dec 2014.

[8] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Birkhäuser Basel, 2013.

[9] M. A. Davenport, M. F. Duarte, Y. Eldar, and G. Kutyniok, "Introduction to compressed sensing," in *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2011.

[10] E. J. Candès, "Compressive sampling," in *Proceedings of the International Congress of Mathematicians*, vol. 3. Madrid, Spain, 2006, pp. 1433–1452.

[11] E. J. Candès and J. Romberg, "Quantitative robust uncertainty principles and optimally sparse decompositions," *Foundations of Computational Mathematics*, vol. 6, no. 2, pp. 227–254, 2006.

[12] E. J. Candès, J. Romberg, and T. 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.

[13] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[14] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[15] M. A. Davenport, "Random observations on random observations: Sparse signal acquisition and processing," Ph.D. dissertation, Citeseer, 2010.

[16] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[17] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[18] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[19] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[20] D. E. Dudgeon and R. T. Lacoss, "An overview of automatic target recognition," *The Lincoln Laboratory Journal*, vol. 6, no. 1, pp. 3–10, 1993.

[21] J. Johnson, "Analysis of image forming systems," *Image Intensifier Symposium*, pp. 244–273, 1958.

[22] A. Mahalanobis, R. R. Muise, S. R. Stanfill, and A. Van Nevel, "Design and application of quadratic correlation filters for target detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 3, pp. 837–850, 2004.

[23] A. N. Vasile and R. M. Marino, "Pose-independent automatic target detection and recognition using 3d laser radar imagery," *Lincoln Laboratory Journal*, vol. 15, no. 1, pp. 61–78, 2005.

[24] F. Hollaus, M. Diem, and R. Sablatnig, *Computer Analysis of Images and Patterns: 16th International Conference, CAIP 2015, Valletta, Malta, September 2-4, 2015, Proceedings, Part II*. Springer International Publishing, 2015, p. 114.

[25] L. Zheng, A. Maleki, X. Wang, and T. Long, "Does ℓ_p-minimization outperform ℓ_1-minimization?" arXiv:1501.03704v1.

[26] T. T. Cai and A. Zhang, "Sharp RIP bound for sparse signal and low-rank matrix recovery," *Applied and Computational Harmonic Analysis*, vol. 35, pp. 74–93, 2013.

[27] Q. Sun, "Recovery of sparsest signals via l^q-minimization," *Applied and Computational Harmonic Analysis*, vol. 32, pp. 329–341, 2012.

[28] B. Millikan, A. Dutta, N. Rahnavard, Q. Sun, and H. Foroosh, "Initialized iterative reweighted least squares for automatic target recognition," in *Proceedings of IEEE Military Communications Conference 2015*, 2015, pp. 506–510.

[29] L. Baldassarre, Y.-H. Li, J. Scarlett, B. Gözcü, I. Bogunovic, and V. Cevher, "Learning-based compressive subsampling," arXiv:1510.06188.

[30] K. Fukunaga and W. L. Koontz, "Application of the Karhunen-Loève expansion to feature selection and ordering," *IEEE Transactions on Computers*, no. 4, pp. 311–318, 1970.

[31] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer Science & Business Media, 2010.

[32] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. John Wiley & Sons, 2012.

[33] S. Theodoridis and K. Koutroumbas, "Chapter 5 - Feature Selection," in *Pattern Recognition*, 4th ed. Boston: Academic Press, 2009, pp. 275–276.

[34] SENSIAC, "ATR algorithm development image database," http://www.sensiac.org/external/products/list_databases.jsf.

[35] E. Y. Lam and J. W. Goodman, "A mathematical analysis of the dct coefficient distributions for images," *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1661–1666, 2000.

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

[37] J. Romberg, "ℓ_1-MAGIC reference site," http://users.ece.gatech.edu/justin/l1magic/, 2005.