An Improved Fourier-Mellin Transform-Based Registration Used in TDI-CMOS

QINPING FENG [1,2,3], **SHUPING TAO** [1,3], **CHUNYU LIU** [1,3], **AND HONGSONG QU** [1,3]

[1]Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China
[2]College of Materials Science and Opto-Electronic Technology, University of Chinese Academy of Sciences, Beijing 100049, China
[3]Key Laboratory of Space-Based Dynamic and Rapid Optical Imaging Technology, Chinese Academy of Sciences, Changchun 130033, China

Corresponding author: Shuping Tao (taoshuping11@sina.com)

ABSTRACT Time-delayed integration (TDI) is a common imaging mode used in airborne cameras to compensate for image motion. The paper presents an improved Fourier-Mellin transform (FMT)-based registration method, which can be used to realize the registration-based TDI. It first analyzes the noise influences on the FMT spectrum by comparing the curve changes in the standard-deviation versus column plots under the various intensity of noise, then designs a special filter called WCSDF according to the amount of variation of each column in the FMT spectrum, which can reduce the noise affection on registration process to the larger extent. And an improved FMT-based registration method is finally proposed to form a framework of registration-based TDI. In computer simulations, the proposed method shows a significant improvement in robustness to noise (noise level up to 40). Compared with the existing method, its low computational complexity makes the method easy to be implemented in hardware and can estimate the larger relative shifts among dim noisy images, moreover, the TDI images generated by the proposed framework have higher quality in the index of information entropy, average gradient, and spatial frequency response.

INDEX TERMS Cross-power spectrum, image registration, TDI-CMOS, correlation function, Fourier-Mellin transform.

I. INTRODUCTION

While remote sensing cameras acquiring scene images at a rapid motion, time delay and integration (TDI) is a typical imaging mode that used to reduce image motions. Recently, many types of sensors with TDI are demonstrated [1], [2], and their verifications and applications were also carried out [3]–[5]. For M-stages TDI, it first captures the moving signal shortly-exposed from a scene for M times, then accumulates these signals to obtain high-contrast, reduced motion-effects TDI images, the signal-noise ratio (SNR) of which is increased by $\sqrt{M}$ times. The TDI is generally used in the line-scan applications and easily applied on CCD. Recently, many methods for implementing TDI-CMOS are presented, both in the analog and digital domain. Reference [6] reviewed several TDI-CMOS sensors with analog accumulator, which were designed at the cost of silicon area for analog accumulator. Reference [7] implemented a

128-stage TDI-CMOS image sensor with on-chip digital accumulator, which reduce the cost of silicon area. Reference [8] proposed an off-chip digital TDI-CMOS, which need no silicon area, and various functions can be implemented to satisfy different applications. Reference [9] proposed a 32-stage 15-b digital TDI-CMOS, where the data prediction switching technique is included. According to the process of TDI, because of the constantly changes in the attitude of camera, the motion compensation is required to ensure that each pixel acquires exactly the same object. Currently, Methods for compensating image motion can be classified into three categories: Mechanical, optical, and electronic. The mechanical or optical methods need high-precision mechanical or optical structure to compensate image motion, [10] studied a smart structure to control the vibration. However, this additional structure itself adds additional complexity and reduce reliability [11], [12]. The electronic methods compensate image motion by adjusting transfer time, which can only compensate the along-track-motion of very limited range. There are also some literatures about anti-vibration

The associate editor coordinating the review of this manuscript and approving it for publication was Sudipta Roy [ID].

TABLE 1. Acronyms Used in the Article.

TDI	Time delay and integration
CCD	Charge-coupled Device
CMOS	Complementary Metal Oxide Semiconductor
SNR	Signal-noise ratio
AWGN	Additive white gaussian noise
STD	Standard deviation
WSD	Weighted standard deviation
WCSDF	Weighted column standard deviation based filter
FMT	Fourier-Mellin transform
LPT	Log-polor transform
SFR	Spatial frequency response
ET	Exposure time, or integrating time
SIFT	Scale-invariant feature transform
SURF	Speeded-Up Robust Features
FPN	Fixed Pattern Noise
RBM	row-band mean
RBSD	row-band standard deviation
RBRSD	row-band relative standard deviation
GWSD	Gaussian-distributed WSD
SSRG	sum of squared row gradient
SSCG	sum of squared column gradient
AG	average gradient

technology on TDI. Reference [13] proposed a dynamic path accumulation method of digital-domain for TDI-CMOS image sensors, which can compensate for the image motion introduced by vibration, but it cannot compensate the image motion in two-dimensions.

In reality, the image motion is not only composed of along-track and cross-track motion, but also include the rotational shifts induced by constantly changing in satellites' attitude(such as roll, yaw or pitch). However, the existing solutions do not address the issues. To solve these problems, registration accuracy becomes a critical factor to obtain a high-SNR TDI images. In order to align images, scholars have proposed various kinds of image registration algorithms from different aspects, which can be roughly classified into 3 categories [14], [15]: 1) feature-based, 2) area-based, 3) deep learning. Feature-based methods need to detect a series of blob-like or edge-like features, such as SIFT [16], SURF [17], and GLOH [18], followed by feature description and matching. However, its time-consuming nature makes the methods of this kind not suitable for real-time application such as TDI technology. As a method of supervised learning, deep learning needs complex, handcrafted training data, to prepare such a training data, considerable expert knowledge and manpower are needed [19], [20]. Methods of area-based directly exploit the intensity information of overlapping area to ensure the accuracy and the stability [21]. These methods generally use similarity metrics such as normalized correlation function, mutual information, etc.
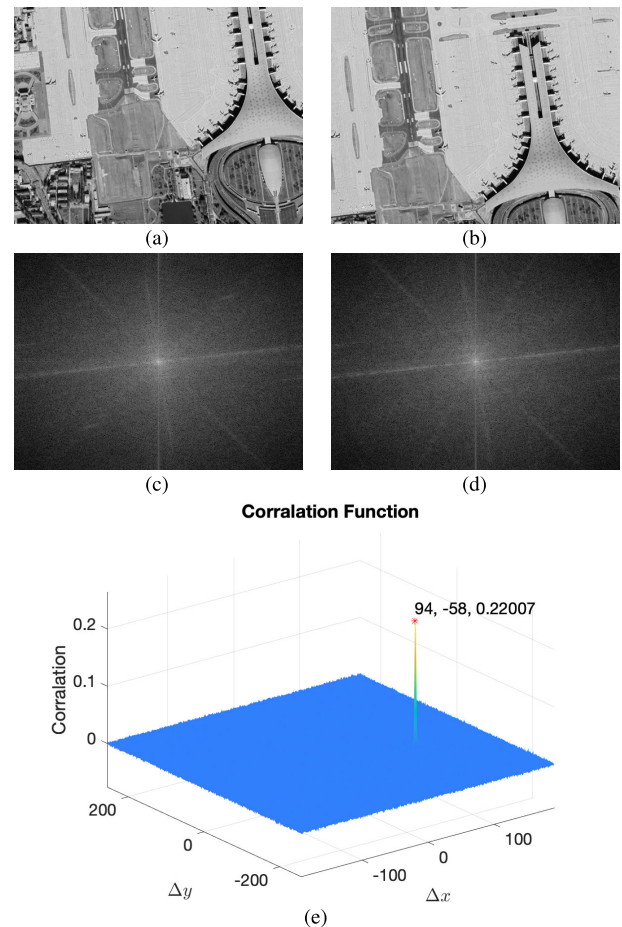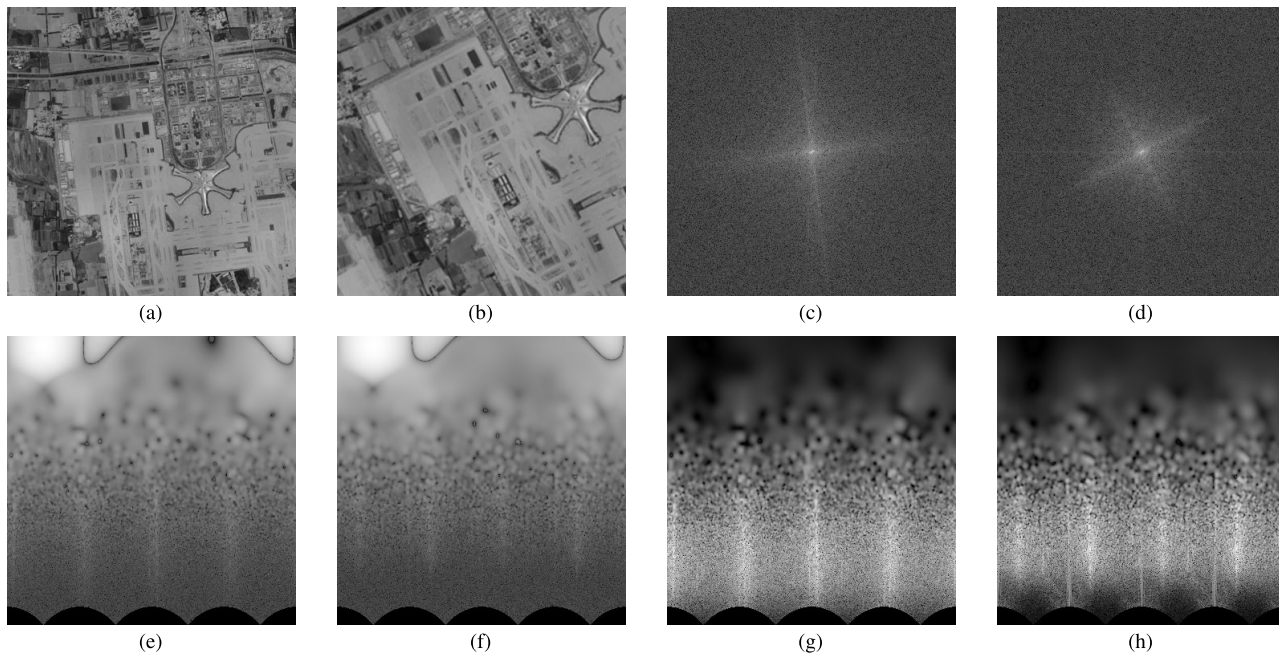

FIGURE 1. Two images with translational shift and theirs Correlation Function: (a) Reference image; (b) Sensed Image; (c) Fourier transform of Fig. 1(a); (d) Fourier transform of Fig. 1(b); (e) the correlation function between these two Fourier spectrum. The amplitudes in Fig. 1(c) and Fig. 1(c) are shown in logarithmic scale, and in Fig. 1(e), the indicator peak shows the translational shift of (94, −58) pixels.

