An Effective Technique for Subpixel Image Registration Under Noisy Conditions

IEEE Transactions on Systems, Man, and Cybernetics - Part A
vol. 38, no. 4, Jul. 2008
Li Chen, and Kim-Hui Yap
Presented by Ho-Gun Ha

School of Electrical Engineering and Computer Science
Kyungpook National Univ.
Abstract

◆ Proposed method
  – Conventional method
    • Power spectrum-based techniques
      – Second-order statistics
    • Susceptible to noise
      – Significant performance deterioration
  – Effective higher order statistics method
    • Utilizing the characteristics of bispectrum
      – Suppressing Gaussian noise
Introduction

◆ Image registration
  – Definition
    • Process to establish the point-to-point correspondence between multiple images of the same scene
      – Medical imaging
      – Super-resolution
      – Video surveillance
  – Classification of image registration methods
    • Feature-based techniques
    • Gradient approaches
    • Fourier methods
• Feature based methods
  – Detecting a salient and distinctive features
    » Edges, corners, and contours
  – Matching the correspondence between the observed images

• Gradient approaches
  – Estimation of a parameter using a system of linear equation
    » Optical flow

• Fourier methods
  – Using Fourier shift property
    » Phase factor of exponential
  – Phase correlation method
    » Taking inverse discrete Fourier transform (IDFT) of the normalized cross power spectrum
• Main advantages of frequency-domain algorithms
  – Decoupling the estimation of translation from the estimation of rotation and scaling
  – Computational efficiency due to FFT

More attractive alternative for estimating global motion than spatial-domain techniques
– Subpixel shift estimation
  • Direct subpixel registration method
    – Comparing interpolated image to reference image
      » Minimum error between image pair
      » Fast estimation using ML optimization
  • Indirect subpixel registration method
    – Feature based methods
      » Finding the congruence that minimize the error between features and targets
    – Fourier technique
      » Estimating the best fit phase plane in frequency domain
      » Evaluating the dominant peaks of the IDFT
Problem Formulation

◆ Subpixel shift estimation
  – Noise and aliasing free condition

\[ f_k(i, j) = s(i + \delta_{x,k}, j + \delta_{y,k}), \quad k = 1, 2 \]  \hspace{1cm} (1)

where \( s \) is the original image.

\( f_k \) is two images that are shifted versions of \( s \).

and \((\delta_x, \delta_y) = (\delta_{x,2} - \delta_{x,1}, \delta_{y,2} - \delta_{y,1})\) is relative translation between images, which is restricted to subpixel level \([0, 1)\).

– Fourier transform of the images
  • Shift property

\[ F_1(\omega_1, \omega_2) = F_2(\omega_1, \omega_2)e^{-j(\omega_1\delta_x + \omega_2\delta_y)} \]  \hspace{1cm} (2)

where \( F_k(\omega_1, \omega_2) = \mathcal{F}[f_k(i, j)] \) is the FT of \( f_k(i, j) \) \((k = 1, 2)\).
– Normalized cross power spectrum
  • Phase information
    – Critical role in the estimation of subpixel translation
    
    \[
    P(\omega_1, \omega_2) = \frac{F_1(\omega_1, \omega_2) F_2(\omega_1, \omega_2)^*}{|F_1(\omega_1, \omega_2) F_2(\omega_1, \omega_2)^*|} = e^{-j(\omega_1 \delta_x + \omega_2 \delta_y)}
    \]  
    (3)

– Relationship between DFT and continuous FT (CFT)
  • Nyquist-Shannon sampling theorem
    – Equivalent process in band limited or no aliasing image
    
    \[
    F^D(\omega_1, \omega_2) = \frac{1}{T_x T_y} \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} F^c \left( \frac{\omega_1 + 2\pi i}{T_x}, \frac{\omega_2 + 2\pi j}{T_y} \right)
    \]  
    (4)

where \( F^D \) and \( F^C \) denote the DFT and CFT of the image.

and \( T_x \) and \( T_y \) are the sampling periods along the vertical and horizontal axes, respectively.
Noise effect

- Various sources of noise
  - Common factors of noise
    - Photoelectric
    - Film grain
    - Quantization noises
  - Imaging devices (CCD)
    - Thermal noises
    - Shot noises

- Conditions of evident noisy effects
  - Under low lighting conditions
    - High gain of the camera
– Additive noise condition
  • Not ignoring noise term

\[ f_k(i, j) = s(i + \delta_{x,k}, j + \delta_{y,k}) + \omega_k(i, j) \quad k = 1, 2 \]  

where \( \omega_k \) is the additive noise that arises during the image formation process.

