12-1997

Joint MAP Registration and High Resolution Image Estimation Using a Sequence of Undersampled Images

Russell C. Hardie
University of Dayton, rhardie1@udayton.edu

Kenneth J. Barnard
Air Force Wright Laboratory

Ernest E. Armstrong
Technology/Scientific Services Inc.

Joint MAP Registration and High Resolution Image Estimation Using a Sequence of Undersampled Images[1]

Russell C. Hardie†, *Kenneth J. Barnard*‡ and *Ernest E. Armstrong*‡

†Department of Electrical and Computer Engineering
University of Dayton
300 College Park Avenue
Dayton, OH 45469
Phone: (513) 229-3178
Fax: (513) 229-2471
Email: rhardie@engr.udayton.edu

‡Sensors Technology Branch
Wright Laboratory WL/AAJT
Building 622
3109 P Street
WPAFB OH 45433-7700
Phone: (513) 255-9609
Fax: (513) 255-6489

Abstract

In many imaging systems, the detector array is not sufficiently dense to adequately sample the scene with the desired field of view. This is particularly true for many infrared focal plane arrays. Thus, the resulting images may be severely aliased. This paper examines a technique for estimating a high resolution image, with reduced aliasing, from a sequence of undersampled frames. Several approaches to this problem have been investigated previously. However, in this paper a maximum *a posteriori* (MAP) framework for jointly estimating image registration parameters and the high resolution image is presented. Several previous approaches have relied on knowing the registration parameters *a priori* or have utilized registration techniques not specifically designed to treat severely aliased images. In the proposed method, the registration parameters are iteratively updated along with the high resolution image in a cyclic coordinate-descent optimization procedure. Experimental results are provided to illustrate the performance of the proposed MAP algorithm using both visible and infrared images. Quantitative error analysis is provided and several images are shown for subjective evaluation.

List of Figures

List of Tables

1 Introduction

If a detector array used for image acquisition is not sufficiently dense, so as to meet the Nyquist criterion, the resulting images will be degraded by aliasing effects. Since the optics of the imaging system will serve to effectively bandlimit the image on the detector array, it is possible to acquire an image which is free of aliasing. However, this requires the appropriate combination of optics and detector array. Generally a broad instantaneous field of view is desired which requires optics with a short focal length. To prevent aliasing in this case requires a dense detector array which may be very costly or simply unavailable. Thus, many imaging systems are designed to allow some level of aliasing during image acquisition. This is particularly true for staring infrared imagers because of fabrication complexities. Some visual charge coupled device (CCD) cameras also suffer from undersampling. The goal of this work is to develop a technique for estimating an unaliased high resolution image from the aliased images acquired from such an imaging system. We also wish to combat additive noise and blurring due to the finite detector size and optics.

Here the problem is approached from the framework of image sequence processing [1]. Thus, the high resolution image will be estimated from a sequence of low-resolution aliased images. This is possible if there exists sub-pixel motion between the acquired frames. Thus, a unique "look" at the scene is provided by each frame. In particular, we consider the scenario where an imager is mounted on a moving or vibrating platform, such as an aircraft, and is imaging objects in the far field. Thus, the line-of-sight jitter will generally provide the necessary motion between the focal plane array and the scene at each acquisition time with minimal occlusion effects. This process is referred to as uncontrolled microscanning [2, 3]. The key to exploiting these multiple frames is accurate knowledge of the sub-pixel registration parameters for each frame. If the images are severely undersampled, we have found that traditional motion estimation techniques, such as block matching, may not provide the desired sub-pixel accuracy. This has motivated the development of the approach presented here.

This basic problem of high resolution image recovery using multiple frames was first addressed in the literature by Tsai and Huang [4]. Their observation model is based on the shift property of the Fourier transform. Each aliased observation provides a set of equations in the frequency domain. Provided that enough frames are available, the unaliased discrete spectrum, and hence unaliased image, can be solved for. However, one must know the shifts in order to solve for the high resolution image in this fashion. A method for estimating the shifts using the multiple aliased frames is proposed in [4] for the case where the images are bandlimited. While this is an insightful solution, it may be impractical in many applications because of prohibitively high computational complexity. Furthermore, it requires a set minimum number of frames to operate, which may not be available. The image recovery algorithm in [4] is extended in [5] for the case where noise is considered by using

a recursive least squares solution for the set of frequency domain linear equations. This is extended again for the case where blurring in considered in [6]. However, the estimation of the global frame shifts is not addressed in [5] or [6]. A technique which uses the least squares solution, similar to that in [5], along with a fast suboptimal scheme for estimating the global frame shifts is described in [7].

Another approach to the high resolution image reconstruction problem uses a projection onto convex sets (POCS) algorithm [8]. The POCS approach has been extended to treat motion blur and noise in [1, 9, 10]. Block matching or phase correlation is suggested in [9, 10] as a means of estimating the required motion parameters. A related multiframe technique which considers global translational shift and rotation is presented in [11]. In [12], this technique is extended to treat the more general case of a perspective projection of a plane. All of these methods rely on a two stage estimation procedure whereby the registration is done independently of the high resolution image reconstruction.

The problem has also been approached from a statistical estimation framework. Specifically, a maximum *a posteriori* (MAP) estimator is developed in [13, 14] which is an extension of a single frame image expansion algorithm proposed in [15]. The MAP estimator in [13, 14] uses an edge preserving Huber-Markov random field for the image prior. The motion is estimated by a block matching algorithm applied to the individual frames. These individual frames are first expanded using the single frame algorithm in [15] to allow for sub-pixel estimates. This provides a useful solution and it can treat the case where the image motion is not global, which is critical if the scene contains moving objects. The block matching technique, however, does not exploit *a priori* information about the motion if any is available. It also does not exploit all the observed frames when estimating the motion parameters for each image. Another related multiframe MAP technique, which uses an approach similar to that described here, can be found in [16].

A statistical estimation approach to multiframe resolution enhancement was first applied to staring infrared imager data in [17]. In particular, a maximum likelihood (ML) technique using the expectation maximization (EM) algorithm is developed which seeks to estimate translational motion and a high resolution image [17] in the case of Poisson image statistics. As in [17], here we view the problem as that of blind image restoration. That is, the undegraded high resolution image must be estimated from a sequence of noisy, blurred and undersampled images jointly with certain parameters of the observation model. Namely, these observation model parameters are the registration parameters associated with each observed frame. Unlike most blind deconvolution applications, however, here we assume we know the system point spread function (PSF). What we do not know is the relative position of each observed low resolution pixel with respect to a fixed high resolution grid. Therefore, in this paper we seek to minimize a MAP cost function

with respect to the high resolution image and the registration parameters simultaneously using a cyclic coordinate-descent optimization procedure [18]. In this iterative technique, the registration parameter estimates are updated using the current best estimate of the high resolution image. In this way, information from all the observed frames is used in updating the registration parameters. We believe this approach represents an effective use of all of the observed data.

The organization of the rest of the paper is as follows. In Section 2, the observation model is described. The MAP estimation problem is formulated in Section 3 and the prior statistical models are described. Also in Section 3, the cyclic coordinate-descent optimization procedure for minimizing the MAP cost function is developed. Experimental results are provided in Section 4. These include results obtained using simulated data and using a staring infrared imaging sensor. Quantitative error analysis is provided and several images are shown for subjective evaluation. Finally, some conclusions are given in Section 5.

2   Observation Model

Consider the desired high resolution image of size $L_1 N_1 \times L_2 N_2$ written in lexicographical notation as the vector $\mathbf{z} = [z_1, z_2, \ldots, z_N]^T$, where $N = L_1 N_1 L_2 N_2$. That is, $\mathbf{z}$ represents the ideal undegraded image or underlying scene values sampled at or above the Nyquist rate. Let the parameters L_1 and L_2 represent the down sampling factors in the observation model in the horizontal and vertical directions, respectively. Thus, each observed low resolution frame is of size $N_1 \times N_2$. Let the k'th low resolution frame be denoted in lexicographical notation as $\mathbf{y}_k = [y_{k,1}, y_{k,2}, \ldots, y_{k,M}]^T$ for $k = 1, 2, \ldots, p$ and where $M = N_1 N_2$. Let the full set of p observed low resolution images be denoted

$$\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_p^T]^T = [y_1, y_2, \ldots, y_{pM}]^T. \tag{1}$$

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

Next we wish to define an appropriate relationship between the undegraded high resolution image and the observed low resolution images. This observation model is the basis for the estimation technique developed in the next section. We use a straightforward but general observation model where the low resolution pixels are defined as a weighted sum of the appropriate high resolution pixels with additive noise. The weighted sum models the blurring of the underlying scene values due to the finite detector size and the PSF of the optics. An important factor in specifying the weights is the position of each low resolution pixel with respect to the fixed high resolution grid (i.e., the registration parameters). Specifically, the observed low resolution pixels from frame k are related to the high resolution image through the following model

$$y_{k,m} = \sum_{r=1}^{N} w_{k,m,r}(\mathbf{s}_k) z_r + \eta_{k,m}, \tag{2}$$

3

for $m = 1, 2, \ldots, M$ and $k = 1, 2, \ldots, p$. The weight $w_{k,m,r}(\mathbf{s}_k)$ represents the "contribution" of the r'th high resolution pixel to the m'th low resolution observed pixel of the k'th frame. The vector $\mathbf{s}_k = [s_{k,1}, s_{k,2}, \ldots, s_{k,K}]^T$, contains the K registration parameters for frame k. Depending on the application, these parameters may represent global translational shift in the horizontal and vertical directions, rotation, affine transformation parameters, or other motion parameters. This motion is measured in terms of some reference on a fixed high resolution grid. The term $\eta_{k,m}$ in (2) represents an additive noise samples which will be assumed to be independent and identically distributed (i.i.d.) Gaussian noise samples with variance σ_η^2.