In addition, there is another type of area-based algorithm that uses Fourier transform, it has theoretical high-accuracy, low time complexity, and robustness to brightness differences and frequency-related noise. It has been developed rapidly and applied in many fields, especially in real-time applications.

In this paper, our contributions can be summarized as follow:

1) We first analyze the noise influences in FMT spectrum by observing the changes of the curves in standard-deviation vs. column plots, then

2) a weighted column standard deviation based filter (WCSDF) is designed to handle the noise influences during the process of estimating the relative shifts, finally

3) an improved Fourier-Mellin cross-power spectrum based registration is proposed to form a framework of registration-based TDI, which can directly estimate the relative shifts among a series of short-exposure captures, and align them to generate high-SNR TDI images.

FIGURE 2. Two images with translational, rotational as well as scaling parallax: the parallax between a reference image $f(x, y)$ (a) and a sensed image $g(x, y)$ (b) contain relative translation of $(30, -50)$, rotation of 20 degrees, and scaling by $1.4\times$; (c) $m^{(F)}(u, v)$ and (d) $m^{(G)}(u, v)$ are the Fourier transform of (a)(b); (e)$m_{LP}^{(F)}$ and (f)$m_{LP}^{(G)}$ are the LPT of (c)(d); (g)(h) are obtained from (c)(d) by applying Gaussian high-pass filter, followed by LPT; The intensities in (c) to (h) are in logarithmic scale.

Unlike existing methods, the TDI framework proposed in this paper estimates the relative shifts directly from a series of short-exposures, then registers and accumulates them to generate high-contrast TDI images, whereas the existing methods, such as [8], which calculate the shifts by calculating the motion velocity.

The following sections are organized as follow: Section II reviews some related works, and their existing problems are also described; Some improvements and the proposed algorithm are explained in Section III; In Section IV, we conduct some experiments—both computational and physical—to verify the effectiveness of the proposed algorithm. Finally, some conclusions is described in section V.

II. RELATED WORKS
A. THE BASIS
Fourier based registration, especially the well-known phase-correlation method, was first proposed in [22], and expanded in [23]. This kind of methods have been employed in many applications, such as image registration, image fusion, motion tracking, etc. Assuming that we have two images with a translational shift: one is viewed as reference image $f(x, y)$ and the other, $g(x, y)$, called sensed image that to be registered, then their relation is:

$$g(x, y) = f(x - \Delta x, y - \Delta y) \quad (1)$$

where $(\Delta x, \Delta y)$ is their relative translation. In Fourier transform domain, the relation become:

$$G(u, v) = \exp[-2\pi i(u\Delta x + v\Delta y)] \cdot F(u, v) \quad (2)$$

where $i = \sqrt{-1}$, $F(u, v)$ and $G(u, v)$ are the Fourier transform of $f(x, y)$ and $g(x, y)$, respectively. Then the cross-power spectrum of those two images can be obtained by:
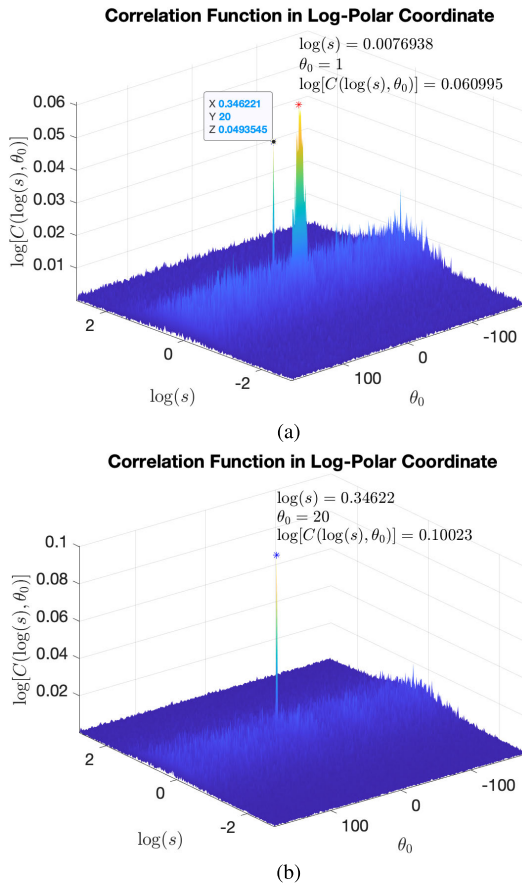
$$Q(u, v) = \frac{F^*(u, v) \cdot G(u, v)}{|F^*(u, v) \cdot G(u, v)|} = \exp[-i(u\Delta x + v\Delta y)] \quad (3)$$

Then, their correlation function can be obtained by calculating the inverse Fourier transform of $Q(u, v)$,

$$C(\Delta x, \Delta y) = \mathcal{F}^{-1}\{Q(u, v)\} \approx \delta(x - \Delta x, y - \Delta y) \quad (4)$$

Theoretically, it should be an impulse function. Because of the discrete nature, in ideal, the maximum value of 1 appears only at $(\Delta x, \Delta y)$ (called indicator peak), and the remaining are all zeros. However, there must exist some noise, and area that does not overlap in any two non-identical images, which will weaken the indicator peak. Nevertheless, the peak is still significantly higher than elsewhere and always can accurately reflect the relative displacement. An illustration is shown in Fig. 1.

The Fourier-based algorithm has an extension that uses Fourier-Mellin transform (FMT) based cross-power spectrum, which can deal with the rotational and scaling differences by utilizing the log-polar transform (LPT) on regular Fourier spectrum, as the LPT has the property that the rotation and scaling in Cartesian coordinate are equivalent to the translation in log-polar coordinate. According to equation (2), the spatial translational shifts between images appears as a phase shift in their Fourier transforms, therefore it can

Correlation Function in Log-Polar Coordinate

$\log(s) = 0.0076938$
$\theta_0 = 1$
$\log[C(\log(s), \theta_0)] = 0.060995$

X 0.346221
Y 20
Z 0.0493545

(a)

Correlation Function in Log-Polar Coordinate

$\log(s) = 0.34622$
$\theta_0 = 20$
$\log[C(\log(s), \theta_0)] = 0.10023$

(b)

FIGURE 3. FMT Correlation function: (a) obtained from Fig. 2(e) and 2(f); (b) obtained from Fig. 2(g) and 2(h). In (a) there exist two peaks, the shorter peak (called indicator peak) is showing the correct amount of shifts between two images, with a rotation of 20°, and a scaling of exp(0.34622) ≈ 1.4×, but it is interfered by a higher peak near to the center, which is caused by the densely distributed low-frequency components. By applying the high-pass filter, the correlation function will indicate the amount of shifts correctly, as shown in (b).

be separated by applying norm operation on their Fourier transforms. For any two images $f(x, y)$ and $g(x, y)$ with rotational shift $\Delta\theta$ and scaling factor s, the relation between their Fourier transform amplitude spectrums $m^{(F)}(u, v)$ and $m^{(G)}(u, v)$:

$$m^{(G)}(u, v) = s^{-2}m^{(F)}[s^{-1}(u \cos \Delta\theta + v \sin \Delta\theta),$$
$$s^{-1}(-u \sin \Delta\theta + v \cos \Delta\theta)] \quad (5)$$

can be expressed in log-polar form:

$$m_{LP}^{(G)}(\lambda, \theta) = s^{-2}m_{LP}^{(F)}(\lambda - \Delta\lambda, \theta - \Delta\theta) \quad (6)$$

where $\Delta\lambda = \log s$. Then, the rotational and scaling shifts can be estimated directly by the cross-power spectrum Q_{LP} and the correlation function C_{LP}:

$$Q_{LP}(\alpha, \beta)$$
$$= \frac{\overline{M_{LP}^{(F)}(\alpha, \beta)} \cdot M_{LP}^{(G)}(\alpha, \beta)}{\left| \overline{M_{LP}^{(F)}(\alpha, \beta)} \cdot M_{LP}^{(G)}(\alpha, \beta) \right|}$$
$$= \exp\left[-2\pi i(\alpha \Delta\lambda + \beta \Delta\theta)\right] \quad (7)$$

$$C_{LP}(\lambda, \theta) = \mathcal{F}^{-1}\{Q_{LP}(\alpha, \beta)\} \approx \delta(\lambda - \Delta\lambda, \theta - \Delta\theta) \quad (8)$$

where $M_{LP}^{(F)}$ and $M_{LP}^{(G)}$ are the Fourier transform of $m_{LP}^{(F)}$ and $m_{LP}^{(G)}$, also called the FMT of $f(x, y)$ and $g(x, y)$; $Q_{LP}(\alpha, \beta)$ is the FMT-based cross-power spectrum; $C_{LP}(\lambda, \theta)$ is the FMT-based correlation function, the over-line notation such as $\overline{M_{LP}^{(F)}}(\alpha, \beta)$ denotes its conjugate. The image details is shown in Fig. 2(a)—(f) and Fig. 3(a). After eliminating the relative rotation and scaling, the remaining translational shifts can be estimated by a regular correlation function.

B. OTHER VARIETIES

Further improvements and refinements of FMT-based registration have been proposed to address different issues. Reference [24] shows some error sources that affecting the precision of estimating the shifts, such as densely distributed low-frequency, and demonstrated constructive techniques for improving precision and SNR in registration process, as well as removing spurious high frequency by Blackman window. [23] handled the issue by applying high-pass filter, as shown in Fig. 2(g)(h) and Fig. 3(b). Reference [25] presented a pseudopolar Fourier transform to increase the performance of the registration algorithms, and expanded the range of recoverable scales (up to 4). Reference [26] developed exact algorithms to solve the issue of obtaining discrete Fourier transform for polar and spherical grids. Reference [27] constructed a special grid composed of multiple polar grids with different scaling factors to further reduce or eliminate interpolation error in calculating the FMT-based correlation function. In [28], the rotational and scaling shifts are estimated using the Radon transform, as well as sub-pixel estimation.

C. THE ESTIMATION ERROR CAUSED BY NOISE

Those aforementioned papers, however, did not address the problems of registration errors caused by noise, which make the methods fail to estimate the relative shifts—especially rotational and scaling—between any two noisy dim images. Since the probability of photons entering detectors obey Poisson distribution, there exists Poisson noise in short-exposure captures. In the light of this phenomenon, we simulate two short-exposure captures that have large poisson noise, as shown in Fig. 4, their FMT-based correlation function cannot correctly reflect the amount of the shift as the indicator peak is submerged by the noise. Thus, it is necessary to suppress the noise influence in estimating the relative shifts among images.

