Improved digital processing method used for image motion measurement based on hybrid opto-digital joint transform correlator

Hongwei Yi (易红伟), Hui Zhao (赵 惠)*, Yingcai Li (李英才), and Desheng Wen (汶德胜)

Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China

**E-mail: zhaohui@opt.ac.cn*

Received April 2, 2010

Hybrid opto-digital joint transform correlator (HODJTC) is effective for image motion measurement, but it is different from the traditional joint transform correlator because it only has one optical transform and the joint power spectrum is directly input into a digital processing unit to compute the image shift. The local cross-correlation image can be directly obtained by adopting a local Fourier transform operator. After the pixel-level location of cross-correlation peak is initially obtained, the up-sampling technique is introduced to relocate the peak in even higher accuracy. With signal-to-noise ratio $\geq$ 20 dB, up-sampling factor $k \geq 10$ and the maximum image shift ≤ 60 pixels, the root-mean-square error of motion measurement accuracy can be controlled below 0.05 pixels.

OCIS codes: 100.2000, 110.2960.

doi: 10.3788/COL20100810.0989.

As is widely known, the relative motion between object and camera during the exposure time can lead to image blurring, so does the camera jittering. Generally speaking, two ways can be used to eliminate motion blur. One is to compensate the motion or jittering through specific hardware, for example, high-speed steering mirror; the other is to restore the blurred images using popular algorithms such as inverse filtering, Wiener filtering, Richard-Lucy algorithm, maximum entropy, expectation-maximization (EM) algorithm, and so on[1]. No matter which way is used, motion measurement is important, because it is the first and necessary step for hardware compensation and also helps improve the software restoration quality. So, image motion measurement is not only a hot topic in academic field, but also has potentials for real applications.

Compared with other methods used for image motion measurement, optical joint transform correlation (JTC)[2,3] has many advantages, but the most significant one is that it can provide very high-precision result of motion measurement in a very high speed even with a very low signal-to-noise ratio (SNR). The input image generated by combining two consecutive images is sent to JTC; then, the correlation image is obtained with two successive optical Fourier transforms. The image shift is computed by searching the cross-correlation peak in the correlation image. In the past years, with a deeper understanding to JTC, its performance has been better analyzed and improved. Ge *et al.* analyzed the influences of under-exposure on JTC and proved its insensitiveness to a very low SNR[4]. Li *et al.* used joint wavelet transform to extract edge features from the input image to eliminate the influence of noise[5]. Fielding *et al.* improved the system capability of identifying cross-correlation peaks by adopting binary transform to the joint power spectrum (JPS) image[6]. Chen *et al.* introduces discrete cosine transform (DCT)[7] and high-pass filter[8] respectively to preprocess the JPS image to further improve the system performance. Feng *et al.* brought in amplitude-modulation method to strengthen discrimination ability of JTC[9]; but the elimination of power spectra of reference and object images, as proposed by Alam *et al.*[10], can make Feng's method better and more stable.

Up to now, different types of JTC come up, such as binary nonlinear JTC[11], amplitude-modulated JTC[12], pseudo-random phase encoded JTC[13], planar integrated JTC[14], and so on. However, the basic configuration of

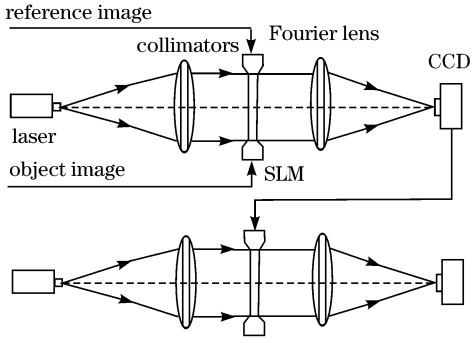

Fig. 1. Traditional system scheme of JTC. SLM: spatial light modulator.

