A Motion Detection Approach based on UAV Image Sequence

Hong-Xia Cui[1*], Ya-Qi Wang[1*], FangFei Zhang[2], TingTing Li[1]
[1]College of Information Science and Technology, Bohai University
Jinzhou, 121013, China
[2]Institute of Foresty, Beijing Foresty University
Beijing, China,100083
[*]Corresponding author: Hong-Xia Cui, Ya-Qi Wang

Received October 8, 2016; revised October 18 , 2017; accepted December 28, 2017; published March 31, 2018

Abstract

Aiming at motion analysis and compensation, it is essential to conduct motion detection with images. However, motion detection and tracking from low-altitude images obtained from an unmanned aerial system may pose many challenges due to degraded image quality caused by platform motion, image instability and illumination fluctuation. This research tackles these challenges by proposing a modified joint transform correlation algorithm which includes two preprocessing strategies. In spatial domain, a modified fuzzy edge detection method is proposed for preprocessing the input images. In frequency domain, to eliminate the disturbance of self-correlation items, the cross-correlation items are extracted from joint power spectrum output plane. The effectiveness and accuracy of the algorithm has been tested and evaluated by both simulation and real datasets in this research. The simulation experiments show that the proposed approach can derive satisfactory peaks of cross-correlation and achieve detection accuracy of displacement vectors with no more than 0.03pixel for image pairs with displacement smaller than 20pixels, when addition of image motion blurring in the range of 0~10pixel and 0.002variance of additive Gaussian noise. Moreover,this paper proposes quantitative analysis approach using tri-image pairs from real datasets and the experimental results show that detection accuracy can be achieved with sub-pixel level even if the sampling frequency can only attain 50 frames per second.

Keywords: Fourier transform, fuzzy edge detection, motion detection, UAV, photogrammetric mapping and remote sensing

The work is sponsored by Chinese Natural Science Funds (No.41371425) and Provincial Natural Science Foundation of Liaoning Province,China under Grant no.20170540015

http://doi.org/10.3837/tiis.2018.03.014 **ISSN : 1976-7277**

1. Introduction

With rapid development of unmanned aerial vehicles(UAVs), more and more UAVs are equipped with devices such as still or video cameras, communication tools and other sensors to conduct many activities across different fields such as surveillance and object tracking, environmental monitoring, agriculture operation, photogrammetric mapping and remote sensing, scientific research and so on[1-5]. In photogrammetric and remote sensing applications, it is essential to conduct motion detection aiming at motion compensation and even navigation using real time or off-line processing techniques. For traditional aerial or satellite remote sensing and photogrammetry applications, accurate motion detection can be realized by expensive POS(Position and Orientation System) integrated with GPS and IMU sensors onboard flying platforms[6-8]. With development of images sensors, vision-based approaches have been attractive alternatives for motion measurement[9-15]. As one of the most robust and effective vision-based methods, Joint Transform Correlation(JTC) based algorithms have shown remarkable promise for motion detection from consecutive image frames and proved by simulation tests in airborne and satellite remote sensing as well as some other applications[9,10,11,12]. However, the computability, accuracy and robustness of conventional JTC algorithm may be influenced by images with degraded image quality, complex background or low-contrast texture, etc. Previous research works on improvement of conventional JTC algorithm can be divided into two categories, namely, preprocessing of input images with conventional edge detection methods[12]in spatial domain and joint power spectrum with enhancement , subtraction or binarization methods in frequency domain[16][17]. To overcome the problems of huge amount of calculations, specified optical devices known as Joint Transform Optical Correlators (JTOCs)[9-12,17,18] have been developed and even incorporated with the above improvement techniques in simulation, laboratory and ground tests.

Due to limitation of payload , small size capility of most UAV systems and complex installation of hardware devices, the existing JTOCs can't be installed on UAV-based camera systems for in-flight motion detection. Indeed, it is also suitable to realize off-line motion detection for off-line motion analysis with low altitude images for aerial photogrammetric and remote sensing applications. However, platform motion and complex vibrations during aerial photographing may cause image blurring, what results in degradation of the modulation transfer function (MTF) and geometrical distortions of the obtained images from high resolution cameras, which allow in principle a ground pixel resolution of less than 0.10 meter at 500m above ground [1-3][19]. Such a topic, due to illumination fluctuation, motion blurring and background disturbance in low-altitude images, conventional and the conventional improved JTC may be invalid .

To deal with above problems, this paper proposed an alternative modified JTC method by employing the following preprocessing methods. Firstly, due to low-altitude images generally may contain edges with different levels of blurring, conventional Gradient-based methods of edge detection [11][20][21] face the problem due to edge localization in the images that exhibit smooth transition in gray level. To deal with this ambiguity and vagueness in edge structures, a novel fuzzy edge detection method is proposed and applied to preprocess the input images for joint transform correlation in this paper. Secondly, cross-correlation power spectrums are extracted by eliminating the disturbance of self-correlation items from joint power spectrum in frequency domain and transformed by inverse two-dimensional discrete Fourier transform so as to produce quite sharp cross-correlation peaks.

Then, location of peaks at integer pixel level can be refined at sub-pixel level with the weighted centroid descriptor .Experiments with simulation and realistic datasets are carried out using stereo-image pairs and tri-image pairs, respectively. Quantitative analysis is given to evaluate the performance of the developed method. The rest of the paper is organized as follows. Section 2 briefly reviews some works related to displacement detection from joint transform correlation principle. Section 3 presents conventional joint transform correlation principle and proposes an improved joint transform correlation algorithm. Section 4 evaluates the performance of the improved algorithm through experiments with simulation and real low-altitude images. Section 5 analyzes the proposed motion detection method. At the end, Section 6 concludes the whole article and presents outlooks of our future works.

2. Related Work

Over the past two decades, considerable research efforts have been devoted to improve JTC techniques and develop sensors of JTOCs with extremely fast processing ability in many research fields and industrial areas aiming at motion measurement by detailed investigations and simulation testing [9-12,16,17,18,22,23]. In[9,12,18],the techniques of in-situ focal plane image motion measurement based on joint transform correlation have been proposed and proved by detailed investigations and experimental airborne testing. In [10], the authors proposed to determine image motion with an onboard optical correlator and used the image motion data as visual feedback signal for navigation purpose and presented preliminary performance results based on a high fidelity simulation of the complete optical processing chain. In[24], measurement of vibration using images was performed in real-time way with a embedded optical and electronic device based on a refined phase-correlation method and verified in laboratory tests. In principle, the vector of mutual shift between two images can be determined by locating the bright correlation peaks in the correlation image. To generate sharp peaks even for low SNR (signal to noise ratio) dark images, many modified JTC-based approaches have been focused on preprocessing of input images and joint power spectrums.

