

High Resolution Image Reconstruction From a Sequence of Rotated and Translated Frames and its Application to an Infrared Imaging System

Russell C. Hardie^a, Kenneth J. Barnard^b, John G. Bogner^c,
Ernest E. Armstrong^c and Edward A. Watson^b

^a Department of Electrical and Computer Engineering
University of Dayton, 300 College Park Avenue
Dayton, OH 45469-0226

^b Air Force Wright Laboratory
Electro-Optics Sensor Technology Branch
3109 P Street, Wright-Patterson Air Force Base
Dayton, OH 45433-7700

^c Technology/Scientific Services, Inc. P.O. Box 3065
Dayton, OH 45437-3065

ABSTRACT

Some imaging systems employ detector arrays which are not sufficiently dense so as to meet the Nyquist criteria during image acquisition. This is particularly true for many staring infrared imagers. Thus, the full resolution afforded by the optics is not being realized in such a system. This paper presents a technique for estimating a high resolution image, with reduced aliasing, from a sequence of undersampled rotated and translationally shifted frames. Such an image sequence can be obtained if an imager is mounted on a moving platform, such as an aircraft. Several approaches to this type of problem have been proposed in the literature. Here we extend some of this previous work. In particular, we define an observation model which incorporates knowledge of the optical system and detector array. The high resolution image estimate is formed by minimizing a new regularized cost function which is based on the observation model. We show that with the proper choice of a tuning parameter, our algorithm exhibits robustness in the presence of noise. We consider both gradient descent and conjugate gradient optimization procedures to minimize the cost function. Detailed experimental results are provided to illustrate the performance of the proposed algorithm using digital video from an infrared imager.

Keywords: high resolution, image reconstruction, aliasing, infrared, Nyquist, minimization

*Submitted to Optical Engineering April 1997.
Revised July 1997.*

Send correspondence to Russell C. Hardie. Email: rhardie@engr.udayton.edu; Telephone: 513-229-3178; Fax: 513-229-2756.

1. INTRODUCTION

In imaging systems where the focal plane detector array is not sufficiently dense so as to meet the Nyquist criterion, the resulting images will be degraded by aliasing effects. This undersampling is a common problem among staring infrared imaging systems. This is due to fabrication complexities associated with manufacturing dense infrared focal plane arrays (FPAs). While it is possible to equip such systems with optics to properly bandlimit the input, this generally means employing optics with a very small instantaneous field of view (IFOV). This may be highly undesirable in some applications. Furthermore, charge coupled device (CCD) cameras may also employ detector arrays which are not sufficiently dense for the desired optics. Thus, the goal of this work is to obtain high resolution images, with reduced aliasing, from such systems. In addition, we wish to remove the blurring effects of the finite size detectors and the optics.

One way to overcome the undersampling problem, without low pass filtering the input or sacrificing IFOV, is to exploit multiple frames from an image sequence. This is possible if there is relative motion between the scene and the FPA during image sequence acquisition. In this case, a unique “look” at the scene (i.e., a unique set of samples) may be provided by each frame. The desired image sequence can be obtained if an imager is mounted on a moving platform, such as an aircraft. We refer to this as uncontrolled microscanning.¹ It is also possible to introduce a controlled mirror in the optical path to translate the scene intensity image across the FPA. This is referred to as controlled microscanning.^{1,2} Here we focus on uncontrolled microscanning and we consider both rotational and translational motion of the scene relative to the FPA. The key to the high resolution image recovery algorithm is accurate knowledge of the sub-pixel translation and rotation of each frame. If these parameters are unknown, as in the uncontrolled case, they must be estimated from the observed images. Thus, we must consider both image registration and high resolution image reconstruction.

Several approaches to the high resolution image reconstruction problem have been proposed in the literature.³⁻¹¹ A maximum likelihood (ML) technique using the expectation maximization (EM) algorithm has been developed and applied to forward looking infrared (FLIR) data.¹² This method seeks to jointly estimate translational shifts and a high resolution image. In addition, a joint registration and high resolution reconstruction technique using maximum *a posteriori* (MAP) estimation has been proposed.¹³ These ML and MAP algorithms do not explicitly treat the case of rotational motion which is the focus of this paper.

The reconstruction algorithm presented here can be viewed as an extension of the basic approach presented by Irani and Peleg.¹¹ The technique proposed by Irani and Peleg seeks to minimize a specified cost function using an iterative algorithm.¹¹ This cost function is the total squared error between the observed low resolution data and the predicted low resolution data. The predicted data is the result of projecting the high resolution image estimate through the observation model. More will be said about this shortly. Here we employ the same registration technique and we seek to minimize a related cost function. However, our approach includes a number of differences.

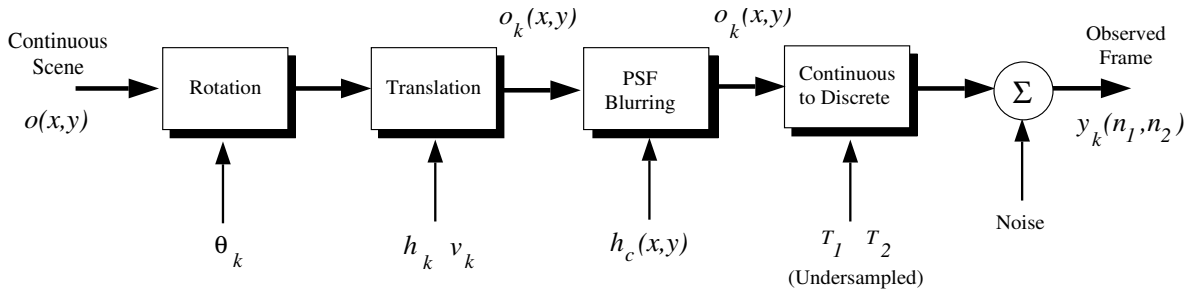


Figure 1. Continuous observation model resembling the physical process of image acquisition.

In particular, the observation model uses information about the optical system and FPA to form a theoretical point spread function (PSF). Also, the new cost function defined here includes a regularization term. This extra term gives the cost function the desirable property of having a unique global minimum. We show that this term can add robustness, particularly when the fidelity of the data is low. We employ a gradient descent and a conjugate gradient technique for minimizing the cost function to form our estimate. We show that the conjugate gradient technique, in particular, provides rapid convergence. Finally, here we investigate the application of high resolution image reconstruction to a real-time infrared imaging system.

The organization of the rest of the paper is as follows. In Section 2 the observation model is described. Both a continuous and discrete model are developed. Image registration is addressed in Section 3. The high resolution image reconstruction algorithm is developed in Section 4. In particular, the regularized cost function is defined and two optimization procedures are described. Experimental results are provided in Section 5. These include results obtained using simulated data and using FLIR images acquired from a real-time system. Quantitative error analysis is provided and several images are shown for subjective evaluation. Finally, some conclusions are given in Section 6.

2. OBSERVATION MODEL

In this section, the observation model is presented. This model is the basis for the high resolution reconstruction algorithm developed in Section 4. We begin with a continuous model which closely follows the physical image acquisition process. An equivalent discrete model is then presented. It is the discrete model which is utilized in the reconstruction algorithm. We conclude this section with the characterization of the system point spread function, since this represents a key element in the observation model.

2.1. Continuous Model

A block diagram of the continuous observation model is shown in Fig. 1. The true scene intensity image is defined as $o(x, y)$. The motion of the imager that occurs between image acquisitions is modeled as a pure rotation and translation of the scene intensity image. For a moving imager and a stationary scene in the far field, this is a fairly accurate model, since occlusion effects and perspective changes are minimal. Thus, the k 'th observed frame in a

sequence can be expressed as

$$o_k(x, y) = o(x\cos\theta_k - y\sin\theta_k + h_k, y\cos\theta_k + x\sin\theta_k + v_k), \quad (1)$$

for $k = 1, 2, 3, \dots, p$. Note that θ_k represents the rotation of the k 'th frame about the origin (i.e., $x = 0, y = 0$). The parameters h_k and v_k represent the horizontal and vertical shift associated with the k 'th frame.

The blurring effect of the optics and finite detector size is modeled by a convolution operation yielding

$$\tilde{o}_k(x, y) = o_k(x, y) * h_c(x, y), \quad (2)$$

where $h_c(x, y)$ is the continuous system PSF. More will be said about the PSF in Section 2. Finally, the blurred, rotated and translated image is sampled below the Nyquist rate and corrupted by noise. This yields the k 'th low resolution observed frame

$$y_k(n_1, n_2) = \tilde{o}_k(n_1T_1, n_2T_2) + \eta_k(n_1, n_2), \quad (3)$$

where T_1 and T_2 are the horizontal and vertical sample spacings and $\eta_k(n_1, n_2)$ is an additive noise term. Ideally, the FPA performance is limited by the photon or shot noise. In this case, the signal follows Poisson statistics.¹² For high light levels, however, we believe that signal independent additive Gaussian noise is a sufficiently accurate and tractable model.

Let the dimensions of the low resolution image $y_k(n_1, n_2)$ be $N_1 \times N_2$. These data in lexicographical notation will be expressed as $\mathbf{y}_k = [y_{k,1}, y_{k,2}, \dots, y_{k,M}]^T$, where $M = N_1N_2$. Finally, let the full set of p observed low resolution images be denoted

$$\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \dots, \mathbf{y}_p^T]^T = [y_1, y_2, \dots, y_{pM}]^T. \quad (4)$$

Thus, all observed pixel values are contained within \mathbf{y} .

2.2. Discrete Model

While the continuous model provides insight into the physical process, we require a discrete observation model to develop the high resolution reconstruction algorithm. That is, we need a model relating a discrete high resolution image to the low resolution observed frames \mathbf{y} . Figure 2 illustrates such a discrete model which is equivalent to that in Fig. 1. The difference here is that we first define $z(n_1, n_2)$ to be an intensity image sampled at or above the Nyquist rate with no blur or noise degradation. It is this discrete image we wish to estimate from the observed frames. Let this high resolution image be of size $L_1N_1 \times L_2N_2 = N$, where L_1 and L_2 are positive integers. More will be said about these parameters shortly. In later analysis it will be convenient to express this image in lexicographical notation as the vector $\mathbf{z} = [z_1, z_2, \dots, z_N]^T$.