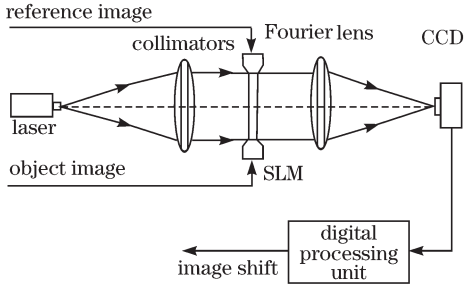

Fig. 2. System scheme of HODJTC.

JTC for measuring image motion, as shown in Fig. 1, remains nearly unchanged in a long period. But a very different configuration was proposed in 2002[15], as shown in Fig. 2. Unlike the traditional configuration, the one in Fig. 2 only contains one optical transform and the JPS data are directly sent to a digital processing unit to obtain the correlation image and image shift, and therefore it is called hybrid opto-digital joint transform correlator (HODJTC). Obviously, HODJTC is superior to traditional JTC from the viewpoint of system scheme because it has the potential in reducing system size, weight, and cost. However, in Ref. [15], a global discrete Fourier transform operator is applied to the JPS data to obtain the whole correlation image, which is believed unnecessary and thus there are spaces for further improvements.

The principle of JTC is given by

$$\begin{aligned} I_{\rm i}(x,y) =& f_{\rm r}(x,y-a) \\ &+ f_{\rm o}(x+\delta x, y+a+\delta y), \end{aligned} \tag{1}$$

$$\begin{aligned} I_{\rm c}(x,y) =& f_{\rm r}(x,y)\otimes f_{\rm r}(x,y) + f_{\rm o}(x,y)\otimes f_{\rm o}(x,y) \\ &+ f_{\rm r}\otimes f_{\rm o} * \delta(x-\delta x, y-2a-\delta y) \\ &+ f_{\rm o}\otimes f_{\rm r} * \delta(x+\delta x, y+2a+\delta y), \end{aligned} \tag{2}$$

where $I_{\rm i}$ and $I_{\rm c}$ are input and correlation images, respectively; x and y denote spatial coordinates in the horizontal and vertical directions; $f_{\rm r}$ and $f_{\rm o}$ are two consecutive images (reference and object images) with relative motion of δx and δy in two directions, and the distance between them equals a; $\otimes$ and $*$ denote the correlation and convolution operator, respectively. In Eqs. (1) and (2), we neglect the intermediate two Fourier transforms and only give the input and output results.

Clearly, the ideal location of cross-correlation peak (IPL) corresponding to the case of zero image motion can be predicted because the arrangement of two consecutive images in input image is known, as Eq. (1) tells. At the same time, the maximum image shift (MIS) between $f_{\rm r}$ and $f_{\rm o}$ can be estimated to a certain extent as well. So, with the JPS data obtained through the optical setup shown in Fig. 2, an improved digital processing procedure containing two steps can be described as follows.

Step 1:

$$\begin{aligned} F(u,v) &= {\rm DFT}_{\rm window1}(f) \\ &= W^u_{\rm window1} \times f \times W^v_{\rm window1}, \end{aligned} \tag{3}$$

where the dimension of window1 is MIS×MIS;

Step 2:

$$\begin{aligned} F(u,v) &= {\rm DFT}_{\rm window2}(f) \\ &= W^u_{\rm window2} \times f \times W^v_{\rm window2}, \end{aligned} \tag{4}$$

where the initial dimension of window2 is 3×3, but its real dimension equals 3k×3k.

In Eqs. (3) and (4), f is the JPS image whose dimension is $M\times N$; u and v are spatial frequencies; k is the up-sampling factor; DFT is the discrete Fourier transform operator and F is the correlation image computed through DFT. The dimension of DFT is restricted by the dimensions of window1 and window2, and thus DFT becomes a local Fourier transform operator here, as discussed below.