On one hand, Qian *et al.*[11] proposed a JTC-based system by extracting the edge of the input images based on Canny operator and has verified with simulation data and achieved satisfactory results with precise edge location and strong reliability for satellite images. However, it is difficult to derive suitable Gauss parameters and dual threshold needed by Canny operator[20] for edge detection of images with blurring. The false edges may be produced and details for true edges may disappear, which causes failure in motion detection from images. In order to extract image edges with blurring, it is suitable to apply fuzzy theory based edge detection method which own the ability to handle the edge with the ambiguity caused by blurring. In [25], Pal and King proposed fuzzy edge detection algorithm based on fuzzy sets theory and it has been applied to many applications in image processing. The central idea is to increase the use of fuzzy Enhancement contrast between different regions, which can extract fuzzy edges. Many research works are still going on in this area to make improvements of conventional Pal-King algorithm. Pushpa *et al.* [26]proposed an adaptive approach of contrast enhancement with simple membership function for fuzzy transformation. Madasu *et al.* [27] proposed Gaussian membership function for fuzzification and general intensification operator (GINT) for modification of property plane. Loo *et al.* [28] proposed a fuzzy edge detection method by adopting logarithmic function for mapping pixel image to fuzzy image and trigonometric function for property domain modification. Considering that edge detection results may vary by using different fuzzy theory based methods and it is difficult to be

evaluated, this paper proposed a modified Pal-King algorithm to deal with edge extraction of low-altitude images and applied in the proposed virtual sensor system.

On the other hand, to preprocess joint power spectrum, various techniques, such as Fourier plane enhancement[16], correlation plane subtraction and phase-shift power spectrum subtraction, have been found to be able to avoid disturbances from self-correlation and produce better correlation performance when compared to conventional JTCs or JTOCs. In [28],the fringe-adjusted filter devices are incorporated in JTOCs for the case of noise-free single and multiple target detection under varying illumination conditions of the input scene. As far as the virtual sensor is considered in this paper, self-correlation disturbance can be easily eliminated by only using the power spectrum of cross-correlation parts. Although many research have investigated a wide range of modifications for JTC or JTOCs , the majority have achieved satisfactory performance with simulation datasets . Very limited research for real-time motion detection using JTC based algorithms has been conducted with images recorded by UAVs , simply because the limited size payload ,power energy and communication chains of UAVs .

In this paper, to ensure the image quality for photogrammetric and remote sensing applications, this paper developed a modified JTC algorithm to conduct motion measurement along forward flight direction for supplying reliable techniques needed by further motion compensation.

3.JTC-based Methods

Supposing that two images with overlapping areas taken with a camera at two close time instants during motion, the second image will be shifted with respect to the first by a displacement vector. In principle, the vector can be effectively determined by JTC approach using two subsequent 2D-Fourier transforms. As described in above two sections, JTC approach can be realized by optical and electronic devices, namely, JTOC sensors. In this paper, conventional JTC algorithm and modified JTC algorithm were realized through software and used in off-line photogrammetric and remote sensing applications, which were named as virtual sensors. After introduction of conventional JTC algorithm, the modified strategies would be proposed and applied in motion detection with low-altitude images from UAV systems in this section.

3.1 Conventional JTC algorithm

Conventional JTC algorithm can be realized in two steps. In spatial domain, supposing that input image is composed of reference image $f_1(x,y)$ and target image $f_2(x,y)$ centered at positions $(-a,0)$ and $(a,0)$,respectively.

Let $(\Delta x, \Delta y)$represent displacement vector of $f_2(x,y)$ relative to $f_1(x,y)$, joint input image $f(x,y)$ can be expressed as :

$$f(x,y) = f_1(x+a,y) + f_2(x-a-\Delta x, y-\Delta y) \tag{1}$$

The Fourier transform of $f(x,y)$ can be derived as follows:

$$\begin{aligned} T(u,v) &= \iint \left[f_1(x+a,y) + f_2(x-a-\Delta x, y-\Delta y) \right] \exp\left[-j2\pi(xu+yv) \right] dxdy \\ &= T_1(u,v) \cdot \exp(j2\pi ua) + T_2(u,v) \cdot \exp\left\{ -j2\pi \left[u(a+\Delta x) + v\Delta y \right] \right\} \end{aligned} \quad (2)$$

Where (u, v) are frequency coordinates, * denotes convolution operation; $T_1(u,v)$, $T_2(u,v)$ and $T(u,v)$ represent discrete Fourier transform of image $f_1(x,y)$, $f_2(x,y)$ and $f(x,y)$ respectively. In frequency domain ,the joint power spectrum (JPS) of Equation (2) can be calculated as following:

$$|T(u,v)|^2 = |T_1(u,v)|^2 + |T_2(u,v)|^2 + |T_s(u,v)|^2 \quad (3)$$

Where

$$|T_1(u,v)|^2 = T_1(u,v) \cdot T_1^*(u,v)$$

$$|T_2(u,v)|^2 = T_2(u,v) \cdot T_2^*(u,v)$$

$$|T_s(u,v)|^2 = T_1(u,v) \cdot T_2^*(u,v) \exp\left\{ j2\pi \left[u(2a+\Delta x) + v\Delta y \right] \right\} + T_1^*(u,v) \cdot T_2(u,v) \exp\left\{ -j2\pi \left[u(2a+\Delta x) + v\Delta y \right] \right\}$$

Then, joint correlation output $c(x,y)$ which consists of pairs of cross-correlation output and self-correlation output can be derived in spatial domain by the second Fourier transformation of JPS as follows:

$$\begin{aligned} c(x,y) &= \iint |T_1(u,v)|^2 + |T_2(u,v)|^2 + |T_s(u,v)|^2 \exp\left[-j2\pi(ux+vy) \right] dudv \\ &= c_1(x,y) + c_2(x,y) + c_3(x,y) + c_4(x,y) \end{aligned} \quad (4)$$

Where $c_1(x,y) = f_1(x,y) \otimes f_1(x,y)$

$c_2(x,y) = f_2(x,y) \otimes f_2(x,y)$

$c_3(x,y) = f_1(x,y) \otimes f_2(x,y) * \delta(x-2a-\Delta x, y-\Delta y)$

$c_4(x,y) = f_2(x,y) \otimes f_1(x,y) * \delta(x+2a+\Delta x, y+\Delta y)$

in which δ is the delta-like function;$\otimes$ denotes correlation operation; * denotes convolution operation; the first two items represent the self-correlation parts centered at the origin point; the third and the fourth items represent the cross-correlation output centered at $(2a+\Delta x, \Delta y)$ and $(-2a-\Delta x, -\Delta y)$,respectively .

Finally, the motion vector between two images can be determined by finding the displacement of location of the maximum cross-correlation peak. Then, a virtual sensor based on conventional JTC algorithm can be realized through software, namely, VCJTC. However, as shown in **Fig.1**, the lower quality of input images may produce correlation output with broad self-correlation parts and a pair of cross-correlation output without obvious peaks from conventional JTC algorithm , which can affect accurate detection of cross-correlation peaks.

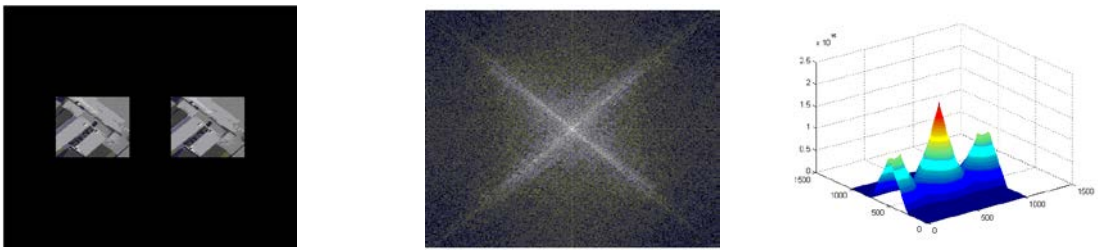


Fig. 1. Workflow of VCJTC. (a)Original input image;(b) joint power spectrum;(c) 3-dimensional output of self-correlation and cross-correlation.

