

9-2012

Fast Super-Resolution Using an Adaptive Wiener Filter with Robustness to Local Motion

Russell C. Hardie

University of Dayton, rhardie1@udayton.edu

Kenneth J. Barnard

Air Force Wright Laboratory

Follow this and additional works at: https://ecommons.udayton.edu/ece_fac_pub

 Part of the [Computer Engineering Commons](#), [Optics Commons](#), [Other Electrical and Computer Engineering Commons](#), and the [Other Physics Commons](#)

eCommons Citation

Hardie, Russell C. and Barnard, Kenneth J., "Fast Super-Resolution Using an Adaptive Wiener Filter with Robustness to Local Motion" (2012). *Electrical and Computer Engineering Faculty Publications*. 61.

https://ecommons.udayton.edu/ece_fac_pub/61

This Article is brought to you for free and open access by the Department of Electrical and Computer Engineering at eCommons. It has been accepted for inclusion in Electrical and Computer Engineering Faculty Publications by an authorized administrator of eCommons. For more information, please contact frice1@udayton.edu, mschlangen1@udayton.edu.

Fast super-resolution using an adaptive Wiener filter with robustness to local motion

Russell C. Hardie^{1,*} and Kenneth J. Barnard²

¹ *Department of Electrical and Computer Engineering, University of Dayton, 300 College Park, Dayton, Ohio 45469-0232, USA*

² *Air Force Research Laboratory, AFRL/RYSM, Building 620, 2241 Avionics Circle, Wright-Patterson AFB, Ohio 45433, USA*

[*rhardie@udayton.edu](mailto:rhardie@udayton.edu)

Abstract: We present a new adaptive Wiener filter (AWF) super-resolution (SR) algorithm that employs a global background motion model but is also robust to limited local motion. The AWF relies on registration to populate a common high resolution (HR) grid with samples from several frames. A weighted sum of local samples is then used to perform nonuniform interpolation and image restoration simultaneously. To achieve accurate subpixel registration, we employ a global background motion model with relatively few parameters that can be estimated accurately. However, local motion may be present that includes moving objects, motion parallax, or other deviations from the background motion model. In our proposed robust approach, pixels from frames other than the reference that are inconsistent with the background motion model are detected and excluded from populating the HR grid. Here we propose and compare several local motion detection algorithms. We also propose a modified multiscale background registration method that incorporates pixel selection at each scale to minimize the impact of local motion. We demonstrate the efficacy of the new robust SR methods using several datasets, including airborne infrared data with moving vehicles and a ground resolution pattern for objective resolution analysis.

© 2012 Optical Society of America

OCIS codes: (100.6640) Superresolution; (280.4991) Passive remote sensing; (110.3080) Infrared imaging.

References and links

1. S. C. Park, M. K. Park and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Mag.* **20**, 21–36 (2003).
2. R. C. Hardie, R. R. Schultz and K. E. Barner, "Super-resolution enhancement of digital video," *EURASIP J. Adv. Signal Process.* **2007**, 19–19 (2007).
3. B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of International Joint Conference on Artificial Intelligence* (Vancouver, 1981), pp. 674–679.
4. M. S. Alam, J. G. Bognar, R. C. Hardie and B. J. Yasuda, "Infrared image registration using multiple translationally shifted aliased video frames," *IEEE Trans. Instrum. Meas.* **49** (2000).
5. M. Irani and S. Peleg, "Improving resolution by image registration," *CHIP: Graph. Models Image Process.* **53**, 231–239 (1991).
6. R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong and E. A. Watson, "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Opt. Eng.* **37**, 247–260 (1998).
7. R. C. Hardie, K. J. Barnard and R. Ordonez, "Fast super-resolution with affine motion using an adaptive wiener filter and its application to airborne imaging," *Opt. Express* **19**, 26208–26231 (2011).

8. M. D. Robinson and P. Milanfar, "Fundamental performance limits in image registration," *IEEE Trans. Image Processing* **13**, 1185–1199 (2004).
9. M. D. Robinson and P. Milanfar, "Statistical performance analysis of super-resolution," *IEEE Trans. Image Processing* **15**, 1413–1428 (2006).
10. R. C. Hardie, "Super-resolution using adaptive wiener filters," in *Super-Resolution Imaging*, P. Milanfar, ed. (Taylor & Francis/CRC Press, 2010), pp. 35–61.
11. R. C. Hardie, "A fast super-resolution algorithm using an adaptive wiener filter," *IEEE Trans. Image Processing* **16**, 2953–2964 (2007).
12. B. Narayanan, R. C. Hardie, K. E. Barner and M. Shao, "A computationally efficient super-resolution algorithm for video processing using partition filters," *IEEE Trans. Circuits Syst. Video Technol.* **17**, 621–634 (2007).
13. M. Shao, K. E. Barner and R. C. Hardie, "Partition-based interpolation for color filter array demosaicking and super-resolution reconstruction," *Opt. Eng.* **44**, 107003–1–107003–14 (2005).
14. T. R. Tuinstra and R. C. Hardie, "High resolution image reconstruction from digital video by exploitation on non-global motion," *Opt. Eng.* **38** (1999).
15. M. Tanaka and M. Okutomi, "Towards robust reconstruction-based superresolution," in *Super-Resolution Imaging*, P. Milanfar, ed. (Taylor & Francis/CRC Press, 2010), pp. 219–246.
16. A. W. M. van Eekeren, K. Schutte and L. J. van Vliet, "Multiframe super-resolution reconstruction of small moving objects," *IEEE Trans. Image Processing* **19**, 2901–2912 (2010).
17. M. Kim, B. Ku, D. Chung, H. Shin, D. Han and H. Ko, "Robust video super resolution algorithm using measurement validation method and scene change detection," *EURASIP J. Adv. Signal Process.* **2011**, 1–12 (2011). 10.1186/1687-6180-2011-103.
18. Z. Zhang and R. Wang, "Robust image superresolution method to handle localized motion outliers," *Opt. Eng.* **48**, 077005 (2009).
19. J. Dijk, A. W. M. van Eekeren, K. Schutte, D.-J. J. de Lange and L. J. van Vliet, "Superresolution reconstruction for moving point target detection," *Opt. Eng.* **47**, 096401 (2008).
20. N. A. El-Yamany and P. E. Papamichalis, "Robust color image superresolution: an adaptive m-estimation framework," *J. Image Video Process.* **2008**, 16:1–16:12 (2008).
21. M. K. Park, M. G. Kang and A. K. Katsaggelos, "Regularized high-resolution image reconstruction considering inaccurate motion information," *Opt. Eng.* **46**, 117004 (2007).
22. Z. A. Ivanovski, L. Panovski and L. J. Karam, "Robust super-resolution based on pixel-level selectivity," *Proc. SPIE* **6077**, 607707 (2006).
23. S. Farsiu, D. Robinson, M. Elad and P. Milanfar, "Fast and robust multi-frame super-resolution," *IEEE Trans. Image Processing* **13**, 1327–1344 (2004).
24. S. Farsiu, S. Farsiu, S. Farsiu, D. Robinson, D. Robinson, M. Elad, M. Elad, P. Milanfar and P. Milanfar, "Advances and challenges in super-resolution," *Int. J. Imag. Syst. Tech.* **14**, 47–57 (2004).
25. F. O. Baxley, K. J. Barnard, R. C. Hardie and M. A. Bicknell, "Flight test results of a rapid step-stare and microscan midwave infrared sensor concept for persistent surveillance," in *Proceedings of MSS Passive Sensors* (Orlando, FL, 2010).
26. R. D. Fiete, "Image quality and λ FN/ p for remote sensing systems," *Opt. Eng.* **38**, 1229–1240 (1999).
27. R. Franzen, "Kodak lossless true color image suite," <http://r0k.us/graphics/kodak>.
28. P. Burt and E. Adelson, "The laplacian pyramid as a compact image code," *IEEE Trans. Communications* **31**, 532–540 (1983).
29. A. C. Bovik, *The Essential Guide to Video Processing* (Academic Press, 2009), 2nd ed.
30. S. Coles, *An introduction to statistical modeling of extreme values*, Springer Series in Statistics (Springer-Verlag, London, 2001).

1. Introduction

Multiframe super-resolution (SR) techniques have proven to be highly effective for image restoration with imaging systems limited by detector sampling [1]. Such detector SR algorithms [2] fuse multiple temporal frames to produce a high resolution output with reduced aliasing, blur, and noise. Detector SR algorithms generally seek to produce imagery with the highest resolution afforded by the optics. In contrast, optical SR methods [2], seek to produce imagery with spatial frequency content above the diffraction limited cut-off frequency of the optics. To treat aliasing, detector SR methods rely on accurate subpixel registration. The challenge is that this registration must generally be performed using the observed imagery that is corrupted by aliasing and noise. These degradations negatively impact ones ability to obtain highly accurate subpixel registration. However, using a parametric motion model with relatively few parameters and a large image region to estimate those parameters, one can usually

obtain useful registration results with a suitable regression method. For example, the Lucas-Kanade [3] registration method employs least squares and has proven to be highly effective for SR with translation [4], rotation [5, 6], and affine motion models [7]. A study of the performance limits for this type of image registration can be found in [8] and a related statistical performance analysis of SR is provided in [9].

For interframe motion that includes small moving objects, deformable motion, motion parallax, and other complex optical flow, accurate subpixel registration from noisy aliased data can be elusive. Consequently, achieving significant aliasing reduction with such data using SR is very difficult. However, in many imaging applications, the bulk of the imagery does closely follow a relatively simple background motion model. For example, with airborne imagery such as that in considered in [7], it is shown that an affine motion model can be effective for SR on the static portions of the scene. Local motion effects from motion parallax and moving objects are not considered in that paper.