• Normalized cross power spectrum of Eq.(5)
  – Assuming that the noise \( \omega_k \) is independent of \( s \)

\[ P(\omega_1, \omega_2) = \frac{\Im \left[ R_{f_2,f_1}(\tau_1, \tau_2) \right]}{\Im \left[ R_{f_1,f_1}(\tau_1, \tau_2) \right]} \]

\[ = \frac{\Im \left[ R_{ss}(\tau_1 - \delta_x, \tau_2 - \delta_y) \right] + \Im \left[ R_{\omega_k \omega_k}(\tau_1, \tau_2) \right]}{\Im \left[ R_{ss}(\tau_1, \tau_2) \right] + \Im \left[ R_{\omega_k \omega_k}(\tau_1, \tau_2) \right]} \]

\[ = \frac{e^{-j(\omega_1 \delta_x + \omega_2 \delta_y)} + F_{\omega_k \omega_k}(\omega_1, \omega_2) / F_{ss}(\omega_1, \omega_2)}{1 + F_{\omega_k \omega_k}(\omega_1, \omega_2) / F_{ss}(\omega_1, \omega_2)} \]  

where \( R_{xy}(\tau_1, \tau_2) \equiv E[x(i, j)y(i + \tau_1, j + \tau_2)] \) is the correlation function between \( x \) and \( y \).
– Approximation of Eq.(6)

• Preconditions of the approximation
  – Neglecting $F_{\omega_1\omega_1}(\omega_1, \omega_2)$ and $F_{\omega_1\omega_2}(\omega_1, \omega_2)$
    » High SNR ($F_{\omega_1\omega_1}(\omega_1, \omega_2) \approx 0$)
    » Uncorrelated noise ($F_{\omega_1\omega_2}(\omega_1, \omega_2) \approx 0$)

• Same result of conventional Fourier method

$$P(\omega_1, \omega_2) \approx e^{-j(\omega_1\delta_x + \omega_2\delta_y)}$$

• Not satisfying the preconditions in some applications
  – Sonar images
    » High noise level
    » Correlated noise sources
Subpixel Registration in the Bispectrum Domain

- Proposed cross-bispectrum method
  - Higher order spectra
    - Higher order cumulants contain additional information that is not conveyed by the signal’s correlation or power spectrum
  - Suppressing additive Gaussian noise
    - Equal to zero of all joint cumulants of the order of >2

- Assumption
  - Nontrivial bispectrum of the original image
    - Non-Gaussian distribution with nonzero skewness
  - zero-mean Gaussian noise
    - Signal-independent random noise
– Third-order auto- and cross-cumulants of observed signals $f_1$ and $f_2$

- Zero of all joint cummulants of the order of $>2$
  - Zero-mean Gaussian noise of $w_1$ and $w_2$
  - No interference of Gaussian noise

$$R_{f_i f_i f_i} (\tau_1, \tau_2, v_1, v_2)$$

$$\begin{align*}
\Box E \left[ f_i(i,j) f_i(i+\tau_1, j+\tau_2) f_i(i+v_1, j+v_2) \right] \\
= R_{s s s} (\tau_1, \tau_2, v_1, v_2)
\end{align*}$$

$$R_{f_2 f_2} (\tau_1, \tau_2, v_1, v_2)$$

$$\begin{align*}
\Box E \left[ f_2(i,j) f_1(i+\tau_1, j+\tau_2) f_2(i+v_1, j+v_2) \right] \\
= R_{s s s} (\tau_1 - \delta_x, \tau_2, v_1, v_2)
\end{align*}$$

where $R_{s s s} (\tau_1, \tau_2, v_1, v_2) \Box E \left[ s(i,j)s(i+v_1, j+\tau_2)s(i+v_1, j+v_2) \right]$ is the third-order autocumulant of the desired signal $s$. 

