Shift Estimation Algorithm for Dynamic Sensors With Frame-to-Frame Variation in Their Spectral Response

Todd M. Giles, Majeed M. Hayat, *Senior Member, IEEE*, and Sanjay Krishna, *Senior Member, IEEE*

***Abstract*—This study is motivated by the emergence of a new class of tunable infrared spectral-imaging sensors that offer the ability to dynamically vary the sensor's intrinsic spectral response from frame to frame in an electronically controlled fashion. A manifestation of this is when a sequence of dissimilar spectral responses is periodically realized, whereby in every period of acquired imagery, each frame is associated with a distinct spectral band. Traditional scene-based global shift estimation algorithms are not applicable to such spectrally heterogeneous video sequences, as a pixel value may change from frame to frame as a result of both global motion and varying spectral response. In this paper, a novel algorithm is proposed and examined to fuse a series of coarse global shift estimates between periodically sampled pairs of nonadjacent frames to estimate motion between consecutive frames; each pair corresponds to two nonadjacent frames of the same spectral band. The proposed algorithm outperforms three alternative methods, with the average error being one half of that obtained by using an equal weights version of the proposed algorithm, one-fourth of that obtained by using a simple linear interpolation method, and one-twentieth of that obtained by using a naïve correlation-based direct method.**

***Index Terms*—Dot-in-a-well (DWELL) quantum-dot detectors, focal plane arrays, image registration, infrared, motion estimating, spectral imager.**

I. Introduction

THIS study is motivated by the emergence of a new class of quantum-dot mid-IR focal-plane arrays (FPAs) that offer the ability to dynamically vary the sensor's intrinsic spectral response from frame to frame in an electronically controlled fashion. The variation in the spectral response is achieved by changing the applied bias voltage on the detectors [1]. A manifestation of this is when a sequence of dissimilar spectral responses is periodically realized, whereby in every period of acquired imagery, each frame is associated with a distinct spectral band. One such operational detector has been recently demonstrated [2].

With the new capability of spectral tunability come many algorithmic challenges to exploit the data that these sensors provide. Some prior algorithmic work has already been done at a single-pixel level; an example is the implementation of an algorithmic spectrometer for target recognition [3]. An algorithm for algorithmic spectral-matched filtering for the purpose of hyperspectral feature selection and classification has also been developed and tested using this type of sensor [4]. However, there remain many other unexplored areas of algorithm development that use the DWELL sensor; one such outstanding area is that of global shift estimation.

Global shift estimation is used in a variety of applications including electronic stabilization [1], improving image quality using overlapping motion compensated frames [6], resolution enhancement [7]–[11], and nonuniformity correction in IR FPAs [7], [12]–[14]. The performance of each of these applications heavily relies on the accuracy of estimates of the frame-to-frame motion. This motion is typically determined by selecting one frame as the reference frame, and then estimating the spatial translation and rotation to the next frame in order to bring it to alignment with the first.

Most shift estimation algorithms assume no frame-to-frame variation in the detectors' spectral response. However, when such variation is present, the motion estimates may become highly inaccurate. This is because of loss of correlation between the initial reference image and the newly shifted frame that is generated by a detector with a different spectral response. Clearly, when the spectral response at the detector changes, what is seen at the detector will also change—even in the absence of any motion. Having a strong correlation between sequential images is a key factor in many shift estimation algorithms.

Different from the case when the spectral response of a sensor is stationary in time, in the dynamic DWELL sensor described earlier, the scenes from different spectral responses of the sensor will no longer be fully correlated to one another. An illustrative (perhaps extreme) example of this is shown in Fig. 1, where two images of the same scene have a contrast reversal effect due to differences in the reflectance of the materials in the scene as a function of wavelength. The cross correlation between these two images is minimal at the point of optimal shift, which is precisely when we expect it to be a maximum. Thus, we expect

Manuscript received September 01, 2008; revised December 22, 2008; accepted January 15, 2009. Current version published February 24, 2010. This work was supported in part by the National Science Foundation under Award IIS-0434102 and Award ECS-401154, by the National Consortium for MASINT Research, by Los Alamos National Laboratories, and by Sandia National Laboratories. The associate editor coordinating the review of this paper and approving it for publication was Dr. Patti Gillespi.

T. Giles was with the Center for High Technology Materials and the Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131-1356 USA. He is now with Sandia National Laboratories, Albuquerque, NM 87123-3453 USA (e-mail: tgiles@sandia.gov).