As before, this image is rotated by θ_k and shifted by h_k and v_k producing $z_k(n_1, n_2)$. Here we will define the shifts h_k and v_k in terms of low resolution pixel spacings for convenience. Note that this step requires interpolation

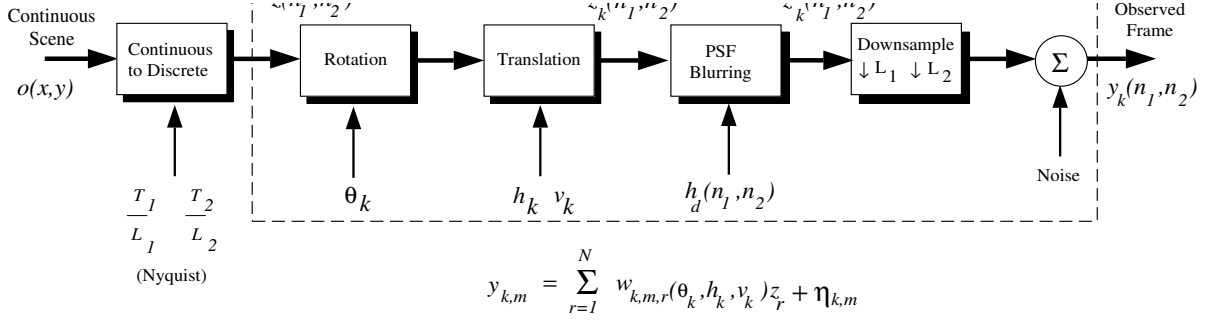


Figure 2. Equivalent discrete observation model illustrating the relationship between the ideally sampled image \mathbf{z} and the observed frames \mathbf{y} .

since the sampling grid changes in the geometric transformation. Theoretically, one could use ideal interpolation since $z(n_1, n_2)$ is defined to be sampled above the Nyquist rate. However, in practice, simpler interpolation methods such as nearest neighbor and bilinear interpolation can be used.¹⁴ For large values of L_1 and L_2 , the high resolution grid is so dense that simple interpolation methods can be reasonably accurate.

Next, the system PSF is accounted for yielding

$$\tilde{z}_k(n_1, n_2) = z_k(n_1, n_2) * h_d(n_1, n_2), \quad (5)$$

where $h_d(n_1, n_2)$ is the equivalent discrete system PSF. Note that the blurring is performed after the geometric transformation. If the PSF is circularly symmetric, then the blurring can be equivalently introduced prior to the rotation. This saves on computations since the blurring is performed only once in this case. Finally, the transformed image is subsampled down to the resolution of the observed frames yielding

$$y_k(n_1, n_2) = \tilde{z}_k(n_1 L_1, n_2 L_2) + \eta_k(n_1, n_2). \quad (6)$$

Aliasing is generally introduced in this final step.

This discrete model can be rewritten in a simple form where the low resolution pixels are defined as a weighted sum of the appropriate high resolution pixels with additive noise. This generalized form can account for the PSF blurring, the geometric transformation and the subsampling of the high resolution image. Specifically, the observed low resolution pixels from frame k are related to the high resolution image as follows

$$y_{k,m} = \sum_{r=1}^N w_{k,m,r}(\theta_k, h_k, v_k) z_r + \eta_{k,m}, \quad (7)$$

for $m = 1, 2, \dots, M$ and $k = 1, 2, \dots, p$. The weight $w_{k,m,r}(\theta_k, h_k, v_k)$ represents the “contribution” of the r 'th high resolution pixel to the m 'th low resolution observed pixel of the k 'th frame. The parameters θ_k , h_k and v_k represent the rotation, horizontal and vertical translational shifts, respectively, of the k 'th frame with respect to

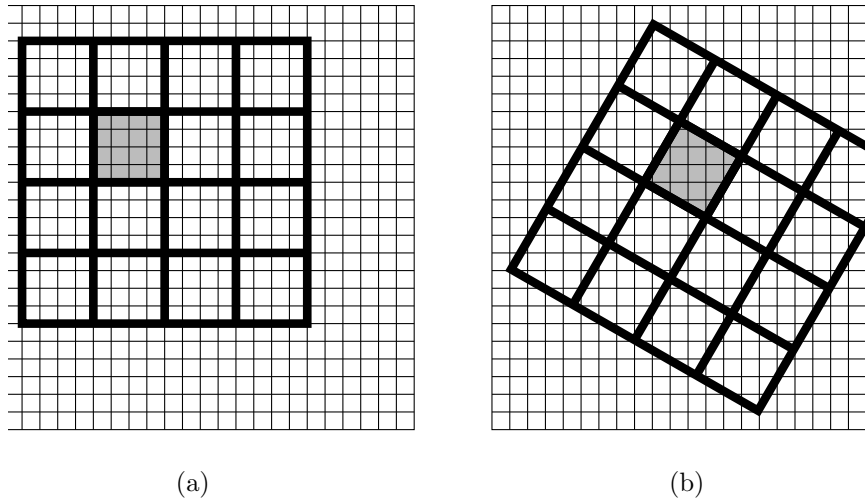


Figure 3. Discrete detector model showing those “virtual” high resolution pixels that contribute to a low resolution pixel for two different registration positions.

some reference on the high resolution grid. The term $\eta_{k,m}$ in (7) represents an additive noise sample. To further compact the notation, the model in (7) can be expressed in terms of the entire set of low resolution pixels as

$$y_m = \sum_{r=1}^N w_{m,r} z_r + \eta_m, \quad (8)$$

for $m = 1, 2, \dots, pM$ and where $w_{m,r}$ is simply the contribution of the r 'th high resolution pixel in \mathbf{z} to the m 'th low resolution pixel in \mathbf{y} . It is assumed that the underlying scene, \mathbf{z} , remains constant during the acquisition of the multiple low resolution frames. Furthermore, we assume here that the only frame to frame differences in the weights result from rotation and translational of each low resolution frame relative to the high resolution grid.

A simple way to visualize the form of the observation model in (8) is to consider only the blur from the finite detector size. This scenario is illustrated in Fig. 3. Here each low resolution pixel is obtained by summing the “virtual” high resolution pixels within the support of that low resolution detector. One low resolution detector in Fig. 3(a) is shaded to illustrate this point. This discrete detector model simulates the integration of light intensity that falls within the span of the low resolution detector. As the low resolution grid shifts relative to the fixed high resolution grid, as in Fig. 3(b), a different set of high resolution pixels contribute to each low resolution pixel. This yields a new set of linearly independent equations from (8). Clearly, some type of interpolation is required for any non-integer shift on the high resolution grid or any non-trivial rotation. This interpolation can be accounted for by modifying the weights in (8). This simple detector model can give good results. However, a more realistic PSF model is described in the following section.

2.3. System Point Spread Function

For most systems, there are two main contributors to the system PSF. The primary contributor is generally the finite detector size as illustrated in Fig. 3. This effect is spatially invariant for a uniform detector array. The second contributor is the optics. Here we assume an isoplanatic model for the optics.¹⁵ We derive and use a theoretical PSF because, for the type of systems considered in this paper, direct measurement of an *unaliased* system PSF is not possible.

Let us begin by considering a system with a uniform detector array. The effect of the integration of light intensity over the span of the detectors can be modeled as a linear convolution operation with a PSF determined by the geometry of a single detector. Let this PSF be denoted $d(x, y)$. Applying the Fourier transform to $d(x, y)$ yields the effective continuous frequency response resulting from the detectors

$$D(u, v) = \mathcal{FT}\{d(x, y)\}, \quad (9)$$

where $\mathcal{FT}\{\cdot\}$ represents the continuous Fourier transform. Next, define the incoherent optical transfer function (OTF) of the optics to be $H_o(u, v)$. The overall system OTF is given by the product of these, yielding

$$H(u, v) = D(u, v)H_o(u, v). \quad (10)$$

The overall continuous system PSF is then given by

$$h_c(x, y) = \mathcal{FT}^{-1}\{H(u, v)\}, \quad (11)$$

where $\mathcal{FT}^{-1}\{\cdot\}$ represents the inverse Fourier transform. Finally, the impulse-invariant discrete system PSF¹⁶ on the high resolution grid is obtained by sampling the continuous PSF such that

$$h_d(n_1, n_2) = \frac{T_1 T_2}{L_1 L_2} h_c\left(\frac{n_1 T_1}{L_1}, \frac{n_2 T_2}{L_2}\right). \quad (12)$$

This accurately represents the continuous blurring when the effective sampling frequency L_1/T_1 exceeds two times the horizontal cutoff frequency of $H(u, v)$ and L_2/T_2 exceeds two times the vertical cutoff frequency.¹⁶

Let us now specifically consider a system with uniform rectangular detectors. An illustration of such a detector array with critical dimensions labeled is provided in Fig. 4. The shaded areas represent the active region of each detector. The detector model PSF in this case is given by

$$d(x, y) = \frac{1}{ab} \text{rect}\left(\frac{x}{a}, \frac{y}{b}\right) = \begin{cases} 1 & \text{for } |x/a| < 1/2 \text{ and } |y/b| < 1/2 \\ 0 & \text{otherwise} \end{cases}. \quad (13)$$

Let the active region dimensions, a and b , be measured in millimeters (mm). Thus, the effective continuous frequency response resulting from the detectors is

$$D(u, v) = \text{sinc}(au, bv) = \frac{\sin(\pi au)\sin(\pi bv)}{\pi^2 aubv}, \quad (14)$$

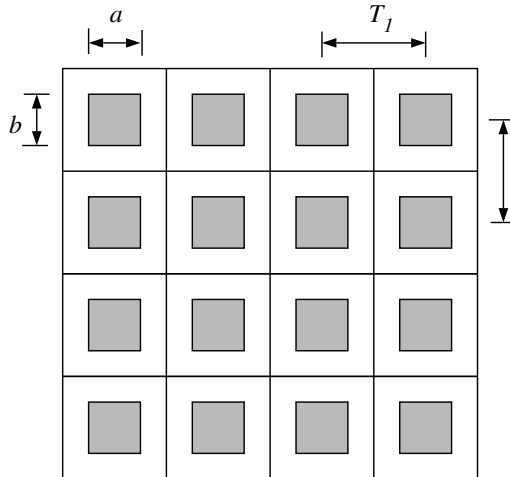


Figure 4. Uniform detector array illustrating critical dimensions.

where u and v are the horizontal and vertical frequencies measured in cycles/mm.

