A Bispectrum Technique to Subpixel Image Registration under Noisy Conditions

Kim-Hui Yap and Li Chen

School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore,
Email: {ekhyap, pg05136025}@ntu.edu.sg

ABSTRACT

This paper proposes an effective higher-order statistics method to address subpixel image registration. Conventional power spectrum-based techniques employ second-order statistics to estimate subpixel translation between two images. They are, however, susceptible to noise, thereby leading to significant performance deterioration in low signal-to-noise (SNR) environments. In view of this, we propose a bispectrum-based approach to alleviate this difficulty. The new method utilizes the characteristics of bispectrum to suppress Gaussian noise. It develops the phase relationship between the image pair, and estimates the subpixel translation by solving a set of nonlinear equations derived from the bispectrum. Experimental results show that the proposed method is effective in identifying subpixel translations under different noise levels and environments.

1. INTRODUCTION

Image registration is the process of establishing point-to-point correspondence including translations, rotation and scaling between two images of the same scene. This process is needed in various imaging applications, such as stereo depth perception, object recognition, and image fusion, among others. Image registration methods can be classified into several categories including feature-based techniques, gradient approaches, and Fourier methods [1].

This work focuses on high-accuracy image registration, namely subpixel translation estimation, where pixel-level translation estimation is assumed to have been performed in coarse registration step. The direct subpixel registration methods interpolate the shifted image with different subpixel parameters, and compare the result with the reference image [2]. The minimum error between the image pair will provide a sound estimation of the actual motion parameters.

The idea behind Fourier methods is quite simple as that the cross-correlation between the delayed signal and the reference signal will have a peak at the delayed time. Hence, phase correlation method identifies the translations by taking the inverse discrete Fourier transform (IDFT) of the normalized cross power spectrum [3]. Pixel-level Fourier techniques have been extended to address subpixel registration using two different approaches. The first approach is by estimating the best-fit phase plane in the frequency domain, of which the slope of the plane is used to determine the subpixel translation [4]. The aliasing error due to inadequate sampling is masked out by spectrum cancellation [5]. The second approach is to evaluate the dominant peaks of the IDFT of the normalized cross power spectrum, and proceed to estimate the subpixel translations [6]. These methods require the assumption that the observed images have high SNR, or in other words, in low noise environments. The precondition, however, is impractical in some applications such as sonar imaging where the SNR is low. Under these circumstances, the performance of these techniques will degrade significantly.

In view of this, we propose a new bispectrum technique to address the problem of subpixel image registration. The proposed method is motivated by the observation that the n-th joint cumulants of the Gaussian random process are equal to zero for n>2. Therefore, the phase information between the image pair can be estimated reliably from the higher-order statistics such as bispectrum, as they are robust towards correlated Gaussian noise and low SNR environments.

2. PROBLEM FORMULATION

Let $\boldsymbol{s}$ be the original image, and $\boldsymbol{f}_i\{i=1,2\}$ be two images that are shifted versions of $\boldsymbol{s}$:

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

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, which is restricted to subpixel level [0, 1) in this work. In the absence of noise and aliasing, the *shift property* of Fourier transform gives:

$$\tilde{f}_1(\omega_x, \omega_y) = \tilde{f}_2(\omega_x, \omega_y) e^{-j(\omega_x \delta_x + \omega_y \delta_y)} \tag{2}$$

where $\tilde{\boldsymbol{f}}_i = \mathfrak{F}[\boldsymbol{f}_i]$ is the Fourier transform of $\boldsymbol{f}_i$.

In most imaging applications, noise exists in the captured images up to a certain level. The noise is commonly due to factors such as photoelectric noise, film grain noise, and quantization noise, among others. Moreover, the imaging sensor array, usually the charged-coupled device (CCD), is subject to various sources of noise, including thermal noise and shot noise. The noisy effect is particularly evident under low lighting condition when the camera gain is high.