We believe that a Gaussian noise model is useful for a variety of imaging systems and scenarios. In most practical imaging applications, detected photons follow Poisson statistics [19]. The measurable manifestation of photon-noise depends on the particular detector type. For photon detectors that detect light through the conversion of photons into charge carriers, those carriers also follow Poisson statistics [19, 20]. A system is photon-noise limited when this photon noise is the dominant noise source in the system. When there exists a large number of photon generated charge carriers, the Poisson statistics can be approximated by a Gaussian distribution [21]. Furthermore, if the imaging system is not photon-noise limited, then the superposition of the contributing events from all noise sources may result in a probability density function which approaches a Gaussian as defined by the central limit theorem [21, 22]. Notwithstanding this, Poisson image statistics are valid provided the system is photon-noise limited. A ML high resolution image estimation approach is developed in [17] under the assumption of Poisson image statistics.

The observation model in (2) assumes that the underlying scene samples, $\mathbf{z}$, remain constant during the acquisition of the multiple low resolution frames, except for any motion allowed by the motion model. Thus, the only frame-to-frame differences of the model weights in (2) result from the motion of each low resolution pixel relative to the high resolution grid. It is these minute sub-pixel motions we wish to exploit to form our high resolution image estimate. One simple model for specifying the weights, which does not include the PSF of the optics, is illustrated in Fig. 1. Here each low resolution pixel (shown on the right) is obtained by summing the high resolution pixels within the span of that low resolution detector (shown on the left). In this example, only $L_1 L_2$ high resolution pixels contribute to a specific low resolution pixel. To represent uniform detector sensitivity, the weights corresponding to those $L_1 L_2$ high resolution pixels would be set to $1/(L_1 L_2)$. The other weights would be set to zero. This discrete detector model simulates the integration of light intensity that falls within the span of the low resolution detector. If the entire low resolution grid moves relative to the fixed high resolution grid (i.e., global rigid motion), a different set of high resolution pixels contribute to each low resolution pixel. This yields a new set of linearly independent equations from (2).
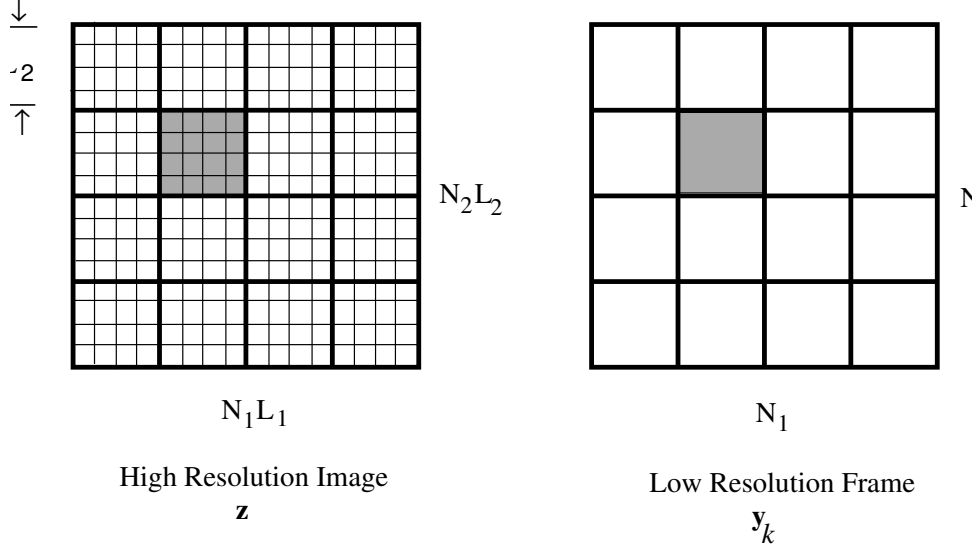
Figure 1: Discrete detector model showing those high resolution pixels (on the left) that contribute to a low resolution pixel (on the right). The image $\mathbf{z}$ represents the true underling high resolution image we wish to estimate and $\mathbf{y}_k$ is the k'th observed low resolution frame. Note the different grid sizes for $\mathbf{z}$ and $\mathbf{y}_k$.

Alternatively, the model in (2) can be expressed in terms of the entire set of low resolution pixels as

$$y_m = \sum_{r=1}^{N} w_{m,r}(\mathbf{s})z_r + \eta_m, \tag{3}$$

for $m = 1, 2, \ldots, pM$ and where $w_{m,r}(\mathbf{s})$ is the "contribution" of z_r to y_m. Here the entire set of motion parameters is contained in $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \ldots, \mathbf{s}_p^T]^T$. In many practical imaging situations, these parameters are not known *a priori*. Therefore, we consider them to be random parameters to be estimated along with the high resolution image $\mathbf{z}$. In the experimental results presented in Section 4, we focus on translational motion because of our specific experimental setup and application. However, the following algorithm development is done in general terms to allow for a variety of image acquisition scenarios. In the case of translational motion $K = 2$ and $\mathbf{s}_k = [h_k, v_k]^T$, where h_k and v_k are the horizontal and vertical shifts for frame k, respectively. These will be most conveniently measured in terms of high resolution detector spacings.

For some of the later analysis, it will be convenient to represent the observation model in matrix notation. Thus, rewriting (3) in matrix notation yields

$$\mathbf{y} = \mathbf{W_s}\mathbf{z} + \mathbf{n}, \tag{4}$$

where element (m, r) in $\mathbf{W_s}$ is $w_{m,r}(\mathbf{s})$ and $\mathbf{n} = [\eta_1, \eta_2, \ldots, \eta_{pM}]^T$. Note that since the elements of

n are i.i.d. Gaussian samples, the multivariate probability density function (pdf) of **n** is given by

$$Pr(\mathbf{n}) = \frac{1}{(2\pi)^{\frac{pM}{2}} \sigma_\eta^{pM}} exp\left\{ -\frac{1}{2\sigma_\eta^2} \mathbf{n}^T \mathbf{n} \right\} = \frac{1}{(2\pi)^{\frac{pM}{2}} \sigma_\eta^{pM}} exp\left\{ -\frac{1}{2\sigma_\eta^2} \sum_{m=1}^{pM} \eta_m^2 \right\}. \qquad (5)$$

3 MAP Estimator

We wish to form a MAP estimate of the high resolution image **z** and the registration information **s** simultaneously, given that we observe **y**. The estimates can be computed as

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \begin{array}{c} \arg\max \\ \mathbf{z}, \mathbf{s} \end{array} Pr(\mathbf{z}, \mathbf{s}|\mathbf{y}). \qquad (6)$$

Using Bayes rule, this can alternatively be expressed as

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \begin{array}{c} \arg\max \\ \mathbf{z}, \mathbf{s} \end{array} \frac{Pr(\mathbf{y}|\mathbf{z}, \mathbf{s})Pr(\mathbf{z}, \mathbf{s})}{Pr(\mathbf{y})}. \qquad (7)$$

Noting that the denominator is not a function of **z** or **s**, and that **z** and **s** are statistically independent, the estimates can be written as

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \begin{array}{c} \arg\max \\ \mathbf{z}, \mathbf{s} \end{array} Pr(\mathbf{y}|\mathbf{z}, \mathbf{s})Pr(\mathbf{z})Pr(\mathbf{s}). \qquad (8)$$

Since we later define the form of these densities to be Gaussian in nature it is more convenient, and equivalent, to minimize the minus log of the functional in (8). This yields

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \begin{array}{c} \arg\min \\ \mathbf{z}, \mathbf{s} \end{array} L(\mathbf{z}, \mathbf{s}) = \begin{array}{c} \arg\min \\ \mathbf{z}, \mathbf{s} \end{array} \left\{ -log[Pr(\mathbf{y}|\mathbf{z}, \mathbf{s})] - log[Pr(\mathbf{z})] - log[Pr(\mathbf{s})] \right\}. \qquad (9)$$

Thus, we must specify the prior image density $Pr(\mathbf{z})$, the prior motion density $Pr(\mathbf{s})$, and the conditional density $Pr(\mathbf{y}|\mathbf{z}, \mathbf{s})$. Then, a method for jointly optimizing (9) with respect to **z** and **s** must be found. First let us consider the prior density models.

3.1 Prior Models

The problem of estimating **z** from **y** is generally an ill-posed inverse problem. This can lead to estimates with excessive noise magnification if not treated. However, an appropriate choice of $Pr(\mathbf{z})$ in (9) can serve to regularize the problem. Such a prior statistical image model should accurately reflect the characteristics of the random process from which **z** is viewed as a realization.

While the selection of an image prior is application dependent, we believe that a simple Gibbs distribution [23, 24] can provide a useful solution for a wide variety of imaging systems. As will be seen, this image prior model used with a Gaussian noise model has the desirable property of yielding a convex MAP cost function with a unique global minimum. This cost function is readily differentiable with respect to the high resolution image **z** allowing for a closed-form solution.

The probability density function for the Gaussian prior has the form

$$Pr(\mathbf{z}) = \frac{1}{(2\pi)^{\frac{N}{2}}|C_z|^{1/2}} exp\left\{-\frac{1}{2}\mathbf{z}^T C_z^{-1}\mathbf{z}\right\}, \tag{10}$$

where C_z is the $N \times N$ covariance matrix of $\mathbf{z}$. The exponential term in (10) which includes the image covariance can be factored into a sum of products yielding

$$Pr(\mathbf{z}) = \frac{1}{(2\pi)^{\frac{N}{2}}|C_z|^{1/2}} exp\left\{-\frac{1}{2\lambda}\sum_{i=1}^{N}\mathbf{z}^T \mathbf{d}_i \mathbf{d}_i^T \mathbf{z}\right\}, \tag{11}$$

where $\mathbf{d}_i = [d_{i,1}, d_{i,2}, \ldots, d_{i,N}]^T$ is a coefficient vector and λ can be viewed as a "tuning" parameter. Thus, the prior can be rewritten as

$$Pr(\mathbf{z}) = \frac{1}{(2\pi)^{\frac{N}{2}}|C_z|^{1/2}} exp\left\{-\frac{1}{2\lambda}\sum_{i=1}^{N}\left(\sum_{j=1}^{N}d_{i,j}z_j\right)^2\right\}. \tag{12}$$

The coefficient vectors $\mathbf{d}_i$ for $i = 1, 2, \ldots, N$ effectively express *a priori* assumptions about the local relationship between pixel values in $\mathbf{z}$. As is commonly done, here these parameters are selected to provide a higher probability for smooth random fields. The parameter λ can be used empirically to control the "penalty" for discontinuities and rapidly varying features in $\mathbf{z}$. Finally, by equating (10) and (12), the elements of the inverse covariance matrix can be written in terms of $d_{i,j}$. Let the (i,j)'th element in C_z^{-1} be denoted $C_{i,j}^{-1}$ which is given by

$$C_{i,j}^{-1} = \frac{1}{\lambda}\sum_{r=1}^{N}d_{r,i}d_{r,j}. \tag{13}$$

For the experimental results presented in Section 4, we have selected the following values for the coefficient vectors:

$$d_{i,j} = \begin{cases} 1 & \text{for } i = j \\ -1/4 & \text{for } j : z_j \text{ is a cardinal neighbor of } z_i. \end{cases} \tag{14}$$

Figure 2 illustrates the four cardinal neighbors of a pixel z_i. To account for border effects on our finite lattice, the coefficients are modified at the edges where fewer than four cardinal neighbors are available. Note that the prior in (12) can be viewed as a Gibbs distribution where the exponential term is a sum of clique potential functions [23]. For the coefficients in (14), the pixel cliques with non-zero clique potential functions look like the set of shaded pixels in Fig. 2 including the center sample labeled z_i. Such cliques can be derived from a third order neighborhood system [25, 26].