3.2 A modified JTC algorithm

In order to overcome the limitation of conventional JTC algorithm, a modified JTC algorithm is proposed by introducing improved strategies. In spatial domain, the input images are preprocessed by extracting the edge of input images with a proposed fuzzy edge extraction approach. In frequency domain, to verify the sharpness of cross-correlation peaks, only cross-correlation power spectrum are extracted by eliminating the disturbance of self-correlation items from joint power spectrum. Finally, location of cross-correlation peaks at integer pixel level can be refined at sub-pixel level by weighted centroid descriptor. The modified JTC algorithm can also be realized with software ,namely,VMJTC(see **Fig. 2)**.

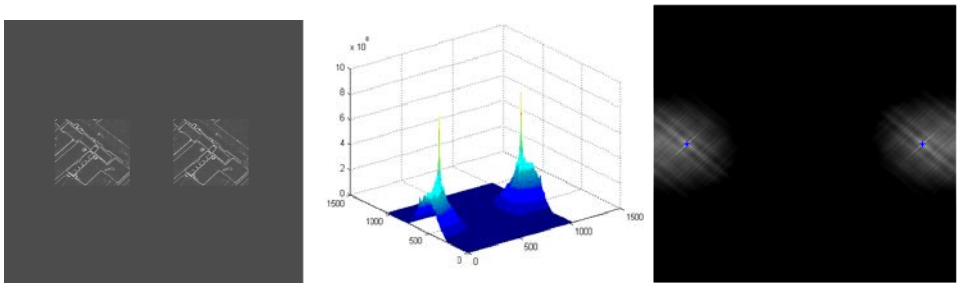


Fig. 2. Workflow of VMJTC. (a)input image;(b) 3-dimensional output of correlation peaks; (c)detection of central points for cross- correlation peaks

3.2.1 An impoved fuzzy edge detection method

Main steps in image processing using the concept of fuzzy sets includes the fuzzy processing from spatial domain to fuzzy domain by fuzzy membership function , enhancement processing on the fuzzy plane and inverse transform from the fuzzy plane to pixel plane.

Let X be an image of size M×N with gray levels and x(i,j) represents intensity element of the (i,j) th pixel which has to be mapped to fuzzy characteristic plane in the first step. Image X can be expressed as follows:

$$X = \left\{ \mu_{ij}, x_{ij} \right\}$$

Where i = 1,2,…,M, j = 1,2,…,N , $0 \leq \mu_{ij} \leq 1$, and μ_{ij} represents the membership grade of the (i, j) th pixel intensity x_{ij} in the input image. Thus, in the fuzzy image plane representation, pixel values between 0 to 255 are mapped between 0 to 1. In order to perform transformation, different membership functions have been developed in[25-28].

In [25], S.K. Pal and R.A. King firstly applied the fuzzy logic theory to the image edge detection and proposed the membership function as follows:

$$\mu_{ij} = G\left(x_{ij}\right) = \left[1 + \frac{x_{\max} - 1 - x_{ij}}{F_d}\right]^{-Fe} \tag{5}$$

Where F_d and F_e represent the adjusted parameters, $x_{\max}$ is the maximum gray level and x_{ij} the (i, j) th pixel intensity. It can be seen that the gray values in spatial domain from $x_{\min}$ to $x_{\max}$ are transformed to the values of fuzzy domain between "*a*" to 1 by the membership Equation 5, in which "a" approximates to zero .In this way, some low gray levels would be roughly neglected and defined as zero. In [26],a simple function is adopted for the conversion of spatial domain into the fuzzy domain in Equation 6.

$$\mu_{ij} = G\left(x_{ij}\right) = \frac{x_{ij}}{x_{\max}} \tag{6}$$

In [29], another simple S-shape function for fuzzification is introduced as follows

$$\mu_{ij} = G\left(x_{ij}\right) = \begin{cases} \frac{1}{2}\left[\frac{x_{ij} - x_{\min}}{T - x_{\min}}\right]^2 \\ 1 - \frac{1}{2}\left[\frac{x_{ij} - x_{\max}}{T - x_{\max}}\right]^2 \end{cases} \tag{7}$$

Where T represents the threshold . However, the simpler membership functions may decrease the hierarchical structure of complex objects in aerial images, which cannot meet the demand of different types of aerial images. Thus, the current study proposed the modified Pal-king fuzzy membership function with equation 8.

$$\mu_{ij} = G(x_{ij}) = [\log_2(1 + (\frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}})^{\alpha})]^{\alpha} \qquad x_{\min} \le x_{ij} \le x_{\max} \tag{8}$$

Where μ_{ij} represents the membership grade of pixel (i, j) intensity value x_{ij} ; $i = 1, 2..M$; $j = 1, 2N$; $x_{\min}$ is the minimum gray level; $x_{\max}$ is the maximum gray level; α is the exponential factor, where $\alpha \ge 2$. It can be seen that pixel values between $x_{\min}$ to $x_{\max}$ in spatial domain are mapped between 0 to 1 in fuzzy domain. The relationship between spatial domain and the corresponding fuzzy domain from four membership transformation functions are shown in **Fig. 3**.

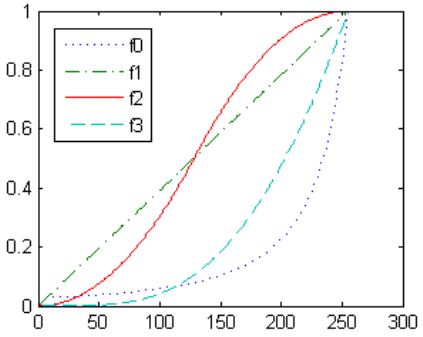


Fig. 3. Relationship between spatial domain and fuzzy domain

Where f_0 represent the function proposed by S.K. Pal and R.A. King [25], f_1 the simple ramp membership function [26] , the simple S-shape membership function f_2 [29] and f_3 the membership function proposed in this paper. It can be seen that the shape of proposed function is more similar to that of Pal-king algorithm, so the advantages of Pal-King membership function can be acquired . In the same time, the value domain of the proposed membership function ranges from 0 to 1 , the drawbacks of Pal-King fuzzy membership function can be avoided in ensuring all gray levels being transformed to fuzzy domain. Then, the same as Pal-King algorithm , transformation processing is built up by applying iteration function and intensification operator[25]. The iteration function can increase the enhancement with parameter *r* as following .

$$\mu_{ij}^{'} = T_r\left(\mu_{ij}\right) = T_1\left(T_{r-1}\left(\mu_{ij}\right)\right) \qquad r = 1,2,3...$$

where $\mu_{ij}^{'}$ is the membership function which had been enhanced, r is the iteration times of the fuzzy enhancement function T_r .Aiming at enhancement of the contrast of gray level, the fuzzy enhancement model can be defined as follows:

$$\begin{cases} \mathrm{T_r}\left(\mu_{ij}\right) = \mu_{ij}\dfrac{\mu_{ij}}{\mu_e} & ,0 \le \mu_{ij} \le \mu_e \\ T_r\left(\mu_{ij}\right) = \dfrac{\mu_{ij}(2-\mu_{ij})-\mu_e}{1-\mu_e} & ,\mu_e \le \mu_{ij} \le 1 \end{cases} \tag{9}$$

where μ_e represents the crossover parameter in fuzzy domain and can be derived by $\mu_e = G(x_e)$,in which x_e represents the crossover point of an image in spatial domain and can be determined based on the region segmentation of Otsu[30] algorithm. Second, the inverse model is proposed to transform fuzzy feature space into space domain as follows.

$$x_{ij}' = \mathrm{G}^{-1}(\mu_{ij}^{'}) = x_{\min} + (x_{\max} - x_{\min})[(2\mu_{ij}^{'1/\alpha} - 1)^{1/\alpha}] \tag{10}$$

Where x_{ij}' represents the enhanced gray value of image pixel (i, j). Finally, the edges are extracted by "Min" or "Max" operator [25].After edge detection processing, the images will be transformed by joint transform correlation method.