In the presence of additive noise, the images $\boldsymbol{f}_i\{i=1,2\}$ in (1) are modeled as

$$f_i(x,y) = s(x+\delta_{x,i}, y+\delta_{y,i}) + n_i(x,y) \tag{3}$$

where $\boldsymbol{n}_i$ is the additive noise that arises during the image formation process.

Conventional Fourier method of (2) ignores the noise in their formulation, and hence will experience performance deterioration in noisy environments. Assuming $\boldsymbol{n}_i$ is independent of $\boldsymbol{s}$, the normalized cross power spectrum can be expressed as:

$$\begin{aligned}
\tilde{P}(\omega_x,\omega_y) &= \frac{\mathfrak{F}[R_{f_2 f_1}(\tau_x,\tau_y)]}{\mathfrak{F}[R_{f_1 f_1}(\tau_x,\tau_y)]} \\
&= \frac{\mathfrak{F}[R_{ss}(\tau_x-\delta_x,\tau_x-\delta_y)] + \mathfrak{F}[R_{n_1 n_2}(\tau_x,\tau_y)]}{\mathfrak{F}[R_{ss}(\tau_x,\tau_y)] + \mathfrak{F}[R_{n_1 n_1}(\tau_x,\tau_y)]} \\
&= \frac{e^{-j(\omega_x\delta_x+\omega_y\delta_y)} + \tilde{R}_{n_1 n_2}(\omega_x,\omega_y)/\tilde{R}_{ss}(\omega_x,\omega_y)}{1+\tilde{R}_{n_1 n_1}(\omega_x,\omega_y)/\tilde{R}_{ss}(\omega_x,\omega_y)}
\end{aligned} \tag{4}$$

where $R_{ab}(\tau_x,\tau_y) \triangleq E[a(x,y)b(x+\tau_x, y+\tau_y)]$ is the correlation function between matrix $\boldsymbol{a}$ and $\boldsymbol{b}$.

It is observed that $\tilde{P}(\omega_x,\omega_y)$ can be approximated as $\tilde{P}(\omega_x,\omega_y) \approx e^{-j(\omega_x\delta_x+\omega_y\delta_y)}$ only when $\tilde{R}_{n_1 n_2}(\omega_x,\omega_y)$ and $\tilde{R}_{n_1 n_1}(\omega_x,\omega_y)$ are negligible compared to $\tilde{R}_{ss}(\omega_x,\omega_y)$. This requires the SNR to be high ($\tilde{R}_{n_1 n_1}(\omega_x,\omega_y) \approx 0$) and the two noise processes to be uncorrelated ($\tilde{R}_{n_1 n_2}(\omega_x,\omega_y) \approx 0$). These preconditions, however, are restrictive in some applications, such as sonar imaging where the noise level is high and their sources are correlated. In view of this, we develop a bispectrum algorithm to address this difficulty.

3. SUBPIXEL REGISTRATION UNDER NOISY CONDITIONS

3.1. Proposed Cross Bispectrum Method

Higher-order spectra defined in terms of higher-order cumulant contains additional information that is not conveyed by the signal's correlation or power spectrum. It is useful in suppressing additive Gaussian noise (white and color noise) because all joint cumulants of order>2 are equal to zero for Gaussian random processes [7]. Hence, in circumstances where the observed signal consists of non-Gaussian signal of interest and is corrupted by additive Gaussian noise, there are clear advantages in estimating the desired signal through higher-order spectra.

In this work, we assume that the original image $\boldsymbol{s}$ in (3) follows non-Gaussian distribution with nonzero skewness (i.e. $E[\boldsymbol{s}^3] \neq 0$, which indicates the existence of non-trivial bispectrum). This assumption is in agreement with most observations that images are indeed non-Gaussian distributed. We further assume that $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are zero-mean Gaussian, signal independent random noise, which are potentially cross-correlated. It is noted that these assumptions are valid in many general applications [8].