In this paper, we extend the work in [7, 10–13] and propose a novel adaptive Wiener filter (AWF) SR method that is robust to limited local motion. The AWF SR method relies on registration to populate a common high resolution (HR) grid with samples from several frames. A weighted sum of local samples is then used to perform nonuniform interpolation and image restoration simultaneously. Previously, only global motion has been considered with AWF SR. By adding robustness to local motion, the applicability and utility of the AWF SR method is greatly enhanced. In our proposed approach, we begin with robust global registration using a suitable parametric background motion model. This generally allows for accurate subpixel registration for much of the image. We then employ local motion detection. Pixels found to be inconsistent with the background motion model are excluded from contributing to the AWF filter output to prevent distortions in the output. Note that in some cases, larger rigid objects can be segmented and registered in a manner similar to the background. This has been demonstrated using other SR methods in several papers including [14–16]. Smaller objects obeying well defined motion trajectories have also been treated in [16]. With local motion owing to motion parallax, deformable motion, object pose variations, and small objects with unpredictable trajectories, we believe the detect-and-exclude approach is practical and can keep computational complexity low.

An alternative approach to addressing local motion is to use a much more complex motion model in an attempt to account for all motion present in the video sequence. However, we shall show that subpixel registration accuracy is inversely proportional to the block size used. Thus, if we break up the image into smaller blocks to account for complex local motion, the ability to achieve the true subpixel registration necessary for aliasing reduction is reduced or lost (without further constraining assumptions). Thus, small objects moving in aliased imagery may simply defy attempts to register them with subpixel accuracy. Furthermore, local motion presents a problem for nonuniform interpolation SR algorithms like the AWF that assume commutation of the motion and the blurring operators in the observation model. Unless the motion is very limited, this assumption may be violated with the presence of local motion. So rather than incurring likely registration and restoration errors from local motion, we tackle the somewhat easier problem of detecting and segmenting regions exhibiting deviations from the background motion model.

The problem of adding local motion robustness to SR algorithms is receiving increased attention in recent years as the research community seeks to develop practical fielded SR systems that may encounter a variety of complex imaging conditions [15, 17–24]. Much of this work has focused on iterative SR algorithms with a relatively high computational complexity. This paper is novel in that we are focusing on adding robustness to the fast AWF SR method and we are exploring both terrestrial and airborne imaging applications [7]. We believe this paper

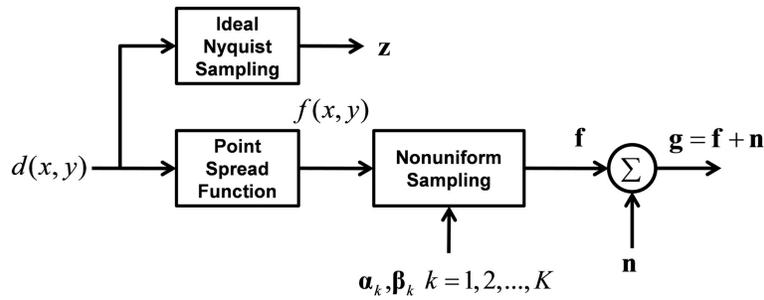


Fig. 1. Assumed observation model relating a desired 2-D continuous scene, $d(x, y)$, to the observed LR image samples, \mathbf{g} , from a set of K frames. This model is the basis for the nonuniform interpolation based SR algorithms, including the AWF SR method.

makes several contributions. First, we present an analysis of the Lucas-Kanade [3] registration method accuracy as a function of the block size used for different levels of aliasing. This provides insight into the problem of registering small moving objects in the presence of aliasing. We propose a variation of Lucas-Kanade that uses a multiscale approach with pixel selection at each scale to minimize the impact of local motion on the background estimation. We also provide a novel analysis of several local motion detection methods for operating in aliased imagery. We evaluate the performance of the methods using a receiver operating characteristic (ROC) curve analysis on simulated data. We incorporate several of these local motion detection algorithms into the AWF SR framework to giving rise to new and robust AWF SR methods. We demonstrate the efficacy of the robust AWF SR methods using several datasets, including real airborne infrared data with moving vehicles and a ground resolution pattern for objective resolution analysis [25].

The remainder of this paper is arranged as follows. In Section 2, we present an overview of the proposed new robust AWF SR algorithms. Robust registration is considered in Section 3 and local motion detection is explored in Section 4. Robust SR results are provided in Section 5. Finally, conclusions are offered in Section 6.

2. Robust adaptive Wiener filter based super-resolution

In this section, we provide an overview of the new AWF SR method that is robust to local motion. We begin with the observation model and then we present the top level AWF SR algorithm. Sections 3 and 4 examine the robust registration and local motion detection components, respectively.

2.1. Observation model

Nonuniform interpolation SR algorithms [1], including the AWF SR method, are based on an assumed observation model similar to that shown in Fig. 1. Here the desired continuous image is given by $d(x, y)$ and the ideally sampled image is represented by the vector $\mathbf{z} = [z_1, z_2, \dots, z_N]^T$ using lexicographical notation. The continuous image $f(x, y)$ represents the desired image after convolution with the point spread function (PSF). Observing K low resolution (LR) frames with global and local motion gives rise to a set of samples that in general are nonuniformly distributed when placed on a common grid. Let these samples be represented by $\mathbf{f} = [f_1, f_2, \dots, f_M]^T$. Here the background interframe motion for frame k is described by the parameters in β_k , and local motion is described by α_k . Local motion will be characterized simply by a detection mask showing pixels in each observed frame that do not obey the background motion model. More

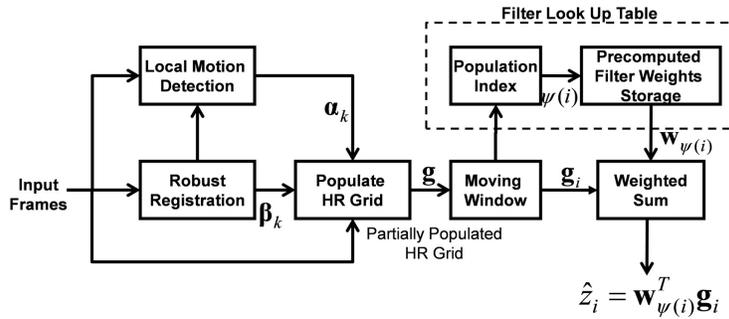


Fig. 2. Robust AWF SR with local motion detection. Only pixels consistent with the motion model (i.e., valid pixels) are used to populate the HR grid. Samples on the HR grid are filtered to provide the final AWF SR output.

will be said about how local motion is detected and treated in Section 4. Finally, additive noise is assumed in Fig. 1 yielding $\mathbf{g} = \mathbf{f} + \mathbf{n}$, where $\mathbf{n} = [n_1, n_2, \dots, n_M]^T$ is an array of noise samples. We shall assume zero-mean independent and identically distributed Gaussian noise with a variance of σ_η^2 .

The model in Fig. 1 is convenient as it gives justification for the fast nonuniform interpolation SR methods, where a uniform set of samples of $f(x, y)$ are estimated from \mathbf{g} and some form of image restoration is applied to deconvolve the PSF blur and reduce noise. However, the model effectively incorporates the PSF prior to the interframe motion (which is embedded in the nonuniform sampling block). The physical image acquisition process would have the motion before the PSF. Thus, the validity of the model in Fig. 1 hinges on the commutation of the PSF and motion models. This issue is treated in detail in [7] for global affine motion. It is shown in [7] that for limited zoom and shear and typical PSFs, the commutation error is negligibly small. This important result opens the door for fast nonuniform interpolation based SR methods to be applied beyond simple global translational interframe motion. We shall rely on this result when employing an affine background motion model for airborne applications. With local motion, like that from a fast moving object, the commutation of the PSF and motion model cannot be justified using [7]. In this case, the order of the PSF blur and motion can make a significant difference in the immediate vicinity of the moving object (i.e., within the span of the PSF). To address this issue, our approach is to use only samples from one frame in the immediate vicinity of detected local motion when performing the AWF SR filtering process. Since the AWF SR method is inherently a local moving-window operation, this does not present a major additional computational burden.

With regard to the PSF model, we follow the approach in [7] and model diffraction and detector integration. Other blurring sources could also be incorporated. With a diffraction limited optical system, the spatial cut-off frequency is given by $\omega_c = 1/(\lambda \mathcal{N})$, where λ is the wavelength of light used and \mathcal{N} is the f-number of the optics. To characterize the level of undersampling in such a system, we shall use the parameter $Q = \lambda \mathcal{N} / p$ [26], where p is the detector pitch. Note that the sampling frequency is given by $1/p$ and the Nyquist criterion dictates that $1/p > 2\omega_c$. Therefore, when $Q = 2$, the imaging sensor is sampling at the Nyquist rate. In most imaging systems, a much lower Q is employed [26]. Resolution in these undersampled systems may be thought of as limited by the detector (as opposed to optically limited) [2].