The choice of a prior statistical model for the registration parameters will be highly application specific. Furthermore, in many cases, the registration parameters may be significantly over-determined by the data and a prior is not necessary to yield a useful solution. One exception is the case where the high resolution image exhibits strong periodicity. A registration estimator with no
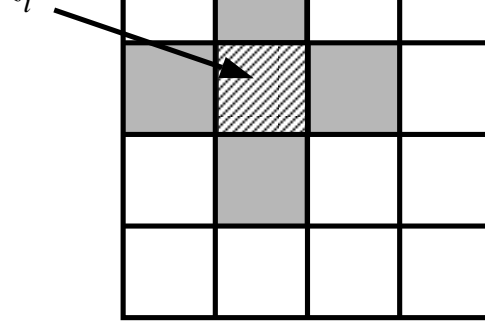
Figure 2: The high resolution image prior neighborhood model showing the cardinal neighbors of a pixel z_i. In this case, $d_{i,j}$ would be non-zero only for j such that z_j is an immediate spatial neighbor of z_i (those shaded pixels).

a priori information may easily report incorrect parameters because it has been trapped in a local minima. Cases where the signal-to-noise ratio is very low or numerous registration parameters must be estimated may also benefit from a valid prior statistical model if one can be identified. If the signal-to-noise ratio is high and there are a relatively small number of registration parameters to estimate, we believe that a "no preference" prior can yield a useful solution. Thus, the estimation of the registration parameters reduces to a maximum likelihood estimate.

Using the Gibbs image prior and a "no preference" registration prior, we now begin to formulate the MAP solution for the high resolution image. First we derive a closed form solution then a gradient descent procedure is described. The latter provides a more practical implementation for large images.

3.2 Matrix MAP Solution

Let us begin by expressing the conditional density from (9) in matrix form. Given the observation model in (4) and the noise density in (5), the conditional density can be written as

$$Pr(\mathbf{y}|\mathbf{z}, \mathbf{s}) = \frac{1}{(2\pi)^{\frac{N}{2}} \sigma_\eta^N} exp\left\{-\frac{1}{2\sigma_\eta^2}(\mathbf{y} - \mathbf{W_s}\mathbf{z})^T(\mathbf{y} - \mathbf{W_s}\mathbf{z})\right\}. \tag{15}$$

Using (9), (10), and (15) and ignoring terms which are not functions of $\mathbf{z}$ or $\mathbf{s}$, the MAP estimates can be expressed as

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \begin{array}{c}\arg\min\\ \mathbf{z}, \mathbf{s}\end{array} L(\mathbf{z}, \mathbf{s}), \tag{16}$$

where

$$L(\mathbf{z}, \mathbf{s}) = \frac{1}{2\sigma_\eta^2}(\mathbf{y} - \mathbf{W_s}\mathbf{z})^T(\mathbf{y} - \mathbf{W_s}\mathbf{z}) + \frac{1}{2}\mathbf{z}^T C_z^{-1}\mathbf{z}. \tag{17}$$

Thus, we have to minimize the cost function in (17) with respect to $\mathbf{z}$ and $\mathbf{s}$ simultaneously, where $\mathbf{y}$ represents the observed data. This provides a somewhat challenging optimization problem. This

8

is partly due to the fact that (17) is not readily differentiable with respect to $\mathbf{s}$ for many motion models. However, given $\mathbf{z}$ is fixed, it is possible to perform a search over a finite set of discrete motion parameters to minimize (17) with respect to $\mathbf{s}$. Also, (17) forms a quadratic function in $\mathbf{z}$ and can be minimized readily with respect to $\mathbf{z}$ if the registration information $\mathbf{s}$ is fixed. Thus, the approach will be to use a cyclic coordinate-descent optimization procedure [18]. That is, we descend the cost function in (17) with respect to $\mathbf{s}$ and $\mathbf{z}$ separately and alternately. These alternating minimizations will continue until the algorithm adequately converges or a set number of iterations is reached.

Consider beginning the optimization procedure using some initial estimate of the high resolution image $\hat{\mathbf{z}}^0$. This estimate can be obtained, for example, by interpolating the first low resolution image in the observed sequence. At each iteration n, the motion parameters can be obtained by minimizing (17) with respect to $\mathbf{s}$, given $\mathbf{z}$ is the current estimate of the high resolution image $\hat{\mathbf{z}}^n = [\hat{z}_1^n, \hat{z}_2^n, \ldots, \hat{z}_N^n]^T$. Specifically, the motion estimates for $n = 0, 1, 2, \ldots$ can be computed as

$$\hat{\mathbf{s}}^n = \overset{\arg\min}{\mathbf{s}} \; L(\hat{\mathbf{z}}^n, \mathbf{s}) = \overset{\arg\min}{\mathbf{s}} \left\{ (\mathbf{y} - \mathbf{W_s}\hat{\mathbf{z}}^n)^T (\mathbf{y} - \mathbf{W_s}\hat{\mathbf{z}}^n) \right\}. \tag{18}$$

To derive the update procedure for the high resolution image, we begin by computing the gradient of the cost function in (17) with respect to the vector quantity $\mathbf{z}$. This gradient is given by

$$\nabla_{\mathbf{z}} L(\mathbf{z}, \mathbf{s}) = \frac{1}{\sigma_\eta^2} (\mathbf{W_s}^T \mathbf{W_s} \mathbf{z} - \mathbf{W_s}^T \mathbf{y}) + C_z^{-1} \mathbf{z}, \tag{19}$$

where

$$\nabla_{\mathbf{z}} L(\mathbf{z}, \mathbf{s}) = \begin{bmatrix} \frac{\partial L(\mathbf{z},\mathbf{s})}{\partial z_1} \\ \frac{\partial L(\mathbf{z},\mathbf{s})}{\partial z_2} \\ \vdots \\ \frac{\partial L(\mathbf{z},\mathbf{s})}{\partial z_N} \end{bmatrix}. \tag{20}$$

At iteration n, we fix the motion parameters in (19) such that $\mathbf{s} = \hat{\mathbf{s}}^n$. Then by setting $\nabla_{\mathbf{z}} L(\mathbf{z}, \mathbf{s})|_{\mathbf{s}=\hat{\mathbf{s}}^n}$ equal to zero and solving for $\mathbf{z}$ we obtain the the next image estimate

$$\hat{\mathbf{z}}^{n+1} = \left[\mathbf{W}_{\hat{\mathbf{s}}^n}^T \mathbf{W}_{\hat{\mathbf{s}}^n} + \sigma_\eta^2 C_z^{-1} \right]^{-1} \mathbf{W}_{\hat{\mathbf{s}}^n}^T \mathbf{y}. \tag{21}$$

This process of updating the motion parameters and the image continues for $n = 0, 1, 2, \ldots$ until $L(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n)$ stabilizes or $||\hat{\mathbf{z}}^{n+1} - \hat{\mathbf{z}}^n||/||\hat{\mathbf{z}}^n|| < \epsilon$, where ϵ is a predetermined threshold. It is possible that the functional in (17) contains local minima and the proposed optimization procedure can get trapped by these. Thus, it may be important to begin the procedure with the the best estimate of $\mathbf{z}$ possible.

A practical problem may arise in implementing (21), even for modest sized images, due to the large matrix dimensions. Although the matrices $\mathbf{W_s}$ and C_z^{-1} are generally sparse, it is often more practical to use a gradient descent procedure rather than perform the matrix inverse in (21). This approach is described in the following section.

3.3 Gradient Descent Optimization

In this section, we derive an iterative gradient descent minimization procedure with respect to the high resolution image $\mathbf{z}$. A pointwise notation is used because we believe it may provide the reader with additional insight into the mechanics of the algorithm.

To begin, we first write all the densities in (9) in terms of individual high resolution pixel values. Consider first the conditional density obtained using (3) and (5)

$$Pr(\mathbf{y}|\mathbf{z},\mathbf{s}) = \frac{1}{(2\pi)^{\frac{pM}{2}} \sigma_\eta^{pM}} exp \left\{ -\frac{1}{2\sigma_\eta^2} \sum_{m=1}^{pM} \left(y_m - \sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r \right)^2 \right\}. \qquad (22)$$

Using (9) with the conditional density in (22) and the prior in (12), the MAP estimates can be formulated as

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \begin{matrix} \arg\min \\ \mathbf{z}, \mathbf{s} \end{matrix} L(\mathbf{z},\mathbf{s}), \qquad (23)$$

where the MAP cost function is now expressed as

$$L(\mathbf{z},\mathbf{s}) = \frac{1}{2\sigma_\eta^2} \sum_{m=1}^{pM} \left(y_m - \sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r \right)^2 + \frac{1}{2\lambda} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} d_{i,j} z_j \right)^2. \qquad (24)$$

Clearly, the cost function in (24) balances two types of errors. The first term will be referred to as the linear equation error. This error is minimized when $\mathbf{z}$, projected through the observation model, matches the observed data. Minimization of this term alone can lead to excessive noise magnification in some applications due to the ill-posed nature of this inverse problem. The second term will be referred to as the image prior error which serves as a regularization operator. This is generally minimized when $\mathbf{z}$ is smooth. The weight of each of these competing "forces" in the cost function is controlled by σ_η^2 and λ. For example, if the fidelity of the observed data is high (i.e., σ_η^2 is small), the linear equation error dominates the cost function. If the observed data is very noisy, the cost function will emphasize the image prior error. This will generally lead to smoother image estimates. Thus, the image prior term can be viewed as a simple penalty term controlled by λ which biases the estimator away from noisy solutions which are inconsistent with the Gibbs image prior assumption.