M. M. Hayat and S. Krishna are with the Center for High Technology Materials and the Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131-1356 USA (e-mail: hayat@ece.unm.edu; skrishna@chtm.unm.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSEN.2009.2037805

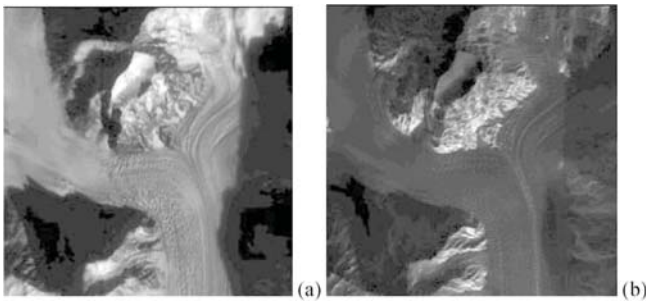

Fig. 1. Two images of a glacier taken from the U.S. Department of Energy's Multispectral Thermal Imager (MTI) satellite. (a) Spectral response centered at 2.2 μm. (b) Spectral response centered at 8.6 μm.

an erroneous shift estimate when using, for example, a cross-correlation-based method to compute the shift estimates.

There are many existing shift estimation algorithms for motion estimation between frames generated by a static (single band) FPA. Most common are cross-correlation-based techniques [15], [16]. Several methods attempt to improve the performance of cross-correlation methods by preprocessing the image. Examples of these include gradient-based techniques [17], [18], and object [19], or point matching [20]. All of these methods can perform well in estimating shifts from consecutive frames of similar spectral response at the detector, but all of them rely on the presence of strong correlation between corresponding images, something that may be lost when the spectral response at the detector changes from frame to frame.

In this paper, a novel algorithm is proposed to fuse a series of coarse global shift estimates between periodically sampled pairs of nonadjacent frames that correspond to nonadjacent frames of the same spectral band. The coarse global shift estimates are made using a projection-based shift estimator (PBSE) [16], though they could be made using any existing shift estimator. The overlapping coarse estimates are then combined, using a weighted average scheme, to compute an estimate of the frame-to-frame motion. The weights are calculated taking into account the band-dependent variability in the accuracy of shift estimation for each band. The main contribution in this paper is centered in the idea of combining overlapping, coarse motion estimates to determine frame-to-frame shifts without requiring correlation between consecutive frames.

The remainder of this paper is organized into four sections. In Section II, we develop a mathematical model of an image sequence generated by the spectrally dynamic DWELL quantum-dot sensor. In Section III, we describe the proposed shift estimation algorithm in detail. In Section IV, we compare results obtained using the proposed algorithm with that of two simplistic alternative approaches. Finally, our conclusions are summarized in Section V.

II. Mathematical Model

A. DWELL Sensor Image Sequence

The spectral response of the DWELL quantum-dot sensor is electronically controllable, allowing us to periodically capture a sequence of images with dissimilar spectral responses. Therefore, in every period of acquired imagery, each frame is associated with a distinct spectral band. We thus have an image sequence corresponding to a sensor whose spectral response varies dynamically from frame to frame, albeit in a periodic fashion. If we assume that there is negligible variation (nonuniformity) in the spectral response across pixels of the same frame, we can then define the spectral response simply in terms of the wavelength λ and the discrete time n at which the image was acquired. Moreover, since the spectral response repeats itself periodically, say every N frames, we have

$$S(n,\lambda) = S(n+N,\lambda). \tag{1}$$

An image sequence taken from a sensor with this type of varying spectral response will contain a periodic sequence of images from N distinct spectral bands. We can define this image sequence $\mathbf{I}$ in terms of the actual intensity values, J, radiating from the scene, and subsequently detected on the sensor at each pixel (x,y). Each intensity value is effectively filtered by the spectral response of the sensor at the discrete time n. This relationship is

$$\mathbf{I}_n(x,y) = \int_\lambda J_n(x,y,\lambda) S(n \mod N,\lambda) d\lambda \tag{2}$$

where $\mathbf{I}_n$ and J_n refer to the nth image in the image sequence and nth set of intensity values incident on the sensor, respectively. Note that each consecutive image is filtered by the periodically varying spectral response from frame to frame—leading to consecutive images that may have reduced correlation between them because of variation in the spectral response.

Note that (2) is describing one of a myriad ways that one could possibly generate an image sequence from a sensor that allows for controlled spectral variation from frame to frame. Other applications may lend themselves to different types of image sequences, but the one here is chosen for its generality. Next, let us compare this new type of spectrally periodic image sequence to a collection of image sequences from multiple sensors with fixed (static) spectral responses.

The image sequence $\mathbf{I}$ can be thought of as an interleaved collection of image sequences from multiple sensors with nonvariable spectral responses. If we take every Nth frame of $\mathbf{I}$, we note that the spectral response becomes a function of wavelength only and is no longer dependent on the discrete index n, as in (2). By varying the starting value of n from 0 to $N-1$, we obtain a collection of N image subsequences, each subsequence corresponding to a distinct, fixed spectral response.