3.2.2 Detection of motion displacement with sub-pixel accuracy

To eliminate the disturbances caused by the zero-frequency components, Fourier plane image subtraction method is often employed for joint power spectrum in frequency domain. For the proposed modified joint transform algorithm in this paper, only the cross-correlation power spectrum items were converted back to spatial domain as equation(11) directly:

$$c(x,y) = f_1(x,y) \otimes f_2(x,y) * \delta(x-2a-\Delta x, y-\Delta y) + f_2(x,y) \otimes f_1(x,y) * \delta(x+2a+\Delta x, y+\Delta y) \quad (11)$$

Then, the location of the cross-correlation peak can be acquired by with integer pixel level by searching the maximum value in the cross-correlation output matrix . As shown in **Fig. 2**, depending on the above proposed edge detection method and elimination of zero-order correlation items, the peaks of cross-correlation can be quite sharp and without disturbances from self-correlation. Because the zero order self-correlation parts have been removed, pixels in the local region of (*M/2,N/2*) around the peak point (i_k, j_k) in cross-correlation output plane are choosed to compute the centriod (x_k, y_k) as follows.

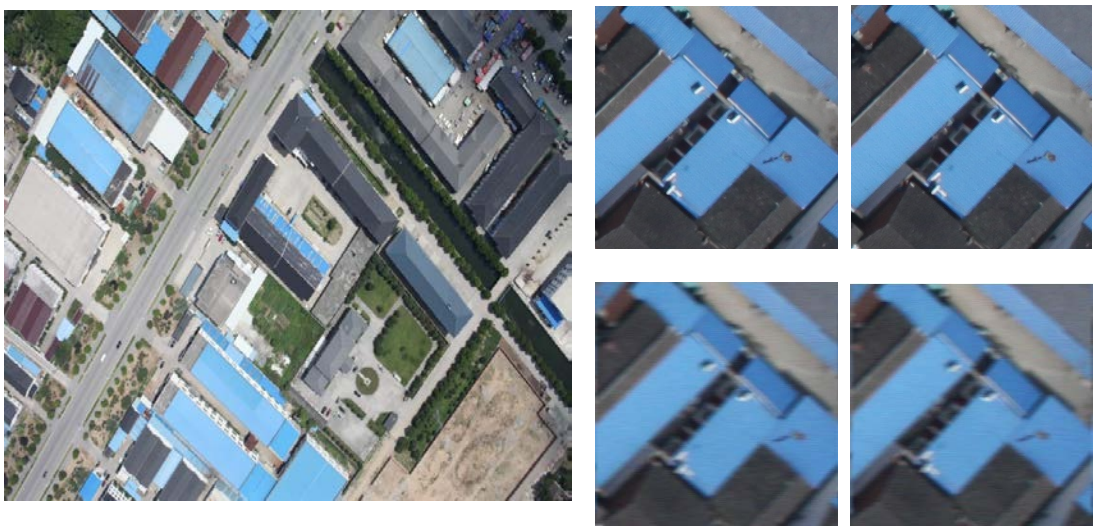
$$x_k = \frac{\sum_{i=i_k-H/2}^{i_k+H/2} \sum_{j=j_k-W/2}^{j_k+W/2} c(i,j) * j}{\sum_{i=i_k-H/2}^{i_k+H/2} \sum_{j=j_k-W/2}^{j_k+W/2} c(i,j)}, y_k = \frac{\sum_{i=i_k-H/2}^{i_k+H/2} \sum_{j=j_k-W/2}^{j_k+W/2} c(i,j) * i}{\sum_{i=i_k-H/2}^{i_k+H/2} \sum_{j=j_k-W/2}^{j_k+W/2} c(i,j)} \quad (12)$$

Where M and N are the width and height of the kth($k=1,2$) cross-correlation output, respectively; $c(i,j)$ is the cross-correlation output value of the (i,j)th pixel; (i_k, j_k) and (x_k, y_k) represent the location of the peak at integer pixel level and sub-pixel level ,respectively. Finally, the sub-pixel level displacement vector can be eventually obtained from a pair of cross-correlation peak points as follows:

$$\Delta x = \frac{1}{2}(x_1 + x_2) \ , \Delta y = \frac{1}{2}(y_1 + y_2) \quad (13)$$

4 Experiments and discussions

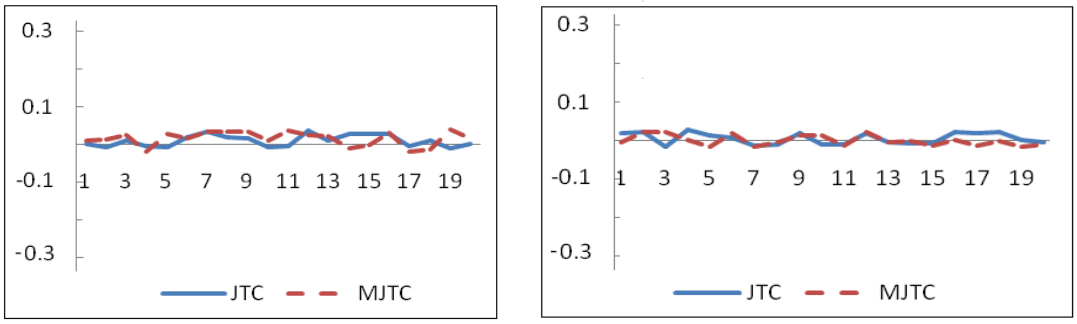
Experiments on simulation and real data were carried out to evaluate the performance of the proposed the virtual sensor of VMJTC and VCTJC. As shown in **Fig. 4**, an aerial image was derived from the UAV-based aerial camera with size of 4368×2912pixel. The simulation images were achieved with a fixed-size bounding box (window) sliding over the entire image. Every time the bounding box moved, the portion of the image within its domain was extracted as a sub-image. The bounding box started with a selected area of the image and moved 1.0 pixel horizontally at each step until 40 sequential sub-images were achieved as simulation images .



