Registration method for infrared images under conditions of fixed-pattern noise

Chao Zuo *, Qian Chen, Guohua Gu, Xiubao Sui

National Defense Key Laboratory of Optoelectronic Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

ABSTRACT

This paper proposes a new registration method for infrared images under conditions of fixed-pattern noise (FPN). Conventional registration techniques are susceptible to FPN and it is therefore very desirable to have a registration algorithm that is tolerant to FPN. For this purpose, we utilize the difference of the cross-power spectrum of two discrete shifted images to suppress the noise power spectrum while the shifts information is well preserved. In particular, we show that the phase of the cross-power spectrum difference is a periodic two-dimensional binary stripe signal with the exact shifts determined to subpixel accuracy by the number of periods of the phase difference along each frequency axis. Robust estimates of shifts can be obtained by transforming its discontinuities to Hough domain. Experimental results show that the proposed method exhibits robust and accurate registration performance even for the noisy images that could not be handled by conventional registration algorithms. We have also incorporated this technique to a registration-based nonuniformity correction (NUC) framework, indicating that our registration technique is able to estimate motion parameters reliably, leading to satisfactory NUC result.

Crown Copyright © 2012 Published by Elsevier B.V. All rights reserved.

1. Introduction

Image registration is the process of establishing point-to-point correspondence between two images of the same scene [1,2]. This process is commonly used in infrared imaging systems, such as electronic stabilization of sensors, image fusion, multi-frame super-resolution and nonuniformity correction in infrared focal-plane arrays (IRFPA) [3–9]. In many situations, the performance of algorithms that employ image registration information depends heavily on the accuracy of the shift estimates. However, little attention has been devoted to the registration methods for infrared images in which fixed-pattern noise (FPN) rather than temporal noise is the dominant noise component [4,5]. FPN is perceived as a superimposed pattern that is approximately constant for all image frames. Almost all registration methods require the assumption that the noise in observed images is both spatially and temporally independently distributed [1,2]. The precondition, however, is usually impractical for infrared images especially when they are deteriorated by FPN. Under these circumstances, the performance of these techniques will degrade significantly.

FPN can be reduced by calibrating the sensor by means of imaging target scenes with uniform intensities [10]. It can also be reduced from sequences of video by post-processing algorithms, i.e., scene-based nonuniformity corrections (NUC) [6–9]. However, some of these methods inherently rely on accurate image registration of the raw video. These registration-based techniques can recover the clean scene by analyzing

a small number of image frames, using the idea that each detector should have an identical response when observing the same scene point over time. Clearly, the performance of these registration-based nonuniformity methods is often sensitive to the accuracy of registration. If the level of the FPN is high relative to the true scene, then a registration algorithm may mistakenly attempt to register the fixed pattern in the FPN-dominated image while ignoring the motion in the salient true scene. Considering this contradiction, it is therefore very desirable to have a registration algorithm that is tolerant to FPN.

There are many existing shift estimation algorithms for motion estimation between frames. The two-dimensional (2-D) cross-correlation is one of the most commonly used techniques [11]. To improve the performance of these traditional registration techniques, a simple idea is to first blurring each image using spatial low-pass filter, then the registration is performed on the obtained images. This scheme can perform well when the level of FPN is low to moderate. When the level of FPN is relatively high or the FPN mainly exists in the low-spatial frequency domain, the effect of the spatial low-pass filter is very limited. Cain et al. [5] employed a projection-based registration technique which projects each two-dimensional image into two one-dimensional vectors through averaging each image across its rows and columns, resulting in vertical and horizontal image vectors, respectively. It achieves improved performance over the traditional 2-D cross-correlation based techniques in the presence of FPN due to the inherent averaging in the projection. But this method assumes the values of FPN to be independent and identically distributed random variables. Unfortunately, real noise patterns of IRFPA show greatly spatially correlations (such as grids and stripes) [12,13] and they do not fit well with this spatially unstructured model. The real structured FPN may not be well cancelled through averaging

* Corresponding author.
  E-mail address: surpasszuo@163.com (C. Zuo).

each rows and columns, thereby leading to significant performance deterioration.

In view of this, we propose a new technique to address the problem of subpixel registration for infrared images under conditions of FPN. We are interested in investigating the information contained in the Fourier domain and the proposed method is motivated by the observation that the auto-power spectrum of FPN can be cancelled in the difference of the cross-power spectrum of two discrete shifted images while the shift information is well preserved. By deriving an exact model for the phase of the cross-power spectrum difference, we show that this phase matrix is a two-dimensional binary tripe signal and the period of the square signal along each axis determines the shifts along corresponding axis. This in particular leads to a simple solution directly from the phase-difference matrix through transforming its discontinuities to Hough domain. The performance of the proposed method is evaluated with infrared images with both simulated and real structured FPN, showing significantly precise and reliable translation estimates even under a high level of FPN.

The rest of this paper is organized as follows. In Section 2, the problem of image registration in the presence of FPN is formulated. In Section 3, the proposed method is explained and developed. In Section 4, experiential results are given In Section 5, conclusions are drawn.

2. Problem formulation

Let s be the original image, and $f_i(i=1,2)$ be two images that are shifted versions of s:

$$f_i(x,y) = s\left(x + \delta_{x,i}, y + \delta_{y,i}\right), i = 1,2 \tag{1}$$

where $(\delta_x, \delta_y) = (\delta_{x,2} - \delta_{x,1}, \delta_{y,2} - \delta_{y,1})$ is the relative translations between the image pair. In the absence of noise and aliasing, the shift property of Fourier transform gives:

$$\hat{f}_1(u,v) = \hat{f}_2(u,v)e^{-j(u\delta_x + v\delta_y)} \tag{2}$$

where $\hat{f}_i$ is the Fourier transform of f_i and (u,v) are the Fourier domain coordinates. To identify (δ_x, δ_y), one computes a normalized cross-power spectrum between $\hat{f}_1$ and $\hat{f}_2$

$$\hat{c}(u,v) = \frac{\hat{f}_1(u,v)\hat{f}_2(u,v)^*}{\left|\hat{f}_1(u,v)\hat{f}_2(u,v)^*\right|} = e^{-j(u\delta_x + v\delta_y)} \tag{3}$$

where the hat sign as usual indicates the Fourier transform and the asterisk stands for the complex conjugate. Once computed, the approach cited in the literature [14] is to compute the inverse Fourier transform of $\hat{c}(u,v)$ and a Dirac delta function can be recognized as an intensity peak, which can be found by simply scanning for the global maximum value. The coordinate of this peak corresponds directly to the translation vector (δ_x, δ_y).