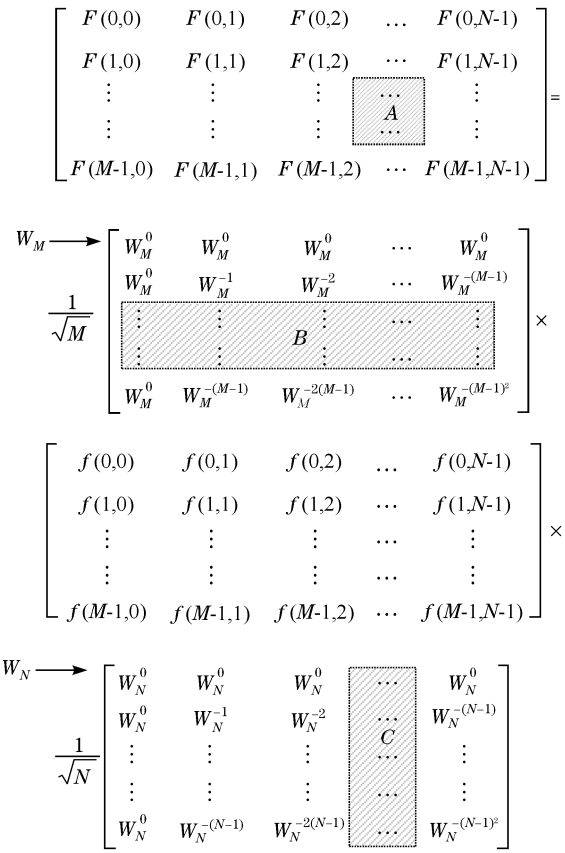

Fig. 3. General explanations of the two-step procedure described by Eqs. (3) and (4). The general expression of components in W_M and W_N can be rewritten as $W^u_M = \exp[{\rm i}2\pi ux/M]$ and $W^v_N = \exp[{\rm i}2\pi vy/N]$, respectively.

From Eqs. (3) and (4), we can see that DFT can be represented as matrix multiplication and all parameters in the two equations can be concretely rewritten as matrices, as shown in Fig. 3, where the general expression of the components in W_M and W_N is given in the corresponding caption. In original HODJTC, Fig. 3 is seen as a whole and all components in the operator matrices W_M and W_N, are used to obtain the entire correlation image F, within which the cross-correlation peak is globally searched to obtain the image shift. But in fact, only the local cross-correlation region, supposed to be A, is of our concern. The matrix representation of DFT gives us a powerful tool to directly obtain this local region. As shown in Fig. 3 clearly, the sub-matrix A can be expressed as

$$A = \left(\frac{1}{\sqrt{M}}B\right) \times f \times \left(\frac{1}{\sqrt{N}}C\right). \tag{5}$$

According to Eq. (5), if we want to obtain the local correlation image, partial components in W_M and W_N, denoted as B and C, are enough. Because the location of cross-correlation peak corresponding to zero image motion and the maximum image shift can be considered as priors, the dimensions of A, B, and C are deterministic before doing the computation. In the first step, the center of A is IPL and its dimension, corresponding to

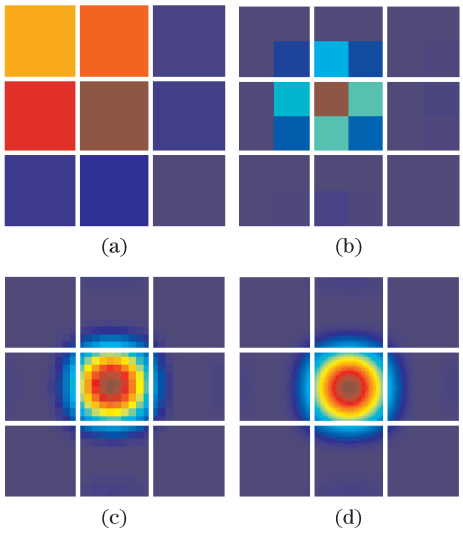

Fig. 4. Up-sampled image of cross-correlation peak with different up-sampling factors k. (a) $k = 1$; (b) $k = 2$; (c) $k = 10$; (d) $k = 50$.