As the spectral response in each of these N subsequences is fixed, it is possible to employ any one of the existing shift estimators to estimate the frame-to-frame shift in each of these N image subsequences. We reiterate that frame-to-frame shift in a subsequence corresponds to a shift in N frames in the original spectrally dynamic sequence. We also remind the reader that the totality of these motion estimates, one for each subsequence, yield a collection of N overlapping and coarse motion estimates in the original image sequence $\mathbf{I}$.

B. Problem Definition: Determine the Frame-to-Frame Motion Estimates From Overlapping Coarse Motion Estimates

Clearly, knowledge of the motion between frames in the image sequence $\mathbf{I}$ in one period allows us to calculate the

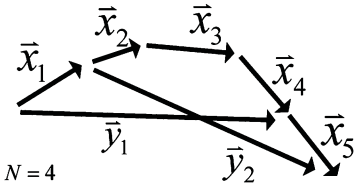

Fig. 2. Visual representation of sample frame to frame motion vectors ($\vec{x}_n$'s) and their relationship to the overlapping coarse motion estimates ($\vec{y}_n$'s) given (3).

coarse motion estimate between two consecutive frames in the first subsequence; it is simply the vector sum of all intermediate frame-to-frame motion estimates. Let us use this simple relationship to model a corresponding inverse problem.

Let us define $\vec{x}_n$ as the 2-D global translational vector (hereafter, termed global motion) between frames n and $n+1$ in the image sequence $\mathbf{I}$. Let us define $\vec{y}_n$ as the global motion between frames n and $n+N$ of the image sequence $\mathbf{I}$. In other words, this is the coarse motion between the frames $\mathbf{I}_n$ and $\mathbf{I}_{n+N}$. We can now relate $\vec{y}_n$ (the motion between frames in a subsequence) in terms of $\vec{x}_n$ (the motion between consecutive frames in the original sequence) as follows:

$$\vec{y}_n = \sum_{i=n}^{n+(N-1)} \vec{x}_i. \tag{3}$$

By using this relationship (also see Fig. 2 for visual representation), we can now describe the problem in the form of an ill-posed inverse mathematical problem. Let us begin by considering a special case.

If we make the oversimplifying assumption that the motion has a constant velocity (i.e., no acceleration), then the solution is trivial: divide any of the coarse motion estimates overlapping the frame-to-frame motion of interest by the number of frame-to-frame transitions contained within the coarse motion estimate (N), i.e.,

$$\vec{x}_n = \frac{\vec{y}_i}{N}, \quad i = n-(N-1), \ldots, n. \tag{4}$$

Note that this will give N different solutions for each individual $\vec{x}_n$. If indeed the motion had constant velocity, and all of the coarse motion estimates calculated were perfectly accurate, then we would end up with the same solution each time. In actuality, however, the motion estimates will not be perfect every time; thus, we will need to intelligently combine the different solutions in an attempt to get the best possible solution for each $\vec{x}_n$.

As will be shown later, using this simplifying assumption of constant velocity, as well as using a novel approach to combine each of the N solutions provides an accurate solution to the problem of determining frame-to-frame motion estimates, given only periodic overlapping coarse motion estimates between every N frames.

III. Algorithm

The rationale of the proposed algorithm is to use coarse and overlapping motion estimates between each pair of frames in every subsequence to calculate frame-to-frame motion estimates by using a weighted average scheme. An enabling idea is that the weights are chosen inversely proportional to the "error" measured for each coarse motion estimate so as to give a higher emphasis on coarse motion estimates with small error. The detailed description of the algorithm, which we term the spectrally dynamic shift estimator (SDSE), is described shortly.

A. Coarse Motion Estimates

Coarse motion estimates between each pair of frames from each subsequence of images from the same band are computed using any traditional shift estimation algorithm. For this study, we used a PBSE for all coarse motion estimates [16]. This gives us a set of coarse motion estimates, which are represented in our math model as the $\vec{y}_n$'s.

B. Error for Coarse Motion Estimates

In order to usefully combine the overlapping coarse motion estimates, it is important to be able to quantify how "good" each estimate is. The quality of each coarse motion estimate is related directly to how well the motion compensated image matches the original. For a perfect estimate—assuming no temporal noise—if you take the frame $\mathbf{I}_n$ and shift it by the vector $\vec{y}_n$, you should end up with an exact copy of the frame $\mathbf{I}_n$ in the overlapping regions of the two images. To calculate the rms error related to each coarse motion estimate, we will simply subtract each pixel in the overlapping region (call this region O) of the shifted version of the source frame with that of the destination frame. We will then square these differences, sum them up, and then take the square root of that sum. This gives us the rms error e_n associated with each motion estimate $\vec{y}_n$

$$e_n = \sqrt{\sum_{x,y \in O} (\mathbf{I}_n(x + (\vec{y}_n)_x, y + (\vec{y}_n)_y) - \mathbf{I}_{n+N}(x,y))^2} \tag{5}$$

where $(\vec{y}_n)_x$ refers to the shift in the x-direction of the motion estimate $\vec{y}_n$ and $(\vec{y}_n)_y$ refers to the shift in the y-direction of the motion estimate $\vec{y}_n$.