In most infrared imaging applications, noise exists in the captured images up to a certain level. In such conditions, the images $f_i(i=1,2)$ in Eq. (1) should be remodeled as

$$f_1(x,y) = s\left(x + \delta_{x,1}, y + \delta_{y,1}\right) + o(x,y) + n_1(x,y) \tag{4}$$

and

$$f_2(x,y) = s\left(x + \delta_{x,2}, y + \delta_{y,2}\right) + o(x,y) + n_2(x,y). \tag{5}$$

where $o(x,y)$ stands for the FPN which is assumed be fixed between two observed images and signal independent. Note that we do not take the assumption that the FPN is spatially independent. This is in agreement with most observations that FPN are indeed spatially

structured distributed [12,13]. The term n_1 and n_2 correspond to the additive temporal noise, which are assumed to be mutually independent. The FPN and additive temporal noise are also assumed mutually independent. It is noted that these conditions are valid in general applications. According to this model, the cross-power spectrum can be expressed as:

$$\begin{aligned}
S_{f_1f_2}(u,v) &= \hat{f}_1(u,v)\hat{f}_2(u,v)^* = \hat{s}(u,v)\hat{s}^*(u,v)e^{-j(u\delta_x + v\delta_y)} \\
&+ \hat{s}(u,v)\hat{o}^*(u,v)e^{-j(u\delta_{x,1} + v\delta_{y,1})} + \hat{s}(u,v)^*\hat{o}(u,v)e^{j(u\delta_{x,2} + v\delta_{y,2})} \\
&+ \hat{s}(u,v)^*\hat{n}_2(u,v)e^{j(u\delta_{x,1} + v\delta_{y,1})} + \hat{s}(u,v)\hat{n}_2^*(u,v)e^{j(u\delta_{x,2} + v\delta_{y,2})} \\
&+ \hat{n}_1(u,v)\hat{o}^*(u,v) + \hat{n}_2^*(u,v)\hat{o}(u,v) + \hat{o}(u,v)\hat{o}^*(u,v) \\
&+ \hat{n}_1(u,v)\hat{n}_2^*(u,v) = \hat{s}(u,v)\hat{s}^*(u,v)e^{-j(u\delta_x + v\delta_y)} + \hat{o}(u,v)\hat{o}^*(u,v).
\end{aligned} \tag{6}$$

Whitening the magnitude is normalized to unity for all frequencies of the cross-power spectrum. In this case, the image data is not cancelled out in the normalized (whiten) cross-power spectrum

$$\hat{c}(u,v) = \frac{S_{f_1f_2}(u,v)}{\left|S_{f_1f_2}(u,v)\right|} = \frac{\hat{s}(u,v)\hat{s}^*(u,v)e^{-j(u\delta_x + v\delta_y)} + \hat{o}(u,v)\hat{o}^*(u,v)}{\left|\hat{s}(u,v)\hat{s}^*(u,v)e^{-j(u\delta_x + v\delta_y)} + \hat{o}(u,v)\hat{o}^*(u,v)\right|}. \tag{7}$$

In such circumstances, inverse Fourier transform of $\hat{c}(u,v)$ is no longer a Dirac delta function (Dirichlet kernel in the discrete case
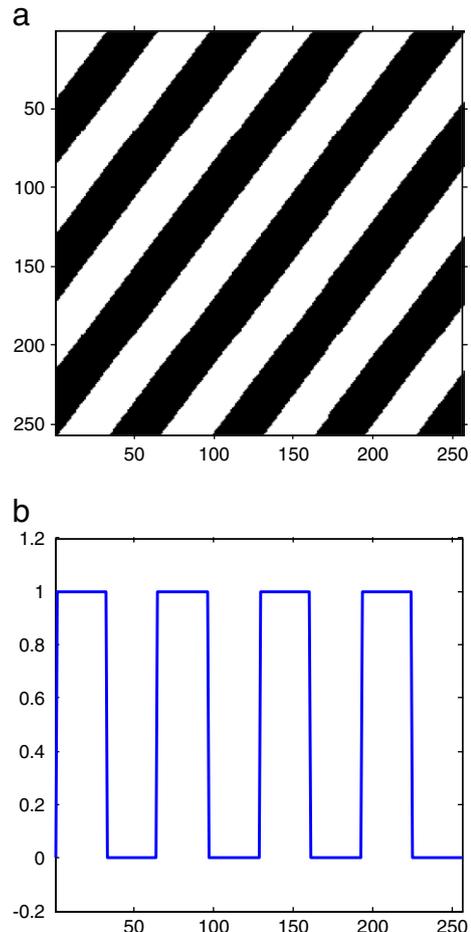

Fig. 1. (a) Phase matrix of discrete cross-power spectrum difference corresponding to shifts of (4,3) pixels, (b) one row of the phase-difference matrices shown in (a).
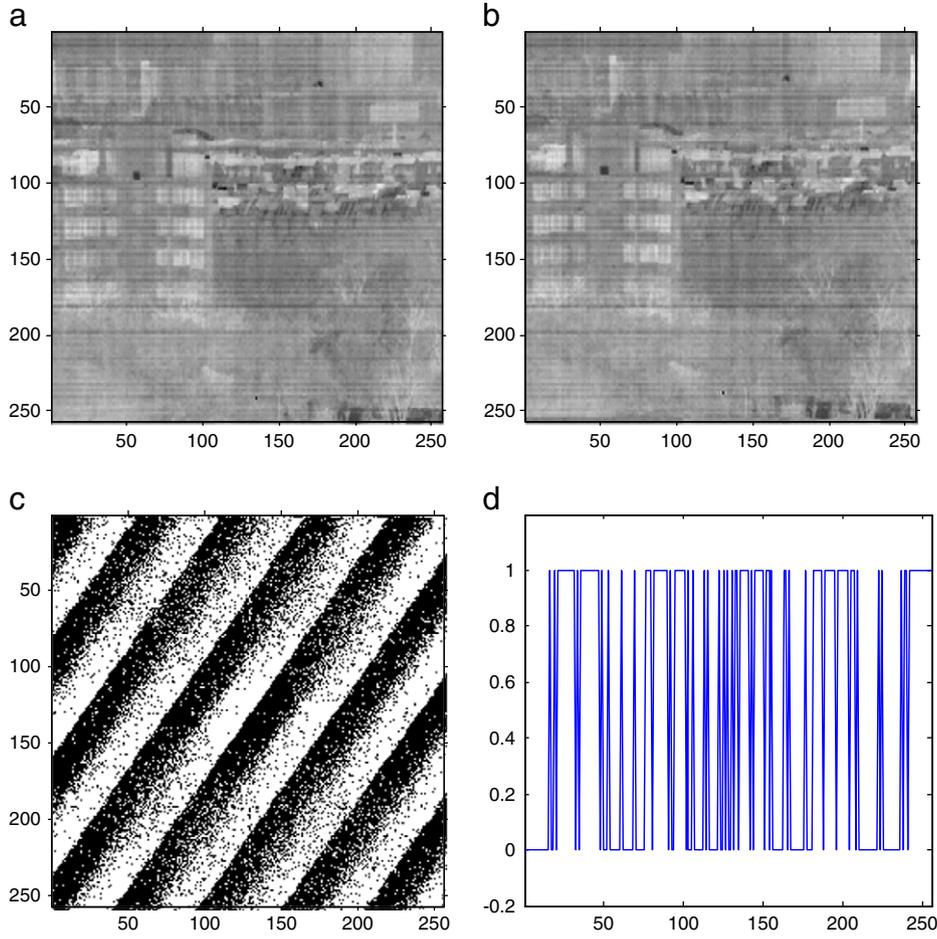
Fig. 2. (a), (b) Two noisy infrared images with shifts of (4.3, 3.2) (c) noisy stripe phase matrix of discrete cross-power spectrum difference between (a) and (b) (d) one row of the phase matrix of discrete cross-power spectrum difference shown in (c).

[15]) but two peaks with magnitudes related to the corresponding frequency content. Therefore, solving in the spatial domain or the 2D fitting method [15,16] is impractical. Other subpixel phase-correlation methods which work directly in the Fourier domain [17–19] are also inapplicable, since it requires fitting a plane to noisy phase-difference data. In view of this, we develop a new algorithm which removes the effect of the auto-power spectrum of FPN to address this difficulty.