（a） （b）

Fig. 4. Low altitude image.(a) original image
(b) a pair of sub-images and the simulated images with noise and forward motion blurring

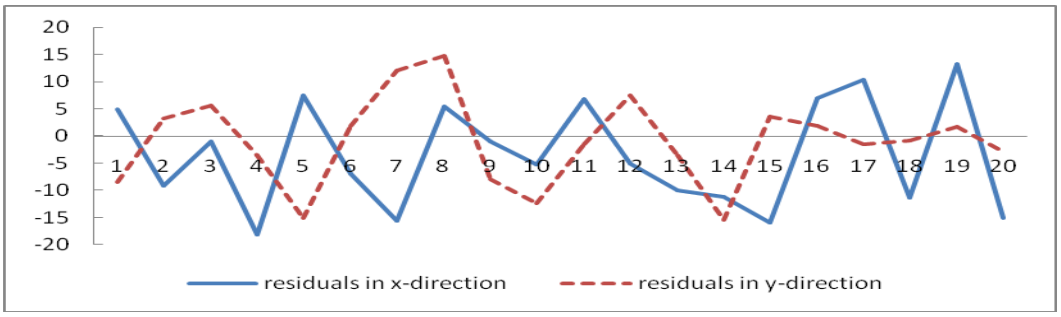
Every two images P_i, P_j ($i = 1, 2..20$; $j = 21, 22..40$)with forward motion displacement of 20 pixels in x-direction were taken as an image pair. Total 20 image pairs were collected and Gaussian noise at 0.002 variances was added to each image. Supposing that the flight direction was along x direction, image blurring caused by forward motion with bl ($bl = 1, 2..10$) pixels was added to all images. In this way, total 200 stereo-image pairs were simulated with different noise and motion blurring. Firstly, two sets of displacement vectors can be computed from 20 stereo-image pairs with ideal (full) image quality(without the addition of Gauss noise and motion blurring) based on VM JTC and VCJTC, respectively. As shown in **Fig. 5**,both of the two virtual sensors clearly obtained similar better results from each stereo-image pair, the measurement residual errors computed from differences between the value of true vector (20.0,0.0) pixel and that of the detected vector was no more than 0.03pixel in x-direction and 0.01pixel in y-direction . As shown in **Fig. 1**(in section 1.1) , experimental results without sharp cross-correlation peaks from VCJTC were produced from an image pair with Gauss noise variance of 0.002 and motion blur of 1.0pixel in x-direction. However, as shown in **Fig. 2**, it can be seen that the cross-correlation peaks for the same image pair were sharp enough to be detected easily and the accuracy of displacement measurement attains 0.026pixel in -direction and 0.01pixel in y-direction ,respectively.As far as 20 stereo-image pairs with the addition of Gauss noise of 0.002 variances and motion blurring of 1.0pixels were considered, residual errors of displacement measurement with VCJTC varied from -16.5pixels to 14.8pixels in x-direction and -14.9pixels to 15.1pixels ,which can be seen in **Fig. 6**(a). However, **Fig. 6**(b) shows that displacement residual errors varied around zero for all image pairs even with Gauss noise variance of 0.002 and motion blurring of 10.0pixels.The accuracy can be still within 0.1pixel in x- , y-direction respectively and did not significantly vary with degraded image quality .



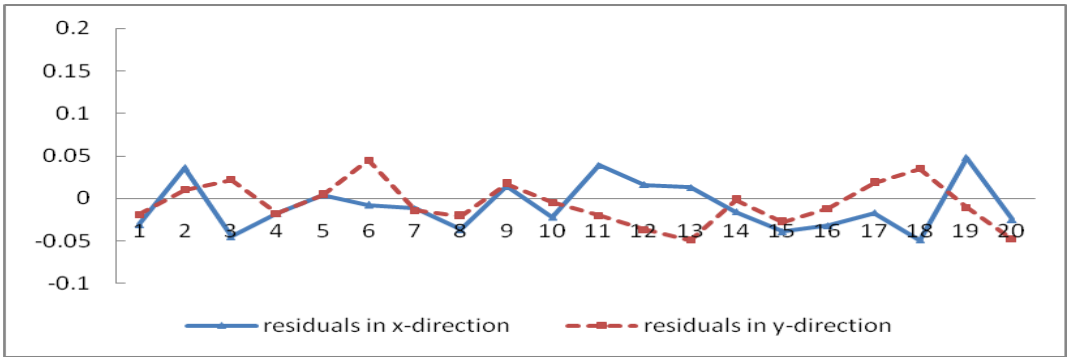
(a) (b)

Fig. 5. Measurement accuracy from VMJTC and VCJTC for low-altitude images without addition of noise and motion blurring. (a)in x-direction （b）in y-direction

Besides, due to the limitation of size and payload, the camera system is installed on the UAV platform with low-weight structure. This brings irregular motion and vibration of low, high and medium frequency during aerial photographing. Therefore, it is more complex and difficult to generate simulation images similar to the low-altitude images with degraded image quality. The feasibility of the proposed VMJTC sensor for displacement detection is further verified by real data.



(a)



(b)

Fig. 6. Measurement accuracy for low-altitude images with addition of noise and motion blurring （a） results from VCJTC（ b） results from VMJTC

Further, experiments were conducted using two aerial digital videos captured by a HD video camera with size of 1920×1080 pixel and sampling frequency of 50 frames per second mounted on a quad-rotor for three flight tests about 200 to 300 m above the ground. As shown in **Fig. 7**, 50 adjacent frame images of each video were transformed and sub-images with size of 256×256pixels at the same location per frame image were extracted

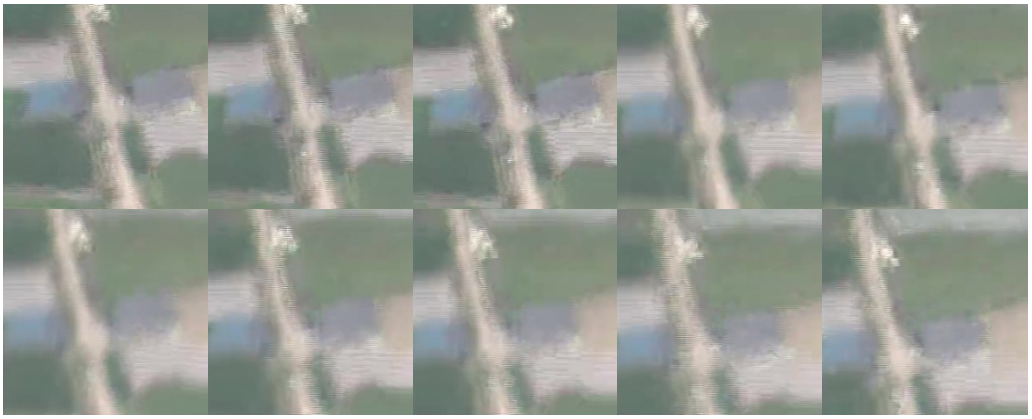
Different from simulation tests, true displacement values were unknown for real images. Therefore, each three adjacent sub-images were combined together to set up a tri-image pair and its relative residual error was defined to assess the accuracy and reliability of VMJTC sensor . We assume that the image $f_j(x, y)$ was moving towards the object image $f_{j+k}(x, y)$ according to the calculated displacements $(\Delta x_{j,j+k}, \Delta y_{j,j+k})$ with $j = 1, 2...49$ and $k = 1, 2$ .The relative residual errors for motion displacement of the j^{th} tri-image pair can be computed based on equation (14) .

$$v_{xj} = \Delta x_{j,j+1} + \Delta x_{j+1,j+2} - \Delta x_{j,j+2}, \quad v_{yj} = \Delta y_{j,j+1} + \Delta y_{j+1,j+2} - \Delta y_{j,j+2} \tag{14}$$

Where v_{xj} and v_{yj} represent the residual errors in x- and y-direction, respectively and would be zero with ideal (full) detection conditions. In this way, total 49 tri-image pairs of each dataset were collected and used for comparison experiments. To further evaluate performance of all tri-image pairs, the root mean squared error(RMSE) was introduced as Equation(15).

$$\sigma_{xy} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(v_{xj}{}^2 + v_{yj}{}^2)} \tag{15}$$

Where σ_{xy} refers to the displacement measurement accuracy of n pairs of tri-images. As seen in **Fig. 7**, the value of accuracy may vary with the image quality and even decrease in long-displacement (the first and the third image in a tri-image pair) measuring conditions. Figure 8 shows displacement measurement results of all tri-image pairs in two datasets using VMJTC sensor.