III. METHODOLOGY
A. PROBLEM ANALYSIS

Due to the nature of short-exposure captures, there are many types of noise that are not negligible: such as photon shot noise, Fano noise, fixed pattern noise (FPN), etc. Since the FPN can be corrected by the previous calibration, other
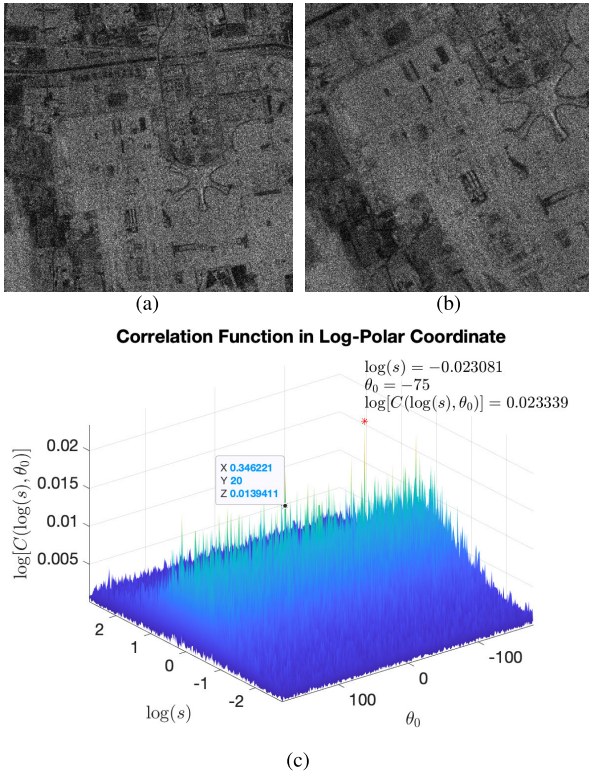
FIGURE 4. Reference and sensed images with poisson noise: (a)Reference image, (b)Sensed image, (c)The FMT correlation function between two images.

types of noise are random in spatial distribution. To quantitatively analyze the effect of the noise on the accuracy of rotation and scaling estimation, we use additive white Gaussian noise (AWGN) to simulate these random noise in computer simulation. In Fig.5, the sensed images with various intensity of AWGN and their log-polar Fourier amplitude spectrums are presented, from which we can conclude that the high-frequency components are prone to be interfered with by the noise compared with low-frequency components, this can be explained that the energy of high-frequency components is fainter than that of low-frequency. Since the high-frequency information is submerged by noise, the accuracy in estimating the relative shift will be greatly affected. By observing the middle row in Fig.5, we can notice that the mid-frequency components have the potential to estimate the relative shifts for the following reason: the mid-frequency components are (1) still have significant intensity variance; (2) less likely to be submerged by noise compared with high-frequency components. To exploit the potential, we first define some concepts in section III-B, wherein the rationality of the proposed method is described.

B. CONCEPTS DEFINATION

For any given original M-by-N image f and its Fourier amplitude spectrum $m^{(F)}$, its Fourier amplitude spectrum in log-polar form, which is noted as $m_{\text{LP}}^{(F)}(i, j)$ is also sampled by the same size, and we define k as the height in $m_{\text{LP}}^{(F)}(i, j)$ after cutting the "wave edge", as shown in Fig.6, if we calculate the column standard deviation (STD) of $m_{\text{LP}}^{(F)}(i, j)$ according to:

$$\sigma(j) = \sqrt{\frac{1}{k} \sum_{i=1}^{k} \left[m_{\text{LP}}^{(F)}(i, j) - \mu(j) \right]^2} \quad (9)$$

where $\mu(j) = \sum_{i=1}^{k} m_{\text{LP}}^{(F)}(i, j)/k$. The STD-vs-column plots according to the middle row in Fig.5 are shown in Fig.7

The unaffected STD-vs-column plot is shown in the left most in Fig.7, but the increasing amount of "spikes" introduced by noise will change the "curve shape".

It is well-known that the high-frequency components have contribution on column STD to some extent. However, the mid-frequency components also have non-negligible distribution on column STD. If an image f is given, we can calculate the row-band mean (RBM), row-band standard deviation (RBSD) and row-band relative standard deviation (RBRSD) from a log-polar Fourier amplitude spectrum $m_{\text{LP}}^{(F)}$, where the RBM, RBSD and RBRSD are calculated according to:

$$\text{RBM}_f(i) = \frac{1}{2Nr} \sum_{\theta=0}^{N-1} \sum_{\lambda=\lambda-r}^{\lambda+r-1} m_{\text{LP}}^{(F)}(\lambda, \theta) \quad (10)$$

$$\text{RBSD}_f(i) = \sqrt{\frac{1}{2Nr} \sum_{\theta=0}^{N-1} \sum_{\lambda=\lambda-r}^{\lambda+r-1} \left[m_{\text{LP}}^{(F)}(\lambda, \theta) - \text{RBM}(i) \right]^2} \quad (11)$$

$$\text{RBRSD}_f(i) = \text{RBSD}_f(i)/\text{RBM}_f(i) \quad (12)$$

where N is the sampled column number, r is the half-bandwidth, as illustrated in Fig. 6(b). In this way, we plot the trend of RBSD/RBRSD-vs-Row from Fig. 2(b), as shown in Fig. 6(c) and 6(d), and the RBSD/RBRSD-vs-Row plots of many other images also all have the similar trend. We can see that the high-frequency components have much higher relative STD (Fig. 6(d)), but Fig. 6(c) indicates that the mid-frequency components have non-negligible contribution to column STD. Therefore we need to introduce "Weighted Standard Deviation" (WSD) to exploit the potential of mid-frequency components.

Inspired by weighted average, the $1/k$ in (9) can be seen as the evenly distributed weights, therefore the WSD is introduced by altering the weights according to the importance of each component, that is:

$$\sigma_{\text{WSD}}(j) = \sqrt{\sum_{i=1}^{k} w_i \left[m_{\text{LP}}^{(F)}(i, j) - \mu(j) \right]^2} \quad (13)$$

with a constraint of $\sum_{i}^{k} w_i = 1$. In practice, there is no need to satisfy the constraint, as it only brings a constant factor, which will be canceled in calculating correlation functions. Because
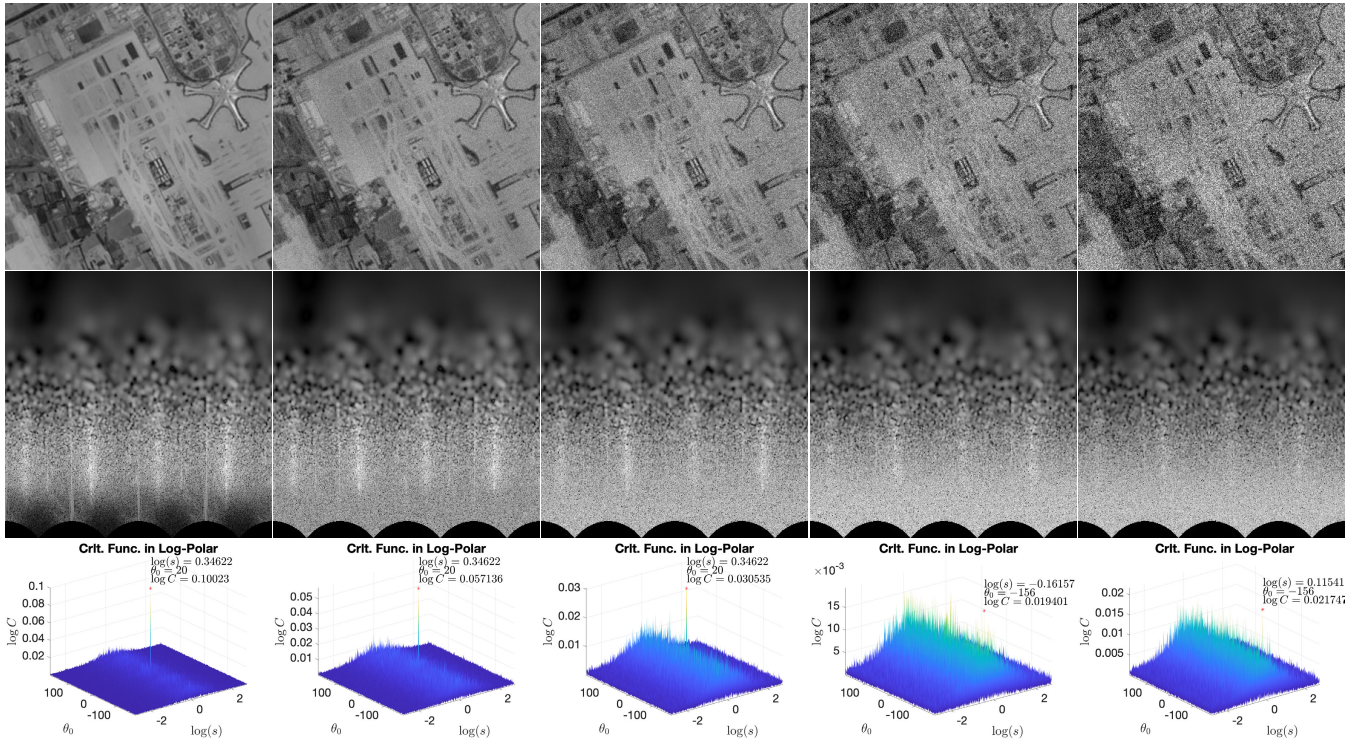
FIGURE 5. The Noise Influence in Estimating Rotational and Scaling Shifts: From left to right column: the Gaussian noise level are 0, 10, 20, 30, 40; First row: Sensed images; Second row: Fourier transform amplitude spectrum in log-polar form (The intensities are in logarithmic scale); Last row: the FMT-based correlation function with reference image (of the same noise level). To save space, their corresponding reference images and the log-polar form FFTs are not shown.
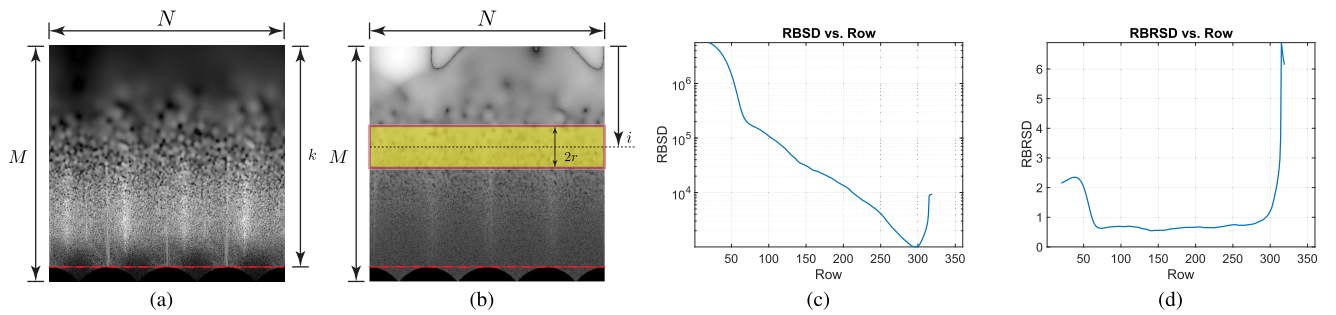