3. Registration method for infrared images under conditions of fixed-pattern noise

From the above discuss we know that due to the presence of auto-power spectrum of FPN, it is difficult to obtain accurate displacement estimates only from the cross-power spectrum of two images. We hope that the noise power spectrum can be removed as much as possible, so as to obtain robust estimates for real translations. So first we calculate the conjugate of cross-power spectrum between $\hat{f}_1$ and $\hat{f}_2$

$$S_{f_2 f_1}(u,v) = S_{f_1 f_2}^*(u,v) = \hat{s}(u,v)\hat{s}^*(u,v)e^{j(u\delta_x + v\delta_y)} + \hat{o}(u,v)\hat{o}^*(u,v). \quad (8)$$

A subtraction between the two cross-power spectrums is performed to eliminate the FPN term

$$\begin{aligned}\Delta S_{ff}(u,v) &= S_{f_1 f_2}(u,v) - S_{f_2 f_1}(u,v) \\ &= \hat{s}(u,v)\hat{s}^*(u,v)e^{-j(u\delta_x + v\delta_y)} - \hat{s}(u,v)\hat{s}^*(u,v)e^{j(u\delta_x + v\delta_y)} \\ &= -2j\hat{s}(u,v)\hat{s}^*(u,v)\sin\left(u\delta_x + v\delta_y\right).\end{aligned} \quad (9)$$

Because the signal's auto-power spectrum is real and non-negative, so the difference of cross-power spectrum between the two images is pure imaginary, and its phase is given by

$$\hat{\varphi}(u,v) = \angle\Delta S_{ff}(u,v) = \begin{cases} \pi/2 & when \quad \text{sign}\left(\Delta S_{ff}(u,v)\right) = 1 \\ -\pi/2 & when \quad \text{sign}\left(\Delta S_{ff}(u,v)\right) = -1 \end{cases} \quad (10)$$

where the sgn($\cdot$) is the signum function. The above discussion is limited to the continuous case. When dealing with digital images, f_1 and f_2 are specified only in finite size discretized arrays. However, replacing the Fourier transform by the discrete Fourier transform (DFT), and also assuming a periodic extension of the images outside their compact support. From the definition of the DFT based on Fourier series, it follows immediately upon substituting $u = \frac{2\pi k}{N}$ and $v = \frac{2\pi l}{M}$ that the difference of the two discrete cross-power spectrum becomes

$$\Delta S_{ff}(k,l) = -2j\hat{s}(k,l)\hat{s}^*(k,l)\sin\left(\frac{2\pi k}{N}\delta_x + \frac{2\pi l}{M}\delta_y\right), \quad (11)$$

where $k = 0,..., M-1$ and $l = 0,..., N-1$. After some manipulations we obtain

$$\hat{\varphi}(k,l) = \angle\Delta S_{ff}(k,l) = \begin{cases} -\pi/2 & \frac{k}{N}\delta_x + \frac{l}{M}\delta_y \in \left[K, \frac{(2K+1)}{2}\right), \quad K \in Z \\ \pi/2 & others \end{cases} \quad (12)$$

Note if we denote $^\pi/_2$ as white (one) and denote $-^\pi/_2$ as black (zero), the phase matrix becomes a discrete 2D periodic binary stripe
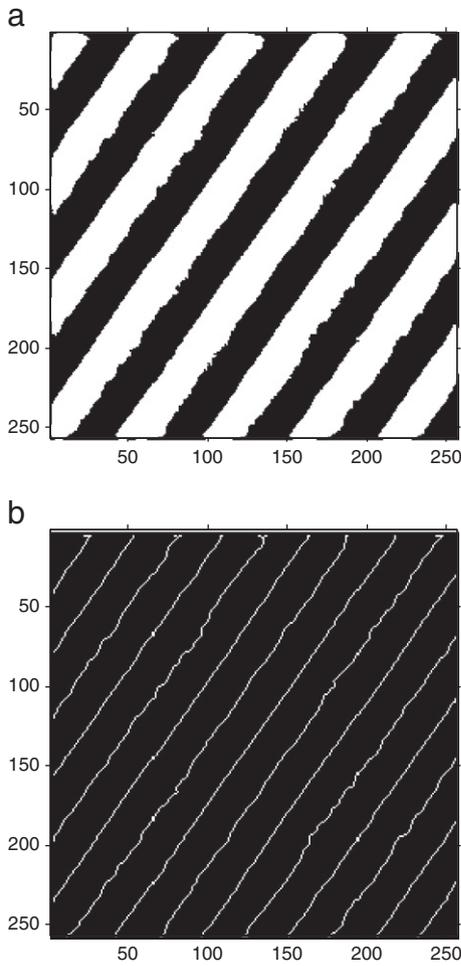
Fig. 3. (a) A denoised version of the phase matrix of discrete cross-power spectrum difference shown in Fig. 2 (c), (b) the discontinuities of the phase matrix of discrete cross-power spectrum difference shown in (a).

signal. Fig. 1 provides an ideal examples of $\hat{\varphi}(k,l)$ when $(\delta_x, \delta_y) = (4, 3)$. Note how the period of the square signal along each axis determines the shifts along corresponding axis: the period along the u axis is $\frac{N}{\delta_x}$, and hence there are δ_x repeated cycles along each row of the phase matrix, where δ_x may or may not be an integer. When δ_x is not an integer, the number of repeated cycles in a row is given by the integer part of δ_x plus a fraction of a cycle defined by the
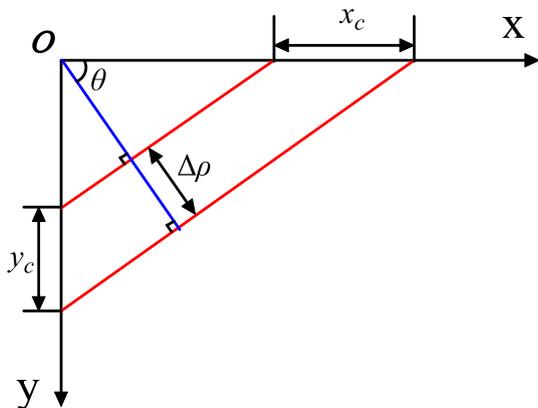

Fig. 4. Schematic diagram of two parallel lines in the phase matrix difference of discrete cross-power spectrum difference.

noninteger portion of δ_x. A similar argument applies to the columns of $\hat{\varphi}(k,l)$. This process of counting the number of cycles along the rows and columns of the phase matrix of discrete cross-power spectrum difference is essentially to determine the shifts. Determining the direction of shifts is much simpler in contrast. If k is fixed at zero and l is close to zero. The value of $\hat{\varphi}(k,l)$ is entirely determined by the sign of δ_y. It is similarly for the direction of δ_x. So for the displacement direction, we only need to examine the values of $\hat{\varphi}(k,l)$ at the neighborhood of $(0,0)$.