2.2. Robust AWF SR overview

The proposed robust AWF SR methodology is illustrated in Fig. 2. For video processing, we use a moving temporal window of K frames to estimate the output video frames. Within each group, we use the most recent frame as the reference and form an SR image aligned with that image. To begin, the robust global registration described in Section 3 is used to determine the position of the pixels in the input frames relative to the reference grid. Next, a local motion detection algorithm is employed to identify LR pixel regions that do not appear consistent with the global motion model. Local motion detection is explored in Section 4. All the LR pixels that appear consistent with the global registration are labeled as valid and used to populate a common HR grid. Note that all pixels from the reference frame are automatically considered valid and placed into the HR grid. The LR pixels from the other frames augment the reference samples only if they are considered consistent with the global motion model (i.e., are valid pixels). Observed pixels that deviate from the global motion model due to moving objects, motion parallax, occlusion, or other factors are simply excluded from populating the HR grid. In the limiting case, where only the reference frame pixels are considered valid, the output corresponds to a single frame AWF SR estimate and does not break down.

Like standard AWF SR [7] and partition weighted sum (PWS) SR [12, 13], the robust AWF SR method employs a finite moving window on the HR grid. The output pixels are formed as weighted sum of the observed samples. In particular, let the samples spanned by the moving window centered about HR pixel i be denoted $\mathbf{g}_i = [g_{i,1}, g_{i,2}, \dots, g_{i,G_i}]^T$. The output for $i = 1, 2, \dots, N$ is given by

$$\hat{z}_i = \mathbf{w}_{\psi(i)}^T \mathbf{g}_i, \quad (1)$$

where \hat{z}_i is the estimate of the i 'th pixel in the ideal image \mathbf{z} , $\psi(i)$ is the population index for window i , and $\mathbf{w}_{\psi(i)} = [w_{\psi(i),1}, w_{\psi(i),2}, \dots, w_{\psi(i),G_i}]^T$ contains the weights. The population index is an integer uniquely specifying the spatial pattern of observed valid pixels and this designates which set of precomputed weights is to be applied at this spatial location. The sampling of the HR image, relative to a LR frame, is increased by a factor of L in both the horizontal and vertical dimensions.

The weights are designed to minimize the mean squared error (MSE) based on the positions of the observed samples in the window and the correlation model used. If the spatially varying correlation model in [11] or the vector quantization partitioning in [12, 13] is used, the weights also depend on the local intensity pattern of the observed samples. This can provide enhanced performance, but adds computational complexity. Following the approach in [7], we focus here on the sample position-based weights using a wide sense stationary (WSS) correlation model. Thus, we have

$$\mathbf{w}_{\psi(i)} = \mathbf{R}_{\psi(i)}^{-1} \mathbf{p}_{\psi(i)}, \quad (2)$$

where $\mathbf{R}_{\psi(i)} = E\{\mathbf{g}_i \mathbf{g}_i^T | \Psi = \psi(i)\}$ is the autocorrelation matrix, $\mathbf{p}_{\psi(i)} = E\{z_i \mathbf{g}_i | \Psi = \psi(i)\}$ is the cross-correlation vector, and Ψ is a random variable representing the population index. By sampling the correlation model according to the spatial arrangement of samples, one is able to fill the autocorrelation matrix and cross-correlation vector. Given a finite number of population patterns, the weights can be computed and stored in a table. All of the correlations needed are derived from the assumed autocorrelation function for $d(x,y)$, which is given by

$$r_{dd}(x,y) = \sigma_d^2 \rho^{\sqrt{x^2+y^2}}, \quad (3)$$

and the observation model in Fig. 1. The parameter σ_d^2 gives the assumed signal power for a zero mean signal and ρ is the one step correlation parameters if x,y are in units of HR pixel spacings. We refer the reader to [7, 10–13] for further details on the correlation models. The

main difference here, compare with this earlier work, lies in how the HR grid is populated based on robust registration and the detection of local motion.

For affine motion, we use the discrete HR grid and look-up table weights based on the reduced observation windows described in [7]. This method is naturally suited to handle local-motion based selective population of the HR grid. This is because the local motion detection does not change the basic filtering operation. The population patterns are found in the look-up table of weights in the same fashion as described in [7]. Examples with a variety of partially populated HR grids and the corresponding AWF SR filter weights are shown in [7]. Inoperable or excessively noisy pixels (i.e., “bad” pixels) in all but the reference frame can be handled by excluding them as well. If a reference frame pixel is “bad” and no valid pixel from another frame fills that HR grid position, the nearest valid reference pixel on the HR grid is substituted. If multiple pixels belong to the same HR grid position, these values are averaged for simplicity. These “redundant” pixels could also be placed into \mathbf{g}_i and given weights according to Eq. (2).

For translational motion, if the non-quantized HR grid approach introduced in [11] is used, excluding some pixels from the HR grid will disrupt the periodic structure of samples. This means that custom weights must be computed to deal with each variation in the local population pattern. This can greatly add to the overall computational complexity. To minimize this for the case of translational motion with a non-quantized HR grid, we propose limiting the population patterns to those that use all the frames and that of just the reference frame. The pattern using all the frames would correspond to areas with no local motion and the pattern using only the reference frame would be for areas impacted by local motion. The one-frame weights can be precomputed prior to processing video, and the other weights are computed using the standard AWF SR method in [11].

3. Robust registration

The key to most multiframe SR algorithms is accurate subpixel registration of the observed frames. A popular choice for SR registration in the presence of global motion is the Lucas-Kanade method [3]. The method uses a truncated Taylor series to express a warped frame in terms of a reference frame plus scaled gradient images. This gives rise to a set of linear equations, one per pixel. Least squares can be used to obtain an estimate of the motion model parameters from these equations.

The theoretical performance limits of the Lucas-Kanade method are explored in [8] for Nyquist sampled imagery. Here we experimentally investigate the subpixel registration accuracy of Lucas-Kanade as a function of image block size and aliasing level as given by Q . We use 4 images from the Kodak lossless image suite [27] and process the 8 bit grayscale versions with the observation model. This includes PSF blurring (with a range of Q values), a one half pixel shift, downsampling by $L = 4$, and noise with $\sigma_\eta = 1$. The results are shown in Fig. 3 where the mean absolute error (MAE) of the estimated frame position is shown versus the linear dimension of the block size used for several values of Q . Note that the blocks to be registered are aligned to within one LR pixel and a one pixel border is excluded from the LS analysis. Hence, a 5×5 block is the smallest considered. As expected, increasing the block size lowers the registration error. Also, we see lower errors for images with less aliasing (i.e., higher Q). The knee in the curve appears at a block size of approximately 25×25 pixels. This result illustrates the challenge of obtaining subpixel registration accuracy on small moving objects in the presence of aliasing. It should be noted that this is a nearly ideal case since the motion is known to be translational, there is low noise, and blocks are prealigned to within 1 LR pixel. With moving objects that exhibit rotational or affine motion, the problem is magnified because of the increased number of motion parameters. Occlusions, pose variations, and deformable motion associated with moving objects present even more complexity and potential registration errors.

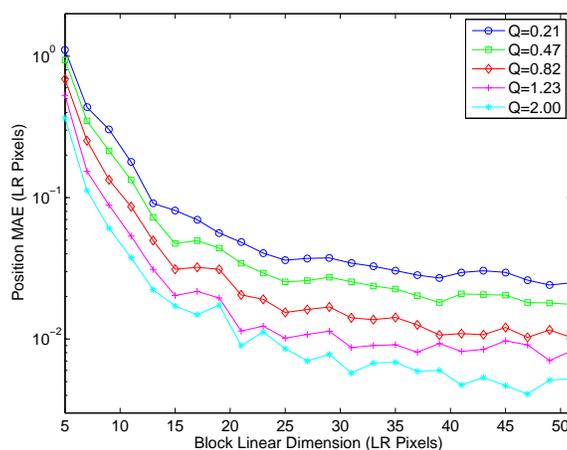


Fig. 3. Shift registration error as a function of the image block sizes under test. The error is mean absolute positional error between the actual shifted block and the estimate. As larger blocks are used, the error goes down. The shift registration accuracy is also impacted by the level of aliasing indicated by the Q factor.

While a large block size aids in overcoming aliasing and noise, the presence of local motion can bias the background motion estimate. To achieve robust background registration, in consideration of these factors, we proposed a modified multiscale Lucas-Kanade method. In this modified approach, we employ a Gaussian pyramid [28]. An example of a three level pyramid is shown in Fig. 4. We begin at the lowest resolution scale. Here, local motion tends to be greatly attenuated. Thus, we are able to get an approximate global registration with an affine model [7]. After aligning the images at this scale, we compute the absolute error image and segment pixels with the largest errors. Pixels within these segmented regions are excluded from the least squares registration at the next scale up. The final exclusion mask at each scale is not required to include the areas segmented at the prior scale. Registration at each level continues in this fashion until the top level is reached and the parameters estimated there are used as the final estimates. Note that the thresholding is done on an independent pixel-by-pixel basis. Morphological operations could be applied to this mask to help mitigate noise and exploit any known spatial characteristics associated with local motion. However, we do not use any such operations for the results presented here.

The segmentation threshold can be selected to achieve desired performance on representative training data or a fixed percentage of pixels can be segmented. For the results in this paper, we exclude the pixels with absolute errors in the top 10%. This fraction can be modified according to the level of local motion expected in an application. An example of the level-specific segmentation mask is shown in Fig. 4 as the red contours. The borders are also excluded at each level because they may not overlap with the reference image. Note that there is a moving vehicle approximately located at coordinates (100,50) at the highest level. This vehicle is excluded at the top levels, allowing us to get a more accurate background registration. Some background pixels are also excluded. However, with 90% of the pixels used for registration at each scale, these false positive (FP) local motion pixels do not present a major concern. The advantage of this method is that the pixels used for registration at each scale are well distributed spatially, which we find is critical for accurate affine registration.