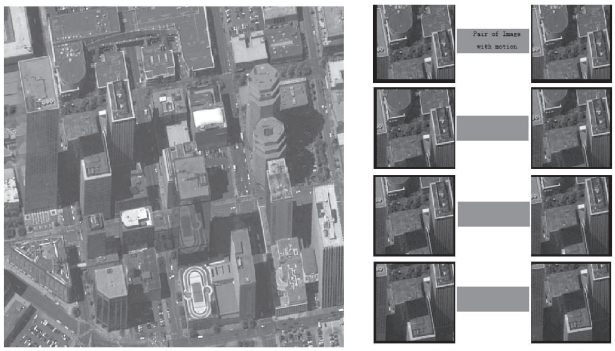

Fig. 5. (a) A high resolution remote sensing image (512×512) and (b) the generated image pairs sent to HODJTC. The dimension of the small image is 128×128; the distance between any pair of images in the input image is 256 pixels, and the size of the simulated input image is 1024×768.

window1, is MIS×MIS. Therefore, the dimensions of B and C are MIS×M and N×MIS, respectively. With the computed local cross-correlation image by using Eq. (5), the real cross-correlation peak (RPL) is searched within A, but only pixel-level accuracy of motion shift determination can be obtained, as Fig. 4(a) tells. In the second step, RPL is selected as the center of A and its initial dimension is quite small, for example, 3×3, corresponding to window2. In order to obtain the local cross-correlation data in a higher resolution, an up-sampling technique is introduced[16]. Let k be the up-sampling factor, then the dimensions of B and C are actually changed from 3×M and N×3 to 3k×M and N×3k. With deterministic M and N, the original number of components in B and C is fixed, so certain method, such as inner-interpolation or zero-padding, is used to obtain the additional components. The additional components are spread along vertical direction for B and along horizontal direction for C. Figure 4 gives the result of up-sampling. Clearly, the bigger the up-sampling factor is, the finer the local cross-correlation image will be. By searching the peak in the up-sampled image, $1/k$ pixel-level accuracy can be obtained and this can greatly improve the performance of HODJTC. Besides the improvement of measurement accuracy, the influence of auto-correlation peak can also be effectively avoided because the computed up-sampled local cross-correlation image is far enough away from the powerful auto-correlation peak.

In order to assess the performance of the method raised above, we design the virtual HODJTC and numerical simulations are carried out. Firstly, a series of image pairs with relative motion in horizontal and vertical directions is generated from a real high resolution image, as displayed in Fig. 5; secondly, the image pairs are combined together to be input into the virtual HODJTC; thirdly, with the new digital processing procedure proposed here, the root-mean-square (RMS) error of image motion determination is computed, three types of analysis are carried out, and the simulation results are shown in Tables 1−3.

In Tables 1 and 2, the introduced image motion is ≤1 pixel and the influences of up-sampling factor and SNR on RMS are provided. Clearly, with the increase of k, the RMS value decreases; but when k is far bigger than 10, the RMS value is less than 0.045 pixels and does not change too much. In other words, it is not necessary to make k very big, 10 is enough for application. At the same time, the RMS value increases with the decrease of SNR. With SNR within the ranges of [20 dB, 50 dB], [10 dB, 20 dB], and [5 dB, 10 dB], respectively, the corresponding RMS gradually increases from 0.045 to 0.19 pixels, which is still acceptable[17]. So, Table 2

Table 1. RMS Error of Image Motion Measurement with Different Up-Sampling Factors k; Simulated Image Motion is Less than 1 Pixel

k	Horizontal Direction RMS Error (pixel)	Vertical Direction RMS Error(pixel)
10	0.041	0.044
20	0.032	0.037
40	0.031	0.036
50	0.030	0.033
80	0.031	0.036
100	0.031	0.036

Table 2. RMS Error of Image Motion Measurement with Different SNRs; the Input Image is Added Gaussian Noise to Simulate Different Requirements of SNR; Simulated Image Motion is Less than 1 Pixel