As before, we will use a cyclic coordinate-descent minimization procedure alternating between the motion parameters and the high resolution image. However, here we use a steepest descent technique to minimize (24) with respect to $\mathbf{z}$, given the motion parameters are fixed. Let us begin with the motion estimation.

At each iteration n, the motion parameter estimates are updated through a search procedure to minimize (24) with respect to $\mathbf{s}$, given that $\mathbf{z} = \hat{\mathbf{z}}^n$. Thus, using only those terms in (24) which are

a function of $\mathbf{s}$, the updated motion parameter estimates become

$$\hat{\mathbf{s}}^n = \underset{\mathbf{s}}{\arg\min} \ L(\hat{\mathbf{z}}^n, \mathbf{s}) = \underset{\mathbf{s}}{\arg\min} \left\{ \sum_{k=1}^{p} \sum_{m=1}^{M} \left(y_{k,m} - \sum_{r=1}^{N} w_{k,m,r}(\mathbf{s}_k) \hat{z}_r^n \right)^2 \right\}. \tag{25}$$

Note that the linear equation error in (25) is written in terms of a sum over individual frames. The cost function in (25) can clearly be minimized by minimizing each of the p terms independently. Thus, the estimate for the registration parameters for frame k at iteration n, $\hat{\mathbf{s}}_k^n$, can be found as follows:

$$\hat{\mathbf{s}}_k^n = \underset{\mathbf{s}_k}{\arg\min} \left\{ \sum_{m=1}^{M} \left(y_{k,m} - \sum_{r=1}^{N} w_{k,m,r}(\mathbf{s}_k) \hat{z}_r^n \right)^2 \right\}, \tag{26}$$

for $k = 1, 2, \ldots, p$.

A search is required to minimize (26) with respect to $\mathbf{s}_k$. This search resembles a type of block matching algorithm. However, here the low resolution frames are not compared to one another, as they might be in a traditional block matching algorithm. Rather, they are compared to the high resolution estimate after it is projected through the observation model with set motion parameters. We believe that this procedure is less susceptible to aliasing than traditional block matching algorithms. This is because no interpolation on the low-resolution aliased grid is required as the motion parameters are varied and tested to find a minimum in (26). Pyramidal search strategies and other efficient techniques designed for traditional block matching can be employed here [1].

To derive the pointwise gradient descent update for the image estimate, we begin by differentiating (24) with respect to some pixel z_k for $k = 1, 2, \ldots, N$. This partial derivative is given by

$$g_k(\mathbf{z}, \mathbf{s}) = \frac{\partial L(\mathbf{z}, \mathbf{s})}{\partial z_k} = \frac{1}{\sigma_\eta^2} \sum_{m=1}^{pM} w_{m,k}(\mathbf{s}) \left(\sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r - y_m \right) + \frac{1}{\lambda} \sum_{i=1}^{N} d_{i,k} \left(\sum_{j=1}^{N} d_{i,j} z_j \right). \tag{27}$$

Note that the first term in (27) is the sum of differences between the "predicted data" minus the actual low resolution data. Each term in the sum is weighted by the contribution of z_k to that low resolution pixel, $w_{m,k}(\mathbf{s})$. The second term (the image prior gradient) is simply a linear combination of the high resolution pixels for each k. This portion of the derivative can be computed for all high resolution pixels via a convolution operation. For the coefficients given in (14), the appropriate convolution kernel is shown in Fig. 3. Finally, the gradient descent update for each pixel estimate, given the motion parameters are fixed as $\mathbf{s} = \hat{\mathbf{s}}^n$, is

$$\hat{z}_k^{n+1} = \hat{z}_k^n - \varepsilon^n g_k(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n), \tag{28}$$

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

$$\hat{\mathbf{z}}^{n+1} = \hat{\mathbf{z}}^n - \varepsilon^n \nabla_{\mathbf{z}} L(\mathbf{z}, \mathbf{s})|_{\mathbf{z}=\hat{\mathbf{z}}^n, \mathbf{s}=\hat{\mathbf{s}}^n}. \tag{29}$$

$$\frac{1}{16\lambda}$$

$$\frac{1}{8\lambda} \quad \frac{-1}{2\lambda} \quad \frac{1}{8\lambda}$$

$$\frac{1}{16\lambda} \quad \frac{-1}{2\lambda} \quad \frac{5}{4\lambda} \quad \frac{-1}{2\lambda} \quad \frac{1}{16\lambda}$$

$$\frac{1}{8\lambda} \quad \frac{-1}{2\lambda} \quad \frac{1}{8\lambda}$$
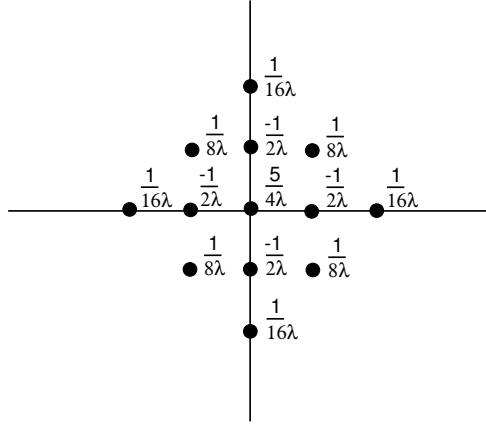
$$\frac{1}{16\lambda}$$

Figure 3: Convolution kernel used to obtain the image prior gradient.

The parameter ε^n in (28) and (29) 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

$$L(\hat{\mathbf{z}}^{n+1}, \hat{\mathbf{s}}^n) = L(\hat{\mathbf{z}}^n - \varepsilon^n \nabla_{\mathbf{z}} L(\mathbf{z}, \mathbf{s})|_{\mathbf{z}=\hat{\mathbf{z}}^n, \mathbf{s}=\hat{\mathbf{s}}^n}, \hat{\mathbf{s}}^n) \tag{30}$$

with respect to ε^n. To do so we begin by writing $L(\hat{\mathbf{z}}^{n+1}, \hat{\mathbf{s}}^n)$ using (24) and (28). Next we differentiate this with respect to ε^n and set the derivative equal to zero. Solving for ε^n yields, after some manipulation,

$$\varepsilon^n = \frac{\frac{1}{\sigma_\eta^2} \sum_{m=1}^{pM} \gamma_m \left(\sum_{r=1}^{N} w_{m,r}(\hat{\mathbf{s}}^n) \hat{z}_r^n - y_m \right) + \frac{1}{\lambda} \sum_{i=1}^{N} \bar{g}_i \left(\sum_{j=1}^{N} d_{i,j} \hat{z}_j^n \right)}{\frac{1}{\sigma_\eta^2} \sum_{m=1}^{pM} \gamma_m^2 + \frac{1}{\lambda} \sum_{i=1}^{N} \bar{g}_i^2}, \tag{31}$$

where

$$\gamma_m = \sum_{r=1}^{N} w_{m,r}(\hat{\mathbf{s}}^n) g_r(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n) \tag{32}$$

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

$$\bar{g}_i = \sum_{j=1}^{N} d_{i,j} g_j(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n) \tag{33}$$

is a weighted sum of neighboring gradient values. A summary of the total estimation procedure is provided in Table 1.

It is worth noting that the conjugate gradient technique [18] cannot be directly applied in place of the gradient descent update in Table 1. Although the conjugate gradient technique generally offers faster convergence, it requires that the cost function remain constant from one iteration to

Table 1: Proposed Iterative MAP Estimation Algorithm.

---

step 1: Begin at $n = 0$ with initial estimate $\hat{\mathbf{z}}^0$ being the interpolated low resolution frame 1.

step 2: For $k = 1, 2, \ldots, p$, find $\hat{\mathbf{s}}_k$ according to (26) to generate $\hat{\mathbf{s}}^n$.

step 3: Compute the gradient $g_k(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n)$ given in (27) for $k = 1, 2, \ldots, N$.

step 4: Compute the optimal step size ε^n using (31).

step 5: Let $\hat{z}_k^{n+1} = \hat{z}_k^n - \varepsilon^n g_k(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n)$ for $k = 1, 2, \ldots, N$, yielding $\hat{\mathbf{z}}^{n+1}$.

step 6: If $||\hat{\mathbf{z}}^{n+1} - \hat{\mathbf{z}}^n|| / ||\hat{\mathbf{z}}^n|| < \epsilon$ or a set number of iterations is reached, stop.

step 7: Let $n = n + 1$ and go to step 2.

---

the next. If the registration parameter estimates change in our iterative procedure, so does the cost function. In light of this, we believe that a straight-forward gradient descent offers a practical solution. However, the conjugate gradient technique, or other similar method, could be applied once the registration parameter estimates have stabilized. Furthermore, such techniques could be used if one applies multiple image update iterations between the motion parameter updates.

4 Experimental Results

In order to demonstrate the performance of the algorithm, several experimental results are presented here. The first set of experiments involve simulated data derived from a single broad band visible image acquired from an airborne platform. These data allow for quantitative error analysis and the weights in the observation model are as shown in Fig. 1. The second set of experimental results have been obtained from a staring infrared imaging system equipped with a microscan mirror to translate the incoming images in real-time. Here a PSF is calculated theoretically based on known optical system parameters and is used in the algorithm. All the results obtained in this section use the Gibbs distribution image prior given by (12) and (14).