As indicated above, the key to solve the problem is to find how many cycles of phase matrix of discrete cross-power spectrum difference fit in the range $[0, 2\pi)$ along each frequency axis. Due to noise and other sources of error, however, counting the number of cycles per 2π may lead to inaccurate results. Besides, δ_x and δ_y may not be
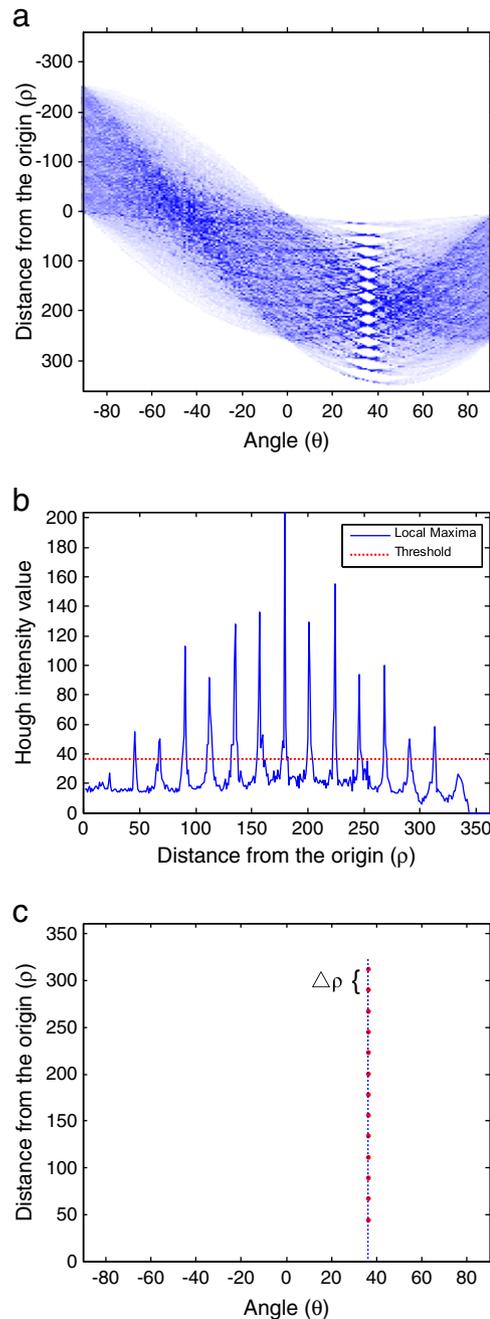

Fig. 5. (a) The Hough transform of the phase discontinuities shown in Fig. 4(b), (b) the local maxima and the threshold, and (c) the peaks detected by thresholding.
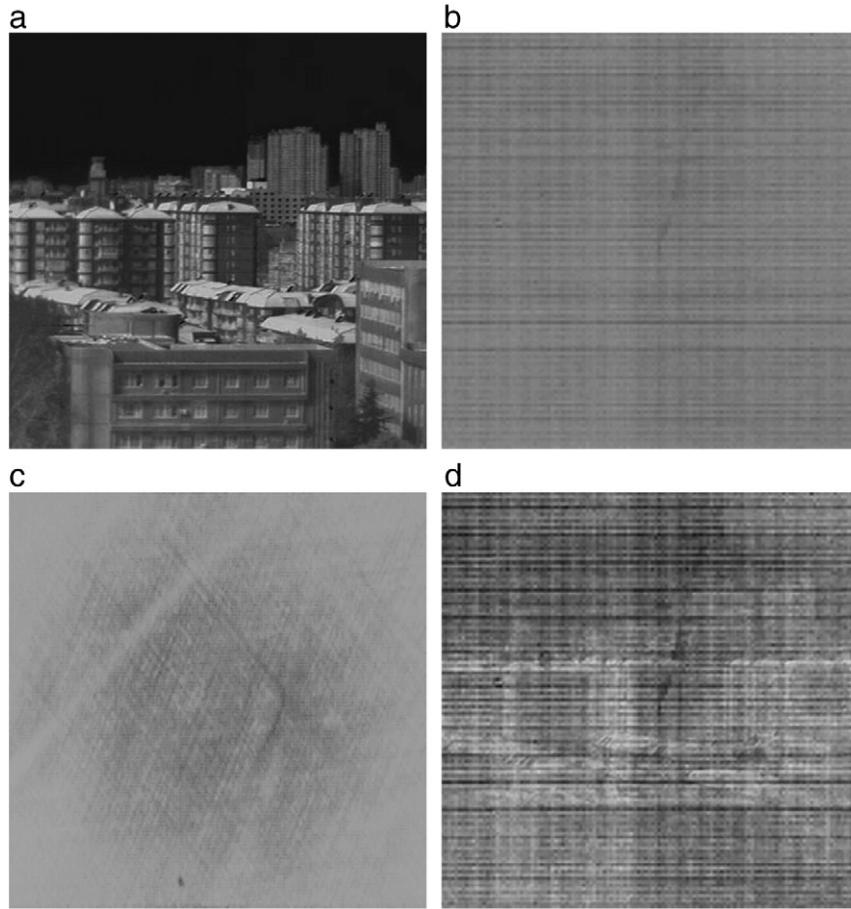
Fig. 6. (a) A 512×512 infrared image acquired from a well corrected HgCdTe IRFPA, (b) ,(c) noise patterns of two IRFPAs,(d) a sample image contaminated by the noise patterns shown in (b) and at a PSNR of 5.

integer. Fig. 2 shows another example of noisy phase matrix of discrete cross-power spectrum difference corresponding to shifts of (4.3, 3.2). Despite the serious FPN, the underlying two-dimensional stripe signals are clearly visible. Figs. 2(d) shows one row of Fig. 2(c). It can be seen that it is difficult to get accurate results simply by examine a single row or column. We must exploit the fact that a total of M×N data points are available for robust estimation.

To overcome this problem, the first thing is to denoise the phase matrix of discrete cross-power spectrum difference. As is shown in Fig. 2(c), the noise consists mainly of the isolated points, which inspires us to adopt a median filter to remove these speckles. Traditional median filter replaces each pixel value by the median of its neighbors, i.e. the value such that 50% of the values in the neighborhood are above, and 50% are below. This can be difficult and costly to implement due to the need for sorting of the values, especially when the filter window is large. However, in our case, the phase matrix is a binary image, so the median filter is in fact a mode filter, which replaces each pixel value by its most common neighbor. So we can omit the sorting operation and just count the number of ones or zeros in the window. Thus the simplified version of median filter is computationally much more efficiently. To protect the original shape of the stripe from being destroyed, the filter size should not exceed $\frac{M}{\delta_s} \times \frac{N}{\delta_s}$, where δ_s is the maximum shift or upper bound of the displacement in pixels between two images. Fig. 3(a) shows a denoised version of the phase matrix of the discrete cross-power spectrum difference shown in Fig. 2(c). The shape of stripe has been basically restored. The challenge now is to determine the exact fractional portion of the repeated cycles. From Fig. 1(a) and Fig. 3(a), we can intuitively find that the discontinuities of phase matrix of discrete cross-power spectrum difference form a set of parallel lines. We

could use these discontinuities as a landmark point of the cycle to count the number of cycles along each row or each column. From Eq. (12) the function of the set of lines can be easily obtained:

$$\frac{k}{N}\delta_x + \frac{l}{M}\delta_y = K. \tag{13}$$

This shows that the family of lines in the phase space parameterized by K-that is, each integer value of K would give a different line along which the phase is at the junction of $\pm\frac{\pi}{2}$. These lines can be easily obtained by the edge detection algorithms (Fig. 3(b)).

The Hough transform is the linear transform for detecting straight lines and basically maps a line to a parameter space of(θ, ρ). Each line itself is parameterized by the angle θ and its distance ρ from the origin. Consider two lines whose parameter equations are as follow

$$\begin{cases} \rho_1 = x\cos\theta_1 + y\cos\theta_1 \\ \rho_2 = x\cos\theta_2 + y\cos\theta_2 \end{cases} \tag{14}$$

Table 1
Results for image pair with shift of (3.50, 4.50) in simulated unstructured pattern noise.