FIGURE 6. Definitions of some notation and the Plots of RBSD/RBRSD vs. Row: (a) A HPF-filtered log-polar Fourier amplitude spectrum (the same as in Fig. 2(h)); (b) An unprocessed log-polar Fourier amplitude spectrum (the same as in Fig. 2(f)); In these figures, M and N are the sampled row and column number, respectively; (c)(d) RBSD/RBRSD vs. Row Plots. In (a), k is the height after cutting the wave-like edge; In (b) r is the half-width of the band to calculate the RBM(i), RBSD(i) or RBRSD(i) for the corresponding i, where i is the center row in the band. The RBSD/RBRSD are calculated from the moving band area in an unprocessed log-polar amplitude spectrum.
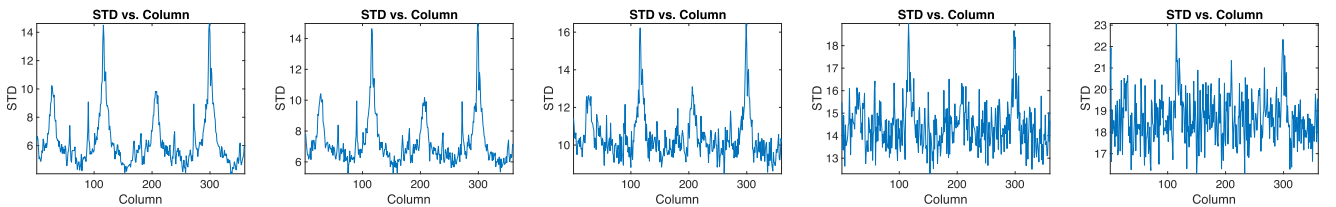

FIGURE 7. STD-vs-column plots according to the middle row of Fig.5: the amount of "spikes" is increasing as the noise level increases.

of the densely distributed low-frequency components and severely submerged high-frequency components in a Fourier amplitude spectrum, we may adopt a Gaussian-like distribution that assign higher weights to mid-frequency components:
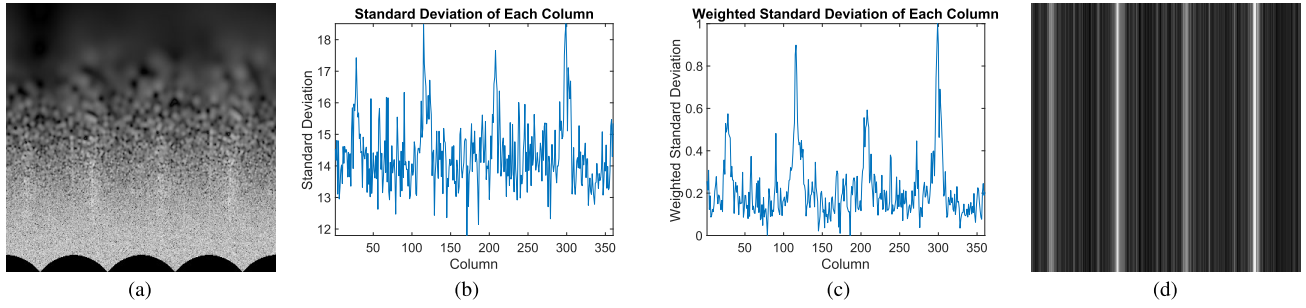
FIGURE 8. Comparison between STD/WSD-vs-column plots: (a)Log-polar Fourier amplitude spectrum of a noisy image (Noise Level of 30); (b)STD-vs-column plot; (c)WSD-vs-column plot; (d)WCSDF according to (c).
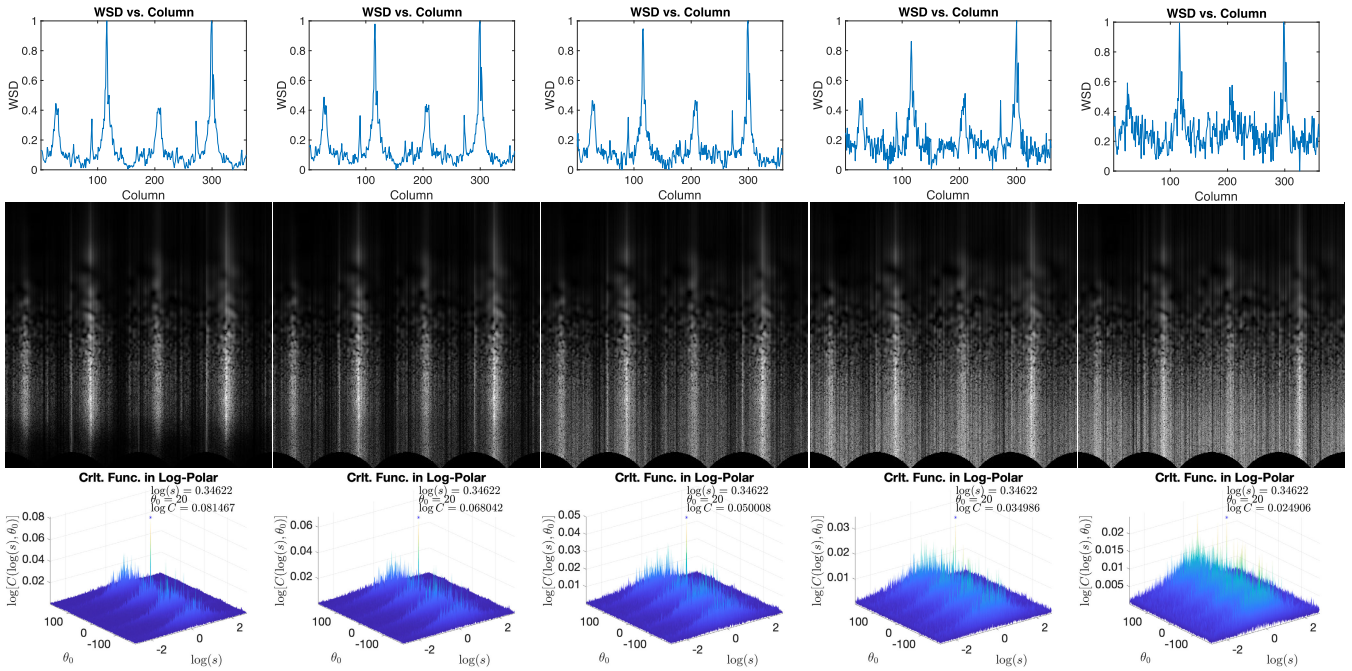
FIGURE 9. The effects of the filter. First row: the column GWSD plots of log-polar Fourier amplitude spectrum from Fig.5; middle row: the log-polar Fourier amplitude spectrum filtered by WCSDF; last row: the corresponding correlation functions in log-polar coordination.

$w_i = \exp[-(i - M/2)^2/S^2]$, then a Gaussian-distributed WSD is defined as follow:

$$\sigma_{\text{GWSD}}(j) = \sqrt{\sum_{i=1}^{k} \frac{\left[m_{\text{LP}}^{(F)}(i,j) - \mu(j)\right]^2}{\exp[(i - M/2)^2/S^2]}} \qquad (14)$$

where M is the height of $m_{\text{LP}}^{(F)}$, as shown in Fig.6; S is the standard deviation of the Gaussian-distribution, which can be adjusted according to noise intensity. Assuming that $S = M/4$, for example, there are less "spikes" in the WSD-vs-column plots compared with the STD-vs-column plots, and the curve shape is maintained, as shown in Fig. 8(b)(c). According to the definition of WSD, it not only exploits the potential of the mid-frequency components, but also some stronger high-frequency components that are not completely submerged by noise.

C. WEIGHTED COLUMN STANDARD DEVIATION BASED FILTER (WCSDF)

With the basis of WSD, we design a filter $w_S(i, j)$ called "weighted column standard deviation based filter (WCSDF)" by repeating the array of WSDs along row direction, as shown in Fig. 8(d), since $w_S(i, j)$ is invariant to i, it can be rewritten as $w_S(j)$. By replacing $M_{\text{LP}}^{(F)}$ and $M_{\text{LP}}^{(G)}$ in equation (7) with $M_{\text{WLP}}^{(F)}$ and $M_{\text{WLP}}^{(F)}$:

$$Q_{\text{WLP}}(\alpha, \beta) = \frac{\overline{M_{\text{WLP}}^{(F)}(\alpha, \beta)} \cdot M_{\text{WLP}}^{(G)}(\alpha, \beta)}{\left| \overline{M_{\text{WLP}}^{(F)}(\alpha, \beta)} \cdot M_{\text{WLP}}^{(G)}(\alpha, \beta) \right|} \qquad (15)$$

where $M_{\text{WLP}}^{(F)} = \mathcal{F}[m_{\text{LP}}^{(F)} \otimes w_S^{(F)}]$ and $M_{\text{WLP}}^{(G)} = \mathcal{F}[m_{\text{LP}}^{(G)} \otimes w_S^{(G)}]$; $w_S^{(F)}$ and $w_S^{(G)}$ are the WCSDF of reference and sensed images, respectively. $\otimes$ denotes element-wise multiplication.

Finally the filtered correlation function is obtained by:

$$C_{\mathrm{WLP}}(\lambda, \theta) = \mathcal{F}^{-1} \{Q_{\mathrm{WLP}}(\alpha, \beta)\} \qquad (16)$$

The first row in Fig.9 illustrates the WSD-vs-column plots of $m_{\mathrm{LP}}^{(G)}$ under different noise levels, from which we can conclude the WSD plots retained the shape of original curve compared with STD plots in Fig.7; the filtered log-polar Fourier amplitude spectrums of sensed images are shown in the middle row of Fig.9, and the last row of Fig.9 shows the correlation functions C_{WLP} between reference and sensed images of various noise levels. Compared with C_{LP} surf-plots in the last row of Fig.5, in the C_{WLP}, the noisy spikes are reduced and the indicator peak become obvious, especially under higher-level of noise. To compare the performance of C_{WLP} with C_{LP}, we perform the following steps:

1) A 1000-by-800 image is selected as a scene image;
2) the center 360-by-360 image of which is cropped as a reference image;
3) a series of 200 images with different parallax (by randomly introducing different shifts and rotates, as well as scales) is generated as sensed images, and a certain level of AWGN is added;
4) estimate the parallax between the sensed images and the reference image;
5) compare the estimated parallax with the actual value of parallax, the estimation accuracy is obtained.