The incoherent optical transfer function (OTF) of diffraction-limited optics with a circular exit pupil can be found¹⁵ as

$$H_o(u, v) = \begin{cases} \frac{2}{\pi} \left[\cos^{-1} \left(\frac{\rho}{\rho_c} \right) - \frac{\rho}{\rho_c} \sqrt{1 - \left(\frac{\rho}{\rho_c} \right)^2} \right] & \text{for } \rho < \rho_c \\ 0 & \text{otherwise} \end{cases}, \quad (15)$$

where $\rho = \sqrt{u^2 + v^2}$. The parameters ρ_c is the radial system cutoff frequency given by

$$\rho_c = \frac{1}{\lambda f/\#}, \quad (16)$$

where $f/\#$ is the f-number of the optics and λ is the wavelength of light considered. Since the cutoff of $H_o(u, v)$ is ρ_c , so is the cutoff of the overall system $H(u, v)$. Thus, the impulse-invariant discrete system defined in (12) will accurately model the continuous system when $L_1 \geq [2\rho_c T_1]$ and $L_2 \geq [2\rho_c T_2]$. This choice of L_1 and L_2 also defines a high resolution sampling grid at or above the Nyquist rate for an arbitrary scene. That is, the effective high resolution sampling rates of L_1/T_1 and L_2/T_2 will be more than twice the OTF cutoff frequency.

Figure 5 shows an example of $D(u, v)$, $H_o(u, v)$, $H(u, v)$ and $h_c(x, y)$ for a particular imaging system. The system considered is the FLIR imager used to collect data for the experimental results presented in Section 5. The FLIR camera uses a 128×128 Amber AE-4128 infrared FPA. The FPA is composed of Indium-Antimonide (InSb) detectors with a response in the $3\mu\text{m} - 5\mu\text{m}$ wavelength band. This system has square detectors of size $a = b = .040\text{mm}$. The imager is equipped with $100\text{mm } f/3$ optics. The center wavelength, $\lambda = .004\text{mm}$, is used in the OTF calculation. Figure 5(a) shows the effective modulation transfer function (MTF) of the detectors, $|D(u, v)|$. The diffraction-limited OTF for the optics, $H_o(u, v)$, is shown in Fig. 5(b). Note that the cutoff frequency is 83.3 cycles/mm. The

overall system MTF, $|H(u, v)|$, is plotted in Fig. 5(c). Finally, the continuous system PSF, $h_c(x, y)$, is plotted in Fig. 5(d).

The detector spacing on the Amber FPA is $T_1 = T_2 = .050\text{mm}$, yielding a sampling frequency of 20 cycles/mm in both directions. Thus, the effective sampling rate must be increased by a factor of 8.33 to eliminate aliasing entirely for an arbitrary scene. This would require that we select $L_1, L_2 \geq 9$. In practice, we find that good results can be obtained with smaller values of L_1 and L_2 .

3. IMAGE REGISTRATION

In most applications, the registration parameters in the observation model, θ_k, h_k and v_k , will not be known *a priori*. Thus, they must be estimated from the observed image sequence. Accurate sub-pixel registration is the key to the success of the high resolution image reconstruction algorithm. A number of image registration techniques have been proposed in the literature. We have found that in some cases a practical and effective method of estimating the sub-pixel translation is using an iterative gradient based technique.^{11,17} For convenience, this algorithm is presented here using the current notation.

To begin, define the first observed frame to be the reference frame and without loss of generality let $\theta_1 = h_1 = v_1 = 0$. According to our model,

$$o_k(x, y) = o_1(x\cos\theta_k - y\sin\theta_k + h_k, y\cos\theta_k + x\sin\theta_k + v_k), \quad (17)$$

for $k = 2, 3, \dots, p$. Note that this assumes that the center of rotation is at the origin (i.e., $x = 0, y = 0$). This is not restrictive however, since we allow any shift h_k and v_k . If the PSF blur is approximately circularly symmetric, then

$$\tilde{o}_k(x, y) \approx \tilde{o}_1(x\cos\theta_k - y\sin\theta_k + h_k, y\cos\theta_k + x\sin\theta_k + v_k). \quad (18)$$

For very small values of θ_k , we can make the following approximations: $\sin\theta_k \approx \theta_k$ and $\cos\theta_k \approx 1$. Using these yields

$$\tilde{o}_k(x, y) \approx \tilde{o}_1(x - y\theta_k + h_k, y + x\theta_k + v_k). \quad (19)$$

Now we use the first three terms of the Taylor series expansion as an approximation for the right side in (19). This yields

$$\tilde{o}_k(x, y) \approx \tilde{o}_1(x, y) + (h_k - y\theta_k)g_x(x, y) + (v_k + x\theta_k)g_y(x, y), \quad (20)$$

where $g_x(x, y) = \frac{\partial \tilde{o}_1(x, y)}{\partial x}$ and $g_y(x, y) = \frac{\partial \tilde{o}_1(x, y)}{\partial y}$.

In light of the relationship expressed in (20), we define the least squares estimates for the registration parameters as follows

$$\hat{\theta}_k, \hat{h}_k, \hat{v}_k = \arg \min_{\theta_k, h_k, v_k} E_k(\theta_k, h_k, v_k), \quad (21)$$

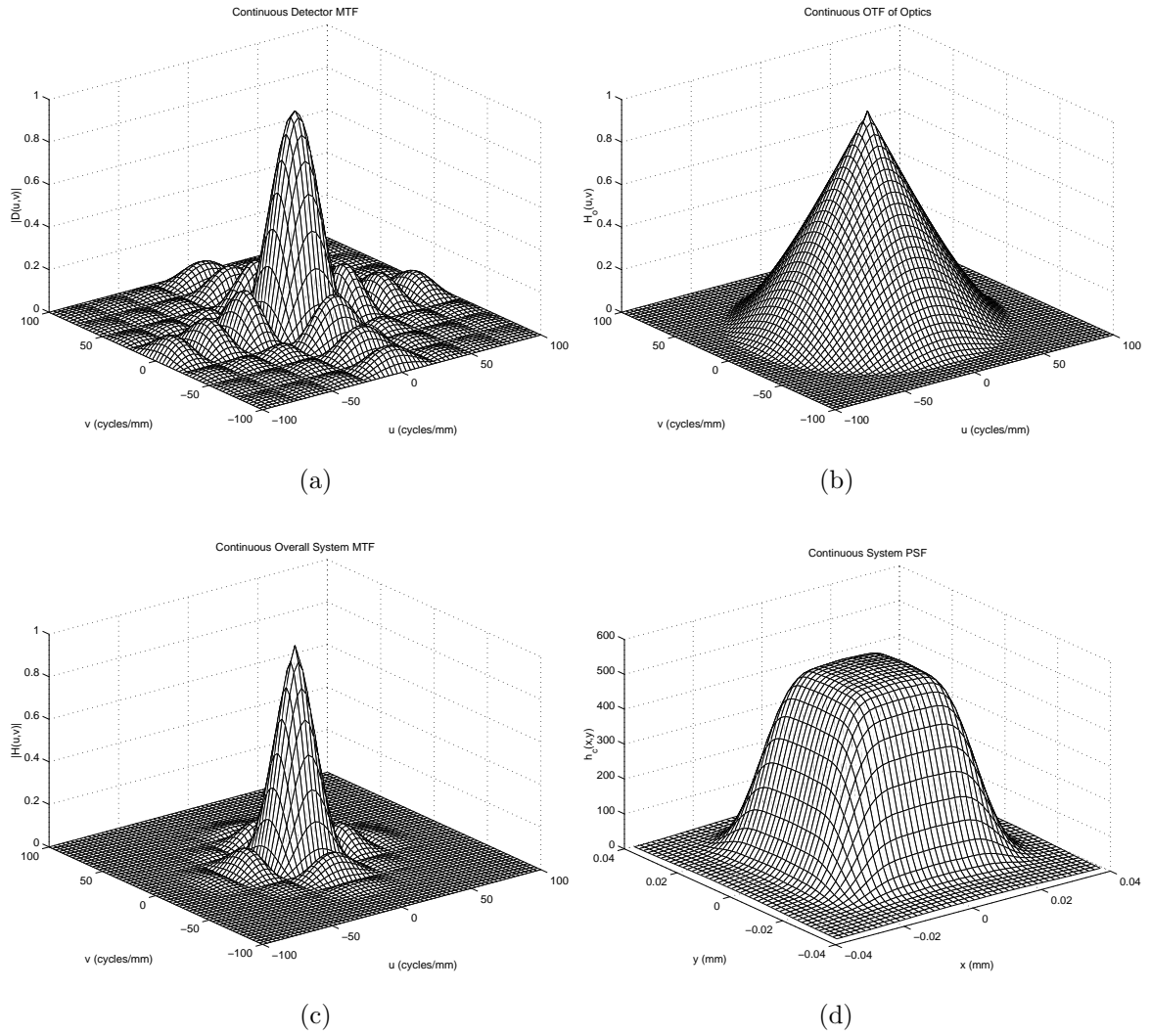


Figure 5. (a) Effective MTF of the detectors in the FLIR imager (b) diffraction-limited OTF of the optics (c) overall system MTF (d) overall continuous system PSF.

where

$$E_k(\theta_k, h_k, v_k) = \sum_{(x,y) \in S} (\tilde{\delta}_k(x,y) - \delta_1(x,y) - (h_k - y\theta_k)g_x(x,y) - (v_k + x\theta_k)g_y(x,y))^2. \quad (22)$$

Note that S represents the grid of points in the R^2 space, defined by x and y , at which we have discrete samples. Rewriting this error in terms of the observed images yields

$$E_k(\theta_k, h_k, v_k) = \sum_{\mathbf{n} \in \mathcal{N}} (y_k(\mathbf{n}) - y_1(\mathbf{n}) - (h_k - n_2 T_2 \theta_k) \hat{g}_x(\mathbf{n}) - (v_k + n_1 T_1 \theta_k) \hat{g}_y(\mathbf{n}))^2, \quad (23)$$

where $\mathbf{n} = [n_1, n_2]$ and \mathcal{N} is the set of indices on the low resolution discrete grid for which we have observations. Note that the center of rotation on the discrete grid is assumed to be at $n_1 = n_2 = 0$. The functions $\hat{g}_x(\mathbf{n})$ and $\hat{g}_y(\mathbf{n})$ are discrete estimates of $g_x(x,y)$ and $g_y(x,y)$, respectively, at location $x = n_1 T_1$ and $y = n_2 T_2$. These can be computed using scaled Prewitt operators,¹⁴ for example.