There is still an issue with the way we are calculating the rms error in (5). If we want to compare rms errors from different overlapping coarse motion estimates, we would like to compare errors that are related to one another in some normalized fashion. As such, we need to normalize the rms error such that regardless of the intensities of the input images, and regardless of the size of the overlapping region O, that the rms errors all approach some minimum value when the motion estimate is most accurate and increase proportionally larger as the accuracy drops.

First, in order to normalize the size of the overlapping region rather than taking the entire region, we will simply define the region O to be the center of the target frame $\mathbf{I}_{n+N}$ with the outer edges stripped off. The amount of the outer edges stripped off can be variable, but only needs to be as large as the largest motion estimate that we expect to see. Thus, region O is a subset of the overlapping region between each pair of frames. By stripping off the outer edges, we are comparing the same number of overlapping pixels when determining the rms error for each of the coarse motion estimates.

Next, we will normalize the rms error to be independent of the intensity values of each frame. For example, if we have one band

with very large intensity values, then a higher rms error value will occur for small motion estimate errors, compared with a band with very small intensity values and the same motion estimate error. To normalize, we simply divide the intensity values by the average of the L_2-norms (or Euclidean vector norms) of the source and destination frames before working with them, i.e., we normalize the images before computing the motion estimates. The rms error equation given in (5) still holds if we first assume that $\mathbf{I}_n$ and $\mathbf{I}_{n+N}$ have both been normalized in this manner prior to being used to calculate the rms error. Let us designate $\hat{\mathbf{I}}_n$ and $\hat{\mathbf{I}}_{n+1}$ the normalized versions of $\mathbf{I}_n$ and $\mathbf{I}_{n+N}$, respectively. Thus, our final normalized rms error $\hat{e}_n$ associated with each motion estimate $\vec{y}_n$ is

$$\hat{e}_n = \sqrt{\sum_{x,y\in O} (\hat{\mathbf{I}}_n(x+(\vec{y}_n)_x, y+(\vec{y}_n)_y) - \hat{\mathbf{I}}_{n+N}(x,y))^2}. \tag{6}$$

C. Frame-to-Frame Motion Estimates Using Weighted Averages

Now that we have all of the overlapping coarse motion estimates, along with their corresponding rms errors, we can estimate each of the frame-to-frame motion estimates as described in (4). We will then have N different estimates for each frame-to-frame estimate (the $\vec{x}_n$'s).

We can improve the final shift estimate if we use an intelligent weighted average scheme. One such set of weights we can use is the normalized set of the inverses of the rms errors associated with each coarse motion estimate. This gives us the benefit of giving a higher bias to those coarse motion estimates that are accurate and a small bias to those that are not. Thus, our final estimate for each $\vec{x}_n$ is given by

$$\tilde{\vec{x}}_n = \frac{1}{N} \sum_{i=n-(N-1)}^{n} w_i \frac{\vec{y}_i}{N} \tag{7}$$

where $\tilde{\vec{x}}_n$ is the weighted average of all the solutions for $\vec{x}_n$ and w_i is the normalized weight associated with each coarse estimate's rms error

$$w_i = \frac{a}{\hat{e}_i}. \tag{8}$$

The term a is the normalization term used to ensure that the sum of all the weights sum to 1, i.e.,

$$\sum_{i=n-(N-1)}^{n} w_i = 1. \tag{9}$$

Solving for a, we obtain

$$a = \frac{1}{\sum_{i=n=(N-1)}^{n} \frac{1}{\hat{e}_i}}. \tag{10}$$

This completes the description of the SDSE algorithm.

In summary, we have computed a frame-to-frame motion estimate using weighted averages of overlapping coarse estimates. In the next section, we will analyze by comparison how well this algorithm performs versus other more trivial alternatives.

IV. Algorithm's Performance Analysis

The algorithm described before, the SDSE algorithm, will be compared to three other algorithms, namely the direct method, the simple interpolation method, and the equal weights method, all of which are described in detail later. Multiple datasets will be used, exhibiting both linear and accelerated motion, as well as giving a variety of spectral scene signatures. All of the raw data used comes from preprocessed datasets from the multispectral thermal imager (MTI) [21].