The results are shown in Table 2 and Fig.10, which indicate that C_{WLP} has significant improvements in robustness to noise.

TABLE 2. Rotation and Scaling Estimation Accuracy (%) Comparisons Between C_{LP} and C_{WLP}.

Noi. Lev.	≤ 20	25	30	35	40
Original	100.0	98.5	82.5	50.0	23.0
Filtered	100.0	100.0	100.0	97.0	88.5

Noi. Lev.	45	50	55	60	≥ 65
Original	10.0	3.0	1.0	0.0	0.0
Filtered	73.0	42.0	0.0	0.0	0.0

D. THE FRAMEWORK OF REGISTRATION-BASED TDI

With the WCSDF, the improved FMT-based registration become more applicable in TDI. Herein, we propose a framework of registration-based TDI, which is composed of registration module and accumulating module. In the registration module, for any given reference image $f(x, y)$ and sensed image $g(x, y)$, it first measures their relative rotation $\Delta\theta_m$ and scale s by calculating their C_{WLP}, during which the quadratic interpolation to reach the accuracy of sub-pixel level. Then to eliminate the rotational and scaling shifts of the sensed image, and half-registered image $g'(x, y)$ is obtained. Further, the remaining translational shifts can be calculated by regular correlation function, and the registered sensed image $g''(x, y)$ is finally obtained.
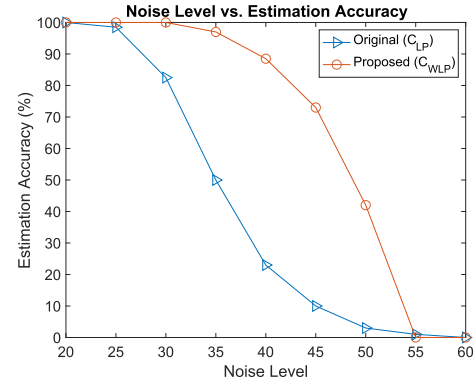
FIGURE 10. Plot of Noise Levels vs. Estimation Accuracy with C_{LP} and C_{WLP}.

The flowchart of registration module is shown in Fig.11, where the process of calculating the filtered FMT-based correlation function is included. In parallel to registration, the accumulating module is working to generate high-SNR TDI images. The processing steps are as follows:

1) While registering a short-exposure image sized $m \times n$, an all-ones matrix $\mathbf{1}_{m \times n}$ of the same size is initialized;
2) Transform the all-ones matrix according to transformation parameters determined from the first step, a coverage indicator matrix $I(x, y)$ is obtained, as shown in Fig. 12(d);
3) When the registration module registered a series of sensed images, accumulating them to generate an overlaid image $A_O(x, y)$(Fig. 12(f)). Meanwhile, coverage indicator matrixes are also accumulated together to obtain an accumulating times matrix $A_T(x, y)$ (Fig. 12(g)).
4) Perform element-wise division between the overlaid image and the accumulating times matrix, and rescale the intensity from 0 to $2^N - 1$, where N is the bit-depth.

After performing the above steps, a final TDI image of high-contrast and high SNR is finally generated, as shown in Fig. 12(h).

In summary, the diagram of the framework is shown in Fig. 13. and the pseudocode is presented in Algorithm 1.

IV. EXPERIMENTS

To verify our algorithm, the computer simulations are firstly conducted in section IV-A to determine the limitations of the improved FMT-based registration method. Finally, physical experiment is also carried out to verify the effectiveness of the proposed method and its improvements compared with other methods.

A. COMPUTER SIMULATIONS

To simulate the continuous acquisition of short-exposure captures by an airborne TDI-CMOS camera, in each experimental setting, a large image of 1000-by-800 is selected to simulate as a whole scene, the middle 360-by-360 sub-image are cropped as a reference image. On this basis, to simulate
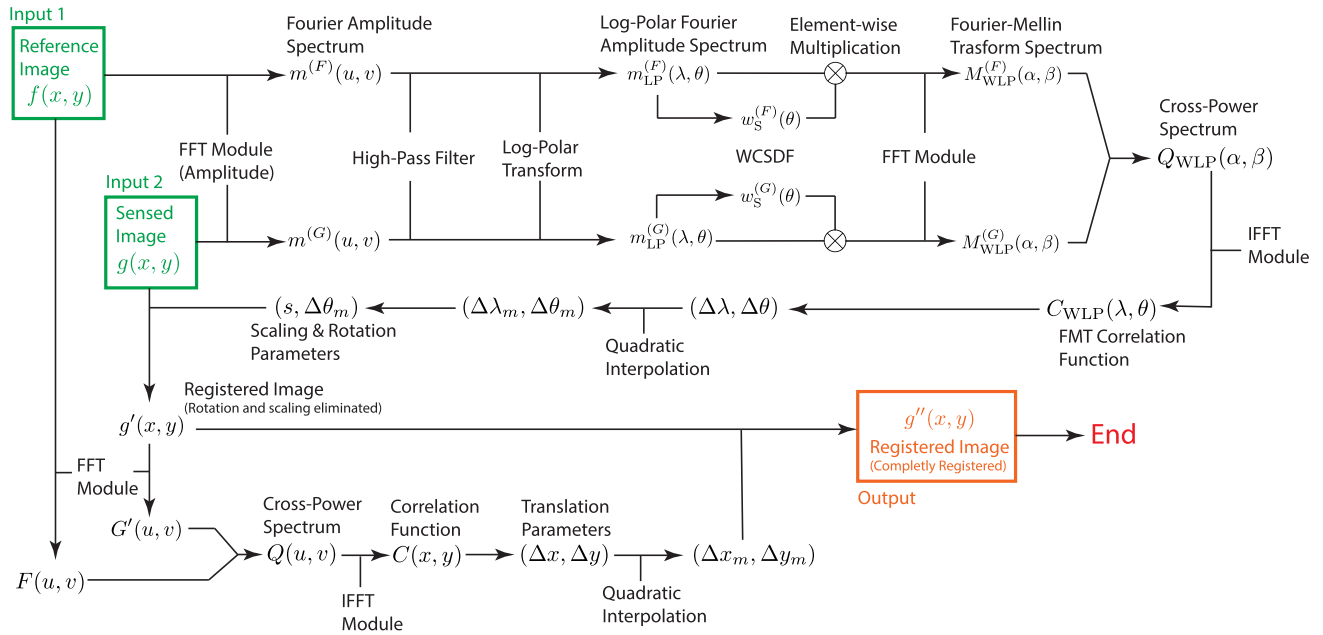
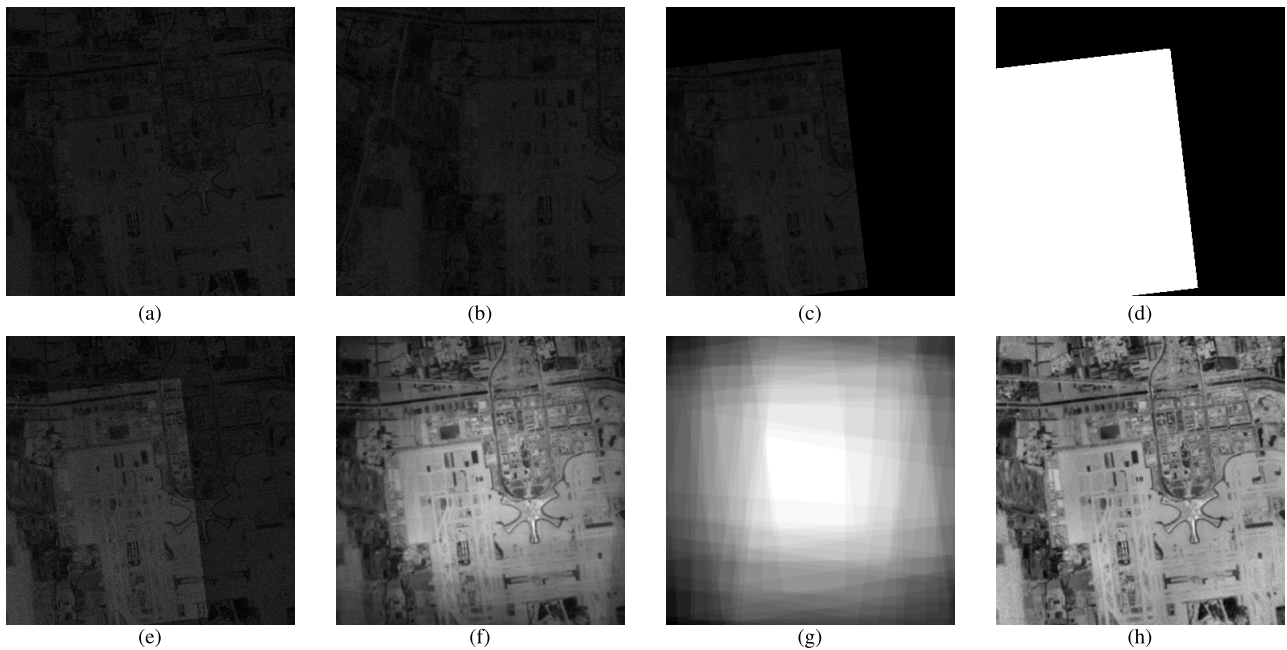**FIGURE 11.** Flow Chart of Registration Module.

FIGURE 12. Example of Registration and Accumulating Process: In a series of noisy dim images, an image (usually the first one) is chose as a reference image (a), the remains are regarded as sensed images to be registered, (b) shows the one of them. (c) shows an registered image from (b), and (d) the corresponding coverage indicator matrix; (e) overlaid image with (a) and (c); (f) an image that overlaid all registered images. (g) the accumulating times matrix; (h) a final TDI image.

the moving of cameras and the various imaging parameters, a series of sensed images with difference parallax is generated by randomly altering relative translation, rotation, and scaling, as well as adjusting their brightness and the AWGN level. The experiment proved that increasing the relative displacement alone has little effects on registration accuracy, as long as there exists an overlapping area between any two images, but the relative rotation and scaling are restricted within $\pm 40\circ$ and $0.7 \sim 1.4\times$ respectively to ensure registration accuracy.

Nevertheless, it still covers the range of the relative shifts between continuously acquired images in a short period, these restrictions still meet the application requirements.