The third-order auto- and cross-cumulants of the observed images $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ are defined as:

$$\begin{aligned}
R_{f_1 f_1 f_1}(\tau_x,\tau_y,v_x,v_y) &\triangleq E[f_1(x,y) f_1(x+\tau_x,y+\tau_y) f_1(x+v_x,y+v_y)] \\
&= R_{sss}(\tau_x,\tau_y,v_x,v_y); \\
R_{f_2 f_1 f_2}(\tau_x,\tau_y,v_1,v_2) &\triangleq E[f_2(x,y) f_1(x+\tau_x,y+\tau_y) f_2(x+v_x,y+v_y)] \\
&= R_{sss}(\tau_x-\delta_x,\tau_y-\delta_y,v_x,v_y)
\end{aligned} \tag{5}$$

where

$$R_{sss}(\tau_x,\tau_y,v_x,v_y) \triangleq E[s(x,y)s(x+\tau_x,y+\tau_y)s(x+v_x,y+v_y)] \tag{6}$$

It is the third-order auto-cumulant of the desired signal $\boldsymbol{s}$.

It is worth mentioning that (5) holds because $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are zero-mean Gaussian noise, hence all their joint cumulants of order>2 are equal to zero. Thus we can estimate the desired signal $\boldsymbol{s}$ from (5) without the interference of Gaussian noise.

The bispectrums are computed by taking the discrete

Fourier transform (DFT) of the cumulant

$$\begin{aligned}\tilde{R}_{f_1 f_1 f_1}(\omega_x,\omega_y,\upsilon_x,\upsilon_y) &\triangleq \mathfrak{F}[R_{f_1 f_1 f_1}(\tau_x,\tau_y,v_x,v_y)] \\ &= F_{sss}(\omega_x,\omega_y,\upsilon_x,\upsilon_y); \\ \tilde{R}_{f_2 f_1 f_2}(\omega_x,\omega_y,\upsilon_x,\upsilon_y) &\triangleq \mathfrak{F}[R_{f_2 f_1 f_2}(\tau_x,\tau_y,v_x,v_y)] \\ &= F_{sss}(\omega_x,\omega_y,\upsilon_x,\upsilon_y)e^{-j(\omega_x\delta_x+\omega_y\delta_y)}\end{aligned} \tag{7}$$

where $\tilde{R}_{sss}(\omega_x,\omega_y,\upsilon_x,\upsilon_y) \triangleq \mathfrak{F}[R_{sss}(\tau_x,\tau_y,v_x,v_y)]$ is the auto-bispectrum of $\boldsymbol{s}$.

Therefore, the phase information of $\omega_x\delta_x+\omega_y\delta_y$ is given by the normalized cross bispectrum:

$$\begin{aligned}\tilde{P}(\omega_x,\omega_y,\upsilon_x,\upsilon_y) &= \frac{\tilde{R}_{f_2 f_1 f_2}(\omega_x,\omega_y,\upsilon_x,\upsilon_y)\tilde{R}_{f_1 f_1 f_1}(\omega_x,\omega_y,\upsilon_x,\upsilon_y)^*}{|\tilde{R}_{f_2 f_1 f_2}(\omega_x,\omega_y,\upsilon_x,\upsilon_y)\tilde{R}_{f_1 f_1 f_1}(\omega_x,\omega_y,\upsilon_x,\upsilon_y)^*|} \\ &= e^{-j(\omega_x\delta_x+\omega_y\delta_y)}\end{aligned} \tag{8}$$

Compared to the cross power spectrum in (4), the normalized cross bispectrum in (8) provides more robust phase information in noisy environments. This is because that the normalized cross bispectrum is independent of the noise distortion terms $\tilde{R}_{n_1 n_1}(\omega_x,\omega_y)$ and $\tilde{R}_{n_1 n_2}(\omega_x,\omega_y)$. This enables reliable subpixel registration when incorporating the *Dirichlet* estimation scheme in [6] to give:

$$\begin{aligned}P(x,y) &= \mathfrak{F}^{-1}[\tilde{P}(\omega_x,\omega_y,\upsilon_x,\upsilon_y)] \\ &= \frac{1}{MN}\frac{\sin(\pi(x-\delta_x))}{\sin(\pi(x-\delta_x)/M)}\frac{\sin(\pi(y-\delta_y))}{\sin(\pi(y-\delta_y)/N)}\end{aligned} \tag{9}$$

where $M\times N$ is the length of IDFT. The *Dirichlet* function in (9) is approximated by *sinc* function as

$$P(x,y) \approx \operatorname{sinc}(x-\delta_x)\operatorname{sinc}(y-\delta_y) \tag{10}$$