PSNR	2DCC	2DCC + LPF	Cain	Proposed
5 dB	(−0.02, 0.17)	(−0.07, 0.26)	(3.36, 4.25)	(3.58, 4.18)
10 dB	(−0.01, 0.14)	(−0.12, 0.36)	(3.35, 4.39)	(3.58, 4.38)
15 dB	(−0.03, 0.06)	(3.01, 4.81)	(3.37, 4.34)	(3.57, 4.42)
20 dB	(−0.01, 0.12)	(3.13, 4.73)	(3.41, 4.42)	(3.52, 4.45)
25 dB	(3.41, 4.59)	(3.27, 4.76)	(3.43, 4.55)	(3.47, 4.47)
30 dB	(3.49, 4.53)	(3.39, 4.62)	(3.42, 4.54)	(3.48, 4.47)
35 dB	(3.49, 4.54)	(3.37, 4.60)	(3.44, 4.52)	(3.48, 4.47)
40 dB	(3.49, 4.52)	(3.41, 4.60)	(3.45, 4.54)	(3.48, 4.47)

Table 2
Results for image pair with shift of (4.50, 3.50) in simulated unstructured pattern noise.

PSNR	2DCC	2DCC + LPF	Cain	Proposed
5 dB	(−0.31, −0.12)	(−0.04, 0.06)	(4.76, 3.13)	(4.23, 3.65)
10 dB	(−0.21, −0.14)	(−0.02, −0.23)	(4.67, 3.31)	(4.38, 3.58)
15 dB	(−0.22, −0.16)	(4.13, 3.31)	(4.57, 3.39)	(4.51, 3.62)
20 dB	(−0.26, 0.22)	(4.22, 3.43)	(4.61, 3.46)	(4.52, 3.55)
25 dB	(4.41, 3.49)	(4.44, 3.42)	(4.55, 3.45)	(4.47, 3.52)
30 dB	(4.49, 3.48)	(4.49, 3.42)	(4.52, 3.48)	(4.48, 3.52)
35 dB	(4.50, 3.49)	(4.47, 3.42)	(4.54, 3.47)	(4.48, 3.52)
40 dB	(4.50, 3.49)	(4.49, 3.40)	(4.54, 3.49)	(4.50, 3.54)

Table 3
Results for image pair with shift of (−3.00, 2.50) in real structured pattern noise.

PSNR	2DCC	2DCC + LPF	Cain	Proposed
5 dB	(−0.31, −0.07)	(−0.47, −0.46)	(−0.72, 0.44)	(−3.17, 2.65)
10 dB	(−0.11, −0.04)	(−0.19, −0.36)	(−2.12, 0.94)	(−3.11, 2.52)
15 dB	(−0.03, −0.06)	(−3.31, 1.91)	(−2.62, 2.14)	(−3.07, 2.42)
20 dB	(−3.31, 2.22)	(−3.13, 2.37)	(−2.78, 2.19)	(−3.04, 2.47)
25 dB	(−3.41, 2.59)	(−3.17, 2.46)	(−2.87, 2.39)	(−3.02, 2.47)
30 dB	(−3.00, 2.53)	(−3.09, 2.42)	(−2.96, 2.44)	(−3.00, 2.48)
35 dB	(−2.99, 2.50)	(−3.07, 2.41)	(−2.97, 2.45)	(−3.00, 2.48)
40 dB	(−2.99, 2.50)	(−3.01, 2.41)	(−2.97, 2.47)	(−3.00, 2.48)

If $\theta_1 = \theta_2$ and $\rho_1 \neq \rho_2$, the two lines are considered to be parallel. Obviously $\theta_1 = \theta_2$ and the absolute value of the difference between ρ_1 and ρ_1 is the space between two parallel lines. Fig. 4 gives a schematic diagram of two parallel lines in the phase matrix of discrete cross-power spectrum difference. It can be seen that the number of cycles, that is, δ_x and δ_y, may or may not be integer values and can be given by

$$\delta_x = \frac{N}{2x_c} = \frac{N\Delta\rho}{2\cos\theta},$$ (15)

and

$$\delta_y = \frac{M}{2y_c} = \frac{M\Delta\rho}{2\sin\theta}.$$ (16)

The two equations inspire us to estimate the shifts in the Hough-transform domain. As can be verified from the above derivations, θ remains invariant among all lines and also $\Delta\rho$ is an invariant parameter. Therefore in the Hough domain, we expect to see a set of peak values situated at equal distances from each other, and parallel to the ρ axis. Fig. 5(a) shows an example of the Hough transform of the phase discontinuities shown in Fig. 4(b), where the peaks can be clearly

Table 4
Results for image pair with shift of (2.50, −3.00) in real structured pattern noise.

PSNR	2DCC	2DCC + LPF	Cain	Proposed
5 dB	(0.31, 0.12)	(0.21, 0.16)	(0.86, −0.74)	(2.35, −3.16)
10 dB	(0.11, 0.12)	(0.22, 0.13)	(1.21, −1.31)	(2.39, −3.08)
15 dB	(0.22, 0.16)	(0.13, 0.14)	(2.07, −2.49)	(2.41, −3.02)
20 dB	(0.26, 0.12)	(2.22, −2.43)	(2.34, −2.75)	(2.44, −3.05)
25 dB	(2.41, −2.94)	(2.44, −2.81)	(2.39, −2.85)	(2.47, −3.02)
30 dB	(2.49, −2.98)	(2.49, −2.92)	(2.41, −2.89)	(2.48, −3.02)
35 dB	(2.50, −2.98)	(2.47, −2.92)	(2.41, −2.92)	(2.48, −3.02)
40 dB	(2.49, −2.98)	(2.49, −2.93)	(2.42, −2.94)	(2.48, −3.02)

identified by a simple thresholding process. As is customary in Hough transform, we used the local maxima for finding a suitable threshold value. In our case, since all the peaks are known to be aligned parallel to the ρ axis, we took the maximum of the Hough matrix for each ρ as the local maximum. This yields a curve similar to the one shown in Fig. 5(b). We then used 1.5 times of the average of the local maxima curve as a threshold. Specifies the size of the suppression neighborhood is also essential to avoiding fake peaks. The $\frac{\min(M,N)}{\delta_s} \times \frac{\min(M,N)}{\delta_s}$ neighborhood around each peak should be set to zero after the peak is identified. Fig. 5(c) shows the peaks detected by the thresholding process.

Estimating δ_x and δ_y from these peak values in the Hough-transform domain is an easy job now. A robust estimates of $\Delta\rho$ and θ can thus be simply obtained by using the median value

$$\Delta\hat{\rho} = \text{med}(\Delta\rho_i), i = 1,...,t-1$$ (17)

and

$$\hat{\theta} = \text{med}(\theta_i), i = 1,...,t,$$ (18)

where t is the total number of peak values. Then δ_x and δ_y can be deduced by Eqs. (15) and (16).