B. PHYSICAL EXPERIMENT

To compare the performance of our method with the classical one, which is proposed by Tao *et al.* [8], a real experimental platform is built to conduct physical simulations (Fig.14), where a scene imageboard is used as a sample
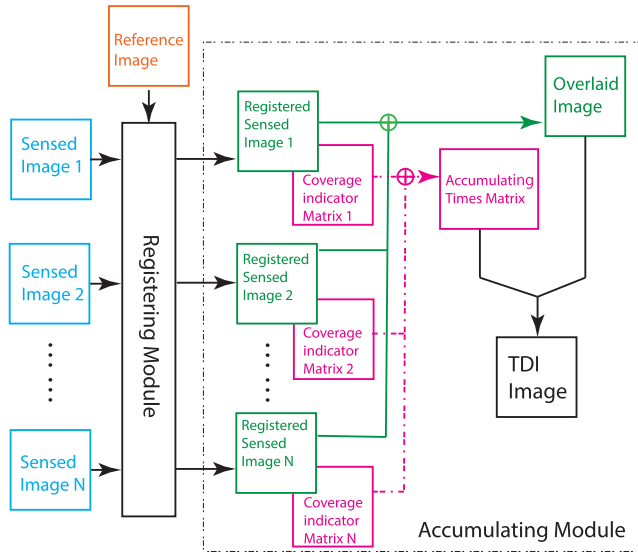
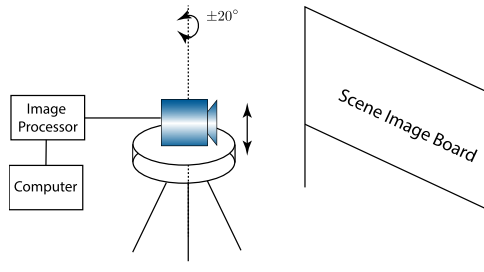**FIGURE 13.** Flow Chart of the Framework of Registration-based TDI.

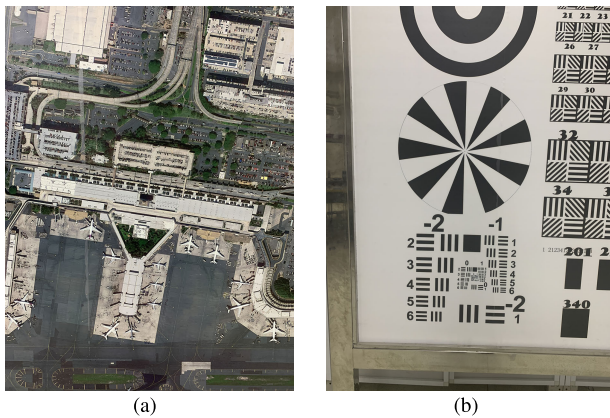**FIGURE 14.** Illustration of Experimental Platform.

(a) (b)

FIGURE 15. Experimental Images: (a)Physical Image Sample; (b)Resolution Chart.

**Algorithm 1** Accumulation of a Series of Short-Exposures

Require: A series of short-exposures $g_i(x, y)$, $(i = 0, 1, 2, \ldots, n)$
**Ensure:** A TDI image $T(x, y)$
1: $f(x, y) \leftarrow g_0(x, y)$;
2: $F(u, v) \leftarrow \mathcal{F}\{f(x, y)\}$;
3: $m^{(F)}(u, v) \leftarrow |F(u, v)|$;
4: Apply HPF and LPT on $m_{(F)}(u, v)$, obtain $m_{\mathrm{LP}}^{(F)}(u, v)$;
5: Obtain WCSDF $w_{\mathrm{S}}^{(F)}(\theta)$ from $m_{\mathrm{LP}}^{(F)}(u, v)$;
6: $M_{\mathrm{WLP}}^{(F)}(\alpha, \beta) \leftarrow \mathcal{F}\left\{w_{\mathrm{S}}^{(F)}(\theta) \otimes m_{\mathrm{LP}}^{(F)}(u, v)\right\}$;
7: Initialize all-ones image $\mathbf{1}_{m \times n}$;
8: $i \leftarrow 1$;
9: Initialize $A_T(x, y)$ and $A_O(x, y)$ with zeros;
10: **while** $i \leq n$ **do**
11:    $g(x, y) = g_i(x, y)$;
12:    $m^{(G)}(u, v) \leftarrow |\mathcal{F}\{g(x, y)\}|$;
13:    Apply HPF and LPT on $m_{(G)}(u, v)$, obtain $m_{\mathrm{LP}}^{(G)}(u, v)$;
14:    Obtain WCSDF $w_{\mathrm{S}}^{(G)}(\theta)$ from $m_{\mathrm{LP}}^{(G)}(u, v)$;
15:    $M_{\mathrm{WLP}}^{(G)}(\alpha, \beta) \leftarrow \mathcal{F}\left\{w_{\mathrm{S}}^{(G)}(\theta) \otimes m_{\mathrm{LP}}^{(G)}(u, v)\right\}$;
16:    $Q_{\mathrm{WLP}}(\alpha, \beta) \leftarrow \dfrac{M_{\mathrm{WLP}}^{(F)}(\alpha,\beta) \cdot M_{\mathrm{WLP}}^{(G)}(\alpha,\beta)}{\left|M_{\mathrm{WLP}}^{(F)}(\alpha,\beta) \cdot M_{\mathrm{WLP}}^{(G)}(\alpha,\beta)\right|}$
17:    $C_{\mathrm{WLP}}(\lambda, \theta) \leftarrow \mathcal{F}^{-1}\{Q_{\mathrm{WLP}}(\alpha, \beta)\}$
18:    Determine the relative shift $(\Delta\lambda, \Delta\theta)$ from $C_{\mathrm{WLP}}(\lambda, \theta)$;
19:    Improve the accuracy to sub-pixel level by quadratic interpolation, obtain $(\Delta\lambda_m, \Delta\theta_m)$
20:    $s \leftarrow \exp(\Delta\lambda_m)$
21:    Eliminate the relative rotation and scaling to obtain $g'(x, y)$;
22:    $G'(u, v) \leftarrow \mathcal{F}\{g'(x, y)\}$;
23:    $Q(u, v) \leftarrow \dfrac{F^*(u,v) \cdot G'(u,v)}{|F^*(u,v) \cdot G'(u,v)|}$;
24:    $C(\Delta x, \Delta y) \leftarrow \mathcal{F}^{-1}\{Q(u, v)\}$;
25:    Determine the relative shift $(\Delta x, \Delta y)$ from $C(\Delta x, \Delta y)$;
26:    Improve the accuracy to sub-pixel level by quadratic interpolation, obtain $(\Delta x_m, \Delta y_m)$
27:    Register $g'(x, y)$ to obtain $g''(x, y)$;
28:    Transform the all-ones images according to the transform parameters to generate coverage indicator matrix $I(x, y)$;
29:    $A_O(x, y) \leftarrow A_O(x, y) + g''(x, y)$;
30:    $A_T(x, y) \leftarrow A_T(x, y) + I(x, y)$;
31: **end while**
32: $T(x, y) \leftarrow A_O(x, y)/A_T(x, y)$

(shown in Fig.15(a)), and our method is implemented in an image processor which is connected to a CMOS camera installed on a three-dimensional turnable. To simulate the unpredictable random motion of an aerial camera (such as changing in roll or pitch), the turntable rotates at the speed of 6 mrad/s in the range of ±20° while the CMOS camera capturing images in a push-broom way. The examples of short-exposure captures are shown in Fig.16. By adjusting the imaging parameters (such as integrating stages L_{I} and exposure time (ET)), the generated TDI images are partly shown in Fig.17.

To quantify the quality of TDI images, we adopt four assessments to evaluate the TDI image quality: 1) Informational entropy; 2) Contrast; 3) Average gradient (AG); 4) Spatial frequency response (SFR).
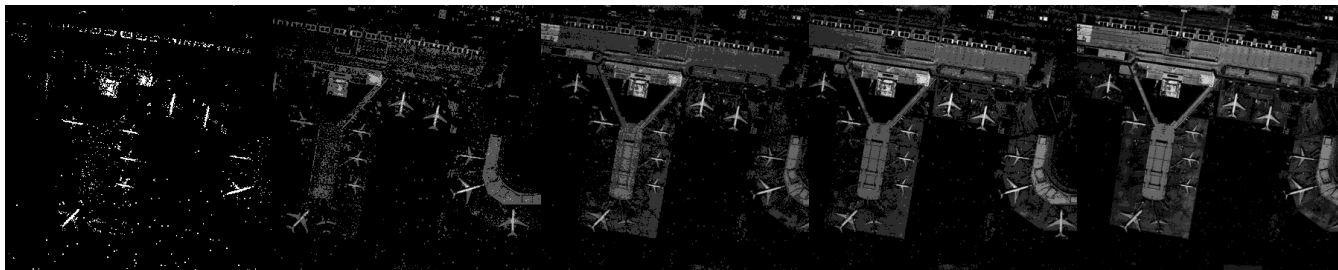
FIGURE 16. Short-exposure captures of different integrating time (exposure time): From left to right, the integrating times are: 0.4, 0.5, 0.6, 0.8, 1.0 ms. The brightness of all these images is rescaled to the range of 0 ∼ 1 for visibility.
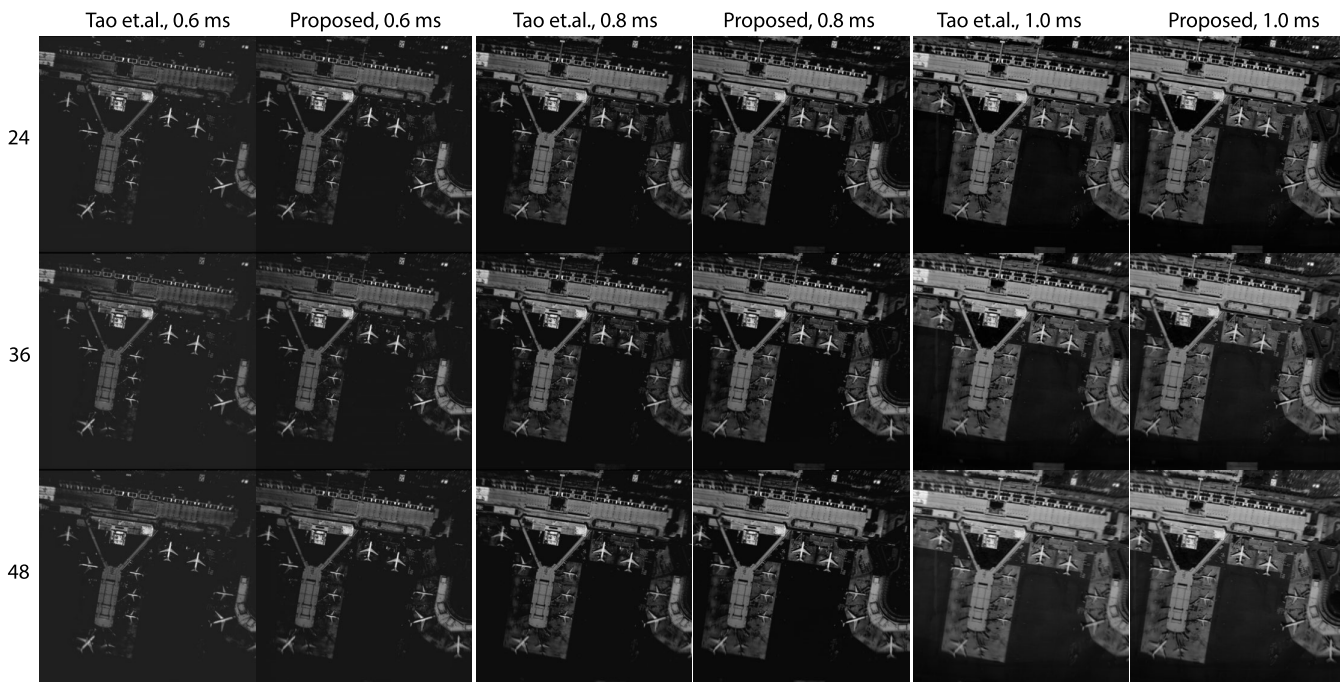

FIGURE 17. TDI images generated in physical simulations: from first row to last row: integrating stages of 24, 36, 48; from left column to right column: exposure time of 0.6, 0.8, 1.0 ms by two different methods (Tao *et al.* [8] and Ours). The brightness of all these images is rescaled to the range of 0 ∼ 1 for visibility.