It is observed that the function of (10) has dominant values centered near $(x,y)=(\delta_x,\delta_y)$ where $0\le\delta_x,\delta_y\le 1$. Therefore, a simple yet effective method of estimating (δ_x,δ_y) can be achieved by solving the nonlinear equations (10) at coordinates of (0,0), (0,1), (1,0) and (1,1), which correspond to the main peaks. The intermediate steps are given as:

$$\begin{aligned}\frac{P(0,0)}{P(1,0)} &= \frac{P(0,1)}{P(1,1)} = \frac{\sin(\pi(1-\delta_x)/M)}{\sin(\pi\delta_x/M)} \approx \frac{1-\delta_x}{\delta_x}; \\ \frac{P(0,0)}{P(0,1)} &= \frac{P(1,0)}{P(1,1)} = \frac{\sin(\pi(1-\delta_y)/N)}{\sin(\pi\delta_y/N)} \approx \frac{1-\delta_y}{\delta_y}\end{aligned} \tag{11}$$

The final estimated motion shift or translations $(\hat{\delta}_x,\hat{\delta}_y)$ can then be deduced directly by using

$$\begin{aligned}\hat{\delta}_x &= \frac{1}{2}\left(\frac{P(1,0)}{P(1,0)+P(0,0)}+\frac{P(1,1)}{P(1,1)+P(0,1)}\right); \\ \hat{\delta}_y &= \frac{1}{2}\left(\frac{P(0,1)}{P(0,1)+P(0,0)}+\frac{P(1,1)}{P(1,1)+P(1,0)}\right)\end{aligned} \tag{12}$$

4. EXPERIMENTAL RESULTS

The effectiveness of the proposed method is demonstrated using images under noisy conditions. The 1024×1024 "Pentagon" image shown in Fig. 1(a) is shifted by different shift sizes and downsampled by a factor 4×4 to produce the image pair. In this study, two sets of subpixel translations are (0.25, 0.75) and (0.75, 0.25). We consider different noise power ranging from 10dB-40 dB SNR environments. In addition, we compare the proposed algorithm with well-known phase correlation method by *Foroosh* et. al. [6]. Table I summarizes the results obtained. It can be observed that the proposed method provides satisfactory performance for subpixel shift estimation under different noisy conditions. Further, it is clear that the proposed method consistently outperforms the *Foroosh* method, especially under low SNR environments.

The 1704×1704 "Castle" image shown in Fig. 1(b) is shifted and downsampled by a factor 8×8 to produce the image pair. In this experiment, however, the noise is correlated across two channels, e.g. $\boldsymbol{n}_1$ is assumed to be AWGN, while $\boldsymbol{n}_2$ is generated from $\boldsymbol{n}_1$ using:

$$n_2(x,y) = \sum_{i=-3}^{3}\sum_{j=-3}^{3} h(i,j)n_1(x+i,y+j) \tag{13}$$

where $h(i,j)=\exp(-(i^2+j^2)/2)$ is a low-pass filter used to ensure that the noise $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are correlated [7]. In this study, we use two sets of subpixel translations (0.125, 0.125) and (0.375, 0.75), and repeat the procedures as before to estimate the subpixel translations. The results obtained are shown in Table II. It is observed that proposed method again provides satisfactory performance in all cases. These results clearly demonstrate the effectiveness of the proposed method in dealing with cross-correlated channel noise.