4. Experimental results

We applied the technique to both synthetic and real data. First, we compared the proposed method with three other registration
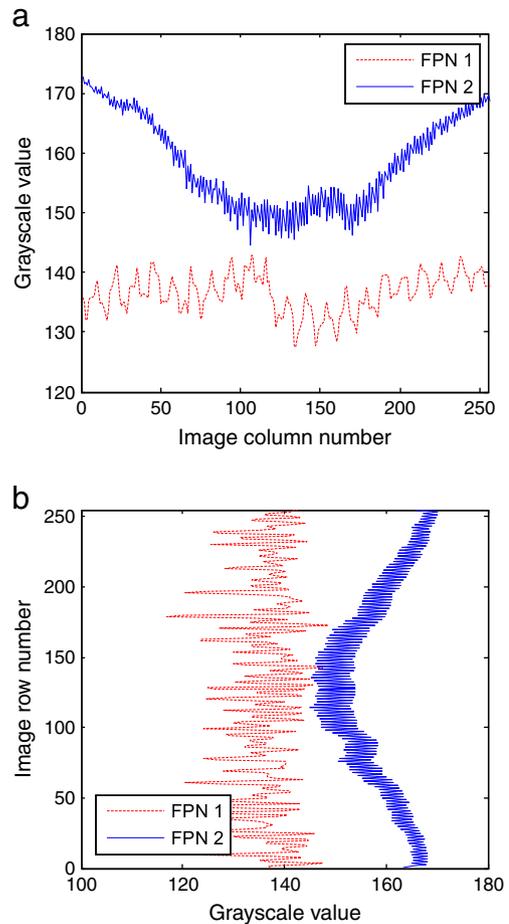

Fig. 7. The projected image vectors of two FPN images shown in Figs. 6(b) and (c). (a) The projected vectors obtained by averaging each FPN image vertically; (b) the projected vectors obtained by averaging each FPN image horizontally.
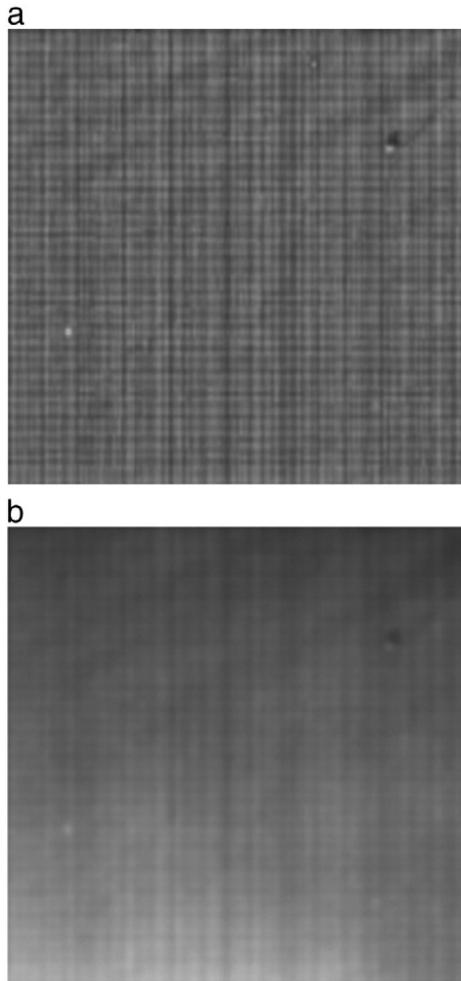
Fig. 8. The offset (a) and gain (b) FPN images obtained from the two point calibration parameters of an uncooled infrared camera.

our technique to real noisy infrared sequence, combining with the registration-based NUC method presented in Ref. [9], to demonstrate its practical effect. In all experiments, the kernel size of the median filter in our method is set as 7×7. We have found that this choice consistently yields very good results in all cases.

4.1. Application to simulated data

In this subsection, the registration of the images was compared with three methods: the 2-D cross-correlation method (2DCC), a combination of the spatial low-pass filtering together with the 2-D cross-correlation method, denoted as the 2DCC plus low-pass filtering (2DCC + LPF), and the projection-based 1-D cross-correlation method proposed by Cain et al. In order to obtain subpixel resolution for the three cross-correlation methods, we interpolate the original images by zero-padding the cross-power spectrum as suggested in Ref. [20]. The 2DCC + LPF uses a Gaussian of low-pass filter of 10×10 pixels for smoothing the images.

For synthetic data, we use the approach described in Ref. [15], to generate images with subpixel shifts, i.e., starting from a real high resolution image, we low-pass-filtered and downsampled integer shifted versions of the image. Then the relative shifts become fractional. The original high resolution image shown in Fig. 6(a) is acquired from a well corrected HgCdTe IRFPA camera.

In the first study, two sets of subpixel translations are (3.5, 4.5) and (4.5, 3.5), and the FPN is considered Gaussian distributed. Clearly, the FPN is considered as spatially unstructured in this case. We consider different noise level ranging from 5 dB to 40 dB peak signal-to-noise ratio (PSNR) environments, which can be achieved by varying the standard deviation of the FPN. Tables 1 and 2 summarize the results obtained.

It can be observed that when the PSNR is relatively high (> 30 dB), the four tested method can all give reliable shift estimates. It seems that, on average, over the range of the PSNR, traditional 2D cross-correlation method performs the best. Our method also performs very well at this lower noise level, but the accuracy is slightly lower than the traditional 2D cross-correlation method. The accuracies of Cain's method and the 2DCC + LPF are inferior comparatively, which partly because the spatial averaging or smoothing may result in information loss of the true scene. When the PSNR reduces to 20 dB, the traditional 2D cross-correlation method fails to give correct estimates. The accuracy of 2DCC + LPF also drops dramatically, and when the
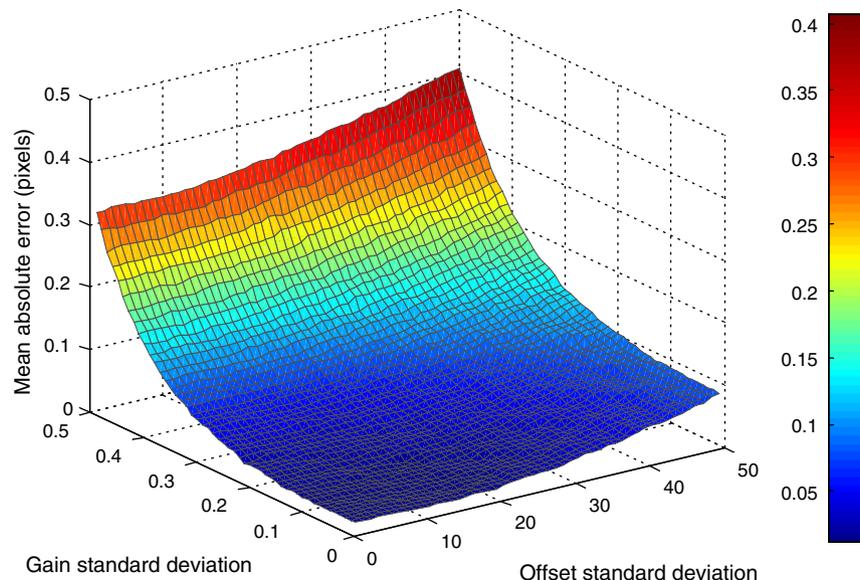
methods using synthetic images with known shifts contaminated by both the simulated and the real FPN. Then, we test our algorithm in both offset and gain nonuniformity conditions. Finally, we applied

Fig. 9. Mean absolute error of translation estimates for various levels of gain and offset FPN.

PSNR reduces to 10 dB, it begins to not work properly, wrongly considering there is no shift between two images. Both the proposed method and Cain's method provides correct estimates for subpixel shift estimation under low PSNR conditions, but it is clear that our method consistently outperforms the Cain's method, under different noise levels and pixels shifts.

In the second experiment, we focus on comparing the four registration algorithms when the FPN is structured. The noise patterns shown in Figs. 6(b) and (c) are real nonuniformities acquired from two IRFPA cameras. We first generated two sets of images with subpixel translations of $(-3.0, 2.5)$ and $(2.5, -3.0)$, using the procedures as before. Then the noise pattern shown in Fig. 6(b) is added to the image pair with translations of $(-3.0, 2.5)$, and the noise pattern shown in Fig. 6(c) is added to the image pair with translations of $(2.5, -3.0)$. We vary the amounts of the added FPN images to produce image pairs with different PSNRs between 5 and 40. The results obtained are shown in Tables 3 and 4.