SNR (dB)	Horizontal Direction RMS Error (pixel)	Vertical Direction RMS Error(pixel)
50	0.036	0.029
40	0.036	0.029
30	0.037	0.031
20	0.041	0.041
10	0.130	0.132
5	0.180	0.189

Table 3. RMS Error of Image Motion Measurement with Different Ranges of Image Motion; Up-Sampling Factor $k = 10$ and SNR = 20 dB

Range of Image Motion (pixel)	RMS Error in Horizontal or Vertical Direction (pixel)
1−10	0.018
10−20	0.032
20−30	0.036
30−40	0.045
40−50	0.057
50−60	0.051
60−70	0.114
70−80	5.240
80−90	13.691
90−100	21.491

proves that the proposed method is still suitable in a very low SNR environment. Factually, 1 pixel image motion is not the usual case, so the motion range within which the method is still effective is also analyzed, as shown in Table 3. As Table 3 tells, when the range of image motion is less than 60 pixels and the SNR is bigger than 20 dB, the corresponding RMS error can be controlled below 0.05 pixels and the maximum allowable image shift can even approach 70 pixels.

In conclusion, we propose a new digital processing method applied for HODJTC and demonstrate its effectiveness with numerical evaluations. HODJTC can simplify the system structure, and our method can even improve its performance because the computation load is reduced a lot and the accuracy of cross-correlation peak computation has also been improved. However, with the increase of up-sampling factor, the calculation load also increases. So, there exists a tradeoff between acceptable measurement accuracy and up-sampling factor k, which should be further studied. At the same time, real experiments should be carried out to further demonstrate the effectiveness of the proposed method.

This work was supported by the National "973" Program of China under Grant No. 2009CB724006.

References

1. H. C. Andrews and B. R. Hunt, *Digital Image Restoration* (Prentice-Hall, Englewood Cliffs, 1977).
2. S. Jutamulia, Proc. SPIE **1812,** 233 (1992).
3. C. S. Weaver and J. W. Goodman, Appl. Opt. **5,** 1248 (1966).
4. P. Ge, Y. Chen, Q. Li, H. Feng, Z. Xu, and Z. Zheng, Acta Opt. Sin. (in Chinese) **29,** 1796 (2009).
5. J. Li, Y. Zhang, and J. Hu, Chinese J. Lasers (in Chinese) **22,** 783 (1995).
6. K. H. Fielding, J. L. Horner, and C. K. Makekau, Opt. Eng. **30,** 1958 (1991).
7. X. Chen, H. Chen, R. Hong, and Z. Chen, Acta Opt. Sin. (in Chinese) **20,** 931 (2000).
8. X. Huang, G. Cao, and H. Lai, Acta Opt. Sin. (in Chinese) **16,** 866 (1996).
9. D. Feng, H. Zhao, and S. Xia, Opt. Commun. **86,** 260 (1991).
10. M. S. Alam and D. Chain, Opt. Eng. **39,** 1203 (2000).
11. B. Javidi and J. Wang, Proc. SPIE **1347,** 385 (1990).
12. M. S. Alam, Proc. SPIE **4470,** 53 (2001).
13. A. Alsamman, Proc. SPIE **5816,** 252 (2005).
14. J. Liu, P. Xu, H. Xu, X. Zhang, B. Wang, and J. Yang, Acta Opt. Sin. (in Chinese) **28,** 1735 (2008).
15. A. Serrano-Heredia, C. M. Hinojosa, P. Hinojosa, R. Rodríguez-Dagnino, L. Molina-Hernández, R. Briones, R. Ponce, and M. Jolibois, Proc. SPIE **4772,** 136 (2002).
16. R. Soummer, L. Pueyo, A. Sivaramakrishnan, and R. J. Vanderbei, Opt. Express **15,** 15935 (2007).
17. H.-S. Wong, Y. L. Yao, and E. S. Schlig, IBM J. Res. Develop. **36,** 83 (1992).