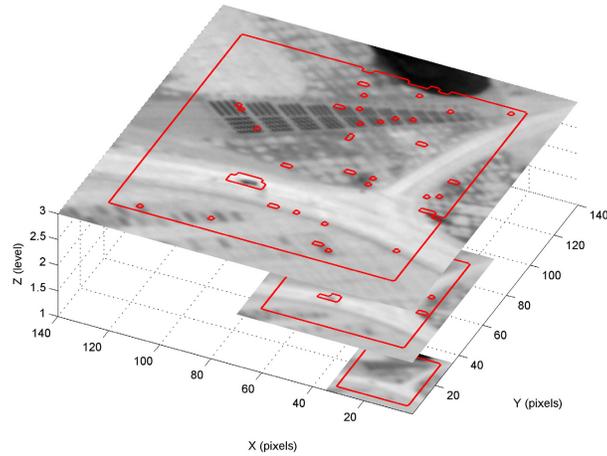


Fig. 4. Robust multiscale registration. A Gaussian pyramid is formed and a Lucas-Kanade based affine registration [3] is applied iteratively at successive scales starting at the lowest resolution. Borders and pixels with the highest registration errors from the preceding level are excluded from the least squares parameter estimation (red contours) to provide robustness to moving objects and local motion.

4. Detecting local motion in the presence of aliasing

In this section, we focus on methods for detecting local motion for use with the robust AWF SR. Note that simply registering the LR frames and computing the error image may not be sufficient for accurate local motion detection. The reason is that aliasing and noise will give rise to errors, even if the registration parameters are ideal. In fact, if the registered frame errors were zero, that would imply that there is no unique information provided by the individual frames. In this case, multiframe SR would be of no value. Given that we are expecting unique information in each frame, the difference frames will be impacted by sampling differences (an expression of aliasing), noise, moving objects, motion parallax, and background registration error. If we treat all errors as local motion, and exclude non-reference frame samples in the vicinity of these errors, we will be unable to successfully eliminate aliasing and optimally reduce noise using multiple frames. On the other hand, if we do not exclude any samples as we populate the HR grid, samples impacted by local motion will not be properly positioned and this leads to severe artifacts as will be demonstrated in Section 5. Thus, a detection algorithm is needed that can minimize the probability of local motion detection error or a specified overall cost. Consider that there is a cost associated with an FP, which would be declaring aliasing or noise to be local motion. There is also a cost associated with a false negative, which would be treating local motion as if it were aliasing or noise.

4.1. Impact of aliasing

Before addressing the detection problem, let us first analyze how aliasing leads to error between registered frames. Consider the theoretical MSE between a pixel in one frame and an estimate of that pixel from a subpixelly shifted frame. We shall assume a simple WSS correlation model and a linear filter for estimation. In particular, let the estimate of i 'th pixel of the reference frame be expressed as $\hat{y}_i = \mathbf{w}^T \mathbf{x}_i$, where \mathbf{x}_i is a corresponding observation vector from the shifted input

frame. The theoretical MSE is given by

$$J(\mathbf{w}) = E \left\{ (y_i - \mathbf{w}^T \mathbf{x}_i)^2 \right\} = E \{ y_i^2 \} - 2\mathbf{w}^T \mathbf{p} + \mathbf{w}^T \mathbf{R} \mathbf{w} \quad (4)$$

where $\mathbf{R} = E \{ \mathbf{x}_i \mathbf{x}_i^T \}$ and $\mathbf{p} = E \{ y_i \mathbf{x}_i \}$. The correlation statistics can be modeled in a fashion similar to that of the AWF SR using the same observation model. The main difference is that here y_i is an LR sample. The minimum MSE Wiener filter weights are given by $\mathbf{w} = \mathbf{R}^{-1} \mathbf{p}$. If bilinear interpolation is used, \mathbf{w} is a simple function of the subpixel shifts with weights for the 4 surrounding samples.

Using Eq. (4), registered pixel MSE values as a function of the subpixel displacement in LR pixel spacings are computed and shown in Fig. 5 for $E \{ y_i^2 \} = 1$, $\rho = 0.7$, and $\sigma_\eta^2 = 0$. We consider both a highly aliased scenario, with $Q = 0.21$, and a Nyquist sampled case where $Q = 2.0$. We show results for both bilinear interpolation and Wiener filter estimation using a 7×7 window. Note that since the zero noise case of shown, the MSE is zero for an integer shift in Fig. 5. However, the error generally goes up as the shift approaches (0.5, 0.5). When there is no aliasing, as is the case for Figs. 5(b) and 5(d), the MSE is relatively small. Since the Wiener filter is informed by the correlation model, it gives a lower MSE in Fig. 5(d) than bilinear in Fig. 5(b). With high levels of aliasing, as is the case for Figs. 5(a) and 5(c), we see higher MSE. The main lesson of these results is that aliasing provides an additional confuser for detecting true scene motion.

4.2. Local motion detection methods

We consider two scenarios, one is where we seek to detect the presence of local motion anywhere within the K frame temporal window at each spatial location. This is appropriate for the fast translational AWF method where we use either all of the frames or just the reference frame at a given spatial location. In this case, multiframe local motion detection methods can be employed. The other scenario, appropriate for the fast affine AWF with quantized HR grid, is where we attempt to detect motion in each individual frame relative to the reference. This way, the number of frames contributing to the HR grid can vary from 1 to K , based on which individual frames within the temporal window contain local motion. This allows one to potentially exploit more of the available frames and minimize the use of single frame processing.

Motion detection is an important problem for many types of video processing applications [29]. However, examining the impact of aliasing on such detection has received more limited attention. We propose and compare a number of methods for detecting local motion for use with robust AWF SR. All of the detection statistics considered are summarized in Table 1. These scalar statistics are computed and thresholded independently, on a pixel-by-pixel basis, to form the appropriate detection mask that governs the population of the HR grid. For individual frame detection, we consider the straightforward frame error (FE) [29]. In an attempt to mitigate error due to aliasing, we also consider frame error with prefiltering (FEP) [15]. The prefilter we use is a Gaussian low pass filter (LPF) and we apply it to the reference and test image prior to interpolation and error computation. The idea is that aliasing and noise are more dominant at high spatial frequencies. However, the prefiltering can attenuate very small moving objects, making them harder to detect. Thus, an appropriate balance between detail preservation and aliasing/noise reduction must be sought. The remaining methods in Table 1 employ multiframe statistics. These include the range (R) of the K aligned frames at each pixel location, and the range with prefiltering (RP). Again, we use a Gaussian prefilter for aliasing/noise reduction. Range is a powerful detection statistic for this application because it represents the maximum error between all pairs of frames.

Another method for mitigating the effects of aliasing is through multiframe AWF processing. The multiframe estimation error (MEE) statistic is the error between the test frame and an AWF

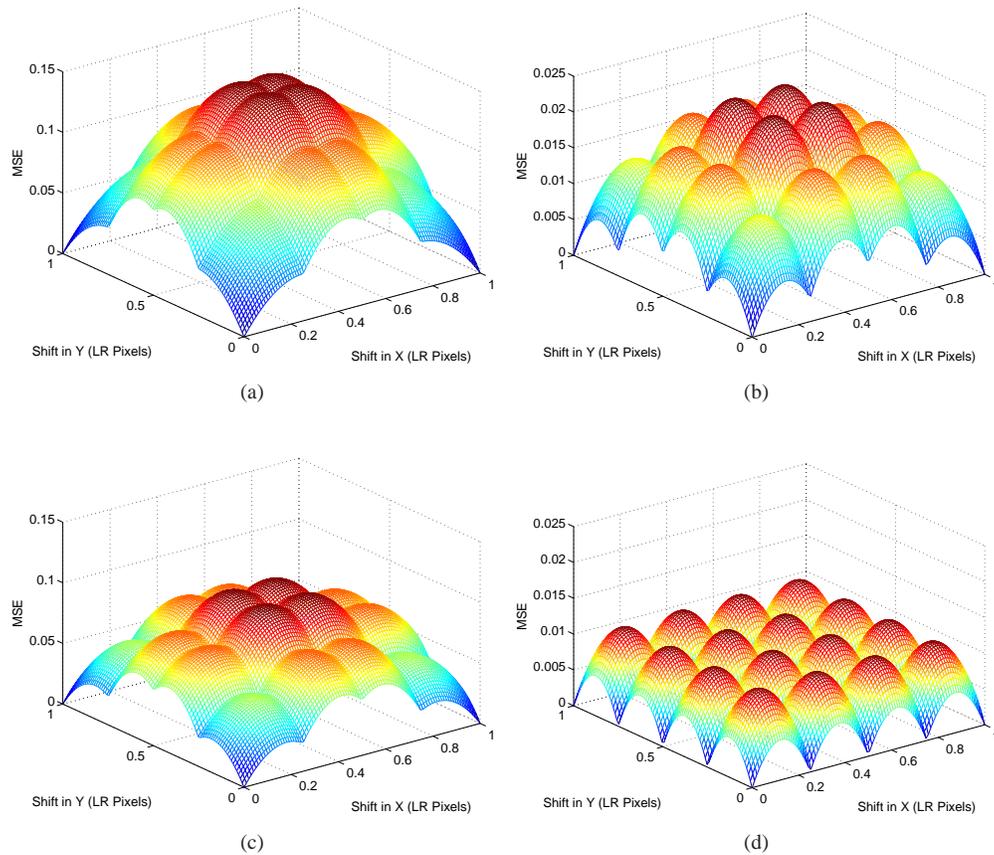


Fig. 5. Theoretical MSE between two registered frames as a function of sub-pixel shift. (a) Highly aliasing imagery ($Q=0.21$) with bilinear interpolation, (b) no aliasing ($Q=2.0$) with bilinear interpolation, (c) highly aliasing ($Q=0.21$) with Wiener filter, (d) no aliasing ($Q=2.0$) with Wiener filter.

prediction of that frame based on the $K - 1$ other frames. The sampling diversity offered by the $K - 1$ frames allows us to make a better prediction in the presence of aliasing and noise than with a single frame comparison. The final detection statistic considered is the forward model error (FME). Here we use all K frames, as if there were no local motion, to form a preliminary AWF SR image. This SR image is then projected through the forward model in Fig. 1 (with no noise). The predicted data are compared to the observed data to produce the FME statistic. This method has the computational advantage that in areas with no local motion, the preliminary SR image can be used as the final output. Only areas determined to be impacted by local motion need to be reprocessed using the reference frame.