While the algorithm developed in Table 1 can be employed with a variety of image motion, the results presented here focus on global translational motion. Such motion may result from an imager mounted on a moving platform, such as an aircraft, with rotational stabilization. The translations may also be introduced by a movable mirror in the optical path. Such platform or mirror motion
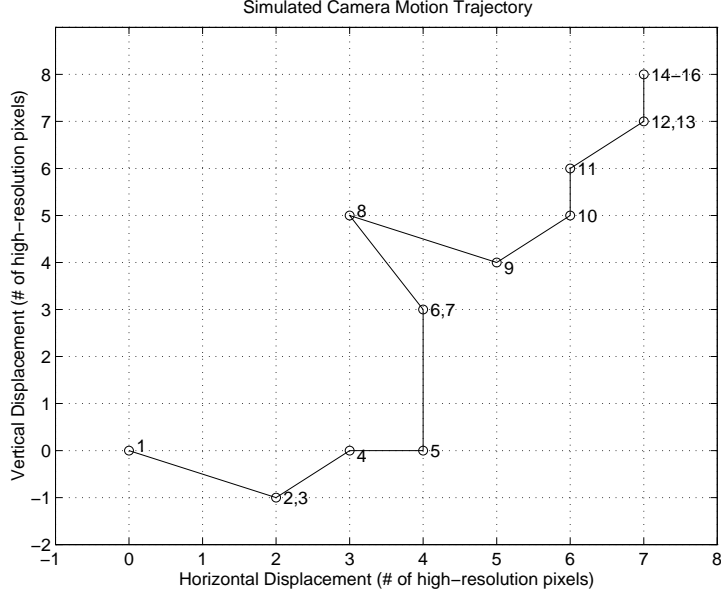
Figure 4: Simulated camera shifts in terms of high resolution pixels for 16 frames.

changes the line-of-sight pointing angle of the imager. For objects in the far field, this effectively produces a translational shift of the scene in the focal plane with minimal occlusion effects.

4.1 Simulated Imagery

A simulated sequence of random translational shifts is shown in Fig. 4. Note that these translational shifts are given in terms of high resolution pixel spacings. Using these shifts, a sequence of translated images is generated from the 8 bit gray-scale image "Aerial" shown in Fig. 5a. This "ideal" image is blurred to simulate the low resolution detector integration and subsampled by a factor of $L_1 = 4$ and $L_2 = 4$. The subsampling is performed with appropriate offsets so as to simulate the translational shifts shown in Fig. 4 and Gaussian noise is added to each low resolution frame. One typical frame of the simulated low-resolution noisy data is shown in Fig. 5b.

The MAP estimate in the case where the noise variance is $\sigma_\eta^2 = 100$ is shown in Fig. 5c. The smoothness parameter for the image prior has been empirically selected to be $\lambda = 150$. Decreasing λ will generally lead to a smoother image estimate. The initial high resolution image estimate is a bilinearly interpolated version of the first frame and 20 iterations of the algorithm are performed. Figure 5d shows the estimate in the case where $\lambda = \infty$, which is effectively a maximum likelihood estimate since the image prior term plays no role in the estimate. Clearly the noise has been magnified in this estimate which demonstrates that the image prior term plays an important role in this case.

It should be noted that the ML estimate in Fig. 5d can be improved by halting the iterations earlier. With fewer iterations, the estimate will tend to retain more of the characteristics of the
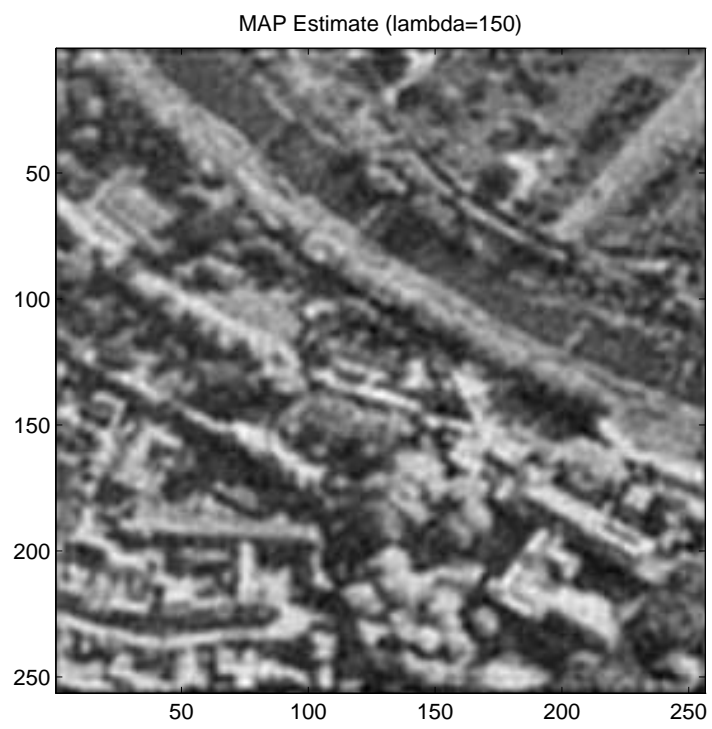
14

Original Image Aerial

(a)

Noisy Low–Resolution Observed Frame 1
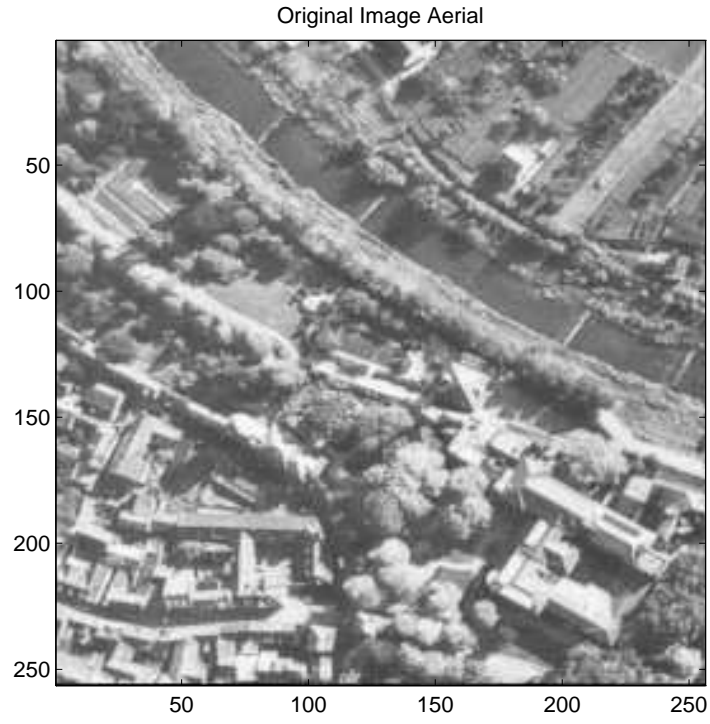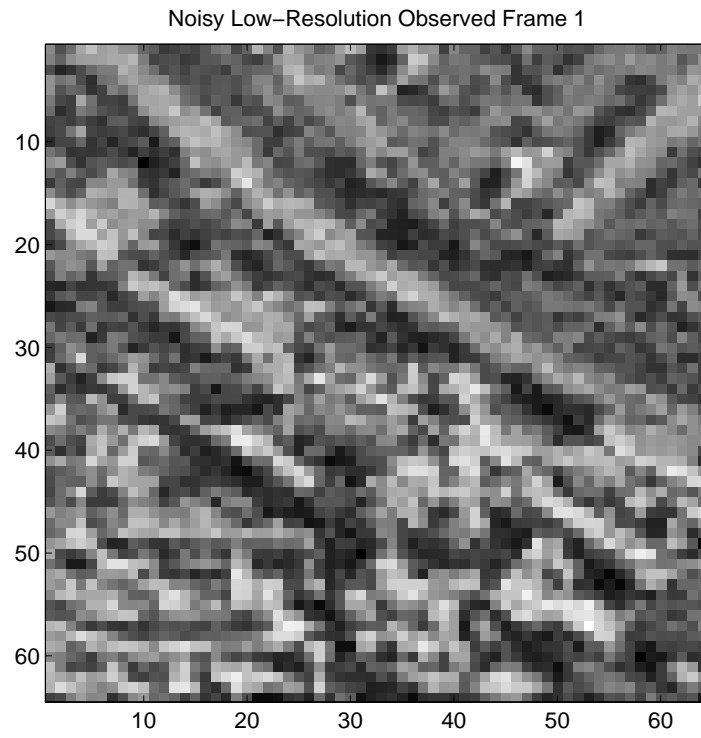
(b)