A. Multispectral Thermal Imager (MTI)

MTI is a Department of Energy (DOE) satellite imaging system designed to collect radiometrically calibrated, medium resolution imagery in 15 spectral bands ranging form 0.45--10.70 μm [21]. MTI itself is not a dynamic sensor with frame-to-frame variation in its spectral response, but is rather a push-broom sensor acquiring all 15 bands simultaneously as the sensor is swept across a scene. Scene imagery from MTI is used as a realistic source of radiometric data for use in testing our algorithm. Four bands from eight different scenes were used as the base data.

In order to simulate a dynamic sensor with frame-to-frame variation in its spectral response, a 256×256 viewpoint was swept across each of the four bands in an MTI scene using simulated motion vectors. Frames from each of the four chosen bands were then interleaved to create periodic image sequences. The same motion vectors were then applied to each subsequent scene, and then each of the algorithms were applied to the resultant image sequences.

Using simulated motion provides us the ability to have truth data for which we can compare our motion estimates to. This truth data are not used in any of the calculations of our motion estimates, i.e., they can run without any knowledge of sensor and/or scene movement, but are only as used a comparison tool after the fact to determine absolute error in the frame-to-frame motion estimates calculated by each method. The error reported next is the sum of the average pixel error in the horizontal direction across all motion estimates in one scene and that of the average pixel error in the vertical direction in one scene. The maximum, minimum, and average errors across the eight scenes are reported.

B. Simulated Motion

Three shift estimation methods were applied to three different simulated motions across all eight MTI scenes.

The first simulated motion was constant linear motion in both the horizontal and vertical directions. A total of 30 frames were generated with a constant motion of two pixels in both the horizontal and vertical directions between each pair of consecutive frames. This motion will be referred to as *linear motion without jitter*.

The second simulated motion was generated by taking the constant linear motion and adding a random jitter on top of that motion. The jitter was either 1 or 0 in the vertical direction, and was 0, 1, or 2 in the horizontal direction. This allowed for a motion of 2–4 pixels in the horizontal direction between frames, and between 2 and 3 pixels in the vertical direction between frames. The jitter was generated randomly one time, and then

that same jitter was reused across each of the shift estimation methods to allow for better comparison. This motion will be referred to as *linear motion with jitter*.

The third and final simulated motion was that of constant acceleration in both the horizontal and vertical directions, with a complete reverse in acceleration after frame the fifteenth frame (again 30 total frames were generated). Motion started at 1 pixel per frame in both the horizontal and vertical directions, increased to eight pixels per frame, and then decreased back to one pixel per frame. This motion will be referred to as *accelerated motion*.

C. Direct Method

The direct method is a naïve method used to calculate frame-to-frame motion estimates in a sensor with frame-to-frame variation in its spectral response. This method simply takes an existing algorithm, in our case, the PBSE algorithm, and applies it directly to two consecutive frames, ignoring the fact that their spectral response has changed.

For linear motion without jitter, the maximum error obtained using this method was 20.4 pixels. The minimum error was 16.5 pixels. The overall average was 18.2 pixels.

For the linear motion with jitter, the maximum error was 21.2 pixels, the minimum error was 13.9 pixels, and the overall average error was 18.5 pixels. For accelerated motion, the maximum error was 23.4 pixels, the minimum error was 17.3 pixels, and the overall average error was 19.8 pixels.

The errors remain fairly consistent across each of the three simulated motions. In fact, the motion estimate error for this method is directly correlated to the size of overlapping images used in the PBSE algorithm. The results for each of the earlier reported estimates all used a 10 pixel buffer area around the images being correlated. When doubling this buffer area to 20 pixels, the errors also doubled, with the mean motion error becoming 37.3 pixels for the linear motion with jitter case, for example. As expected, the PBSE algorithm is completely ineffective at estimating motion between two consecutive frames of different spectral content as shown by these high errors.

D. Interpolation Method

The interpolation method uses a single coarse estimate to determine the frame-to-frame motion estimates. In other words, the weights for all, but one overlapping motion estimate are set to 0, while the weight of a single arbitrary band is set to 1. This is the same as taking a single result from (4) and discarding all the others. This method requires less computation than the direct method, as a single coarse estimate is used to calculate multiple frame-to-frame motion estimates. For example, if four bands were used, this would be four times faster than the direct method.

For linear motion without jitter, the maximum average error was 9.9 pixels. The minimum average error was 0 pixels. The overall average error was 2.8 pixels. The interpolation method works just fine in some cases (perfect even for one), but it has no way to dampen the effect of bad estimates—as shown by the maximum error of nearly 10 pixels in one case.

For linear motion with jitter, the maximum average error was 7.8 pixels, the minimum average error was 0.9 pixels, and the overall average error was 2.0 pixels. The maximum average error for this case was an anomaly among the eight datasets used—taking that one sample out, the overall average error is reduced to 1.1 pixels.