(a)

（b）

Fig. 7. Two test datasets of sub-images extracted from the original low-altitude images. （a） ten images in the first dataset; (b) ten images in the second dataset

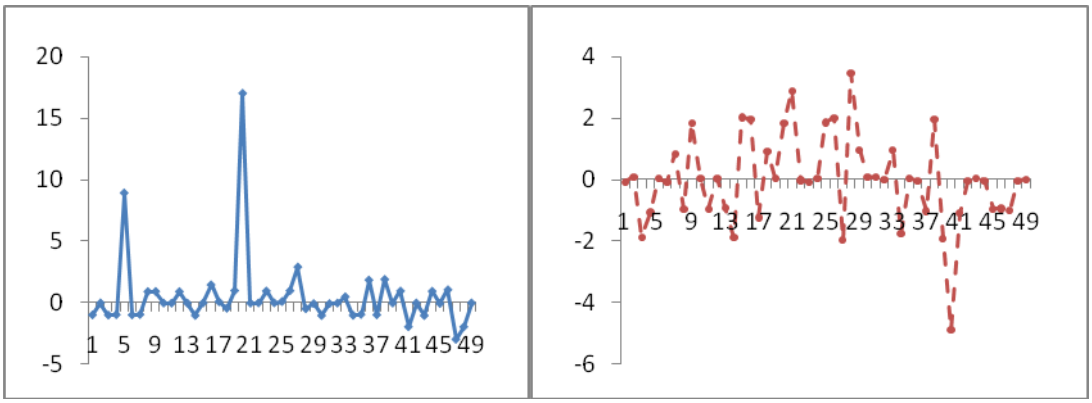


Fig. 8. Residual errors of tri-image pairs in the first dataset (a) in x-direction; (b) in y-direction

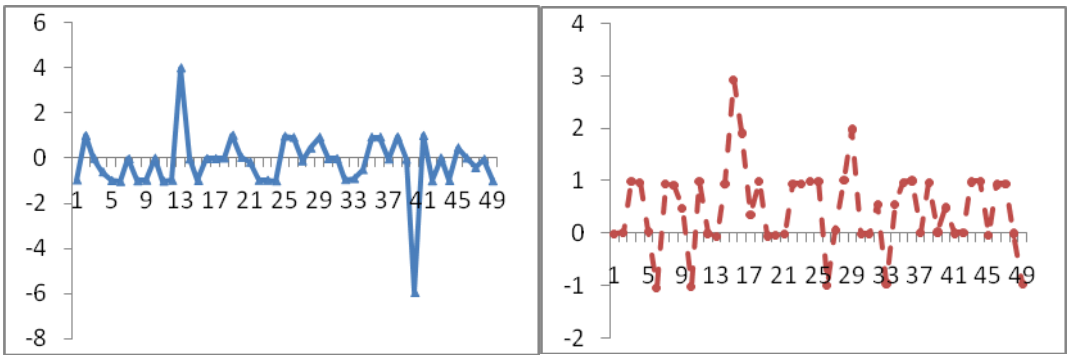
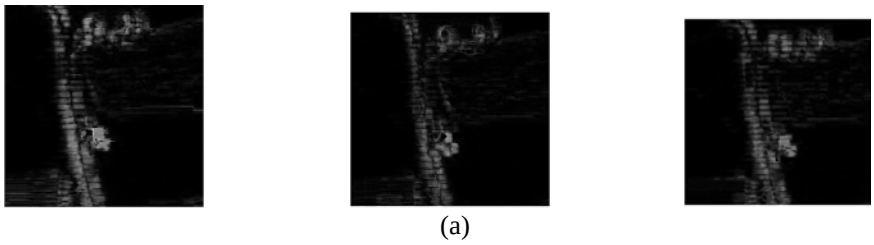
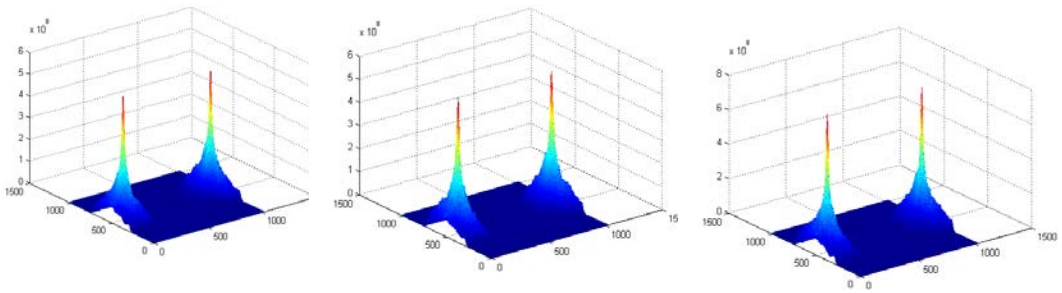


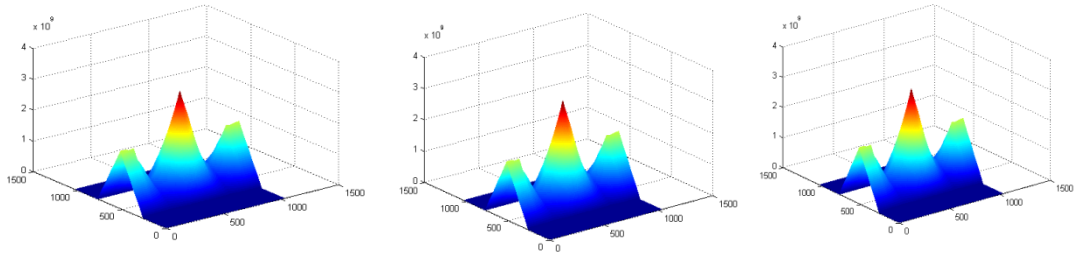
Fig. 9. Residual errors of tri-image pairs in the second dataset (a)in x-direction; (b) in y-direction



(a)



(b)



(c)

Fig. 10. Results of T_1 from two sensors. (a) edge detection using the proposed method; (b) 3D cross-correlation from VMJTC sensor (c) cross-correlation from VCJTC sensor.

The tri-image pairs with residuals twice root mean square errors were considered as outliers caused by large displacement or angular motion. After the gross errors were removed from **Fig. 8** and **Fig. 9**, RMSE of the first dataset achieved $\sigma_{xy} = 1.8151\text{pixels}$， RMSE of the second dataset with $\sigma_{xy} = 1.2107\text{pixels}$. The accuracy can be improved by using images from video cameras with higher sampling frequency in the future work. In fact, without deriving sharp peaks of cross-correlation output , the VJTC sensor have failed to extract displacement information for all image pairs in two datasets. **Fig. 10** illustrates displacement measurement results on two sensors for the tri-image pair T_1 of the first three images（I_1, I_2, I_3）in **Fig. 7(a)** from the first dataset and T_1' of I_1', I_2', I_3' in **Fig. 7(b)** from the second dataset.