TABLE I

RESULTS OF SUBPIXEL REGISTRATION IN AWGN

SNR	(0.25, 0.75)		(0.75, 0.25)	
	Foroosh	Proposed	Foroosh	Proposed
10dB	(0.38, 0.65)	(0.29, 0.68)	(0.64, 0.36)	(0.68, 0.30)
20dB	(0.31, 0.71)	(0.28, 0.74)	(0.71, 0.31)	(0.74, 0.29)
30dB	(0.30, 0.73)	(0.27, 0.74)	(0.74, 0.30)	(0.74, 0.28)
40dB	(0.29, 0.74)	(0.27, 0.74)	(0.74, 0.29)	(0.75, 0.28)

TABLE II

RESULTS OF SUBPIXEL REGISTRATION CORRELATED NOISE

SNR	(0.125, 0.125)		(0.375, 0.75)	
	Foroosh	Proposed	Foroosh	Proposed
10dB	(0.267, 0.255)	(0.151, 0.170)	(0.378, 0.571)	(0.330, 0.609)
20dB	(0.205, 0.199)	(0.135, 0.132)	(0.403, 0.677)	(0.347, 0.736)
30dB	(0.152, 0.150)	(0.132, 0.129)	(0.406, 0.722)	(0.380, 0.770)
40dB	(0.139, 0.137)	(0.128, 0.127)	(0.410, 0.731)	(0.377, 0.774)

5. CONCLUSION

The paper proposes a new bispectrum technique to address subpixel image registration. Its main features include the capability to perform reliable image registration under low SNR environments as well as cross-correlated channel noise. This is because that the method utilizes higher-order spectra of the observed images to suppress Gaussian noise. Experimental results show that the proposed method is effective in identifying subpixel translations under different noise levels and environments.

REFERENCES

[1] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325-376, Dec. 1992.

[2] D. I. Barnea and H. F. Silverman, "A class of algorithms for fast digital image registration," *IEEE Trans. Computers*, vol. 21, no. 2, pp. 179-186,1972.

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

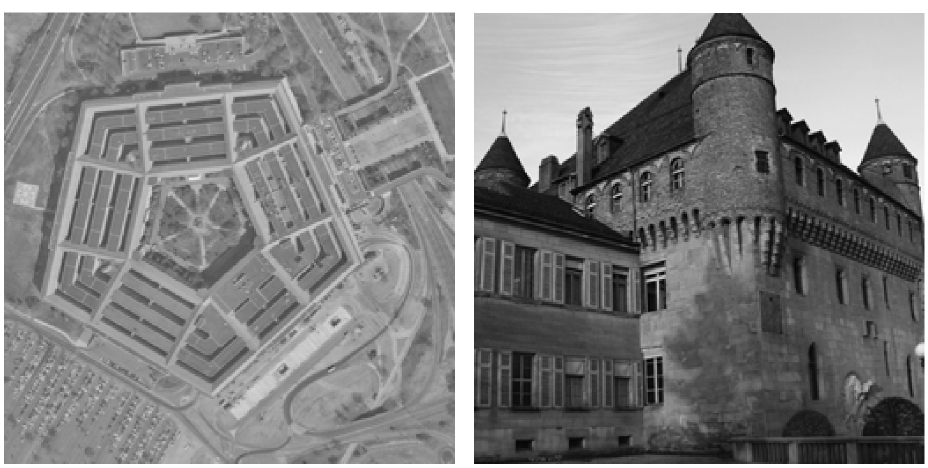
[4] H. S. Stone, M. T. Orchard, E-C. Chang, and S. A. Martucci, "A fast direct Fourier-based algorithm for subpixel registration of images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 39, no. 10, pp. 2235-2243, Oct. 2001.

[5] S. P. Kim and W. Y. Su, "Subpixel accuracy image registration by spectrum cancellation," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pp. 153-156, Apr. 1993.

[6] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Trans. Image Processing*, vol. 11, no. 3, pp. 188-200, Mar. 2002.

[7] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectral Analysis: A Nonlinear Signal Processing Framework*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[8] J. M. M. Anderson and G. B. Giannakis, "Image motion estimation algorithms using cumulants," *IEEE Trans. Image Processing*, vol. 4, no. 3, pp. 346-357, Mar. 1995.



(a) (b)

Fig. 1 Test Images. (a) "Pentagon" image, (b) "Castle" image