Let us now briefly consider the computational complexity of some of the local motion detection methods. One of the main tasks in computing the FE, FEP, R and RP statistics is aligning the $K - 1$ LR frames to the reference frame. This requires $\mathcal{O}((K - 1)M/K)$ flops (multiply and add) for simple interpolation, where M/K is the number of LR pixels in a single frame. We must also perform up to $\mathcal{O}((K - 1)M/K)$ compares for thresholding. The prefiltering, for the FEP and RP methods, can be done with a separable Gaussian filter applied to each new LR

Table 1. Description of the local motion detection methods considered here.

Name	Detection statistic description
Frame Error (FE)	Error between the reference frame and other individual frame that has been aligned using robust global registration. Can detect pixels in specific frames within the K observed exhibiting local motion relative to the reference.
Frame Error with Prefilter (FEP)	Same as FE, but the reference and test frame are prefiltered with a Gaussian LPF. The prefilter is designed to reduce aliasing and noise to reduce FPs due to these factors.
Range (R)	Range of pixel values at each spatial location using all K frames aligned using robust global registration. Can only distinguish if motion is present within the K input frames.
Range with Prefilter (RP)	Same as R statistic, but all frames are prefiltered with a Gaussian LPF.
Multiframe Estimation Error (MEE)	Error between a given frame and a multiframe low-resolution AWF estimate of that frame. Can only distinguish if motion is present within the K input frames at a given spatial location. Mitigates aliasing to some extent due to the sampling diversity provided by the use of multiple frames.
Forward Model Error (FME)	Error between each observed frame and the full K frame AWF SR estimate projected through the forward observation model. Mitigates aliasing by matching the aliasing in the test frame by projecting the SR image through the forward model.

frame and stored. This adds $\mathcal{O}(PM/K)$ flops, where P is the linear dimension of the 2-D filter window. The R detection statistic has the additional burden of the range computation involving K values for each of the M/K LR reference pixel locations, which requires $\mathcal{O}(M)$ compares and exchanges. The complexity of the AWF itself is addressed in [7, 11].

To compare the performance of the local motion detection methods, we use a simulated image sequence with a chirp background pattern and a ruler moving horizontally left to right along the bottom relative to the background. The chirp patterns allows us to highlight any aliasing. Global translational motion is simulated in addition to the ruler motion. We simulate the sequence using $Q = .47$ to match the airborne imaging sensor described in Section 5. The sequence has a peak intensity of 200 digital units (DUs) and the noise has a variance of $\sigma_{\eta}^2 = 1$. Robust multiframe AWF SR is done using $K = 12$ frames with $L = 4$. The detection performance is evaluated using a receiver operating characteristic (ROC) curve. The local motion detection results are shown in Figs. 6 and 7. A single LR reference frame from this study is shown in Fig. 6(a) along with the truth mask contour (Green line) showing the boundary of the local motion over the $K = 12$ frame temporal window. The R, RP, and absolute FME statistic images are shown in Figs. 6(b) - 6(d), respectively. For the RP statistic, the Gaussian prefilter has a standard deviation of 1.5 LR pixels. The red contours represent the detection masks for a probability of detection of $p_d = 0.85$ and the images are scaled so that at the detection threshold they have a grayscale value of 128. We compare methods using a constant p_d in order to illustrate how aliasing impacts FPs. The green contours represent the true local motion mask. Moiré patterns from aliasing can be seen on the input frame in Fig. 6(a). The range statistic in Fig. 6(b) clearly shows high error from the aliasing, and this leads to numerous FPs. As shown in Fig. 6(c), the use of the prefilter dramatically reduces error due to aliasing. The FME in Fig. 6(d) also shows reduction in aliasing error.

The ROC curve analysis for these data, using pixel-level scoring, is shown in Fig. 7. The benefit of the Gaussian prefilter is made clear here for both the FEP and RP detection statistics. Note that FME and MEE are more sensitive at very low FP rates, but the prefilter methods outperform at higher FP rates. It is also interesting to note that the range statistics outperform the frame error statistics. This may be because the frame error methods only consider errors between each frame individually and the reference. On the other hand, the range methods outputs the maximum difference between all pairs of frames. Finally, note that the performance of

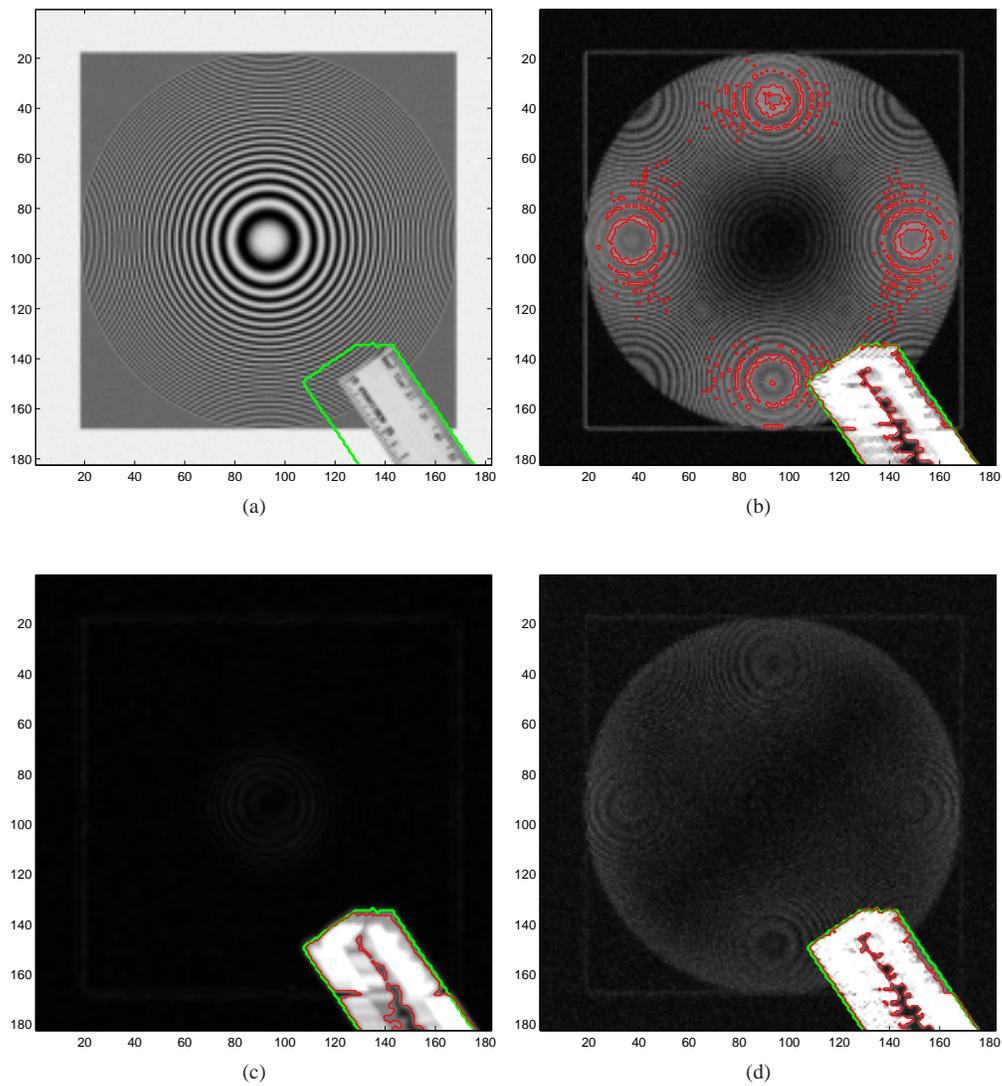


Fig. 6. Simulated local motion detection study. (a) Simulated LR frame of static chirp pattern showing a moving ruler. (b) R statistic with $K = 12$ registered frames, (c) RP statistic (standard deviation of 1.5 LR pixels), (d) absolute FME statistic. The red contours are the detection masks for $p_d = 0.85$ and the green mask is the motion truth.

MEE is very similar to that of FME. As mentioned above, the FME method is more attractive computationally than MEE.