To solve the minimization problem in (21), we begin by differentiating $E(\theta_k, h_k, v_k)$ with respect to θ_k, h_k and v_k and set the derivatives equal to zero. This yields the following three equations

$$\sum_{\mathbf{n} \in \mathcal{N}} (h_k \hat{g}_x^2(\mathbf{n}) + v_k \hat{g}_x(\mathbf{n}) \hat{g}_y(\mathbf{n}) + \theta_k \bar{g}(\mathbf{n}) \hat{g}_x(\mathbf{n})) = \sum_{\mathbf{n} \in \mathcal{N}} \bar{y}_k(\mathbf{n}) \hat{g}_x(\mathbf{n}), \quad (24)$$

$$\sum_{\mathbf{n} \in \mathcal{N}} (h_k \hat{g}_x(\mathbf{n}) \hat{g}_y(\mathbf{n}) + v_k \hat{g}_y^2(\mathbf{n}) + \theta_k \bar{g}(\mathbf{n}) \hat{g}_y(\mathbf{n})) = \sum_{\mathbf{n} \in \mathcal{N}} \bar{y}_k(\mathbf{n}) \hat{g}_y(\mathbf{n}), \quad (25)$$

and

$$\sum_{\mathbf{n} \in \mathcal{N}} (h_k \bar{g}(\mathbf{n}) \hat{g}_x(\mathbf{n}) + v_k \bar{g}(\mathbf{n}) \hat{g}_y(\mathbf{n}) + \theta_k \bar{g}^2(\mathbf{n})) = \sum_{\mathbf{n} \in \mathcal{N}} \bar{y}_k(\mathbf{n}) \bar{g}(\mathbf{n}), \quad (26)$$

where

$$\bar{g}(\mathbf{n}) = n_1 T_1 \hat{g}_y(\mathbf{n}) - n_2 T_2 \hat{g}_x(\mathbf{n}) \quad (27)$$

and

$$\bar{y}_k(\mathbf{n}) = y_k(\mathbf{n}) - y_1(\mathbf{n}). \quad (28)$$

We then simultaneously solve these expressions. To do so let

$$MR_k = V_k, \quad (29)$$

where

$$M = \begin{bmatrix} \sum_{\mathbf{n} \in \mathcal{N}} \hat{g}_x^2(\mathbf{n}) & \sum_{\mathbf{n} \in \mathcal{N}} \hat{g}_x(\mathbf{n}) \hat{g}_y(\mathbf{n}) & \sum_{\mathbf{n} \in \mathcal{N}} \bar{g}(\mathbf{n}) \hat{g}_x(\mathbf{n}) \\ \sum_{\mathbf{n} \in \mathcal{N}} \hat{g}_x(\mathbf{n}) \hat{g}_y(\mathbf{n}) & \sum_{\mathbf{n} \in \mathcal{N}} \hat{g}_y^2(\mathbf{n}) & \sum_{\mathbf{n} \in \mathcal{N}} \bar{g}(\mathbf{n}) \hat{g}_y(\mathbf{n}) \\ \sum_{\mathbf{n} \in \mathcal{N}} \bar{g}(\mathbf{n}) \hat{g}_x(\mathbf{n}) & \sum_{\mathbf{n} \in \mathcal{N}} \bar{g}(\mathbf{n}) \hat{g}_y(\mathbf{n}) & \sum_{\mathbf{n} \in \mathcal{N}} \bar{g}^2(\mathbf{n}) \end{bmatrix}, \quad (30)$$

$R_k = [h_k, v_k, \theta_k]^T$, and

$$V_k = \begin{bmatrix} \sum_{\mathbf{n} \in \mathcal{N}} \bar{y}_k(\mathbf{n}) \hat{g}_x(\mathbf{n}) \\ \sum_{\mathbf{n} \in \mathcal{N}} \bar{y}_k(\mathbf{n}) \hat{g}_y(\mathbf{n}) \\ \sum_{\mathbf{n} \in \mathcal{N}} \bar{y}_k(\mathbf{n}) \bar{g}(\mathbf{n}) \end{bmatrix}. \quad (31)$$

Finally, the estimated registration vector, $\hat{R}_k = [\hat{h}_k, \hat{v}_k, \hat{\theta}_k]^T$, can be computed as

$$\hat{R}_k = M^{-1} V_k. \quad (32)$$

Table 1. Iterative registration algorithm for $y_k(\mathbf{n})$, where $k = 2, 3, \dots, p$.

<p>step 1: Compute $\hat{g}_x(\mathbf{n})$, $\hat{g}_y(\mathbf{n})$ and $\bar{g}(\mathbf{n})$ from $y_1(\mathbf{n})$ and form M matrix.</p> <p>step 2: Let $n = 0$, $y_k^0(\mathbf{n}) = y_k(\mathbf{n})$ and $\hat{R}_k^0 = [0, 0, 0]^T$.</p> <p>step 3: Compute V_k using $\bar{y}_k(\mathbf{n}) = y_k^n(\mathbf{n}) - y_1(\mathbf{n})$ and then let $n = n + 1$.</p> <p>step 4: Compute $\hat{R}_k^n = M^{-1}V_k + \hat{R}_k^{n-1}$.</p> <p>step 5: if $\ \hat{R}_k^n - \hat{R}_k^{n-1}\ /\ \hat{R}_k^{n-1}\ < T$, let $\hat{R}_k = \hat{R}_k^n$ and stop.</p> <p>step 6: Resample $y_k(\mathbf{n})$ towards $y_1(\mathbf{n})$ according to \hat{R}_k^n to create $y_k^n(\mathbf{n})$ and go to step 3.</p>
--

To obtain shifts in terms of low resolution pixel spacings (rather than in mm), we set $T_1 = T_2 = 1$ in (27).

Because of the assumptions made, this technique is only accurate for small shifts and rotations. To treat the case where larger values are expected, we follow an iterative method.^{11,17} In this technique, the initial registration parameters are estimated according to (32). Next $y_k(\mathbf{n})$ is shifted and rotated according to the registration parameter estimates so as to more closely match $y_1(\mathbf{n})$. This modified image is then registered to $y_1(\mathbf{n})$. The process continues whereby $y_k(\mathbf{n})$ is continually modified until the registration estimates become sufficiently small. The final registration estimate is obtained by summing all of the ‘‘partial’’ estimates. The iterative registration procedure for $y_k(\mathbf{n})$, where $k = 2, 3, \dots, p$, is summarized in Table 1. Based on empirical observation, the algorithm appears to converge reliably when the shifts and rotations are not too large (e.g., less than 10 pixel shifts and 10 degree rotations).

Because the three parameters are well over-determined by the data, this least squares estimate is generally accurate. We find that the main source of error lies in the resampling of $y_k(\mathbf{n})$ (step 6 in Table 1), since this requires interpolation on the low resolution aliased grid. Some error is also introduced in the discrete gradient estimates. However, the algorithm appears to provide sufficiently accurate results for the data studied here. In cases where the aliasing is more severe, a joint registration and reconstruction algorithm may be advantageous.^{12,13,10}

4. HIGH RESOLUTION IMAGE RECONSTRUCTION

With estimates of the registration parameters, the observation model can be completely specified. In light of the observation model in (8), we define the high resolution image estimate to be

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z}} C(\mathbf{z}), \quad (33)$$

where

$$C(\mathbf{z}) = \frac{1}{2} \sum_{m=1}^{pM} \left(y_m - \sum_{r=1}^N w_{m,r} z_r \right)^2 + \frac{\lambda}{2} \sum_{i=1}^N \left(\sum_{j=1}^N \alpha_{i,j} z_j \right)^2, \quad (34)$$

and y_m for $m = 1, 2, \dots, pM$ are the observed pixel values. Thus, the estimate $\hat{\mathbf{z}}$ is the \mathbf{z} which minimizes (34).

The cost function in (34) balances two types of errors. The first term on the right hand side is minimized when a candidate \mathbf{z} , projected through the observation model, matches the observed data. To clarify this, note that $\sum_{r=1}^N w_{m,r} z_r$ can be thought of as the “predicted” low resolution pixel m , obtained by projecting a candidate \mathbf{z} through the observation model, and y_m is the actual observed pixel value. However, direct minimization of this term alone can lead to a poor estimate due to the ill-posed nature of the inverse problem. That is, \mathbf{z} is often under-determined by the observed data and many candidate \mathbf{z} ’s will minimize this first term. Thus, the second term serves as a regularization operator. The parameters $\alpha_{i,j}$ are generally selected so that this regularization term is minimized when \mathbf{z} is smooth. This is a commonly applied constraint on image restoration problems since natural scenes often exhibit strong spatial correlation. Here we select

$$\alpha_{i,j} = \begin{cases} 1 & \text{for } i = j \\ -1/4 & \text{for } j : z_j \text{ is a cardinal neighbor of } z_i. \end{cases} \quad (35)$$

The cost function is convex and is readily differentiable. Furthermore, the regularization term provides it with the desirable property of having a unique global minimum which yields a unique optimal estimate image $\hat{\mathbf{z}}$.