For accelerated motion, the maximum average error was 8.4 pixels, the minimum average error was 1.0 pixel, and the overall average error was 4.1 pixels.

The interpolation method outperforms the direct method on all accounts, as expected. Still, the overall average errors are all still fairly large.

E. Equal Weights Method

The equal weights method ignores the error estimates generated by using (6), and assigns equal weights in (7). In this case, all four weights are assigned the value 0.25. If our coarse motion estimates were without any error, then this method would be identical to the proposed method. This method is computationally equivalent to the direct method (i.e., slower than the interpolation method), as it requires the same number of motion estimates to be calculated as the direct method.

For linear motion without jitter, the maximum error was 3.1 pixels. The minimum error was 0 pixels. The overall average was 1.5 pixels. For linear motion with jitter, the maximum error was 2.7 pixels, the minimum error was 0.8 pixels, and the overall average error was 1.8 pixels. For accelerated motion, the maximum error was 4.0 pixels, the minimum error was 0.9 pixels, and the overall average error was 2.3 pixels. It can be concluded that there is a great improvement across the board when equal weights are used, as compared to the interpolation method. The overall average errors were nearly cut in half across the board. Using each of the coarse motion estimates equally instead of just one substantially improves the results.

F. SDSE Method

The SDSE method has been described in detail before in Section III. The same simulated motion was also used to test the performance of the SDSE method. The region O used in calculating the error estimates of (6) was picked to range from 1 to 100 for both x and y. This method is computationally slower than the equal weights method because of the extra error calculation involved for each motion estimate calculated.

For linear motion without jitter, the maximum error was 2.5 pixels. The minimum error was 0 pixels. The overall average was 0.6 pixels. For linear motion with jitter, the maximum error was 2.0 pixels, the minimum error 0.8 pixels, and the overall average error was 1.2 pixels. For accelerated motion, the maximum error was 1.6 pixels, the minimum error was 0.5 pixels, and the overall average error was 0.9 pixels.

It is seen that there is an obvious improvement in using the proposed method over any of the prior methods. An inherent benefit of using this method is its ability to ignore highly erroneous coarse motion estimates. This is very apparent when comparing the maximum errors of the prior methods with that of the proposed method. On average, there is approximately a two times improvement over the equal weights method, and a four times improvement over the interpolation method when using the SDSE method. In the special case with one or more highly

TABLE I
ERROR COMPARISON BETWEEN MOTION ESTIMATION METHODS

	Direct Method	Interpolated Method	Equal Weights Method	Proposed Method
		Linear Motion without Jitter		
Minimum Error	16.5 pixels	0 pixels	0 pixels	0 pixels
Maximum Error	20.4 pixels	9.9 pixels	3.1 pixels	2.5 pixels
Average Error	18.2 pixels	2.8 pixels	1.5 pixels	0.6 pixels
		Linear Motion with Jitter		
Minimum Error	13.9 pixels	0.9 pixels	0.8 pixels	0.8 pixels
Maximum Error	21.2 pixels	7.8 pixels	2.7 pixels	2.0 pixels
Average Error	18.5 pixels	2.0 pixels	1.8 pixels	1.2 pixels
		Accelerated Motion		
Minimum Error	17.3 pixels	1.0 pixel	0.9 pixels	0.5 pixels
Maximum Error	23.4 pixels	8.4 pixels	4.0 pixels	1.6 pixels
Average Error	19.8 pixels	4.1 pixels	2.3 pixels	0.9 pixels

The proposed method performs well in all cases, much better than any of the more naïve methods it is compared against.

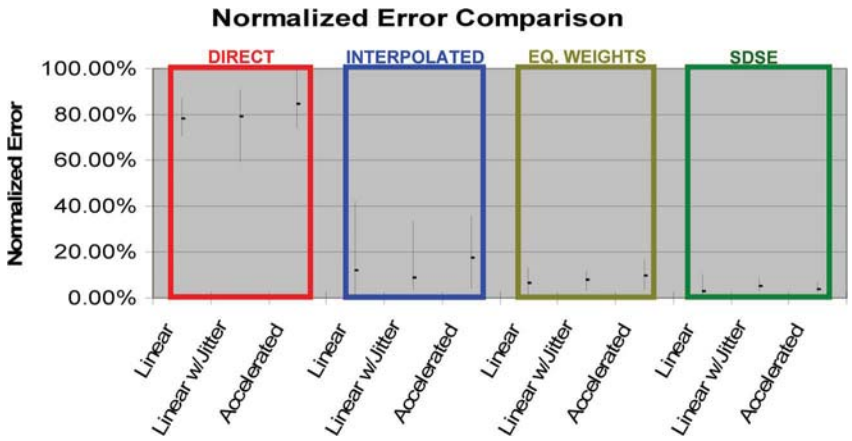

Fig. 3. Graph of the normalized errors for the direct, interpolated, equal weights, and SDSE, respectively. Errors have been normalized to the largest error. Error bars show the range from the minimum average error to the maximum average error, with the overall average error for each simulated motion and motion estimation method marked.

erroneous coarse motion estimates, the SDSE method was observed to perform more than ten times better than the equal weights method, an order of magnitude improvement. These results are summarized in Table I and Fig. 3.