A similar tendency can also be seen in the results of Tables 3 and 4, where the same methods are applied. In general, 2DCC and 2DCC + LPF can determine the translations in high-to-moderate PSNR environments relatively well. They cannot, however, provide satisfactory results under noisy environments. The performance of 2DCC + LPF is worse than when the FPN is unstructured, which can be explained because the structured FPN could not be effectively reduced via spatial filtering, especially when the FPN contains large low-spatial variations. It also can be noted that the accuracy of Cain's method drops rapidly with the decline of the PSNR. Some wrong results are given when the PSNR is under 10 dB since to the noise patterns are spatial uneven distributed. Fig. 7 shows the projected image vectors obtained by averaging each FPN image across its rows and columns. Ideally, the projected vector should have a flat shape. However, since the real FPN is structured and highly spatially correlated, it could not be effectively canceled by averaging. Obviously, when those FPN components dominate the whole projected vectors, the result will become unreliable. However, the proposed method again provides satisfactory performance in all cases since the effect of FPN is eliminated in cross-power spectrum difference, even though the FPN is highly structured. Furthermore, when the PSNR is only 5 dB, our method can also estimate image translation with about 0.2-pixel accuracy. To demonstrate this noise level, we show the noisy image in Fig. 6(d), which is contaminated by the noise patterns shown in Fig. 6(b), at a PSNR of 5. These results clearly demonstrate the effectiveness of the proposed method in dealing with the FPN.

4.2. Performance evaluation in both offset and gain nonuniformity

As is presented in Section 2, in the proposed method, the FPN is modeled as additive, signal independent noise. However, it is widely accepted that the FPN should base on a linear model [10], i.e., it consists of two components: gain and offset. The true scene value is first scaled by a gain factor and then offset by an offset term to produce the observed detector output. Since in the proposed method, the gain is assumed uniform across all detectors. If gain nonuniformity is present, its performance will degenerate. Here we investigate the ability of the proposed registration algorithm to operate in the presence of FPN with both gain and offset components. We first generate 10 image pairs with known shifts using the same procedure as before. Then these images are corrupted by the gain and offset FPN shown in Fig. 8. The gain and offset images are acquired from the two point calibration parameters of an uncooled infrared camera. By linear scaling the two noise patterns, the level of offset and gain FPN can be changed easily. The mean absolute error of the estimate is presented for several levels of nonuniformity (indentified by standard deviations) in Fig. 9.

In Fig. 9, the mean absolute error of the shift estimates both increases nonlinearly with increased bias nonuniformity standard deviation and gain nonuniformity standard deviation. Moreover, the mean absolute error is more sensitive to gain nonuniformity standard deviation as expect. When the gain nonuniformity standard deviation is less than 0.3, its influence on the estimates accuracy is relatively small. According to the linear model of FPN, this level of gain variation can be approximated by bias variation, which our algorithm can handle effectively. However, the effect of gain nonuniformity becomes apparent when its standard deviation is greater than 0.4. In such a case, the error will be expected to increase since the observed images may not fit the model of our algorithm. Fortunately, researches show that gain nonuniformity has a good time stability and offset nonuniformity dominates gain nonuniformity in most applications [7,8,21]. If the level of gain nonuniformity is too high to obtain good registration accuracy, it may be necessary to use calibration-based nonuniformity methods beforehand to reduce its effect.
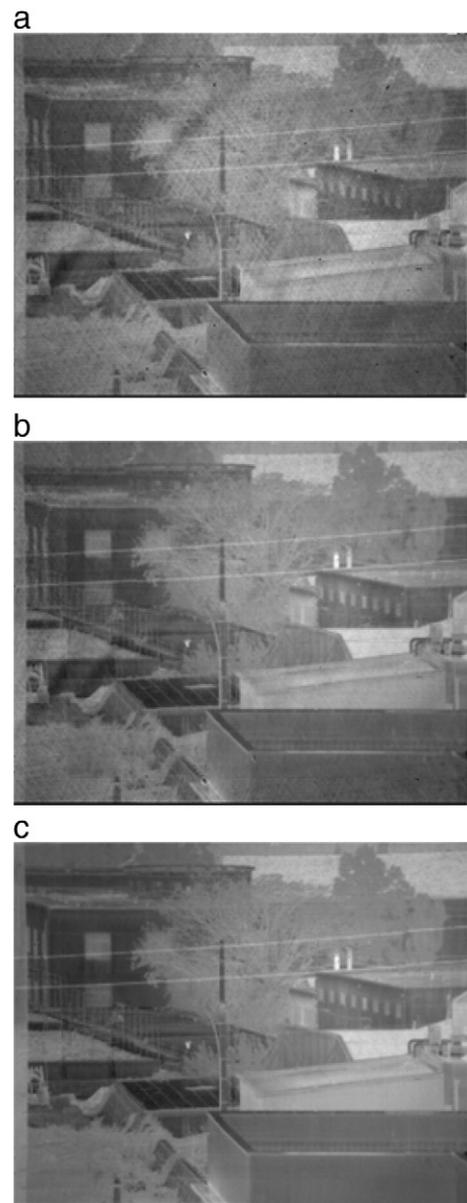

Fig. 10. NUC performance comparison with the two versions of the interframe registration-based NUC method: (a) the raw corrupted frame. (b) Corrected frame when traditional phase-correlation registration is applied. (c) Corrected frame when the proposed registration method is applied.
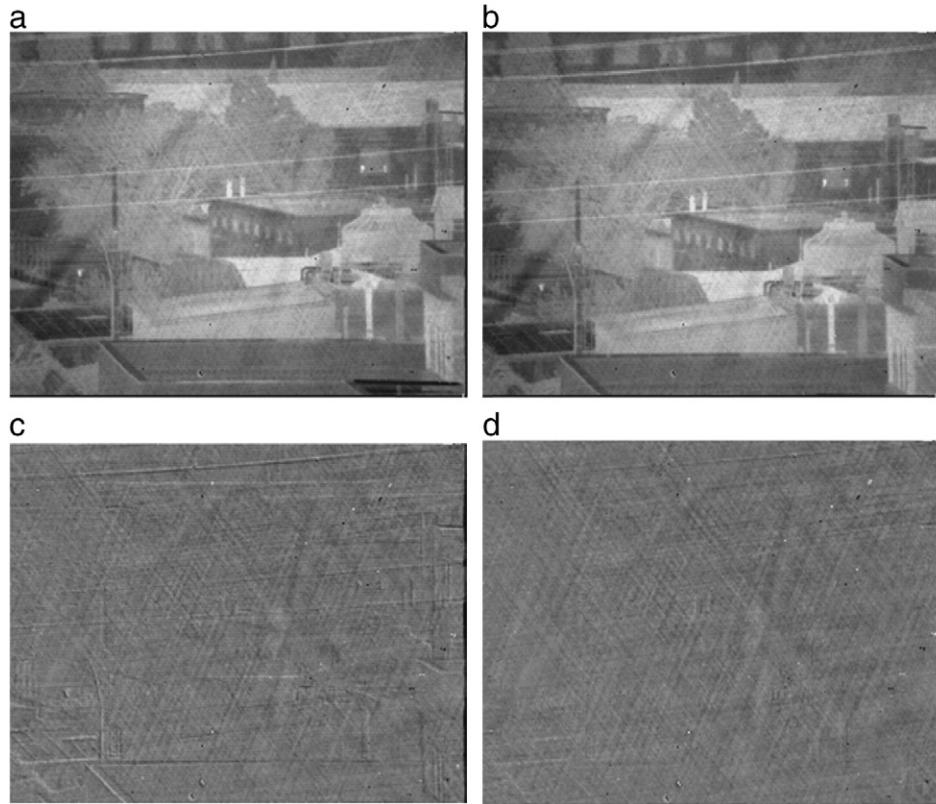
Fig. 11. Registration result of two sample images from the tested noisy video sequence: (a) The 7th frame of the video sequence. (b) The 8th frame of the video sequence. (c) Difference image obtained with traditional phase-correlation registration. (d) Difference image obtained with the proposed method.