The “weight” of the two competing terms in the cost function is controlled by the tuning parameter λ . Larger values of λ will generally lead to a smoother solution. This is useful when only a small number of frames is available or the fidelity of the observed data is low. On the other hand, if λ is too small, the resulting estimate may appear noisy. It is also possible to make λ spatially adaptive.^{18,19} Finally, it can be shown that the estimate defined in (33) and (34) is a MAP estimate in the case of Gaussian noise and where \mathbf{z} is viewed as a realization of a particular Gibbs random field.¹³

Next we consider two unconstrained optimization techniques for minimizing the cost function in (34). First we consider a gradient descent approach and then we present a conjugate gradient method.

4.1. Gradient descent optimization

To derive the gradient descent update procedure²⁰ for the image estimate, we begin by differentiating (34) with respect to some pixel z_k for $k = 1, 2, \dots, N$. This partial derivative is given by

$$g_k(\mathbf{z}) = \frac{\partial C(\mathbf{z})}{\partial z_k} = \sum_{m=1}^{pM} w_{m,k} \left(\sum_{r=1}^N w_{m,r} z_r - y_m \right) + \lambda \sum_{i=1}^N \alpha_{i,k} \left(\sum_{j=1}^N \alpha_{i,j} z_j \right). \quad (36)$$

The iterative procedure begins with an initial estimate of the high resolution image $\hat{\mathbf{z}}^0$. A relatively simple starting point can be obtained by interpolating the first low-resolution frame using standard bilinear or bicubic interpolation.¹⁴ The gradient descent update for each pixel estimate is

$$\hat{z}_k^{n+1} = \hat{z}_k^n - \varepsilon^n g_k(\hat{\mathbf{z}}^n), \quad (37)$$

for $n = 0, 1, 2, \dots$ and $k = 1, 2, \dots, N$. Alternatively, the update can be written as

$$\hat{\mathbf{z}}^{n+1} = \hat{\mathbf{z}}^n - \varepsilon^n \mathbf{g}^n, \quad (38)$$

where

$$\mathbf{g}^n = \begin{bmatrix} g_1(\hat{\mathbf{z}}^n) \\ g_2(\hat{\mathbf{z}}^n) \\ \vdots \\ g_N(\hat{\mathbf{z}}^n) \end{bmatrix}. \quad (39)$$

The parameter ε^n in (37) and (38) represents the step size at the n 'th iteration. This parameter must be selected to be small enough to prevent divergence and large enough to provide convergence in a reasonable number of iterations. The optimal step size can be calculated by minimizing

$$C(\hat{\mathbf{z}}^{n+1}) = C(\hat{\mathbf{z}}^n - \varepsilon^n \mathbf{g}^n) \quad (40)$$

with respect to ε^n . To do so we begin by writing $C(\hat{\mathbf{z}}^{n+1})$ using (34) and (37). Next we differentiate this with respect to ε^n and set the derivative equal to zero. Solving for ε^n yields, after some manipulation,

$$\varepsilon^n = \frac{\sum_{m=1}^{pM} \gamma_m \left(\sum_{r=1}^N w_{m,r} \hat{z}_r^n - y_m \right) + \lambda \sum_{i=1}^N \bar{g}_i \left(\sum_{j=1}^N \alpha_{i,j} \hat{z}_j^n \right)}{\sum_{m=1}^{pM} \gamma_m^2 + \lambda \sum_{i=1}^N \bar{g}_i^2}, \quad (41)$$

where

$$\gamma_m = \sum_{r=1}^N w_{m,r} g_r(\hat{\mathbf{z}}^n) \quad (42)$$

is the gradient projected through the low resolution pixel formation model, and

$$\bar{g}_i = \sum_{j=1}^N \alpha_{i,j} g_j(\hat{\mathbf{z}}^n) \quad (43)$$

is a weighted sum of neighboring gradient values. This iteration continues until the cost function stabilizes or $\|\hat{\mathbf{z}}^{n+1} - \hat{\mathbf{z}}^n\| / \|\hat{\mathbf{z}}^n\| < T$, where T is a specified threshold value. A summary of the gradient descent optimization procedure is provided in Table 2.

4.2. Conjugate gradient optimization

In this section, we describe a conjugate gradient optimization procedure for minimizing the cost function in (34). In particular, we employ the Fletcher-Reeves method.²⁰ We later show that with little additional computational complexity, faster convergence can be achieved using this method compared to gradient descent.

The basic conjugate gradient image update is given by

$$\hat{z}_k^{n+1} = \hat{z}_k^n + \varepsilon^n d_k(\hat{\mathbf{z}}^n), \quad (44)$$

for $n = 0, 1, 2, \dots$ and $k = 1, 2, \dots, N$. Here $d_k(\hat{\mathbf{z}}^n)$ is the conjugate gradient term. Alternatively, the update can be written as

$$\hat{\mathbf{z}}^{n+1} = \hat{\mathbf{z}}^n + \varepsilon^n \mathbf{d}^n, \quad (45)$$

Table 2. Proposed gradient descent iterative estimation algorithm.

<p>step 1: Begin at $n = 0$ with initial estimate $\hat{\mathbf{z}}^0$ being the interpolated low resolution frame 1.</p> <p>step 2: Compute the gradient $g_k(\hat{\mathbf{z}}^n)$ given in (36) for $k = 1, 2, \dots, N$, yielding \mathbf{g}^n.</p> <p>step 3: Compute the optimal step size ε^n using (41).</p> <p>step 4: Let $\hat{\mathbf{z}}^{n+1} = \hat{\mathbf{z}}^n - \varepsilon^n \mathbf{g}^n$.</p> <p>step 5: If $\ \hat{\mathbf{z}}^{n+1} - \hat{\mathbf{z}}^n\ /\ \hat{\mathbf{z}}^n\ < T$, let $\hat{\mathbf{z}} = \hat{\mathbf{z}}^{n+1}$ and stop.</p> <p>step 6: Let $n = n + 1$ and go to step 2.</p>
--

where

$$\mathbf{d}^n = \begin{bmatrix} d_1(\hat{\mathbf{z}}^n) \\ d_2(\hat{\mathbf{z}}^n) \\ \vdots \\ d_N(\hat{\mathbf{z}}^n) \end{bmatrix}. \quad (46)$$

As before, the parameter ε^n is the step size at the n 'th iteration. The optimal step size can be calculated by minimizing $C(\hat{\mathbf{z}}^{n+1}) = C(\hat{\mathbf{z}}^n + \varepsilon^n \mathbf{d}^n)$ with respect to ε^n . This yields

$$\varepsilon^n = - \frac{\sum_{m=1}^{pM} \phi_m \left(\sum_{r=1}^N w_{m,r} \hat{z}_r^n - y_m \right) + \lambda \sum_{i=1}^N \bar{d}_i \left(\sum_{j=1}^N \alpha_{i,j} \hat{z}_j^n \right)}{\sum_{m=1}^{pM} \phi_m^2 + \lambda \sum_{i=1}^N \bar{d}_i^2}, \quad (47)$$

where

$$\phi_m = \sum_{r=1}^N w_{m,r} d_r(\hat{\mathbf{z}}^n) \quad (48)$$

and

$$\bar{d}_i = \sum_{j=1}^N \alpha_{i,j} d_j(\hat{\mathbf{z}}^n). \quad (49)$$

The conjugate gradient vector and its update procedure are defined below. The initial conjugate gradient is set to be

$$\mathbf{d}^0 = -\mathbf{g}^0. \quad (50)$$

This is updated according to

$$\mathbf{d}^{n+1} = -\mathbf{g}^{n+1} + \beta^n \mathbf{d}^n, \quad (51)$$

where

$$\beta^n = \frac{(\mathbf{g}^{n+1})^T \mathbf{g}^{n+1}}{(\mathbf{g}^n)^T \mathbf{g}^n}. \quad (52)$$

Again, the iterations continue until the estimate converges. A summary of the conjugate gradient optimization procedure is given in Table 3.

Table 3. Proposed conjugate gradient iterative optimization algorithm.

<p>step 1: Begin at $n = 0$ with initial estimate $\hat{\mathbf{z}}^0$ being the interpolated low resolution frame 1.</p> <p>step 2: Compute \mathbf{g}^0 and initialize conjugate gradient vector as $\mathbf{d}^0 = -\mathbf{g}^0$.</p> <p>step 3: Compute the optimal step size ε^n using (47).</p> <p>step 4: Let $\hat{\mathbf{z}}^{n+1} = \hat{\mathbf{z}}^n + \varepsilon^n \mathbf{d}^n$.</p> <p>step 5: If $\ \hat{\mathbf{z}}^{n+1} - \hat{\mathbf{z}}^n\ /\ \hat{\mathbf{z}}^n\ < T$, let $\hat{\mathbf{z}} = \hat{\mathbf{z}}^{n+1}$ and stop.</p> <p>step 6: Compute \mathbf{g}^{n+1} and let $\mathbf{d}^{n+1} = -\mathbf{g}^{n+1} + \beta^n \mathbf{d}^n$, where $\beta^n = \frac{(\mathbf{g}^{n+1})^T \mathbf{g}^{n+1}}{(\mathbf{g}^n)^T \mathbf{g}^n}$.</p> <p>step 7: Let $n = n + 1$ and go to step 3.</p>
--

5. EXPERIMENTAL RESULTS

In this section a number of experimental results are presented in order to demonstrate the performance of the proposed algorithm. The first set of experiments use sequential frame data obtained from a FLIR imaging system. The second set of results use simulated data for quantitative performance analysis.

5.1. Infrared image reconstruction

Here we consider the application of the multi-frame algorithm to a FLIR imager*. The optical system parameters of the FLIR are described in Section 2.3. The theoretical discrete PSF of the FLIR system on the high resolution sampling grid is shown in Fig. 6. This is an incoherent PSF based on the assumption of diffraction-limited optics operating at a midband wavelength of $4 \mu\text{m}$ and it includes the effects of the finite detector size.^{15,21} The weights in (8) are determined by this PSF positioned at the appropriate location on the high resolution grid for each low resolution pixel.