Figure 5: (a) Original image "Aerial" (b) simulated low resolution frame 1 ($L_1 = L_2 = 4$, $\sigma_\eta^2 = 100$) (c) MAP estimate with $\lambda = 150$ (d) MAP estimate with $\lambda = \infty$ (e) bilinear interpolation of frame 1 (f) bicubic interpolation of frame 1.
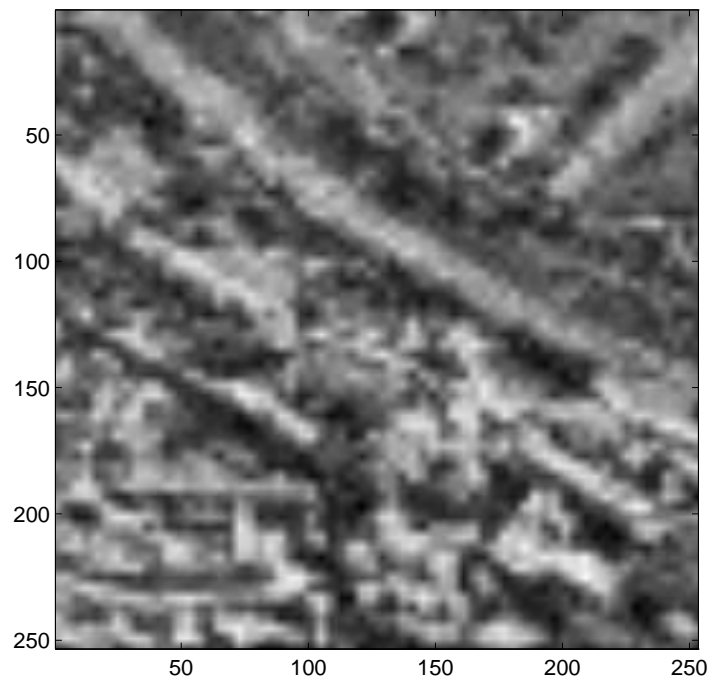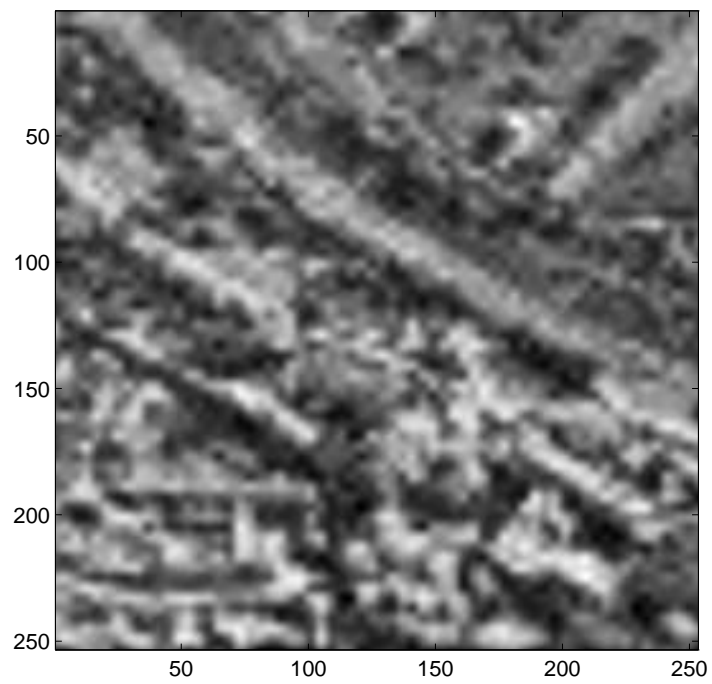
MAP Estimate (lambda=150)

(c)

MAP Estimate (lambda=infinity)

(d)

16

Bilinear Interpolation of Frame 1

(e)

Bicubic Interpolation of Frame 1

(f)

17
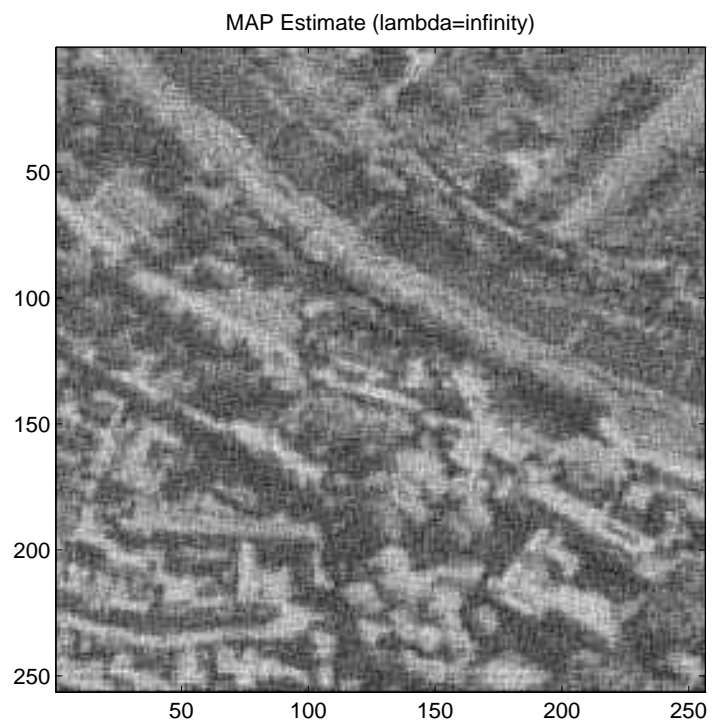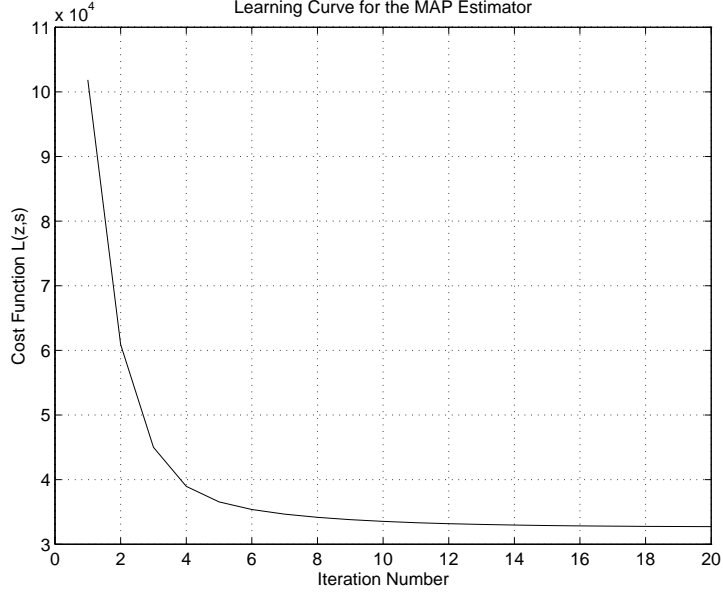
Figure 6: Learning curve showing the MAP cost function, $L(\mathbf{z}, \mathbf{s})$, versus iteration number for the simulated data. Here 16 frames are used with $L_1 = L_2 = 4$, $\sigma_\eta^2 = 100$, and $\lambda = 150$.

starting image (which is generally smooth). Furthermore, constraints can be added such as non-negativity and maximum pixel values which may also improve the ML estimate without the explicit use of a prior term. Thus, Fig. 5d is not intended to be representative of an optimal ML estimate. Rather, it simply shows the effect of eliminating the prior term with all other factors being equal. For comparison with the multiframe estimates, the images formed using bilinear and bicubic interpolation of the single frame in Fig. 5b are shown in Figs. 5e and 5f, respectively.

To show the convergence behavior of the estimator, the MAP cost function in (24) is plotted in Fig. 6 versus iteration number. The use of the optimal step size allows for relatively quick convergence.

Quantitative error analysis is provided in Fig. 7. In particular, the mean absolute error (MAE) per pixel for the MAP estimates with $\lambda = 150$ and $\lambda = \infty$ are shown in Fig. 7a for various noise levels. For each noise level, except the highest, the MAP estimator with $\lambda = 150$ recovers the correct shifts. In the highest noise case, two shift estimates are off by one high resolution pixel. For comparison, the MAEs of images formed by bilinear and bicubic interpolation of a single frame are shown. Note that the MAP estimator provides an estimate with significantly lower error than the single frame interpolators. Also note that the MAE for the multiframe estimate with $\lambda = \infty$ degrades much more rapidly than with $\lambda = 150$ as noise levels increase. Similar results have been observed using the mean squared error (MSE) criterion.

The MAE for the estimators using different numbers of frames is plotted in Fig. 7b. With only one frame, the MAP estimator's performance with $\lambda = 150$ is comparable to that of the bicubic

18
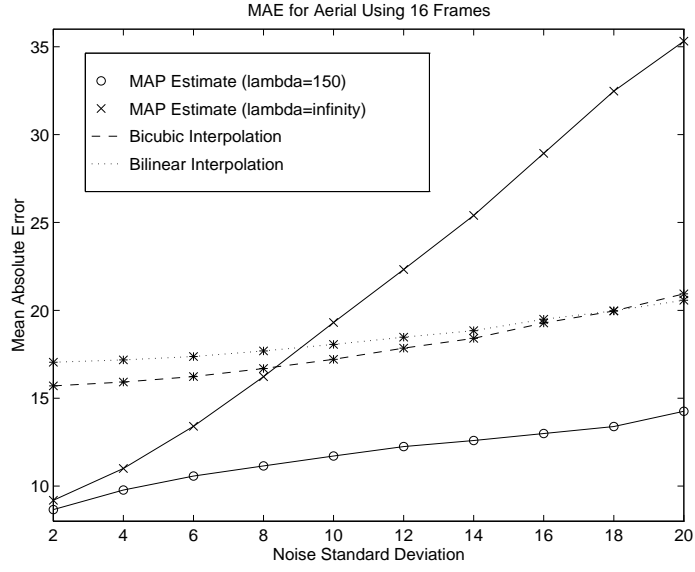
(a)
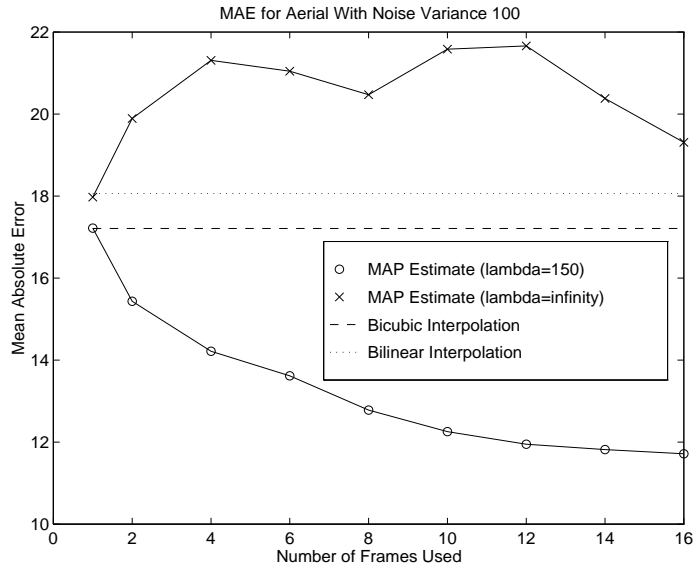
(b)

Figure 7: Mean absolute error for the estimators with (a) varying levels of noise using 16 frames and (b) different number of frames with noise variance $\sigma_\eta^2 = 100$.

interpolator. However, with additional frames, the MAP estimate becomes significantly improved with respect to the single frame interpolators. Similar results are observed with MSE. It is reported in [15] that a single frame MAP interpolation method can outperform traditional interpolators if an edge preserving Hubor-Markov prior is used. However, this is at the expense of optimization simplicity. Note also that the estimate formed with $\lambda = \infty$ has a much higher error which does not decrease monotonically with the number of frames. The algorithm does not appear to be highly sensitive to the choice of λ in this application, but clearly some finite value is desired.