4.3. Application to registration-based nonuniformity correction

In this section, the proposed algorithm is tested on a real infrared sequence data. The data set was acquired at 6 p.m. by using a 320×256 HgCdTe FPA camera operating in the 8–14 µm range and at a rate of 25 fps. An issue encountered in performing registration on real infrared data is the lack of ground truth; hence, the performance cannot be evaluated objectively. Considering that the nonuniformity correction is one of the main applications for our registration method, we have incorporated our method into the recently developed interframe registration-based nonuniformity correction framework [9], so that the quality of the motion estimation can be assessed based on their impact on the results of the nonuniformity corrected images. In Ref. [9], Zuo et al. employed a traditional phase-correlation method [14] to estimate translations between consecutive frames. Here we replace it with our proposed registration method and we use the same parameters as suggested in Ref. [9]. We compared our modified version with the original one. Since only the registration method is changed, the performance difference should only be attributed to the accuracy of the registration methods. The more accurate the registration algorithm is, the higher NUC precision and convergence rate should be achieved.

The NUC results using 30 image frames are shown in Fig. 10. Fig. 10 (a) corresponds to uncorrected 30th frame. Fig. 10(b) corresponds to corrected version of the 30th frame obtained when the traditional phase correlation is used as the registration method in the interframe registration-based NUC method, while Fig. 10(c) corresponds to corrected version of the 30th frame obtained when the proposed registration method is used in the interframe registration-based NUC method. Of course, due to lack of ground truth, the performance evaluation is somewhat subjective. Comparing the results in Fig. 10(b) and (c), it is clear that, to the naked eye, the correction result using the proposed registration method has a better visual effect since the FPN is hardly perceptible. While from the result of the original version, we can see some residual nonuniformity,

especially at the periphery of the image. This is partially because sometimes the traditional phase-correlation method fails to detect the real translation between two dirty images, wrongly considering the motion is insufficient, which slows the correction parameters updating process.

To better explain the difference in NUC performance, two consecutive frames are extracted from tested infrared sequence, shown in Fig. 11(a) and (b). The difference of the two properly registered frames using the traditional phase-correlation method and the proposed method are shown in Fig. 11(c) and (d), respectively. Note the difference images contain only the overlapped area between the two images. Since the objects and scene in the two consecutive images are stationary, the difference image should contain only nonuniformity ideally. Comparing the two difference images, it is clearly that the one registered by the proposed method shows much less signs of the scene information and almost consists of pure FPN. This observable result let us conclude that our registration method gives more accurate shifts estimates than traditional phase-correlation method under conditions of real nonuniformity.

5. Conclusions

The paper proposes a new technique to address subpixel image registration directly in frequency domain by counting the number of cycles of the phase matrix of the cross-power spectrum difference. Its main feature is the capability to perform reliable image registration under fixed-pattern noise environment. This is because this method utilizes the difference of the cross-power spectrum of two discrete shifted images to suppress the noise power spectrum while the shifts information is well preserved. Since no inverse transforming is required, the computational complexity is essentially determined by the Fast Fourier Transform (FFT) (i.e. Nlog(N), for N points). There are of course small overheads associated with computing the phase discontinuities matrix and its Hough transform. Experimental results show that the proposed method is effective in identifying subpixel translations under different

noise levels and environments. Besides, compared with Cain's method, our method provides a better solution for handling structured FPN. Finally, its great practical effect has been shown by applying our registration algorithm to a registration-based NUC method.

It is worth mentioning that the effectiveness of the proposed method is based on an offset-only model for FPN and this rule cannot be applied in every case. The real FPN is not purely additive, but has the gain component. Experimental results show that under certain level of gain nonuniformity, our method can also give reliable results. But our method may not work well in the cases where the gain nonuniformity is the dominate component of FPN. In addition, as in any shift-only registration technique, the performance may degrade if the shift transformation model does not hold. We found that, if the true motion is sufficiently close to translation, then the degradation in performance may be small. Of course, this point needs more detailed investigation.

Acknowledgements

References

[1] L.G. Brown, ACM Comput. Surv. 24 (1992) 325.
[2] B. Zitová, J. Flusser, Image and Vision Computing 21 (2003) 977.
[3] M.S. Alam, J.G. Bognar, R.C. Hardie, B.J. Yasuda, IEEE T. Instrum. Meas. 49 (2000) 915.
[4] E.E. Armstrong, M.M. Hayat, R.C. Hardie, S.N. Torres, B.J. Yasuda, in: G.T. Andrew (Ed.), SPIE, 1999, p. 150.
[5] S.C. Cain, M.M. Hayat, E.E. Armstrong, IEEE T. Image Process 10 (2001) 1860.
[6] W.F. O'Neil, Meeting of the IRIS Specialty Group on Passive Sensors 1, Tucson, AZ, 1997, p. 329.
[7] R.C. Hardie, M.M. Hayat, E. Armstrong, B. Yasuda, Appl Optics 39 (2000) 1241.
[8] B.M. Ratliff, M.M. Hayat, R.C. Hardie, J. Opt. Soc. Am. A 19 (2002) 1737.
[9] C. Zuo, Q. Chen, G.H. Gu, X.B. Sui, J. Opt. Soc. Am. A 28 (2011) 1164.
[10] D.L. Perry, E.L. Dereniak, Opt. Eng. 32 (1993) 1854.
[11] W.K. Pratt, Aerospace and Electronic Systems, IEEE Transactions on, AES-10, 1974, p. 353.
[12] J. Pezoa, O. Medina, in: C. San Martin, S.-W. Kim (Eds.), Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Springer Berlin, Heidelberg, 2011, p. 55.
[13] O.J. Medina, J.E. Pezoa, S.N. Torres, in: D.L. Paul, K.S. Ashok, S.W. Priyalal, R. Manijeh, V. Jose Luis Pau, S. Rengarajan, P.U. Melville, M. Tariq (Eds.), SPIE, 2011, p. 81550H.
[14] C.D. Kuglin, D.C. Hines, IEEE International Conference on Cybernetics and Society, 1975, p. 163.
[15] H. Foroosh, J.B. Zerubia, M. Berthod, IEEE T Image Process 11 (2002) 188.
[16] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, K. Kobayashi, Ieice T Fund Electr, E86a, 2003, p. 1925.
[17] W.S. Hoge, Ieee T Med Imaging 22 (2003) 277.
[18] H.S. Stone, M.T. Orchard, E.C. Chang, S.A. Martucci, Ieee T Geosci Remote 39 (2001) 2235.
[19] H. Foroosh, M. Balci, International Conference on Image Processing, Vols 1–5, 2004, p. 1915.
[20] J.A. Parker, R.V. Kenyon, D.E. Troxel, Medical Imaging, IEEE Transactions on 2 (1983) 31.
[21] E. Gurevich, A. Fein, in: A. Bjorn, F.F. Gabor, S. Marija (Eds.), SPIE, 2003, p. 809.