\[ (7) \]
– Bispectrums

  • Computing by taking the DFT of the cumulants

\[
F_{f_1 f_1 f_1} (\omega_1, \omega_2, \nu_1, \nu_2) = \mathfrak{F} \left[ R_{f_1 f_1 f_1} (\tau_1, \tau_2, \nu_1, \nu_2) \right]
= F_{sss} (\omega_1, \omega_2, \nu_1, \nu_2)
\]

\[
F_{f_2 f_2 f_2} (\omega_1, \omega_2, \nu_1, \nu_2) = \mathfrak{F} \left[ R_{f_2 f_2 f_2} (\tau_1, \tau_2, \nu_1, \nu_2) \right]
= F_{sss} (\omega_1, \omega_2, \nu_1, \nu_2) e^{-j(\omega_1 \delta_x + \omega_2 \delta_y)}
\]

(8)

where \( F_{sss} (\omega_1, \omega_2, \nu_1, \nu_2) \) is the autobispectrum of \( s \).
– Phase information
  * Applying normalized cross bispecturm
  – More robust phase information compared with cross power spectrum in noisy environments
    - Independent of the noise distortion terms
      \[ F_{\omega_1\omega_1}(\omega_1, \omega_2) \text{ and } F_{\omega_1\omega_2}(\omega_1, \omega_2) \]

\[
P(\omega_1, \omega_2, v_1, v_2) = \frac{F_{f_2 f_2}(\omega_1, \omega_2, v_1, v_2) F_{f_1 f_1}(\omega_1, \omega_2, v_1, v_2)^*}{|F_{f_2 f_2}(\omega_1, \omega_2, v_1, v_2) F_{f_1 f_1}(\omega_1, \omega_2, v_1, v_2)^*|}
\]

\[
= e^{-(\omega_1 \delta_x + \omega_2 \delta_y)}
\] (9)
- Incorporating the Dirichlet estimation scheme
  - Reliable subpixel registration
    \[
    D(x, y) = \mathcal{F}^{-1} \left[ P(\omega_1, \omega_2, v_1, v_2) \right] = \mathcal{F}^{-1} \left[ e^{-j(\omega_1\delta_x + \omega_2\delta_y)} \right]
    \]
    \[
    = \frac{1}{MN} \frac{\sin(\pi(x - \delta_x))}{\sin(\pi(x - \delta_x)/M)} \frac{\sin(\pi(y - \delta_y))}{\sin(\pi(y - \delta_y)/N)}
    \] (10)
    where \( M \times N \) is the length of IFT.

  - Approximation of Dirichlet function to sinc function
    \[
    D(x, y) \approx \text{sinc}(x - \delta_x)\text{sinc}(y - \delta_y)
    \] (11)
    where the denominator is approximated by \( \lim_{t \to 0} \sin t \approx t \).

  - Taylor series expansion
    \[
    \sin t = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{t^{2k-1}}{(2k-1)!}
    \] (12)
– Effective method of estimating subpixel

  • Solving the nonlinear equations (11)

    – Intermediate steps

      \[
      \begin{align*}
        \frac{D(0, 0)}{D(1, 0)} &= \frac{D(0, 1)}{D(1, 1)} = \frac{\sin(\pi(1 - \delta_x) / M)}{\sin(\pi\delta_x / M)} \approx \frac{1 - \delta_x}{\delta_x}, \\
        \frac{D(0, 0)}{D(0, 1)} &= \frac{D(1, 0)}{D(1, 1)} = \frac{\sin(\pi(1 - \delta_y) / N)}{\sin(\pi\delta_y / N)} \approx \frac{1 - \delta_y}{\delta_y}.
      \end{align*}
      \]

    – Final estimated motion shift

      \[
      \begin{align*}
        \hat{\delta}_x &= \frac{1}{2} \left( \frac{D(1, 0)}{D(1, 0) + D(0, 0)} + \frac{D(1, 1)}{D(1, 1) + D(0, 1)} \right) \\
        \hat{\delta}_y &= \frac{1}{2} \left( \frac{D(0, 1)}{D(0, 1) + D(0, 0)} + \frac{D(1, 1)}{D(1, 1) + D(1, 0)} \right)
      \end{align*}
      \]
Complexity reduction

- Dividing the image into K segments
  - Obtaining the average cumulant
    - Reducing the computational cost of cumulant

\[
R_{f_1, f_1, f_1}(\tau_1, \tau_2, v_1, v_2) = \frac{1}{K} \sum_{i=1}^{K} \hat{R}_{f_1, f_1, f_1}(\tau_1, \tau_2, v_1, v_2) \tag{15}
\]

where \( \hat{R}_{f_1, f_1, f_1}(\tau_1, \tau_2, v_1, v_2) \) is the cumulant of each segment.