V. CONCLUSION

With the advent of new sensing technologies able to provide new types of multispectral data, we must begin to address the algorithmic needs to effectively process these data. For dynamic sensors with frame-to-frame variation in their spectral response, an effective technique has been demonstrated to compute the frame-to-frame motion estimates accurately. This technique effectively applies any existing shift estimation algorithm to a new type of image sequence with accurate results.

There are several direct paths forward to extend and enhance the work that has been presented in this study. First, the SDSE algorithm needs to be applied to images acquired from a dynamic DWELL FPA. There are differences between the MTI data used to validate this algorithm and that coming from an actual DWELL sensor. These differences could affect the effectiveness of the SDSE algorithm. Next, different regularization methods could be used to solve the inverse problem, instead of assuming linear motion in between each coarse motion estimate. There may be other regularization methods that would improve the SDSE algorithm. Lastly, the DWELL sensor model given within this paper does not take into account noise (e.g., dark current). Modeling the noise and thus taking a statistical approach at solving this problem should help improve the effectiveness of this algorithm applied to data from dynamic DWELL sensors.

REFERENCES

[1] S. Krishna, M. M. Hayat, J. S. Tyo, U. Sakoglu, and S. Raghavan, "Spectrally adaptive quantum dot infrared sensors for focal plane arrays," U.S. Patent 7 217 951, May 2007.

[2] W.-Y. Jang, M. M. Hayat, J. S. Tyo, R. S. Attaluri, T. E. Vandervelde, Y. D. Sharma, R. Shenoi, A. Stintz, E. R. Cantwell, S. Bender, and S. Krishna, "Demonstration of bias controlled algorithmic tuning of quantum dots in a well (DWELL) mid-infrared detectors," *IEEE J. Quantum Electron.*, , to be published.

[3] W.-Y. Jang, R. S. Attaluri, S. Annamalai, M. M. Hayat, J. S. Tyo, and S. Krishna, "Development of an algorithmic spectrometer for target recognition using quantum-dot infrared sensors," presented at the presented at the Nanoelectron. Devices Defense Security (NANO-DDS) Conf., Arlington, VA, Jun. 2007.

[4] B. Paskaleva, M. M. Hayat, Z. Wang, and J. S. Tyo, "Canonical correlation feature selection for sensors with overlapping bands: Theory and application," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 3346–3358, Oct. 2008.

[5] C. Morimoto and R. Chellappa, "Evaluation of image stabilization algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 1998, vol. 5, pp. 2789–2792.

[6] E. Dubois and S. Sabri, "Noise reduction in image sequences using motion-compensated temporal filtering," *IEEE Trans. Commun.*, vol. 32, no. 7, pp. 826–831, Jul. 1984.
[7] S. C. Cain, E. Armstrong, and B. Yasuda, "Joint estimation of image, shift and nonuniformities from IR images," in *Proc. Meeting IRIS Spec. Group Passive Sens. 1*, Tucson, AZ, 1997, pp. 121–132.
[8] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP Graph. Models Image Process.*, vol. 53, pp. 231–239, 1991.
[9] R. R. Schultz and R. L. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Trans. Image Process.*, vol. 3, no. 3, pp. 233–242, May 1994.
[10] S. P. Kim, N. K. Bose, and H. M. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 6, pp. 1013–1027, Jun. 1990.
[11] R. Y. Tsai and T. S. Huang, "Multi-frame image restoration and registration," *Adv. Comput. Vis. Image Process.*, vol. 1, pp. 317–339, 1989.
[12] R. C. Hardie, M. M. Hayat, E. Armstrong, and B. Yasuda, "Scene based nonuniformity correction using video sequences and registration," *Appl. Opt.*, vol. 39, no. 8, pp. 1241–1250, 2000.
[13] W. F. O'Neil, "Experimental verification of dithered scan nonuniformity correction," in *Proc. Meeting IRIS Spec. Group Passive Sens. 1*, Tucson, AZ, 1997, pp. 329–339.
[14] E. Armstrong, M. M. Hayat, R. C. Hardie, S. Torres, and B. Yasuda, "The advantage of nonuniformity correction pre-processing on infrared image registration," *Proc. SPIE*, vol. 3808, pp. 150–161, 1999.
[15] D. I. Barnea and H. F. Silverman, "A class of algorithms for fast digital image registration," *IEEE Trans. Comput.*, vol. C-21, no. 2, pp. 179–186, Feb. 1972.
[16] S. C. Cain, M. M. Hayat, and E. E. Armstrong, "Projection-based image registration in the presence of fixed-pattern noise," *IEEE Trans. Image Process.*, vol. 10, no. 12, pp. 1860–1872, Dec. 2001.
[17] A. Roche, X. Pennec, G. Malandain, and N. Ayache, "Rigid registration of 3-D ultrasound with MR images: A new approach combining intensity and gradient information," *IEEE Trans. Med. Imag.*, vol. 20, no. 10, pp. 1038–1049, Oct. 2001.
[18] Z. Jiang and W. Lu, "A new digital image registration algorithm based on the double spatial intensity gradients using pyramids," *Pattern Recog. Lett.*, vol. 8, pp. 335–340, 1988.
[19] H. Li, B. S. Manjunath, and S. K. Mitra, "A contour based approach to multisensor image registration," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 320–334, Mar. 1995.
[20] J. Ton and A. K. Jain, "Registering landsat images by point matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 27, no. 5, pp. 642–651, Sep. 1989.
[21] R. R. Kay, S. C. Bender, T. D. Henson, D. A. Byrd, J. L. Rienstra, M. L. Decker, N. G. Rackley, R. L. Akau, P. J. Classen, R. E. Kidner, R. B. Taplin, D. M. Bullington, K. D. Marbach, C. E. Lanes, C. Little, B. W. Smith, B. C. Brock, and P. G. Weber, "Multispectral thermal imager (MTI) payload overview," *Proc. SPIE*, vol. 3753, pp. 347–358, 1999.