5. Robust SR results

A number of SR results are provided here to illustrate the efficacy of the proposed method. We begin with results for simulated data that include a quantitative error analysis. Next, real video from a visible camera is used. Finally, we test the algorithm using an airborne midwave infrared (MWIR) camera. All of the datasets include some kind of background resolution pattern and a

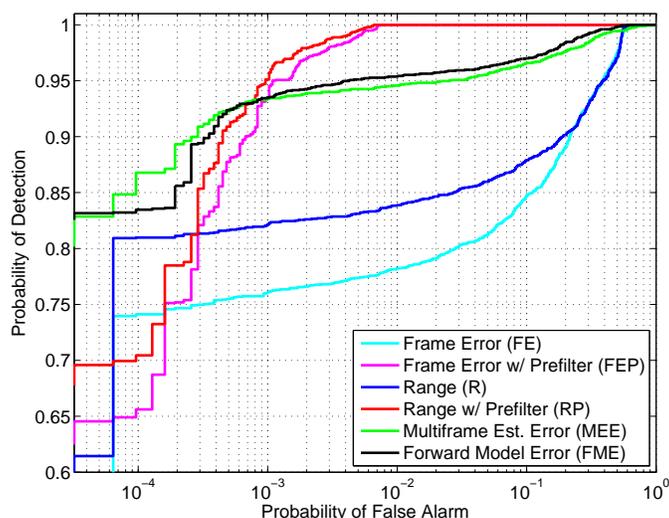


Fig. 7. ROC curve for the local motion detection algorithms using the simulated moving ruler data. The benefit of the Gaussian prefilter is made clear here for both the frame error and range methods. Note that FME and MEE are more sensitive at very low FP rates, but the prefilter methods outperform at higher FP rates.

moving object.

5.1. Simulated data

The SR images for the simulated moving-ruler data from Fig. 6 are shown in Fig. 8. A region of interest (ROI) from the true HR image is shown in Fig. 8(a) and bicubic interpolation of a single LR frame with $L = 4$ is shown in Fig. 8(b). Notice that the single interpolated image has significant Moiré patterns from aliasing. The translational multiframe AWF SR output with no local motion detection [11] is shown in Fig. 8(c). Here the rings on the chirp pattern are fully restored, but the moving ruler is heavily distorted. The distortion occurs because the background motion model is not adequate to describe the positions of the samples from the moving object. The high boost effect of the AWF filter tends to amplify the artifacts. The robust AWF SR results using the R, RP, and FME detection statistics are shown in Figs. 8(d)-8(f), respectively. Thresholds for all of the methods are set for $p_d = 0.85$, based on the ROC analysis shown in Fig. 7. All of the AWF SR methods use $L = 4$ upsampling, a window size of 12×12 HR pixels, $\rho = 0.75$, and $\sigma_d^2/\sigma_n^2 = 100$. Multiframe AWF SR results use $K = 12$ input frames and the prefilter used is a Gaussian LPF with spatial standard deviation of 1.5 LR pixels. With the R statistic, the FPs seen in Fig. 6(b) lead to areas of aliasing in the chirp pattern in Fig. 8(d). Both the RP and FME local motion detection appear to perform well, as can be seen in Figs. 8(e) and 8(f). An aliased “shadow” behind the ruler can be seen because these methods do not use multiple frames to populate the HR grid until no motion is detected anywhere within the K frame temporal window.

Quantitative error results along with processing times are shown in Table 2. The error metrics are MSE, mean absolute error (MAE), and peak signal-to-noise ratio (PSNR). The processing times for the multiframe AWF methods include one-frame incremental registration, local motion detection, populating the HR grid, and the final weighted sum operation. Processing is done using MATLAB on an Intel Core i7 64 bit CPU with a clock speed of 3.07 GHz. The multiscale

Table 2. Quantitative error analysis for the simulated moving ruler data.

Method	MSE (DU)	MAE (DU)	PSNR (dB)	Time (s)
Bicubic	439.74	12.39	19.59	0.005
Single Frame AWF	362.14	10.69	20.43	0.045
Multiframe AWF	70.86	3.81	27.52	0.120
Robust AWF (FE)	321.10	9.58	20.95	0.194
Robust AWF (FEP)	55.56	3.49	28.57	0.205
Robust AWF (R)	197.24	6.70	23.07	0.191
Robust AWF (RP)	55.19	3.48	28.60	0.202
Robust AWF (FME)	56.06	3.50	28.53	0.327

affine registration uses 2 levels with 10 iterations at each level and employs bilinear interpolation. The Robust AWF SR with the RP detection statistic provides the best results, followed closely by FEP and FME. For very small moving objects, the FME may have an advantage, since no smoothing prefilter is required.

5.2. Visible camera video

Here we present results for a real video sequence acquired with an Imaging Source DMK 21BU04 visible camera. This is a 640×480 8-bit grayscale camera employing a Sony ICX098BL CCD sensor with $5.6 \mu\text{m}$ detectors. The camera is fitted with a Computar $F/4$ lens with a focal length of 5mm. Considering a central wavelength of $\lambda = 0.55 \mu\text{m}$, the visible system is theoretically $5.09\times$ undersampled with $Q = 0.39$. The sequence captures a static background of a chirp pattern. Local motion is provided by a tire pressure gauge stick being moved in front of the chirp. The camera is mounted on a tripod with random translational motion applied.

The results are shown in Fig. 9 ([Media 1](#), [Media 2](#)). Bicubic interpolation of a single LR frame with $L = 4$ is shown in Fig. 9(a). As with the simulated data, significant Moiré patterns from aliasing are visible. The translational multiframe AWF output with no local motion detection is shown in Fig. 9(b). The robust AWF SR results using the R and RP detection statistics are shown in Figs. 9(c) and 9(d), respectively. The detection methods are set for a common probability of false alarm of $p_{fa} = 10^{-4}$. Thresholds have been determined using a non-parametric kernel-smoothing probability density function (pdf) estimate for manually selected background regions. The AWF parameters are the same as those used for the simulated data. The R and RP detection statistics are shown in Figs. 9(e) and 9(f), respectively. The statistics images are both scaled to have a value of 128 at the detection threshold, to facilitate comparison. Note that with the low false alarm rate, both the R and RP detectors operate successfully on the background chirp pattern. However, the moving stick is only partially detected with the R statistic and appears fully detected with the RP statistic. These results appear consistent with those from the simulated data. Videos are provided to better illustrate the results. [Media 1](#) shows video corresponding to Figs. 9(a) and 9(d) on the left and right, respectively, while [Media 2](#) shows video of ROIs corresponding to Figs. 9(a) - 9(d).

5.3. Infrared flight data

The flight data used here is acquired with a MWIR imager with a spectral bandwidth of $\lambda = 3 - 5 \mu\text{m}$ (with the exception of the CO_2 absorption band). As in [7], we use $\lambda = 4 \mu\text{m}$ for our PSF model. The system uses $F/2.3$ optics and has a pixel pitch of $p = 19.5 \mu\text{m}$ and we assume

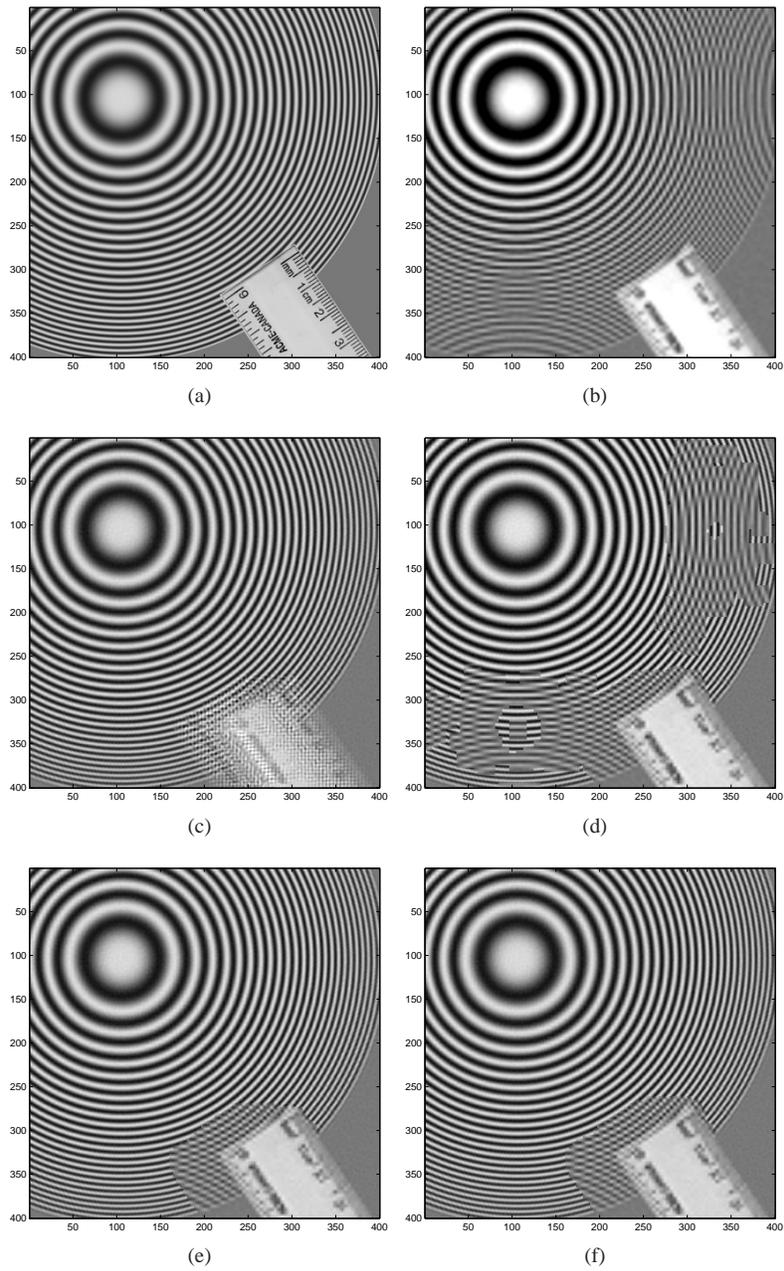


Fig. 8. Robust AWF SR results for the simulated moving ruler image sequence ($L = 4$). (a) True HR image, (b) bicubic interpolation of the single LR reference frame, (c) AWF SR with no local motion detection, (d) robust AWF SR (R), (e) robust AWF SR (RP), (f) robust AWF SR (FME). All of the detection methods are set for $p_d = 0.85$.