1) INFORMATIONAL ENTROPY

Short-exposure captures have a relatively smaller size of quantification, as shown in Fig.16, when accumulating these captures to generate a TDI image, the size of quantification should be improved. We use informational entropy as a metric:

$$\text{Ent}(f) = -\sum_{i=0}^{2^N-1} p(i) \cdot \log_2 p(i) \qquad (17)$$

where $p(i)$ represents the probability of pixels with a gray level of i in an image f. The entropy of TDI images generated under different imaging parameters is shown in Table 3. The entropy of a TDI image would increase as L_I or ET increases before the brightness of TDI images is mostly saturated. When L_I or ET continues to increase, an oversaturated TDI image will be generated, resulting in a decreased entropy. In comparison, the proposed method outperformed the classical one (Tao *et al.*) [8] by an average of 0.4 in terms of entropy.
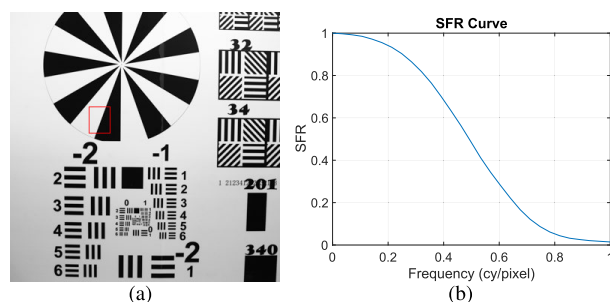
FIGURE 18. (a) A long-exposure capture of resolution chart, note that there is red box labeled on a slant-edge, (b) The SFR Curve calculated from the slant-edge.

2) CONTRAST

The contrast of a TDI image is also an important indicator of image quality. They are calculated and shown in Table.4, from which we can conclude that the contrast can be improved by increasing L_I or ET. However, when the L_I or ET is too

TABLE 3. Entropy of TDI images generated under different imaging parameters.

L_I	Method(ms)/ET	0.4	0.5	0.6	0.8	1.0	1.2	1.5	2.0
24	Tao et.al. [8]	3.463	4.132	4.202	5.154	5.461	5.382	6.208	6.575
	Proposed	3.561	4.172	4.680	5.166	5.738	6.060	6.385	6.806
36	Tao et.al. [8]	3.923	4.491	4.818	5.027	6.184	6.087	6.756	7.018
	Proposed	4.184	4.785	5.294	5.758	6.317	6.639	6.968	7.359
48	Tao et.al. [8]	4.051	4.843	5.499	6.046	6.446	7.028	7.178	7.227
	Proposed	4.653	5.242	5.717	6.163	6.728	7.053	7.366	7.335
96	Tao et.al. [8]	4.760	5.783	6.403	7.096	7.459	6.797	6.440	5.597
	Proposed	5.369	6.079	6.660	7.142	7.547	7.125	6.926	6.093

TABLE 4. Contrast of TDI images generated under different imaging parameters.

L_I	Method(ms)/ET	0.4	0.5	0.6	0.8	1.0	1.2	1.5	2.0
24	Tao et.al. [8]	0.0260	0.0360	0.0430	0.0570	0.0750	0.0979	0.1160	0.1589
	Proposed	0.0258	0.0359	0.0429	0.0571	0.0748	0.0980	0.1157	0.1594
36	Tao et.al. [8]	0.0390	0.0540	0.0640	0.0850	0.1119	0.1459	0.1730	0.2366
	Proposed	0.0388	0.0538	0.0641	0.0853	0.1115	0.1463	0.1725	0.2365
48	Tao et.al. [8]	0.0510	0.0710	0.0850	0.1129	0.1479	0.1937	0.2275	0.2955
	Proposed	0.0514	0.0714	0.0849	0.1131	0.1479	0.1943	0.2284	0.2959
96	Tao et.al. [8]	0.1009	0.1400	0.1658	0.2208	0.2835	0.3197	0.3233	0.3263
	Proposed	0.1006	0.1398	0.1663	0.2209	0.2835	0.3202	0.3242	0.3271

TABLE 5. AG of TDI images generated under different imaging parameters.

L_I	Method(ms)/ET	0.4	0.5	0.6	0.8	1.0	1.2	1.5	2.0
24	Tao et.al.	3.3315	4.9210	5.6659	6.4291	6.9184	7.0285	7.3366	7.4450
	Proposed	3.3473	4.9416	5.7016	6.4709	6.9198	7.0323	7.3686	7.4563
36	Tao et.al.	3.2926	4.8545	5.5804	6.3879	6.8907	6.8983	7.3139	7.3752
	Proposed	3.3255	4.9001	5.6312	6.4430	6.9203	6.9656	7.3364	7.4394
48	Tao et.al.	3.3336	4.8208	5.5255	6.3702	6.8539	6.8805	7.2177	7.2879
	Proposed	3.3358	4.8557	5.5670	6.4276	6.8781	6.9068	7.2469	7.3421
96	Tao et.al.	3.3776	4.7263	5.4171	6.2187	6.6183	6.5908	7.0167	7.1230
	Proposed	3.3824	4.7277	5.4224	6.2797	6.6821	6.6335	7.0861	7.1399

large, the contrast will also decrease due to the oversaturated TDI image, and there is no significant improvement in comparison.

3) AVERAGE GRADIENT

The average gradient (AG) can sensitively reflect the ability of the image to express the contrast of small details. It include the sum of squared row gradient (SSRG) and the sum of squared column gradient (SSCG):

$$\text{SSRG}(f) = \sum_{i=0}^{M-1}\sum_{j=1}^{N-1}[f(i,j)-f(i,j-1)]^2 \quad (18)$$

$$\text{SSCG}(f) = \sum_{i=1}^{M-1}\sum_{j=0}^{N-1}[f(i,j)-f(i-1,j)]^2 \quad (19)$$

Then, we define AG:

$$\text{AG}(f) = \sqrt{\frac{1}{MN}[\text{SSRG}(f)+\text{SSCG}(f)]} \quad (20)$$

The AG of TDI images are calculated and compared in Table 5. As the ET increases, the brightness of the image increases, which leads to more details, so the average gradient also increases. However, the increase of the L_I will cause more motion blur, the AG will decrease. In comparison, the proposed method outperformed the classical one (Tao *et al.*) [8] by an average of 0.05 in terms of AG.

4) SPATIAL FREQUENCY RESPONSE (SFR)

The SFR is an indicator of image clarity, which can be calculated based on the slant-edge analysis [29]. The SFR may decrease as the ET or L_I increases, therefore it is
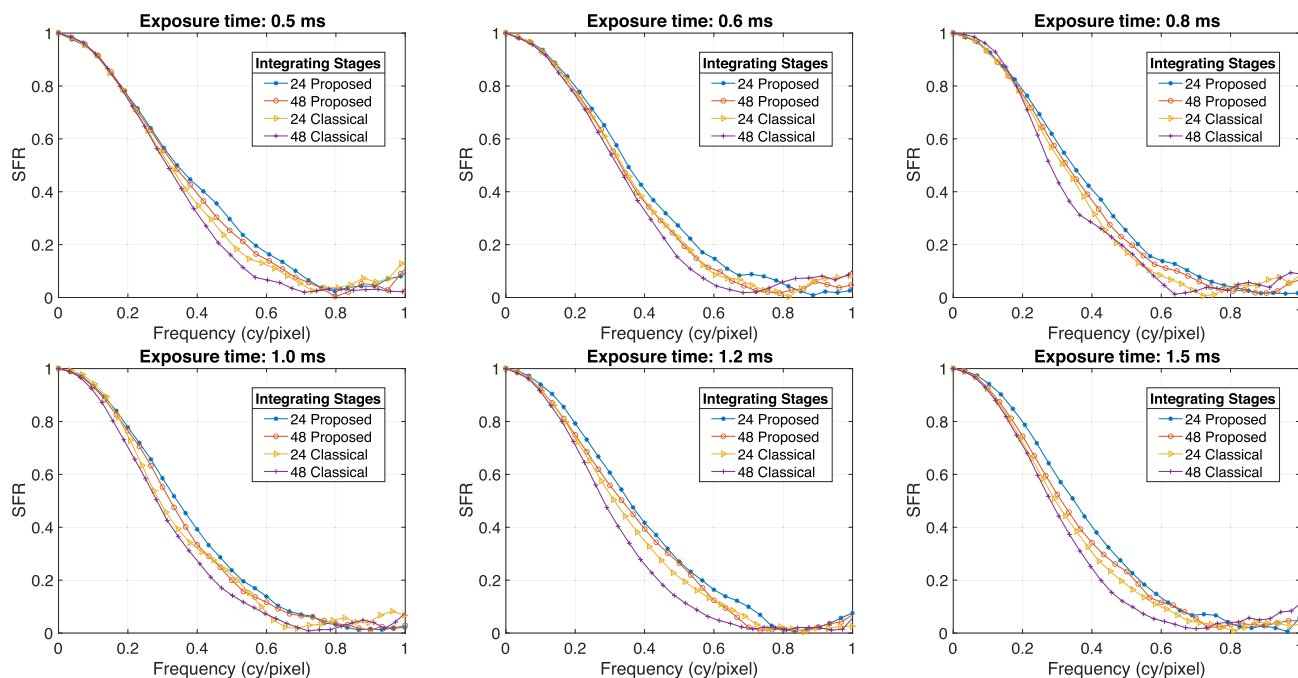
FIGURE 19. SFR Curves of the selected slant-edge in TDI images of resolution chart under different imaging parameters.

necessary to compare the final SFR curve under different imaging parameters. We choose a resolution chart as shown in Fig.15(b). To determine the amount that the clarity decreases, a long-exposure capture is firstly obtained to measure the original SFR of the camera, as shown in Fig.18(a): a proper slant-edge is selected (the red box in Fig.18(a)) for calculating the static SFR curve, as shown in Fig.18(b). Then, several TDI images of the resolution chart is generated under different imaging parameters (exposure time and integrating stages), and theirs corresponding SFR curves are calculated from the same slant-edge, as shown in Fig.19. It can be seen that as L_I increase, the TDI image clarity may be reduced due to (1) the accumulation of registration errors, and (2) the larger range of rolling angle. In comparisons between two different methods, we can conclude that the proposed method has superior performances to the method of Tao *et al.* [8].