Todd M. Giles received the B.S. degree in electrical engineering and the B.S. degree in computer science from New Mexico State University, Las Cruces, in 2001, and the M.S. degree in electrical and computer engineering from the University of New Mexico, Albuquerque, in 2008.

From 2001 to 2005, he was a Software Engineer at Anadigm, Inc., where he was engaged in developing computer-aided design tools for field programmable analog arrays. Since 2005, he has been a Senior Member of the Technical Staff at Sandia National Laboratories, Albuquerque.

Majeed M. Hayat (S'89–M'92–SM'00) was born in Kuwait in 1963. He received the B.S. degree (*summa cum laude*) in electrical engineering from the University of the Pacific, Stockton, CA, in 1985, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Wisconsin–Madison, Madison, in 1988 and 1992, respectively.

From 1993 to 1996, he was with the University of Wisconsin–Madison as a Research Associate and a Co-Principal Investigator of a project on statistical minefield modeling and detection, which was funded by the Office of Naval Research. In 1996, he joined the Faculty of the Electro-Optics Graduate Program and the Department of Electrical and Computer Engineering, University of Dayton, Dayton, OH. He is currently a Professor in the Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, where he is also with the Center of High Technology Materials. His research contributions cover a broad range of topics in statistical communication theory, optoelectronics, signal/image processing, remote sensing and applied probability theory including avalanche photodiodes, optical communication systems, image restoration and enhancement, and queueing models for distributed computing systems and networks.

Dr. Hayat was the recipient of the 1998 National Science Foundation Early Faculty Career Award. He is a Member of the International Society for Optical Engineers and the Optical Society of America.

Sanjay Krishna (M'01–SM'08) received the Master's in physics from the Indian Institute of Technology (IIT), Madras, India, in 1996, the M.S. degree in electrical engineering and the Ph.D. degree in applied physics from the University of Michigan, Ann Arbor, in 1999 and 2001.

In 2001, he joined the University of New Mexico, Albuquerque, as a tenure track faculty member, where he is currently an Associate Professor of electrical and computer engineering at the Center for High Technology Materials. His research interests include growth, fabrication and characterization of self-assembled quantum dots, and type-II InAs/InGaSb-based strain layer superlattices for mid-IR detectors. He is the authored or coauthored of more than 70 peer-reviewed journal articles, more than 40 conference presentations, and two book chapters. He has two issued and five pending patents.

Dr. Krishna was the recipient the Gold Medal from IIT, Madras, in 1996, the Best Student Paper Award at the 16th North American Molecular Beam Epitaxy (NAMBE) Conference, Banff, in 1999, the 2002 Ralph E Powe Junior Faculty Award from Oak Ridge Associated Universities, the 2003 IEEE Outstanding Engineering Award, 2004 Outstanding Researcher Award from the Electrical and Computer Engineering Department, the 2005 School of Engineering Junior Faculty Teaching Excellence Award, the 2007 NAMBE Young Investigator Award, the 2007 National Consortium for MASINT Research–Defense Intelligence Agency Chief Scientist Award for Excellence, and the 2008 Early Career Achievement Award from SPIE and IEEE-Nanotechnology Council.