4.2 Staring Infrared Imagery

Now we consider applying the MAP algorithm to imagery obtained with a staring infrared imaging sensor. The imager considered uses a 128×128 Amber AE-4128 infrared focal plane array (FPA). The FPA is composed of Indium-Antimonide (InSb) detectors with a spectral bandwidth from $3\mu m - 5\mu m$ in wavelength. The infrared imager is set up with a microscanning mirror which allows for the collection of a sequence of frames translated by subpixel amounts in real-time. The mirror is controlled by a piezoelectric positioner. However, calibration of the mirror for high speed operation is very difficult, and therefore, the exact mirror positions are unknown. For our purposes, the mirror effectively simulates an unknown imager platform motion.

The imager is equipped with $F/3$ optics having a focal length of $100mm$. The detector pitch or spacing is $50\mu m$ in both the horizontal and vertical directions. The active area of the detectors is $40\mu m \times 40\mu m$, yielding a fill factor of 80%. Assuming a midband wavelength of $4\mu m$, the theoretical diffraction-limited optical transfer function (OTF) cuts off at 8.33 cyc/mrad [27, 28] and the effective sampling rate of the FPA is 2 cyc/mrad. Thus, the sampling rate must be increased by a factor of 8.33 to guarantee there will be no aliasing effects when imaging an arbitrary scene.

We have selected to perform a reconstruction with $L_1 = L_2 = 5$, although this resolution may not be sufficient to avoid aliasing effects entirely. The theoretical discrete PSF of the imaging system on the high resolution grid is shown in Fig. 8. This is an incoherent PSF based on the assumption of diffraction-limited optics and includes the effects of the finite detector size [27, 28]. The weights in (2) are determined by this PSF positioned at the appropriate location on the high resolution grid for each low resolution pixel.

The MAP algorithm is tested using 16 frames of the infrared imagery. One typical original resolution frame is shown in Fig. 9a. This 64×64 image represents a region of interest from the full 128×128 Amber array. No subsampling has been performed. The scene contains a number of vehicles imaged from a tower at a distance of approximately 1 kilometer. The MAP estimate is shown in Fig. 9b for $\lambda = 200$ and where σ_η^2 is estimated from a small apparently "flat" portion of the original image sequence. The estimate formed with $\lambda = \infty$ is shown in Fig. 9c. Since
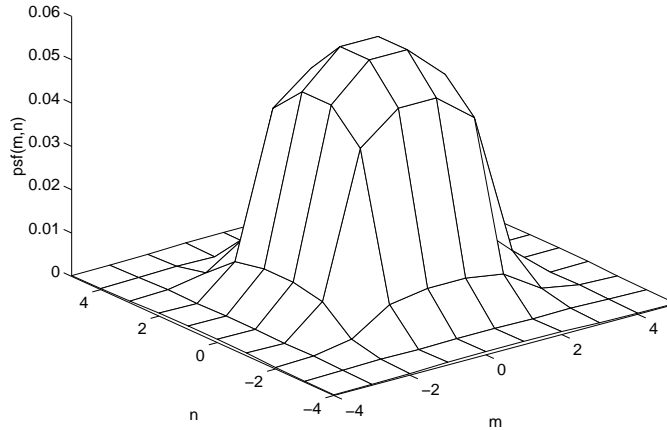
Figure 8: Theoretical imaging system PSF on the high resolution grid where $L_1 = L_2 = 5$. The PSF is based on the assumption of diffraction-limited optics and includes the effects of the finite detector size.

the signal-to-noise ratio is relatively high in these data, the estimates formed with $\lambda = 200$ and $\lambda = \infty$ are similar. Again, decreasing λ will generally produce a smoother image estimate. For comparison, a single frame interpolated using bilinear interpolation is shown in Fig. 9d and using bicubic interpolation in Fig. 9e.

To illustrate the convergence behavior of the algorithm with the infrared data, the cost function is plotted in Fig. 10 versus iteration number for $\lambda = 200$. Here the initial estimate of the high resolution image is a bicubically interpolated version of the first frame. Note the sudden drop between the forth and fifth iterations. Here the shift estimates changed significantly allowing the cost function to reach a lower level. The image shifts recovered by the MAP estimator are shown in Fig. 11. These shifts appear consistent with the programmed mirror pattern.

5 Conclusions

The problem of aliasing reduction and resolution enhancement can be addressed by exploiting multiple frames which offer unique "looks" at a scene. Here we have focused on exploiting frame-to-frame translational shifts which may result from line-of-sight jitter of an imager mounted on a moving platform. However exploiting these sub-pixel motions requires accurate estimates of them. To obtain accurate estimates of the sub-pixel motion, any scene based estimator must have accurate sub-pixel information in some form. We approached this apparently paradoxical problem from the framework of blind image restoration. In particular, a joint MAP estimator for the high resolution image and the registration parameters is described. We believe that this type of simultaneous estimation is

Noisy Low-Resolution Observed Frame 1
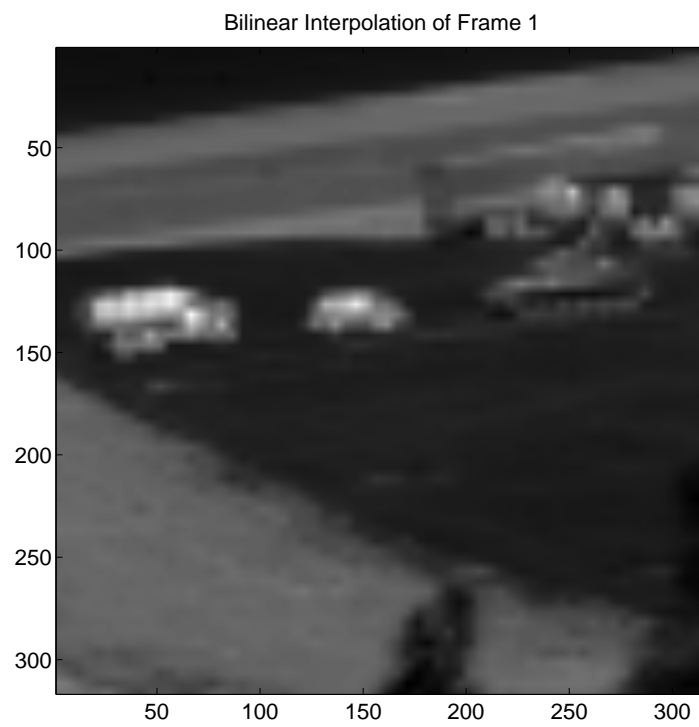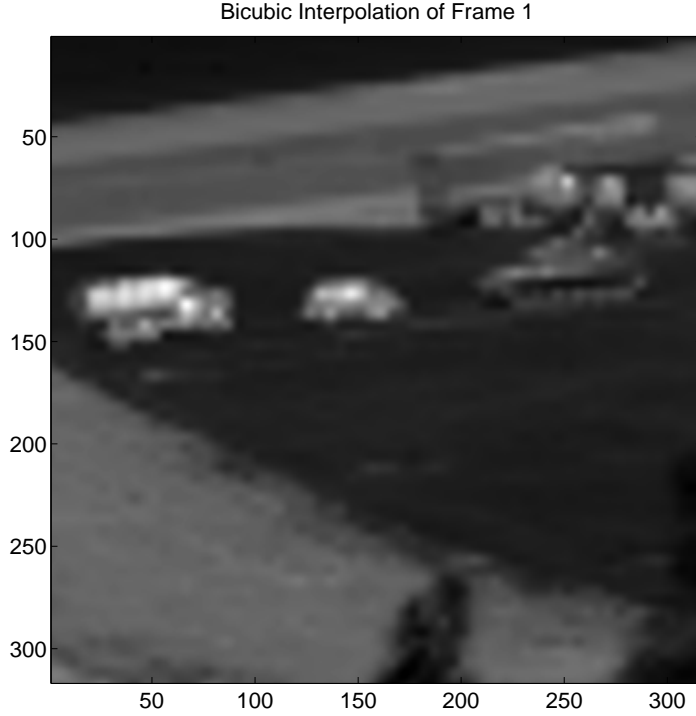
(a)