Twenty frames have been acquired at a 60 frame per second rate. Global rotation and translation is introduced by arbitrarily manipulating the imager during acquisition. One typical original resolution frame is shown in Fig. 7(a). The scene contains a number of small power boats and trailers on a gravel parking lot with a fence in the foreground. The multi-frame estimate is shown in Fig. 7(b) for $L_1 = L_2 = 5$. The image used to initiate the iterative estimation procedure has been obtained by bilinearly interpolating the first frame in the sequence. The image shown in Fig. 7(b) is the result of 10 iterations of the conjugate gradient procedure with $\lambda = .1$. Generally, increasing λ will yield a smoother image estimate. For comparison, a bicubic interpolation of a single frame is shown in Fig. 7(c).

*The data have been acquired at the FLIR facility at Wright Laboratory Sensors Technology Branch (WL/AAJT).

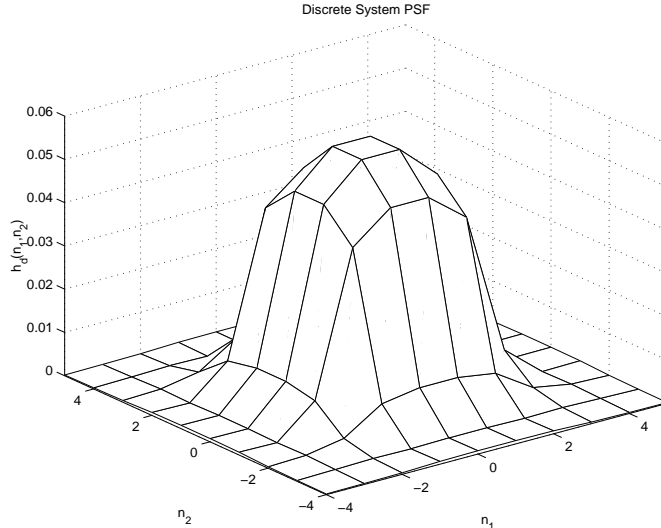


Figure 6. Theoretical discrete system PSF of the FLIR imager for $L_1 = L_2 = 5$.

The multi-frame reconstruction appears to show significantly improved image detail. In addition, note that the aliasing artifacts on the diagonal beam of the gate in the foreground of Figs. 7(a) and 7(c) are virtually eliminated in the multi-frame estimate. The estimated registration parameters for the 20 observed frames are shown in Fig. 8. Rigid translational and rotational motion is assumed yielding three parameters for each frame.

Finally, to illustrate the convergence behavior of the algorithms using the FLIR data, the cost function is plotted in Fig. 9 versus iteration number for both the gradient descent and the conjugate gradient optimization methods. Note, the conjugate gradient algorithm exhibits faster convergence. The algorithm run time is approximately five minutes in MATLAB running on a sun SPARCstation 20. Larger images tend to run proportionately longer.

5.2. Flight data

In this section, results are presented which use digital video obtained by hard mounting the Amber imager to a Cessna 172 aircraft and collecting air-to-ground images[†]. Line-of-sight jitter and optical flow provide the necessary motion. Twenty frames have been arbitrarily extracted from the video sequence. These frames have been acquired at a frame rate of 100 frames/second. The estimated global registration parameters for the 20 frames are shown in Fig. 10.

One standard resolution frame is shown in Fig. 11(a). This image shows a variety of buildings and roads in the Dayton, Ohio area. The multi-frame estimate is shown in Fig. 11(b) for $\lambda = 0.05$ and $L_1 = L_2 = 5$. The initial image estimate is obtained by bilinearly interpolating the first frame and 5 iterations of the conjugate gradient procedure have been performed. For comparison, a bicubic interpolation of a single frame is shown in Fig. 11(c).

[†]The data have been collected courtesy of the Infrared Threat Warning Laboratory, Threat Detection Branch at Wright Laboratory (WL/AAJP).

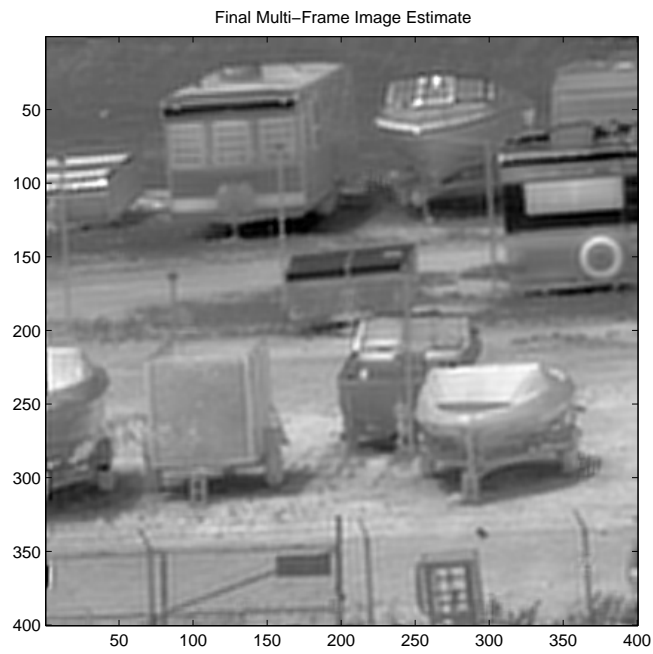
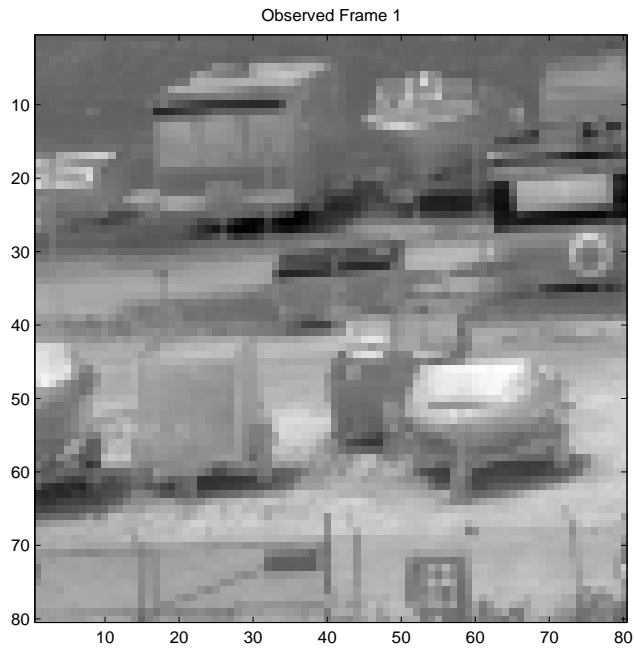
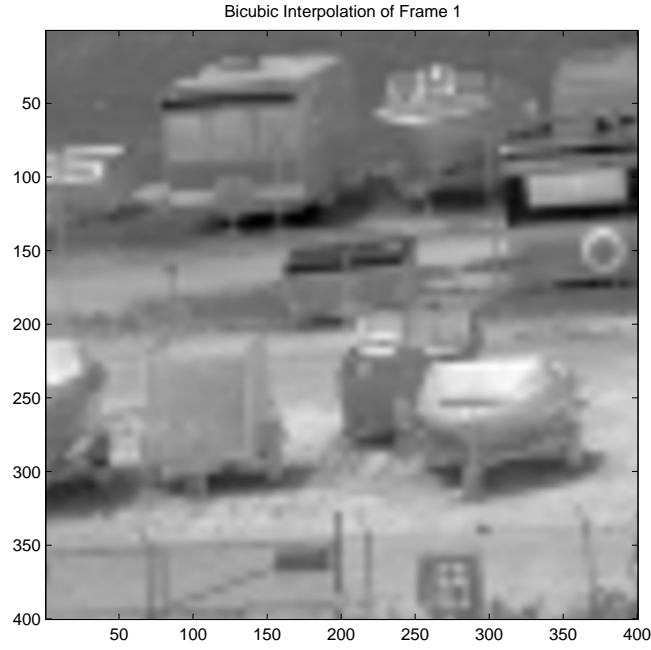


Figure 7. (a) FLIR low resolution frame 1 showing small power boats and trailers on a gravel parking lot (b) multi-frame estimate using 20 frames with $L_1 = L_2 = 5$ and $\lambda = 0.1$ (c) bicubic interpolation of frame 1.



(c)

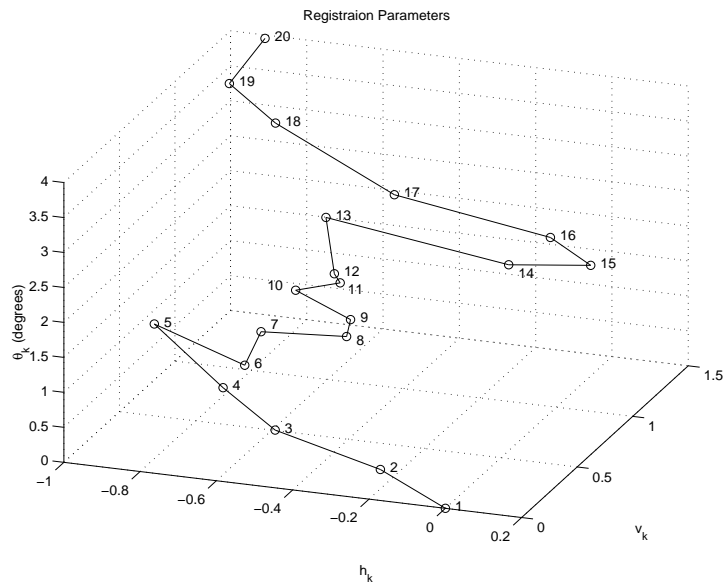


Figure 8. Estimated registration parameters (θ_k , h_k and v_k) for the 20 frames acquired with the FLIR imager.

The multi-frame reconstruction appears to show significantly improved image detail. By processing multiple sets of data, a high resolution video sequence can be generated.