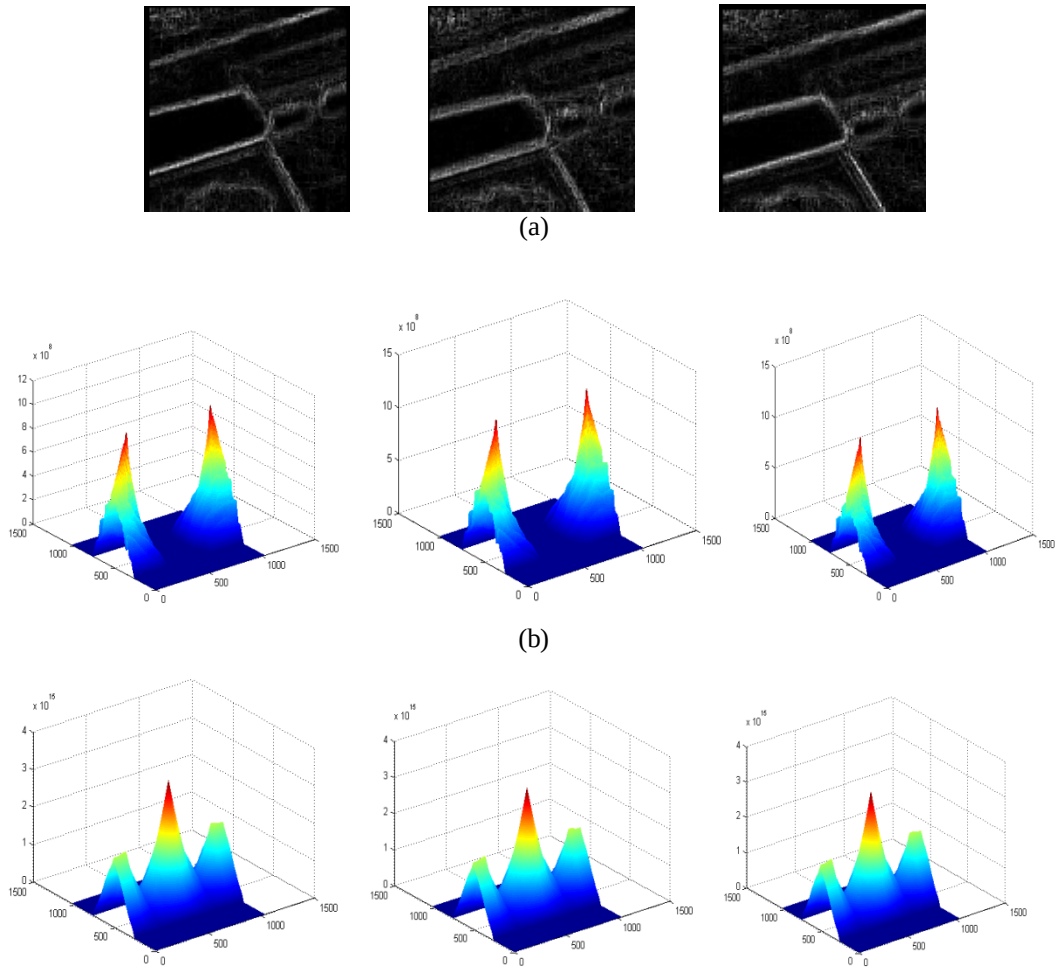


Fig. 11. Results of T_1' from two sensors. (a) edge detection using the proposed method; (b) 3D cross-correlation from VMJTC sensor (c) cross-correlation from VCJTC sensor.

It can be seen that cross-correlation peaks produced by VMJTC sensor were sharp enough to be detected accurately. Then, measurement accuracy can be computed with -0.94876pixel in x-direction and 0.02542pixel in y-direction for T_1 , -0.9162pixel in x-direction and 0.09542pixel in y-direction for T_1'. However, as shown in **Fig. 10** and **Fig. 11**, the obtained cross-correlation output from conventional VCJTC sensor did not posses single and shaper peaks compared with the results from VMJTC sensor. The above experimental results show that the VMJTC sensor can be used to detect motion displacement for real low-altitude images, which is proven to be more suitable ,reliable and accurate than VCJTC sensor based on conventional joint transform correlation algorithm. In this paper,the low-altitude images were taken with vision sensors at a frequency sampling of 50 Hz. Future works will adopt vision sensor of high frequency sampling over 100 Hz to improve the accuracy of the proposed algorithm.

5. Conclusions

This study proposes a modified joint transform correlation algorithm aiming at off-line motion detection with low altitude images in aerial photogrammetry and remote sensing applications. Compared with traditional two-step algorithm, the modified joint transform correlation algorithm can be realized in four steps. In the first step, a novel fuzzy edge detection method was proposed and applied to deal with edge extraction of two adjacent low-altitude images. Then, two images only with edge information were combined to form a joint image. In the second step, similar to conventional joint transform correlation algorithm, the joint image in spatial domain was transformed to frequency domain with Fourier transformation. In the third step, the joint power spectrum was derived and then only cross-correlation items were extracted. Finally, cross-correlation power spectrums in frequency domain were transformed back to spatial domain with Fourier transformation to produce a pair of cross-correlation output. Then, locations of the maximum cross-correlation peaks in local region can be computed with centriod algorithm easily and motion displacement can be achieved at sub-pixel level. From the tests with simulation datasets and experiments with real datasets, it can be seen that the proposed joint transform correlation algorithm presented higher and more reliable displacement measurement results than the traditional algorithm .

In principle,the displacement measurement accuracy using joint transform correlation methods depends on the ability to quantify the degree of similarity between input images. However, when the UAV flies in a complex low-altitude scene with high veclocity, JTC techniques may produce false detections without considering relative larger displacement and rotation factors in the transform equations.Thus, future works will focus on improving the accuracy and robustness by introducing an efficient priori rotation and velocity estimation scheme into the proposed algortithm.