MAP Estimate (lambda=200)
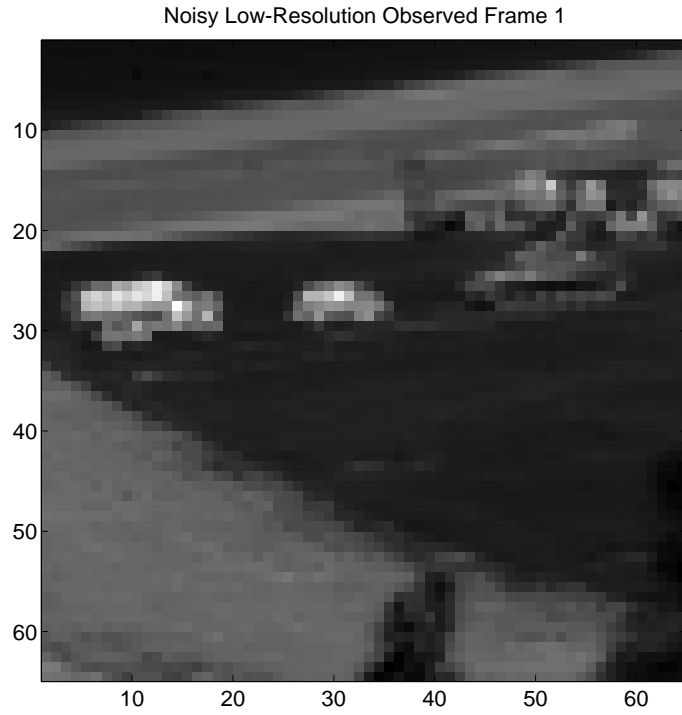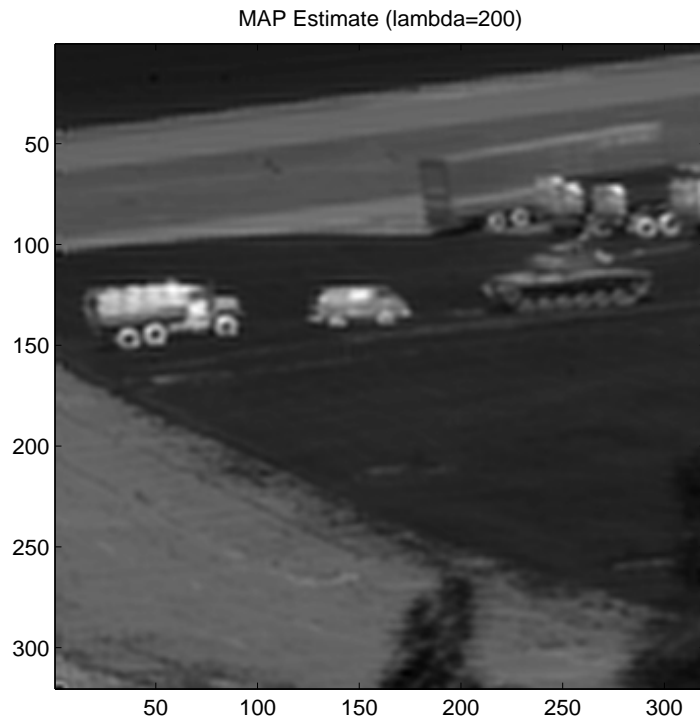
(b)

Figure 9: (a) Low resolution frame 1 showing vehicles imaged from a tower (b) MAP estimate using 16 frames with $L_1 = L_2 = 5$ and $\lambda = 200$ (c) MAP estimate with $\lambda = \infty$ (d) bilinear interpolation of frame 1 (e) bicubic interpolation of frame 1.

22

MAP Estimate (lambda=infinity)

(c)

Bilinear Interpolation of Frame 1

(d)

23

Bicubic Interpolation of Frame 1
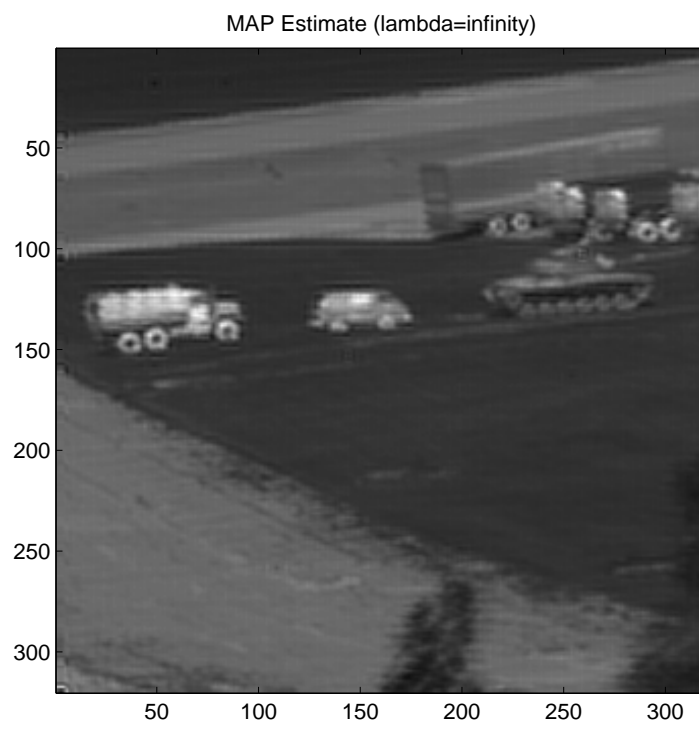
(e)

important if the frames are severely aliased. It has been demonstrated that the MAP algorithm can accurately recover sub-pixel translational shifts, even with significantly aliased imagery. Also, the high resolution image estimates obtained here have significantly lower error than estimates formed by single frame interpolation. Although our experimental results focus on global translational shift, the fundamental approach described here can be applied to more complex scene motion. This will be one area of future work. Future investigation will also explore optimal selection methods for the smoothness parameter λ to avoid empirical searches.
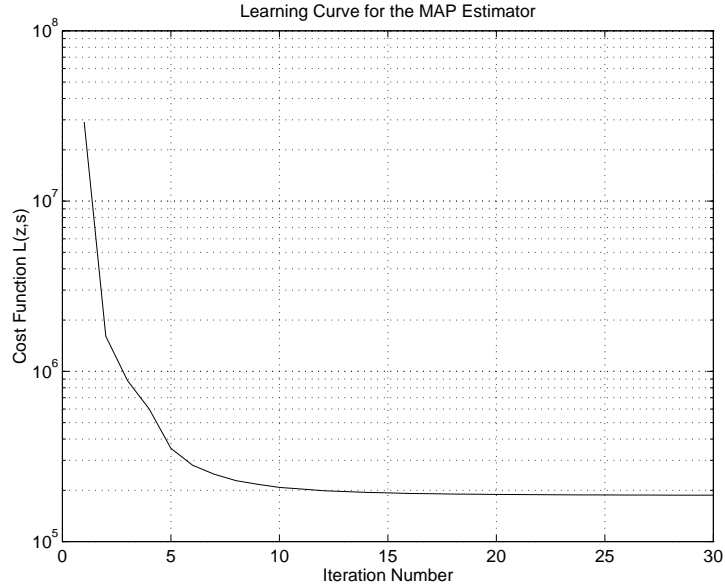
6 Acknowledgments

Figure 10: Learning curve showing the MAP cost function for the infrared data as a function of iteration number where $L_1 = L_2 = 5$ and $\lambda = 200$.

References

[1] A. M. Tekalp, *Digital Video Processing*. Prentice Hall, 1995.

[2] J. C. Gillette, T. M. Stadtmiller, and R. C. Hardie, "Reduction of aliasing in staring infrared imagers utilizing subpixel techniques," *Optical Engineering*, vol. 34, pp. 3130–3137, Nov. 1995.

[3] 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.

[4] 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.

[5] 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*, vol. 38, pp. 1013–1027, June 1990.
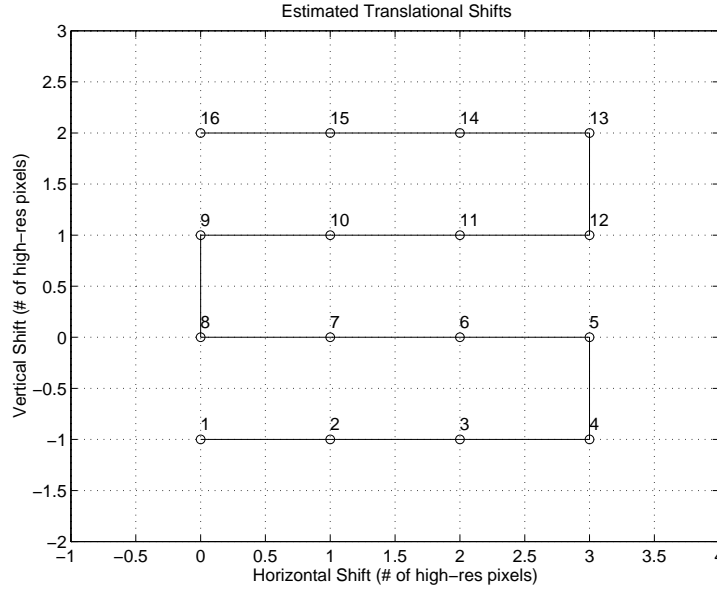
Figure 11: Estimated infrared imager shifts in terms of high resolution pixels for 16 frames.

[6] S. Kim and W.-Y. Su, "Recursive high-resolution reconstruction of blurred multiframe images," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 2, pp. 534–539, Oct. 1993.

[7] E. Kaltenbacher and R. C. Hardie, "High-resolution infrared image reconstruction using multiple low-resolution aliased frames," in *Proceedings of IEEE National Aerospace Electronics Conference (NAECON), Dayton, OH*, May 1996.

[8] H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays, using convex projections," *Journal of the Optical Society of America A*, vol. 6, no. 11, pp. 1715–1726, 1989.

[9] A. Patti, M. Sezan, and A. Tekalp, "High-resolution image reconstruction from a low-resolution image sequence in the presence of time-varying motion blur," in *Proceedings of IEEE International Conference on Image Processing, Austin, TX*, Nov. 1994.

[10] 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.

[11] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing*, vol. 53, pp. 231–239, May 1991.

[12] 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.

[13] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Transactions on Image Processing*, vol. 5, pp. 996–1011, June 1996.

[14] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Detroit MI*, May 1995.

[15] R. R. Schultz and R. L. Stevenson, "A bayesian approach to image expansion for improved definition," *IEEE Transactions on Image Processing*, vol. 3, pp. 233–242, May 1994.

[16] 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.

[17] 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.

[18] D. G. Luenberger, *Linear and Nonlinear Programming*. Addison-Wesley, 1984.

[19] E. L. Dereniak and G. D. Boreman, *Infrared Detectors and Systems*. Wiley, 1996.

[20] W. F. Schreiber, *Fundamentals of Electronic Imaging Systems, Second Edition*. Springer-Verlag, 1991.

[21] C. W. Helstrom, *Probability and Stochastic Processes for Engineers*. Macmillan, 1984.

[22] A. Papoulis, *Probability, Random Variables, and Stochastic Processes, Second Edition*. McGraw-Hill, 1984.

[23] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, Nov. 1984.

[24] C. Bouman and K. Sauer, "A generalized gaussian image model for edge-preserving map estimation," *IEEE Transactions on Image Processing*, vol. 2, pp. 296–310, July 1993.

[25] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Royal Stat. Soc. B*, vol. 36, no. 2, 1974.

[26] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using gibbs random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, Jan. 1987.

[27] J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics.* Wiley, 1978.

[28] J. Goodman, *Introduction to Fourier Optics.* McGraw-Hill, 1968.