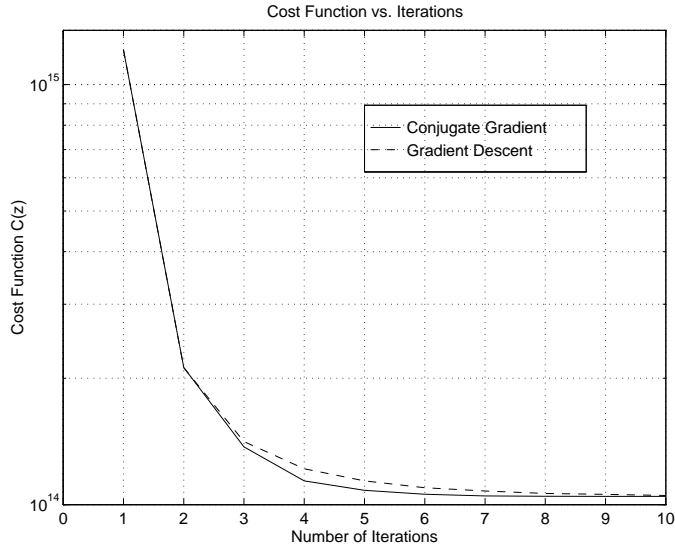


Figure 9. Convergence behavior of the gradient descent and conjugate gradient optimization techniques using the FLIR data. Twenty observed frames have been used with $L_1 = L_2 = 5$ and $\lambda = 0.1$.

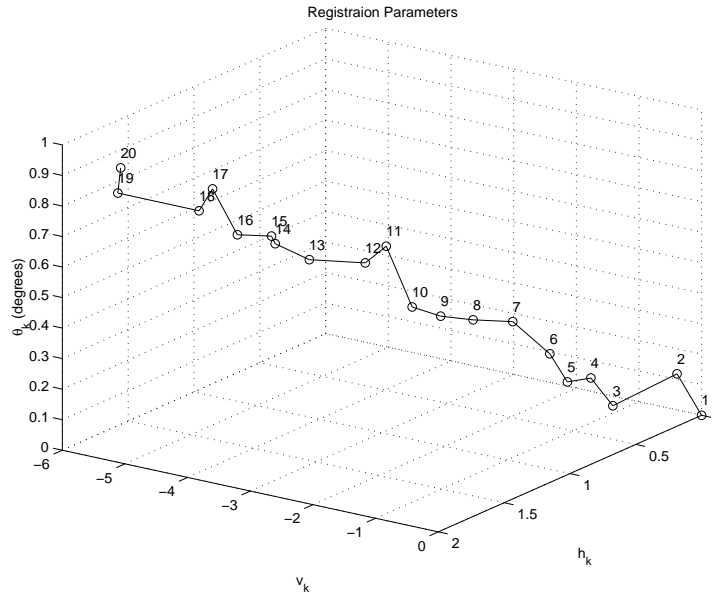
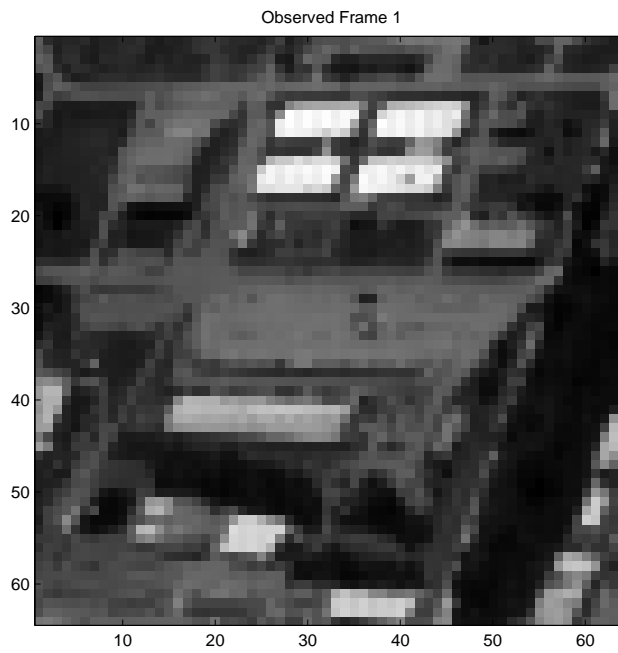


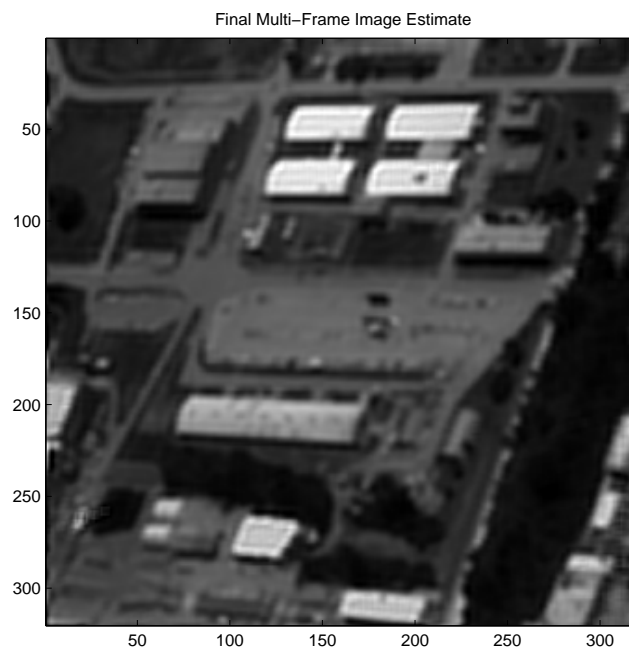
Figure 10. Estimated registration parameters for the 20 frames of flight data with the infrared imager.

5.3. Quantitative analysis

In order to evaluate the algorithms quantitatively, a set of 16 8-bit low resolution images are simulated by rotating, translating, blurring and subsampling an “ideal” image. The original image is of size 250×250 and the down sampling factors are $L_1 = L_2 = 5$. The blurring function is a 5×5 moving average filter which simulates the low resolution detector effects. Finally, additive Gaussian noise of variance $\sigma_\eta^2 = 25$ is introduced in each frame. Twenty iterations of the conjugate gradient optimization have been performed with $\lambda = 0.1$. The mean absolute error (MAE) between

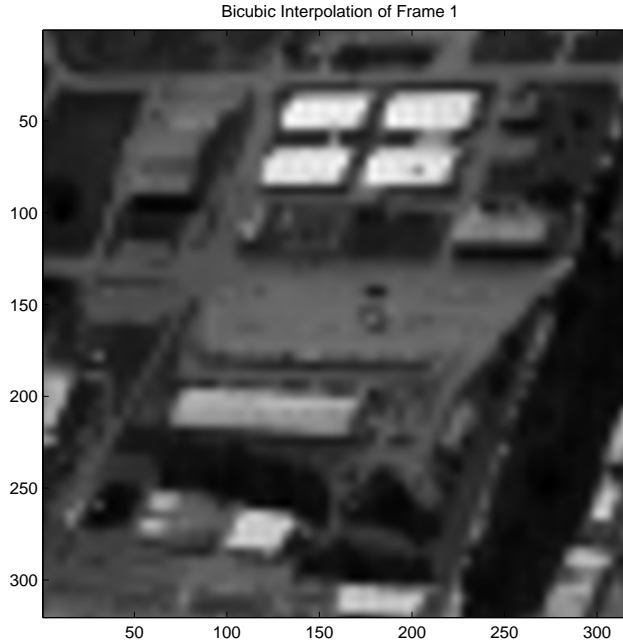


(a)



(b)

Figure 11. (a) Low resolution infrared frame 1 from flight data (b) multi-frame estimate using 20 frames with $L_1 = L_2 = 5$ and $\lambda = 0.05$ (c) bicubic interpolation of frame 1.



(c)

the multi-frame estimate and the “ideal” image is plotted in Fig. 12 as a function of the number of frames used. For comparison, the MAEs of the first frame with bilinear and bicubic interpolation are also shown. With only one frame, the performance of the proposed algorithm is only slightly better than that of the bicubic interpolator. However, with additional frames, the estimate significantly improves with respect to the single frame interpolators. The estimated registration parameters are shown in Fig. 13. The mean absolute error (MAE) between the actual and estimated translational shifts is 0.0414 low resolution pixel spacings and the MAE for the rotation parameters is 0.0250 degrees.

To study the sensitivity of the algorithm to noise, MAEs are computed and plotted in Fig. 14 for various noise levels and three choices of λ . Note that with $\lambda = 0$, the error grows rapidly with the noise level. By choosing a larger value for this parameter (e.g., $\lambda = .1$ or $\lambda = .5$), more robustness is possible. It should be noted that the estimate with $\lambda = 0$ can be improved by halting the iterations earlier. With fewer iterations, the estimate will tend to retain more of the characteristics of the starting image (which is generally smooth). These results simply show the effect of changing λ with all other factors being equal.

6. CONCLUSIONS

Aliasing reduction and resolution enhancement can be achieved by exploiting multiple frames which are rotated and/or translated with respect to one another. This is possible because each frame offers a unique set of discrete samples. For an imager mounted on a moving platform, such as an aircraft, the desired image sequence may arise

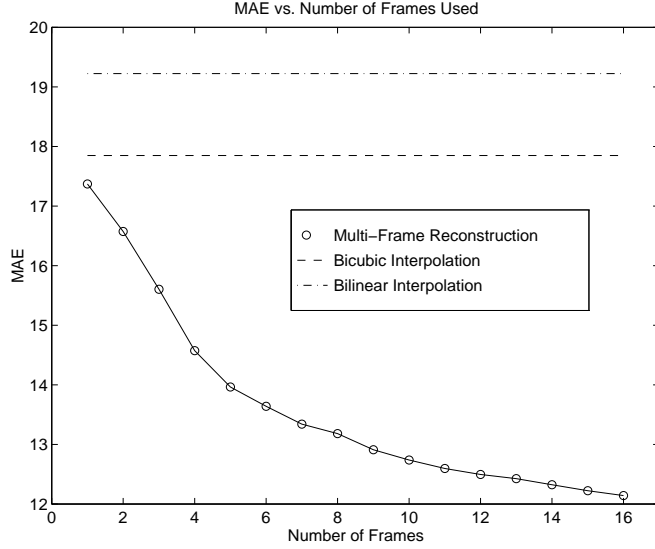


Figure 12. Mean absolute error for the multi-frame estimator using different numbers of frames. The noise variance is $\sigma_\eta^2 = 10$, $L_1 = L_2 = 5$, and $\lambda = 0.1$.