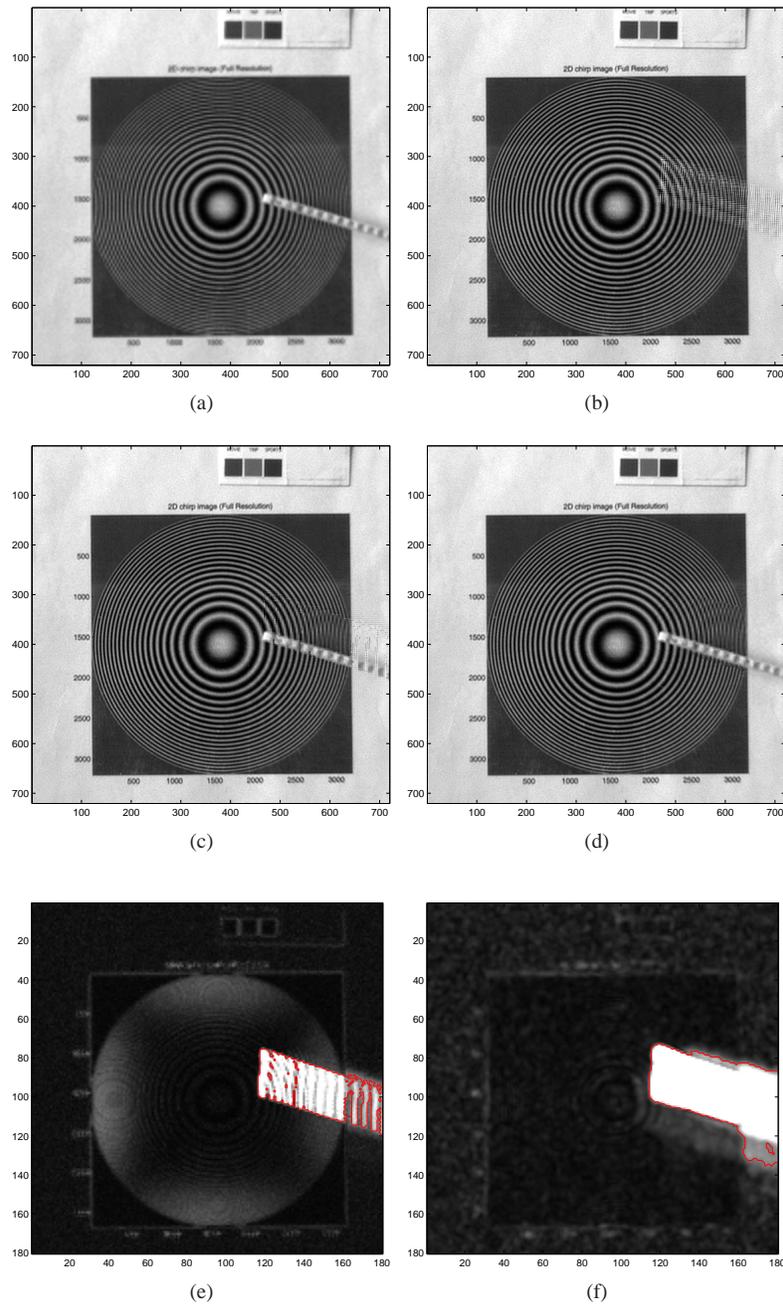


Fig. 9. Robust AWF SR results ($L = 4$) using visible camera sequence using a translational background motion model ([Media 1](#), [Media 2](#)). (a) Bicubic interpolation of the reference frame, (b) AWF SR with no local motion detection, (c) robust AWF SR (R), (d) robust AWF SR (RP), (e) R statistic with detection mask, (f) RP statistic with detection mask. Thresholds are set for $p_{fa} = 10^{-4}$.

100% fill factor rectangular detectors. This system is theoretically $4.24\times$ undersampled with a $Q = 0.47$. The details of the airborne data collection can be found in [25]. On the ground we have a thermal resolution pattern with 13 pairs of 4-bar groups. The bar spacing, which matches the bar width, ranges from 0.25 m to 1.0 m. The scaling factor between bar groups is designed to be $2^{(1/6)}$ [25]. The image sequences considered here show both the bar pattern and a moving car. This allows us the opportunity to evaluate SR objectively using the bar pattern, as well as evaluate robustness to a moving object.

To observe the impact of frame rate on the proposed processing, we include one sequence acquired at a frame rate of 16 Hz, and another at 50 Hz. Both use an integration time of 4 ms to minimize motion blur. All of the AWF parameters are the same as used in the previous results, except that we use $L = 3$, $K = 10$, and $p_{fa} = 10^{-5}$. This temporal window size is chosen because the 50 Hz data includes a step-stare mirror that repositions the field of view every 10 frames. The lower FP rate is chosen because the data have fewer strong aliasing artifacts and we can get a high sensitivity with this level of specificity. As with the visible data, thresholds have been determined using a non-parametric kernel-smoothing probability density estimate for manually selected background regions.

The results with the 16 Hz flight data are shown in Fig. 10 ([Media 3](#), [Media 4](#)). Bicubic interpolation of a single LR frame with $L = 3$ is shown in Fig. 10(a). A vehicle that is moving from left to right can be seen at coordinates (300,175). The bar pattern can be seen below the vehicle, and above is a parking lot with a light post at coordinates (380,50). The affine multiframe AWF SR output with no local motion detection [7] is shown in Fig. 10(b). Here the moving vehicle is heavily distorted due to its motion, and some distortion can also be seen on the light due to motion parallax. The robust AWF SR results using the RP and FEP detection statistics are shown in Figs. 10(c) and 10(d), respectively.

Images showing the number of frames allowed to populate the HR grid from the RP and FEP local motion detection are shown in Figs. 10(e) and 10(f), respectively. Black indicates that only the reference frame is used, white indicates that all $K = 10$ frames are used (no local motion detected), and gray reveals values in between. Both the RP and FEP methods yield good aliasing reduction performance on the bar pattern, but processing around the moving vehicle differs. For the RP detector, we allow only 1 or all K frames at a given area on the HR grid, as can be seen in Fig. 10(e), while FEP allows for 1 to K . Note in Fig. 10(f) that the FEP method starts repopulating the HR grid immediately after the vehicle passes. For RP, we have a much bigger single frame processing “shadow” in the wake of moving objects. Using the robust AWF SR with quantized HR grid, FEP adds no additional filtering complexity and allows us to minimize single frame processing areas. To illustrate the temporal characteristics of the processing, [Media 3](#) shows video of Figs. 10(a) and 10(d) on the left and right, respectively, while [Media 4](#) shows video of ROIs corresponding to Figs. 10(a) - 10(d).

It is interesting to observe the estimated probability density functions (pdfs) of the RP and FEP detection statistics on the manually selected background for the data in Fig. 10. These are shown in Fig. 11. As can be seen in Fig. 11(a), the RP statistic appears to be well modeled with the three parameter generalized extreme value (GEV) pdf [30]. The FEP is approximately Gaussian, as illustrated in Fig. 11(b). These parametric models could aid in setting an operating point in practical applications. When no prefilter is used (i.e., R, FE statistics), we find that these parametric models do not provide heavy enough tails to accurately fit the data.

Results for the 50Hz airborne data are shown in Fig. 12. Single frame bicubic interpolation is shown in Fig. 12(a) and the robust AWF SR (FEP) output is shown in Fig. 12(b). These images show the same ground resolution pattern and include a vehicle at coordinates (140,220) that is moving from left to right in the sequence. The bar pattern is well resolved in Fig. 12(b), compared with that in Fig. 12(a), and the moving vehicle is free from misregistration artifacts.