Acknowledgments

This work was supported by National Natural Science Foundation of China (NSFC) under Grant no.41371425 and Provincial Natural Science Foundation of Liaoning Province,China under Grant no.20170540015

References

[1] Whitehead K., Hugenholtz C., et al. "Remote sensing of the environment with small unmanned aircraft systems," *Journal of Unmanned Vehicle Systems*, vol.2, no.3, pp. 69-85, 2014 Article(CrossRefLink)

[2] Lin Z., "UAV borne low altitude photogrammetry system," in *Proc. of International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXIX-B1, pp. 415-423, 2012. Article(CrossRefLink)

[3] Cui H., Lin Z., et al., "Multiview Photogrammetry Using Low Altitude Digital Images From Unmanned Airship," *Optical Electrical Engineering* , vol.35, no.7, pp.73-78, 2008. Article(CrossRefLink)

[4] Taha Z., Tang Y.R., Yap K. C., "Development of an onboard system for flight data collection of a small-scale UAV helicopter," *Mechatronics*, vol.21, no.1, pp. 132–144, 2011. Article(CrossRefLink)

[5] Šmídl V., Hofman R., "Tracking of atmospheric release of pollution using unmanned aerial Vehicles," *Atmospheric Environment*, vol.67, pp.425-436, 2013. Article(CrossRefLink)

[6] Malihi S.,Maboudi M., Pourmomen M., "Misalignment Calibration of Ultracam D and XP," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Inf. Sci.*, vol XXXIX-B1, pp. 577-579, 2012 Article(CrossRefLink)

[7] Alamús R., Kornus W. and Talaya J., "Studies on DMC geometry," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 60, no. 6, pp. 375–386, 2006. Article(CrossRefLink)

[8] Pejic, M., "Airborne digital sensors: LH Systems ADS 40," *Military Technical Courier*, vol. 52, no.1, pp.87-98, 2004. Article(CrossRefLink)

[9] Tchernykh, V., Harnisch, B., "Compensation of focal plane image motion perturbations with optical correlator in feedback loop," in *Proc.of Sensors, Systems, and Next-Generation Satellites*, vol. 5570, pp.280–288, 2004. Article(CrossRefLink)

[10] Janschek, K., Tchernykh, V., Dyblenko, S, "Performance analysis of opto-mechatronic image stabilization for a compact space camera," *Control Engineering Practice*, vo.15, no.30, pp. 333-347, 2007. Article(CrossRefLink)

[11] Qian, Y.X., Hong, X.T., Jin, W.M., "Motion measurement of dual-CCD imaging system based on optical correlator," *Chinese Journal of Lasers*, vol.40, no.7, pp.160-165, 2013. Article(CrossRefLink)

[12] Fan, C., Li, Y.C., Fu, H.L., Liang, Y., "Research on measurement method of image motion of space camera based on optical correlator," *Acta Optica Sinica*, vol.31, no.7, pp.1213-1216, 2011. Article(CrossRefLink)

[13] Feng, D., Feng, M., Ozer, E., Fukuda, Y., "A vision-based sensor for noncontact structural displacement measurement," *Sensors*, vo.15, no.7, pp. 16557–16575, 2015. Article(CrossRefLink)

[14] Schumacher, T., Shariati, A., "Monitoring of structures and mechanical systems using virtual visual sensors for video analysis: Fundamental concept and proof of feasibility," *Sensors* , vol.13, no.12, pp.16551–16564, 2013. Article(CrossRefLink)

[15] Fukuda, Y., Feng, M.Q., Narita, Y., Kaneko, S., et al. "Vision-based displacement sensor for monitoring dynamic response using robust object search algorithm," *IEEE Sensors Journal*, vol, 13, no.13, pp. 1928–1931, 2010. Article(CrossRefLink)

[16] Li, C., Yin, S., Yu F., "Non zero-order joint transform correlator," *Opt. Eng.* Vol.37, no.1, 58-65, 1998. Article(CrossRefLink)

[17] Javidi, B., Wang, J., "Binary nonlinear joint transform correlation with median and subset median Thresholding," *Applied Optics*, vol.30, no.8, 967-976, 1991. Article(CrossRefLink)

[18] Hong X, Qian Y., "High-accuracy measurement of sub-pixel image motion based on hybrid photoelectirc joint transform correlator," *Acta Optica Sinica*, vol.35, no.2, 321-331, 2015. Article(CrossRefLink)

[19] Heng, L., Lee, G., Pollefeys, M, "Self-Calibration and Visual SLAM with a Multi-Camera System on a Micro Aerial Vehicle," *Autonomous Robots*, vo.39, no.3, pp.259-277, 2015. Article(CrossRefLink)
[20] Canny J., "A computational approach to edge detection," *Readings in Computer Vision*, pp. 184–203, 1987. Article(CrossRefLink)
[21] Haralick, R. M., "Digital step edges from zero-crossings of second directional derivatives," *Readings in Computer Vision*, pp 216-226, 1987. Article(CrossRefLink)
[22] Yi, H., Zhao, H., Li, Y., "Improved digital processing method used for image motion measurement based on hybrid opto-digital joint transform correlator," *Chinese Optics Letters*, 8, 989-992, 2010. Article(CrossRefLink)
[23] Ouchi, K., "Ship detection based on coherence image derived from cross-correlation of multilook SAR images," *IEEE Trans.geosci.remote Sens.lett*, vol.1, no.3, 184-187, 2004. Article(CrossRefLink)
[24] Zhang, D., Guo, J., Lei, X., Zhu, C.G., "A High-Speed Vision-Based Sensor for Dynamic Vibration Analysis Using Fast Motion Extraction Algorithms," *Sensors*, vol. 16, no. 4, 57201-57217, 2016. Article(CrossRefLink)
[25] Pal, S.K., King, R.A., "Image enhancement using fuzzy sets". *IEEE Electronics Letters*, vol.16, no.10, 376-378, 1980. Article(CrossRefLink)
[26] Pushpa, D., Pate,V., Kumar, T., Sadhna, M. A., " Novel Fuzzy Image Enhancement using S-Shaped Membership Function", *International Journal of Computer Science and Information Technologies*, vol.6 ,no.1.564-569,2015. Article(CrossRefLink)
[27] Madasu, H., Devendra, Jha., "An Optimal Fuzzy System for Color Image Enhancement", *IEEE Transactions on Image Processing*, vol.15,no.10,pp. 2956-2966,2006. Article(CrossRefLink)
[28] Loo, C., Alam, M., "Invariant object tracking using fringe-adjusted joint transform correlator," *Optical Engineering*, vol.43, no.9, pp.2175-2183, 2004. Article(CrossRefLink)
[29] Wu, D., Zhang, Y., Cao, L., "Improved algorithm of fuzzy edge detection," *Laser&Infrared*, vol.40, no.12, 1374-1377, 2010. Article(CrossRefLink)
[30] Bhandari A., Kumar A., Singh G., "Modified artificial bee colony based computationally efficient multilevel thresholding for satellite image segmentation using Kapur's, Otsu and Tsallis functions," *Expert Systems with Applications* , vol. 42, no. 3, pp. 1573–1601, 2014. Article (CrossRefLink)

Hongxia Cui : received the B.Sc. degree in Photogrammetry and Remote Sensing from Wuhan University, Hubei Province, China, in 1991, the M.E. degree in Computer Science from Shanxi University, Shanxi, China, in 2002 and the Ph.D. degree in Photogrammetry and Remote Sensing from Wuhan University, Hubei Province, China, in 2006. She is currently a Professor at the College of Computer Science and Technology, Bohai University. Her current research interests include image processing, computer vision, UAV-based photogrammtric system design ,camera calibration ,bundle adjustment and their applications .

Yaqi Wang : received the B.Sc. degree in the College of Computer Science and Technology, Bohai University,Jinzhou,Liaoning Province, China, in 2015. Her current research interests include image computer vision and software applications

Fangfei Zhang : received the B.Sc. degree in Geographic Science from Tian Jin Normal University, China, in 2017.She is currently a undergraduate student at Department of Foresty ,Beijing Foresty University , Beijing, China. She current research interests include UAV-based GIS application research .

Tingting Li: received the B.Sc. degree in the College of Computer Science and Technology, Bohai University,Jinzhou,Liaoning Province, China, in 2016. Her current research interests include image computer vision and software applications