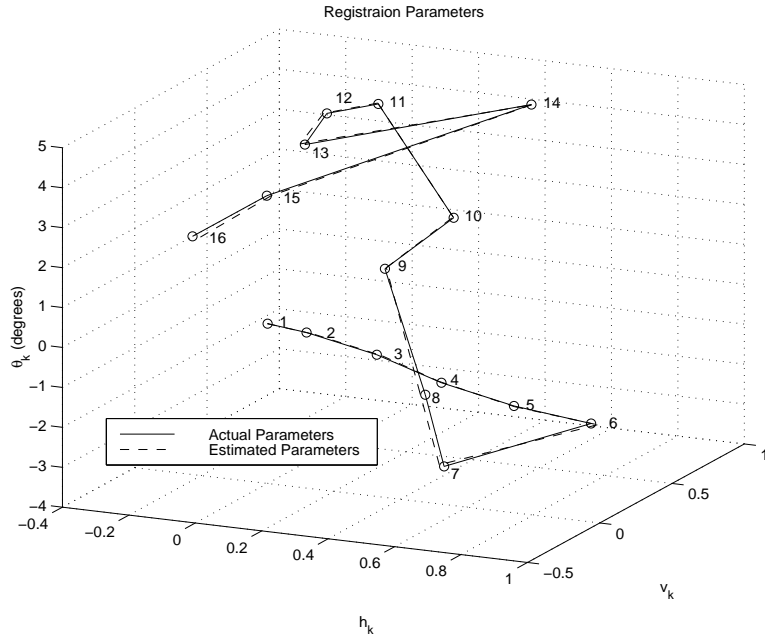


Figure 13. Actual and estimated registration parameters for each of the 16 simulated low resolution frames. The shifts h_k and v_k are measured in terms of low resolution pixel spacings.

from natural line-of-sight jitter and rotation of the platform. With this in mind, it may then be possible to relax image stabilization requirements in some applications and obtain improved resolution images through the proposed algorithm.

The key to the success of the algorithm is having an accurate observation model. This includes the image registration parameters and the system PSF. The observation model proposed here includes information about the

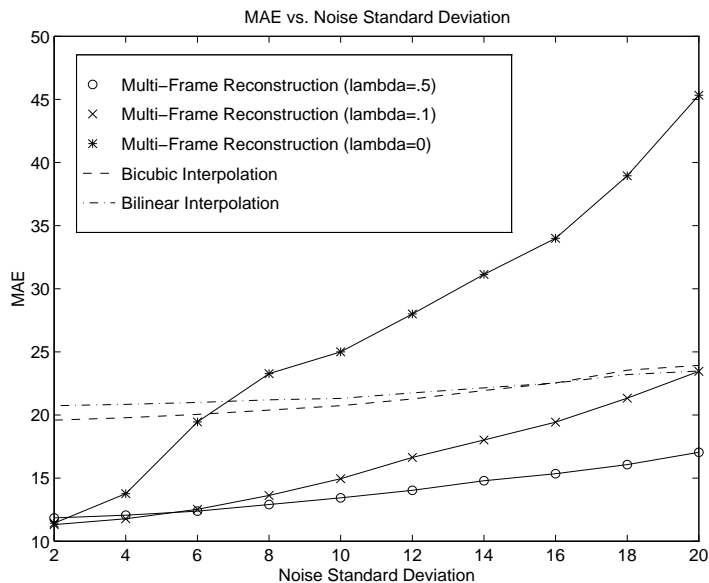


Figure 14. MAE for simulated data with varying levels of noise.

optical system and FPA. A regularized cost function defines the image estimate. Minimization of the cost function is performed using either a gradient descent or conjugate gradient technique.

The quantitative results obtained show that the multi-frame image estimates have significantly lower error than estimates formed by single frame interpolation. Furthermore, we believe that the FLIR results show that the multi-frame estimate has significantly improved image detail. In particular, edges and fine structure emerge in the multi-frame reconstruction that are not visible in the low resolution data. Because these features offer important visual cues, we believe that the utility of the processed image is greatly enhanced.

ACKNOWLEDGEMENTS

This work has been supported by Wright Laboratory's Sensors Technology Branch (WL/AAJT) at Wright Patterson Air Force Base in Dayton Ohio under Air Force contract F33601-95-DJ010. In particular, the authors would like to thank Dr. Mohammad Karim, Dr. Paul McManamon, Don Tomlinson, and Brian Yasuda for supporting this project. Thanks to Eric Kaltenbacher for his assistance in data collection. Thanks are also due to Tony Absi and Art Serrano for their administrative support. Thanks to Steve Cain for suggesting and initially developing a statistical estimation approach to the FLIR resolution enhancement problem. We would also like to gratefully acknowledge the work of those FLIR team members who support the hardware and software which made this work possible, James Glidewell, Bob Gualtieri, Rich Hill, Al Meyer, Joe Richardson, Jeff Sweet, Tim Tuinstra, Carl White and Ron Wiensch. Finally, we would like to thank Dr. Richard B. Sanderson, Joel Montgomery, Mike Hall, David Coates and Frank Baxely, with the Threat Detection Branch at Wright Laboratory (WL/AAJP), for collecting the infrared flight images used here.

REFERENCES

1. J. C. Gillette, T. M. Stadtmiller, and R. C. Hardie, "Reduction of aliasing in staring infrared imagers utilizing subpixel techniques," *Optical Engineering* **34**, pp. 3130–3137, Nov. 1995.
2. E. A. Watson, R. A. Muse, and F. P. Blommel, "Aliasing and blurring in microscanned imagery," in *Proceedings of SPIE Infrared Imaging Systems: Design, Analysis, Modeling and Testing III, Orlando, FL*, vol. 1689, pp. 242–250, Apr. 1992.
3. R. Tsai and T. Huang, "Multiframe image restoration and registration," in *Advances in Computer Vision and Image Processing*, vol. 1, pp. 317–339, JAI Press Inc., 1984.
4. S. Kim, N. Bose, and H. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **38**, pp. 1013–1027, June 1990.
5. H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays, using convex projections," *Journal of the Optical Society of America A* **6**(11), pp. 1715–1726, 1989.
6. A. Patti, M. Sezan, and A. Tekalp, "High-resolution standards conversion of low resolution video," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Detroit, MI*, pp. 2197–2200, May 1995.
7. A. J. Patti, M. I. Sezan, and A. M. Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Transactions on Image Processing* **6**, pp. 1064–1076, Aug. 1997.
8. S. Mann and R. W. Picard, "Virtual bellows: Constructing high quality stills from video," in *Proceedings of IEEE International Conference on Image Processing, Austin TX*, Nov. 1994.
9. R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Transactions on Image Processing* **5**, pp. 996–1011, June 1996.
10. P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "Super-resolved surface reconstruction from multiple images," *NASA Technical Report FIA-94-12, NASA Ames Research Center, Moffett Field CA*, Dec. 1994.
11. M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing* **53**, pp. 231–239, May 1991.
12. S. Cain, R. C. Hardie, and E. E. Armstrong, "Restoration of aliased video sequences via a maximum-likelihood approach," in *Infrared Information Symposium (IRIS) on Passive Sensors, Monterey, CA*, vol. 1, pp. 377–390, Mar. 1996.
13. R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint map registration and high resolution image estimation using a sequence of undersampled images," *IEEE Transactions on Image Processing* **6**, pp. 1621–1633, Dec. 1997.
14. A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1989.
15. J. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, 1968.

16. A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, Prentice Hall, 1989.
17. B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *International Joint Conference on Artificial Intelligence, Vancouver*, Aug. 1981.
18. Y.-L. You and M. Kaveh, "A regularized approach to joint blur identification and image restoration," *IEEE Transactions on Image Processing* **5**, pp. 416–428, Mar. 1996.
19. R. L. Lagendijk, J. Biemond, and D. E. Boeke, "Regularized iterative image restoration with ringing reduction," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **36**, pp. 1874–1888, Dec. 1988.
20. D. G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, 1984.
21. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics*, Wiley, 1978.

List of Figures

1	Continuous observation model resembling the physical process of image acquisition.	3
2	Equivalent discrete observation model illustrating the relationship between the ideally sampled image \mathbf{z} and the observed frames \mathbf{y}	5
3	Discrete detector model showing those “virtual” high resolution pixels that contribute to a low resolution pixel for two different registration positions.	6
4	Uniform detector array illustrating critical dimensions.	8
5	(a) Effective MTF of the detectors in the FLIR imager (b) diffraction-limited OTF of the optics (c) overall system MTF (d) overall continuous system PSF.	10
6	Theoretical discrete system PSF of the FLIR imager for $L_1 = L_2 = 5$	17
7	(a) FLIR low resolution frame 1 showing small power boats and trailers on a gravel parking lot (b) multi-frame estimate using 20 frames with $L_1 = L_2 = 5$ and $\lambda = 0.1$ (c) bicubic interpolation of frame 1.	18
8	Estimated registration parameters $(\theta_k, h_k$ and $v_k)$ for the 20 frames acquired with the FLIR imager.	19
9	Convergence behavior of the gradient descent and conjugate gradient optimization techniques using the FLIR data. Twenty observed frames have been used with $L_1 = L_2 = 5$ and $\lambda = 0.1$	20
10	Estimated registration parameters for the 20 frames of flight data with the infrared imager.	20
11	(a) Low resolution infrared frame 1 from flight data (b) multi-frame estimate using 20 frames with $L_1 = L_2 = 5$ and $\lambda = 0.05$ (c) bicubic interpolation of frame 1.	21
12	Mean absolute error for the multi-frame estimator using different numbers of frames. The noise variance is $\sigma_n^2 = 10$, $L_1 = L_2 = 5$, and $\lambda = 0.1$	23
13	Actual and estimated registration parameters for each of the 16 simulated low resolution frames. The shifts h_k and v_k are measured in terms of low resolution pixel spacings.	23
14	MAE for simulated data with varying levels of noise.	24

List of Tables

1	Iterative registration algorithm for $y_k(\mathbf{n})$, where $k = 2, 3, \dots, p$	12
2	Proposed gradient descent iterative estimation algorithm.	15
3	Proposed conjugate gradient iterative optimization algorithm.	16