V. CONCLUSION

This paper presented a novel framework to generate TDI images. While a CMOS camera acquiring a series of short-exposure captures, the proposed method first calculates the relative shift among these captures, followed by registering and accumulating, a TDI image is finally generated. In the registration module, it adopts cross-power spectrums and correlation functions—both in log-polar and regular—to calculate the relative transform parameters, wherein a WCSDF filter is proposed to reduce the influence of noise in calculating the rotational and scaling shifts among images, which improved the noise robustness (noise level up to around 40). The computational simulations show that the proposed method is applicable when the relative shift is

within $\pm 40°$ in rotation, and $0.7 \sim 1.4\times$ in scaling, as well as a large range of translational shift—as long as there exists a large overlap between any two captures. In reality, while an airborne camera acquiring images, the CMOS detectors may have rotational motion (such as yaw, roll and pitch). However, the ability of de-rotation can only correct the relative shift induced by yaw, but unable to correct the shift induced by roll or pitch, resulting in the remaining registration errors uncorrected, and in turn making the TDI images blurred. Despite that, the physical experiments show that the proposed method can still generate TDI images with desirable clarity. In our physical experiments, we use fluorescent lamps to illuminate the scene image board, since its illumination level is much smaller than that of the sun, the exposure time has to set much longer (0.4 ms $\sim$ 2 ms), resulting in significant image motion during each exposure, causing the TDI image blurred. In sunlight, however, the exposure time need not be set that long, and its corresponding image motion will be much smaller, in that case, the quality of TDI images will be much better.

In summary, the algorithm proposed in this paper enables TDI-CMOS to correct more generic image shifts in the digital domain and perform signal accumulation while increasing some computational complexity. This method directly performs image shift and registration between images, without additional calculation of the image shift speed, and no additional mechanism for compensation, which further promotes the miniaturization of TDI-CMOS. Since it cannot correct the shifts induced by the roll or pitch of the camera, our future works will be concentrated on proposing a novel method that can compensate for more general distortion.

REFERENCES

[1] X. Kong, X. Bu, C. Mao, L. Zhang, H. Ma, and F. Yan, "SPAD sensors with 256×2 linear array for time delay integration demonstration," in *Proc. IEEE SENSORS*, New Delhi, India, Oct. 2018, pp. 1386–1389.

[2] X. Kong, C. Mao, X. Hu, and F. Yan, "Time-delay-integration imaging implemented with single-photon-avalanche-diode linear array," *IEEE Sensors J.*, vol. 21, no. 5, pp. 6012–6023, Mar. 2021.

[3] P.-Y. Huang, W.-Y. Lo, Y.-K. Huang, S.-Y. Lai, J. Ling, D.-C. Chang, and M.-Y. Yeh, "Dynamic image acquisition and verification for a 32-stages time delay and integration CMOS image sensor," in *Sensors, Systems, and Next-Generation Satellites XXII*, vol. 10785, S. P. Neeck, P. Martimort, and T. Kimura, Eds. Bellingham, WA, USA: SPIE, 2018, pp. 153–159.

[4] M. De Stefano, R. Balachandran, and C. Secchi, "A passivity-based approach for simulating satellite dynamics with robots: Discrete-time integration and time-delay compensation," *IEEE Trans. Robot.*, vol. 36, no. 1, pp. 189–203, Feb. 2020.

[5] W. Qiu and C. Xu, "Attitude maneuver planning of agile satellites for time delay integration imaging," *J. Guid., Control, Dyn.*, vol. 43, no. 1, pp. 46–59, Jan. 2020.

[6] K. Nie, S. Yao, J. Xu, J. Gao, and Y. Xia, "A 128-stage analog accumulator for CMOS TDI image sensor," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 7, pp. 1952–1961, Jul. 2014.

[7] K. Nie, J. Xu, and Z. Gao, "A 128-stage CMOS TDI image sensor with on-chip digital accumulator," *IEEE Sensors J.*, vol. 16, no. 5, pp. 1319–1324, Mar. 2016.

[8] S. Tao, X. Zhang, W. Xu, and H. Qu, "Realize the image motion self-registration based on TDI in digital domain," *IEEE Sensors J.*, vol. 19, no. 23, pp. 11666–11674, May 2019.

[9] C. Yin, T. Liao, K.-L. Liu, C.-C. Kao, C.-F. Chiu, and C.-C. Hsieh, "A 32-stage 15-b digital time-delay integration linear CMOS image sensor with data prediction switching technique," *IEEE Trans. Electron Devices*, vol. 64, no. 3, pp. 1167–1173, Mar. 2017.

[10] H. Karagálle, L. Malgaca, and H. F. Öktem, "Analysis of active vibration control in smart structures by ANSYS," *Smart Mater. Struct.*, vol. 13, no. 4, pp. 661–667, Aug. 2004.

[11] J. T. Bosiers, A. C. Kleimann, H. C. van Kuijk, L. Le Cam, H. L. Peek, J. P. Maas, and A. J. P. Theuwissen, "Frame transfer CCDs for digital still cameras: Concept, design, and evaluation," *IEEE Trans. Electron Devices*, vol. 49, no. 3, pp. 377–386, Mar. 2002.

[12] G. G. Olson, "Image motion compensation with frame transfer CCDs," in *Machine Vision and Three-Dimensional Imaging Systems for Inspection and Metrology II*, vol. 4567, K. G. Harding and J. W. V. Miller, Eds. Bellingham, WA, USA: SPIE, 2002, pp. 153–160.

[13] J. Xu, "Digital domain dynamic path accumulation method to compensate for image vibration distortion for CMOS-time-delay-integration image sensor," *Opt. Eng.*, vol. 59, no. 10, Oct. 2020.

[14] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[15] M. Petrou, "Image registration: An overview," *Advances in Imaging and Electron Physic*, vol. 130. Amsterdam, The Netherlands: Elsevier, 2004, pp. 243–291.

[16] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, 1999, pp. 1150–1157.

[17] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," *Revue Praticien-Medecine Generale*, vol. 13, no. 465, pp. 44–45, 1999.

[18] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[19] Y. Fu, Y. Lei, T. Wang, W. J. Curran, T. Liu, and X. Yang, "Deep learning in medical image registration: A review," *Phys. Med. Biol.*, vol. 65, no. 20, Oct. 2020, Art. no. 20TR01.

[20] H. R. Boveiri, R. Khayami, R. Javidan, and A. Mehdizadeh, "Medical image registration using deep neural networks: A comprehensive review," *Comput. Electr. Eng.*, vol. 87, Oct. 2020, Art. no. 106767.

[21] X. Tong, K. Luan, U. Stilla, Z. Ye, Y. Xu, S. Gao, H. Xie, Q. Du, S. Liu, X. Xu, and S. Liu, "Image registration with Fourier-based image correlation: A comprehensive review of developments and applications," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 10, pp. 4062–4081, Oct. 2019.

[22] E. De Castro and C. Morandi, "Registration of translated and rotated images using finite Fourier transforms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vols. PAMI–9, no. 5, pp. 700–703, Sep. 1987.

[23] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Trans. Image Process.*, vol. 5, no. 8, pp. 1266–1271, Aug. 1996.

[24] H. S. Stone, B. Tao, and M. McGuire, "Analysis of image registration noise due to rotationally dependent aliasing," *J. Vis. Commun. Image Represent.*, vol. 14, no. 2, pp. 114–135, Jun. 2003.

[25] Y. Keller, A. Averbuch, and M. Israeli, "Pseudopolar-based estimation of large translations, rotations, and scalings in images," *IEEE Trans. Image Process.*, vol. 14, no. 1, pp. 12–22, Jan. 2005.

[26] S. A. Abbas, Q. Sun, and H. Foroosh, "An exact and fast computation of discrete Fourier transform for polar and spherical grid," *IEEE Trans. Signal Process.*, vol. 65, no. 8, pp. 2033–2048, Apr. 2017.

[27] Y. Dong, W. Jiao, T. Long, G. He, and C. Gong, "An extension of phase correlation-based image registration to estimate similarity transform using multiple polar Fourier transform," *Remote Sens.*, vol. 10, no. 11, p. 1719, Oct. 2018.

[28] T. Fujisawa and M. Ikehara, "High-accuracy image rotation and scale estimation using radon transform and sub-pixel shift estimation," *IEEE Access*, vol. 7, pp. 22719–22728, 2019.

[29] P. D. Burns and D. Williams, "Refined slanted-edge measurement for practical camera and scanner testing," in *Proc. Soc. Imag. Sci. Technol., Image Process., Image Qual., Image Capture, Syst. Conf.*, 2002, pp. 1–8.
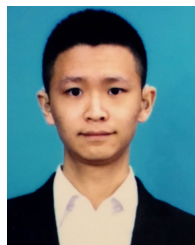
QINPING FENG received the B.S. degree from the Changchun University of Science and Technology, in 2016. He is currently pursuing the joint Ph.D. degree with the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences and the University of Chinese Academy of Sciences. His research interests include digital image processing, remote sensing, and machine learning.
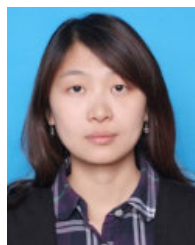
SHUPING TAO received the B.S. degree in communication engineering from Sichuan University, in 2008, and the Ph.D. degree in optics engineering from the University of Chinese Academy of Sciences, in 2013. Since 2013, she has been with the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, China, where she is currently an Associate Research Fellow. Her research interests include digital circuit design and image signal processing.

CHUNYU LIU received the B.S. degree from Jilin University, in 2005, and the Ph.D. degree from the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, in 2011. She is currently a Ph.D. Supervisor and a Researcher with the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, mainly engaged in the research of optical system design and overall design of optoelectronic systems.

HONGSONG QU received the Ph.D. degree from the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, in 2008. He is currently a Professor with the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His research interests include space photoelectric imaging and its application.

● ● ●