Fig. 1. Schematic diagram of the proposed algorithm.
Experimental Results

- Test images
  - Eight test images
  - Resolution of the original images
    - From 512 x 512 to 1704 x 1704

Fig. 2. Test images. (a) “Pentagon” image. (b) “Castle” image. (c) “NTU” image. (d) “Lighthouse” image. (e) “Singapore” image. (f) “Goldhill” image. (g) “House” image. (h) “Window” image.
– Noisy low-quality image pairs
  • Shifting with different pixels
  • Downsampling at different decimation rates
  • Adding different levels of additive white Gaussian noise
– Resolution of noisy low-quality image
  • 128 x 128

Fig. 3. Two samples of the low-quality images used in the experiments.
◆ Image degraded by AWGN
  – Pentagon image
    • 512x512 resolution
      – Downsampling factor (4x4)
    • Different noise level
      – Producing different SNR
    • Subpixel translation

Table 1. Results of sub-pixel registration in AWGN.

<table>
<thead>
<tr>
<th>SNR</th>
<th>Foroosh (0.25, 0.75)</th>
<th>Proposed (0.25, 0.75)</th>
<th>Foroosh (0.75, 0.25)</th>
<th>Proposed (0.75, 0.25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10dB</td>
<td>(0.38, 0.65)</td>
<td>(0.29, 0.68)</td>
<td>(0.64, 0.36)</td>
<td>(0.68, 0.30)</td>
</tr>
<tr>
<td>20dB</td>
<td>(0.31, 0.71)</td>
<td>(0.28, 0.74)</td>
<td>(0.71, 0.31)</td>
<td>(0.74, 0.29)</td>
</tr>
<tr>
<td>30dB</td>
<td>(0.30, 0.73)</td>
<td>(0.27, 0.74)</td>
<td>(0.74, 0.30)</td>
<td>(0.74, 0.28)</td>
</tr>
<tr>
<td>40dB</td>
<td>(0.29, 0.74)</td>
<td>(0.27, 0.74)</td>
<td>(0.74, 0.29)</td>
<td>(0.75, 0.28)</td>
</tr>
</tbody>
</table>
Image degraded by Cross-Correlated channel noise

- Castle
  - 1704x1704 resolution
    - Downsampled factor (8x8)
  - Correlated noise across the channels
    - AWGN of $w_1$
    - Generated from $w_2$ using Eq.(16)

$$\omega_2(m, n) = \sum_{i=-3}^{3} \sum_{j=-3}^{3} b(i, j) w_1(m + i, n + j)$$

where $b(i, j) = \exp(-i^2 + j^2)/2$ is a low-pass filter used to simulate the noise that are correlated.
- Different noise level
  - Producing different SNR
- Subpixel translation

Table 2. Results of sub-pixel registration in correlated noise.

<table>
<thead>
<tr>
<th>SNR</th>
<th>(0.125, 0.125)</th>
<th>(0.375, 0.75)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Foroosh</td>
<td>Proposed</td>
</tr>
<tr>
<td>10dB</td>
<td>(0.267, 0.255)</td>
<td>(0.151, 0.170)</td>
</tr>
<tr>
<td>20dB</td>
<td>(0.205, 0.199)</td>
<td>(0.135, 0.132)</td>
</tr>
<tr>
<td>30dB</td>
<td>(0.152, 0.150)</td>
<td>(0.132, 0.129)</td>
</tr>
<tr>
<td>40dB</td>
<td>(0.139, 0.137)</td>
<td>(0.128, 0.127)</td>
</tr>
</tbody>
</table>

◆ Image registration for image database
  - Large number of experiments based on all eight test images
    - Different pixel shift
    - Different noise degradations
– Pixel and subpixel-level translation
  • Two-stage coarse-to-fine algorithm
    – Identifying the pixel-level shift using conventional phase correlation
    – Identifying the subpixel-level shift using proposed method

Table 3. Average results of pixel and sub-pixel registration for image database.