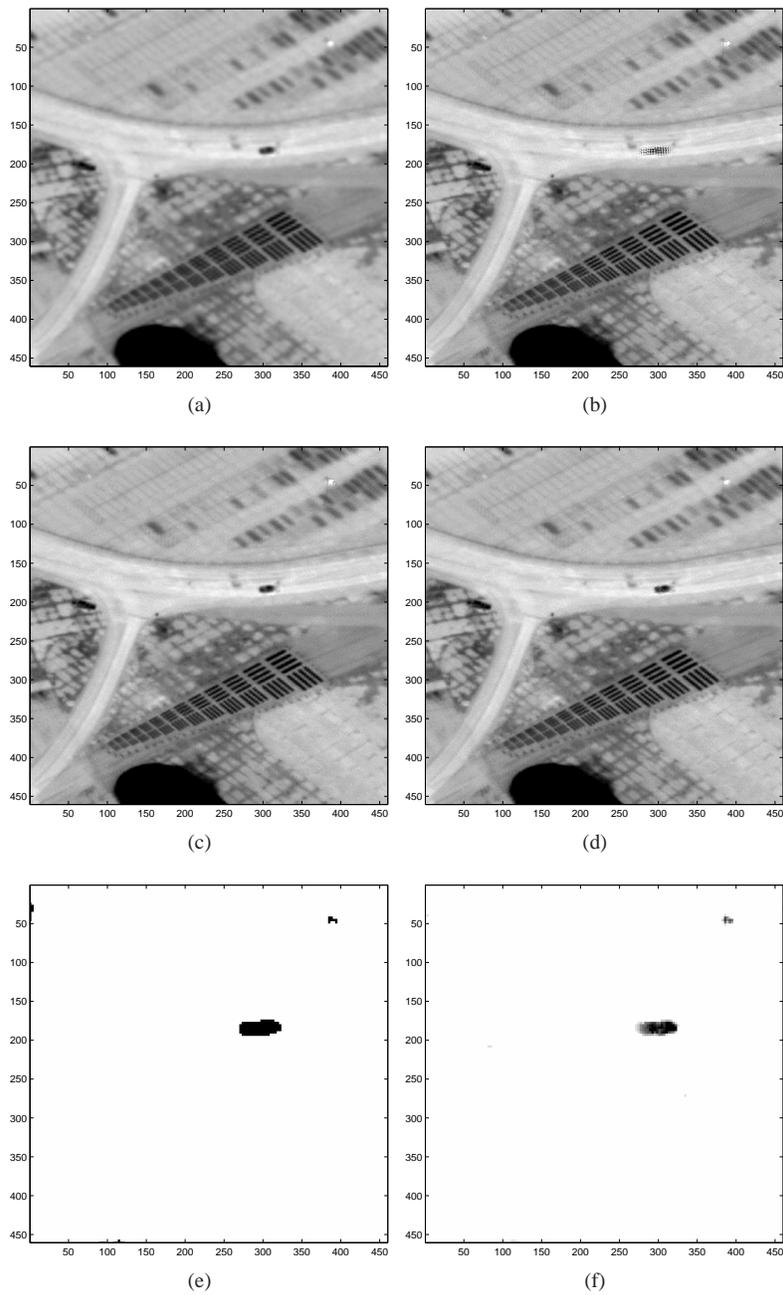


Fig. 10. Robust AWF SR results for airborne data with 16 Hz frame rate using an affine background motion model ($L = 3$) (Media 3, Media 4). (a) Bicubic interpolation of the reference frame, (b) AWF SR with no local motion detection, (c) robust AWF SR (RP), (d) robust AWF SR (FEP), (e) number of input frames contributing in (c), (f) number of input frames contributing in (d). Both detection methods are set for $p_{fa} = 10^{-5}$.

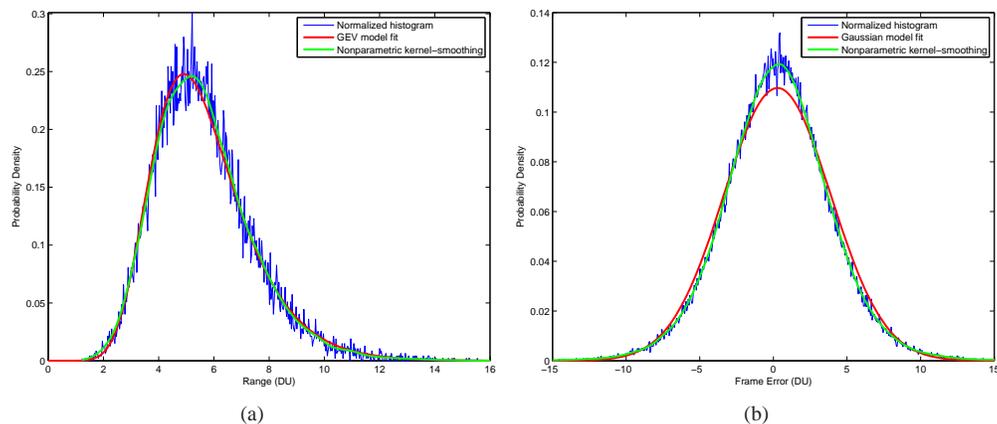


Fig. 11. Probability density function estimation for the (a) RP statistic and (b) FEP statistic for the background in the 16Hz flight data sequence from Fig. 10.

An ROI around the moving vehicle for the AWF SR output with no motion detection is shown in Fig. 12(c). The same ROI for the robust AWF SR (FEP) output is shown in Fig. 12(d). Motion artifacts are visible in Fig. 12(c), but are greatly reduced compared with that seen in Fig. 10(b). This illustrates that another way to help mitigate the problem of local motion is with increased frame rate relative to the moving objects in the scene. A higher frame rate also reduces the camera platform displacement between frames for this airborne application. This has the added benefit of reducing motion parallax.

6. Conclusions

Complex scene motion creates challenges for SR, which generally relies on some form of sub-pixel registration. As we have shown, a parametric background motion model can be used to provide accurate results, provided that large enough image block sizes are used. However, local motion from moving objects and motion parallax will generally not follow the background. Furthermore, commutation of the motion and PSF may not be valid with local motion. This commutation is a key assumption in most fast nonuniform interpolation SR algorithms. Our proposed approach uses local motion detection and only uses samples from the observed frames determined to be consistent with the background motion model. This eliminates the need for accurate subpixel registration for local motion and bypasses the issue of motion/PSF commutation in those areas. To detect local motion in the presence of aliasing, we have proposed and compared several methods. We see that a simple low-pass prefilter can be used to mitigate the effects of aliasing to a great extent with the FEP and RP statistics. However, this may come at the cost of contrast for small moving objects. Multiframe methods can exploit sampling diversity among multiple frames to mitigate aliasing when detection local motion. We have presented two such methods, the MEE and FME detectors.

We have demonstrated that many of the proposed robust methods can work effectively with both ground and airborne video. Also, the airborne results illustrate that the local motion problem is reduced at higher frame rates. This is because the motion differential throughout the scene is reduced, as is the potential for motion parallax. However, fast moving objects, relative to the frame rate, must still be addressed. We have observed that the RP method does appear to provide the best results for larger moving objects and higher FP rates. The FEP method is nearly as good as the RP method and allows one to use a variable number of frames when pop-

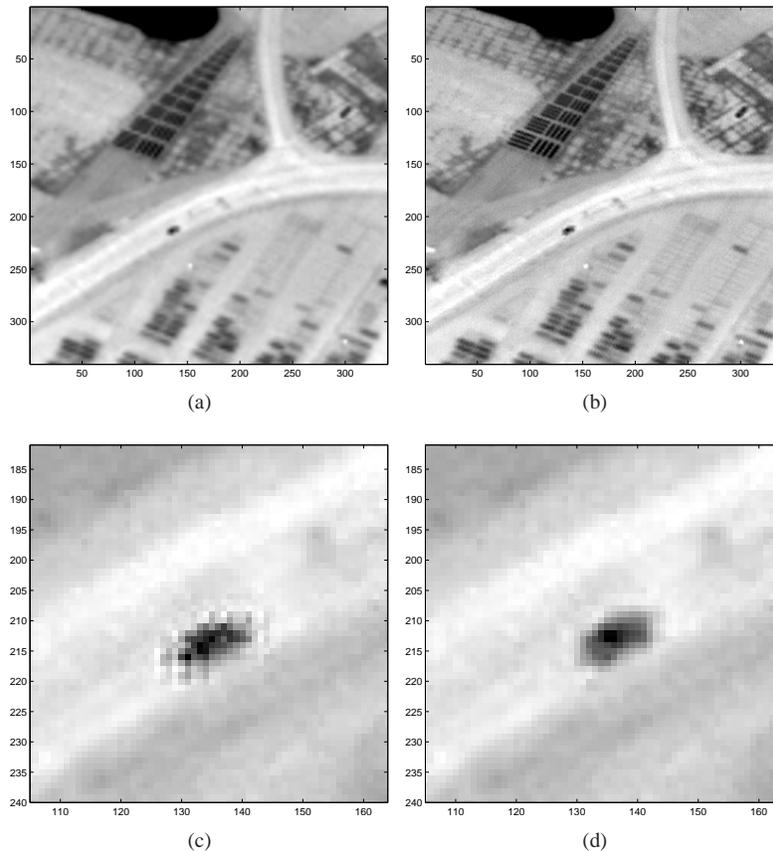


Fig. 12. Robust AWF SR results for airborne data with 50 Hz frame rate using an affine background motion model ($L = 3$). (a) Bicubic interpolation of the reference frame, (b) robust AWF SR (FEP) for $p_{fa} = 10^{-5}$, (c) moving vehicle ROI processed with AWF SR with no local motion detection, (d) moving vehicle ROI processed with robust AWF SR (FEP).

ulating the HR grid, not just 1 or K as with the multiframe methods (i.e., RP, MEE, FME). For small objects and low FP rates, the FME method may be preferred, because it does not rely on a smoothing prefilter, which can attenuate small moving objects.

Acknowledgements

This work is supported in part under Air Force Research Laboratory (AFRL) contract number FA8650-06-2-1081. The airborne data have been collected under AFRL contract FA8650-06-D-1078-0004. The authors would like to thank Dr. Michael Eismann and Mark A. Bicknell for their support of the project. Thanks also to Dr. Raul Ordonez for helpful conversations and Dr. Frank O. Baxley for leading the flight data acquisition effort and providing the MWIR data used here. Engineers from General Dynamics Advanced Information Systems and the University of Dayton Research Institute, including Patrick C. Hytla, Joseph C. French, and Phil Detweiler, configured the MWIR system and NASA Glenn Research Center provided flight support.