<table>
<thead>
<tr>
<th>SNR</th>
<th>(2.25, 1.50) Foroosh</th>
<th>Proposed</th>
<th>(3.66, 4.33) Foroosh</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>10dB</td>
<td>(2.42, 1.69)</td>
<td>(2.39, 1.57)</td>
<td>(3.87, 4.67)</td>
<td>(3.86, 4.47)</td>
</tr>
<tr>
<td>20dB</td>
<td>(2.34, 1.67)</td>
<td>(2.37, 1.56)</td>
<td>(3.79, 4.49)</td>
<td>(3.78, 4.39)</td>
</tr>
<tr>
<td>30dB</td>
<td>(2.33, 1.64)</td>
<td>(2.29, 1.44)</td>
<td>(3.59, 4.47)</td>
<td>(3.68, 4.34)</td>
</tr>
<tr>
<td>40dB</td>
<td>(2.29, 1.57)</td>
<td>(2.26, 1.52)</td>
<td>(3.68, 4.35)</td>
<td>(3.64, 4.30)</td>
</tr>
</tbody>
</table>

– Disadvantage of proposed method
  • Higher computational time
    – 10 times Longer than second-order statistics
◆ Real-world image registration
  – Performing SR framework
    • Lack of ground truth in real-world images
    • Not evaluating an objective performance
  – Compared test methods
    • Scale-up LR image
    • Foroosh’s method
      – Sinc function
    • Effective HOS method
      – Proposed method
Fig. 4. Impact of image registration on image SR. (a) Five LR images. (b) Sample of the scaled-up LR images. (c) Reconstruction HR image using the proposed registration method. (d) Reconstruction HR image using the Foroosh’s registration method.
Conclusion

◆ Proposed method
  – New effective technique to address subpixel image registration
    • Reliable image registration
      – Low SNR environment
      – Cross-correlated channel noise
  – Utilizing higher order spectra of observed images
    • Suppressing Gaussian noise
\[ f_k(i, j) = s(i + \delta_{x,k}, j + \delta_{y,k}), \quad k = 1, 2 \]

\[(\delta_x, \delta_y) = (\delta_{x,2} - \delta_{x,1}, \delta_{y,2} - \delta_{y,1})\]

\[ F_1(\omega_1, \omega_2) = F_2(\omega_1, \omega_2)e^{-j(\omega_1\delta_x + \omega_2\delta_y)} \]

\[ F_k(\omega_1, \omega_2) = \Im[f_k(i, j)] \quad f_k(i, j) \quad (k = 1, 2) \]

\[ P(\omega_1, \omega_2) = \frac{F_1(\omega_1, \omega_2)F_2(\omega_1, \omega_2)^*}{|F_1(\omega_1, \omega_2)F_2(\omega_1, \omega_2)^*|} = e^{-j(\omega_1\delta_x + \omega_2\delta_y)} \]

\[ F^D(\omega_1, \omega_2) = \frac{1}{T_x T_y} \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} F^C \left( \frac{\omega_1 + 2\pi i}{T_x}, \frac{\omega_2 + 2\pi j}{T_y} \right) \]

\[ f_k(i, j) = s(i + \delta_{x,k}, j + \delta_{y,k}) + \omega_k(i, j) \quad k = 1, 2 \]
\[ k = \frac{D(0, 0)}{D(1, 0)} = \frac{D(0, 1)}{D(1, 1)} \]

\[
\approx \frac{\sin(\rho(1 - \delta_x))}{\sin(\rho \delta_x)} \approx \frac{(1 - \delta_x) - \frac{1}{6} \rho^2 (1 - \delta_x)^3}{\delta_x - \frac{1}{6} \rho^2 \delta_x^3}
\]

\[ \rho^2 (1 + k) \delta_x^3 - 3 \rho^2 \delta_x^2 + (3 \rho^2 - 6k - 6) \delta_x + 6 - \rho^2 = 0 \]

\[ \omega_2(m, n) = \sum_{i=-3}^{3} \sum_{j=-3}^{3} b(i, j) w_1(m + i, n + j) \]