Chinese Journal of Electronics
Vol.26, No.4, July 2017

A Sub-pixel Disparity Refinement Algorithm Based on Lagrange Interpolation*

MEN Yubo[1], ZHANG Guoyin[1], MEN Chaoguang[1], LI Xiang[1] and MA Ning[1,2]

(1. *College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China*)

(2. *Harbin Normal University, Harbin 150001, China*)

Abstract — A sub-pixel disparity refinement algorithm based on Lagrange interpolation is proposed to calculate the sub-pixel disparity and the elevation of the target construction on the remote sensing image. The accurate integer matching feature points are obtained by improved affine scale invariant feature transform. A series of interpolation samples can be estimated through the three times magnification results of each feature matching point and its close neighbor pixels. The final sub-pixel deviation can be calculated by the interpolation samples. Experiment results show that a more accurate elevation information can be obtained on the remote sensing image with corrective epipolar line, and an accurate sub-pixel disparity result can be estimated without the epipolar line constraint as well.

Key words — Computer vision, Remote sensing image processing, Sub-pixel disparity, Interpolation.

I. Introduction

Higher precision sub-pixel matching as the key technology in the field of computer vision has been applied to many domains. The traditional sub-pixel matching methods can be divided into three categories: the phase correlation method[1,2], optimization solution method[3,4] and the interpolation method[5,6]. The advantage of phase correlation method is insensitive to gray scale difference and narrow band noise, but it is estimated only in the frequency domain, so the phase correlation method cannot be calculated for the affine transformed image. The optimization solution method usually has a large computational complexity. It needs to improve the convergence probability and the global optimal solution probability. The matching accuracy of interpolation method is much better than phase correlation method and optimization solution method. However, in most cases, the traditional interpolation methods only apply to the image pairs with epipolar constraint.

In the sub-pixel interpolation methods, VC Miclea proposes a sub-pixel interpolation function for accurate real-time stereo matching algorithms[7]. The cost volume calculated with two high-speed Semi-Global matching can obtain dense sub-pixel disparity maps, but the matching speed is slower because of using the global algorithm, and the original interpolation function cannot apply to image with non-corrective epipolar line. Nan proposes a cubic spline interpolation method to obtain sub-pixel result[8]. The cubic spline interpolation method has a simple operation process and a good convergence. However, it can only guarantee the smoothness of each curve segment, but cannot guarantee the smoothness of the whole interpolation curve. Therefore, the sub-pixel position has a certain error, and the uncertain image gray level distortion will reduce the matching accuracy. This paper proposes a sub-pixel disparity refinement algorithm based on the Lagrange interpolation method which is named SDRL. The SDRL algorithm can optimize interpolation samples to avoid the effect of distortion to increase the accuracy of sub-pixel matching, and it can be used on the image with non-corrective epipolar line.

II. The SDRL Algorithm

1. The basic idea of SDRL algorithm

The accuracy of sub-pixel disparity depends on the accuracy of integer level pixel matching directly. First, we use the improved Affine scale invariant feature transform (ASIFT) to conduct the integer level pixel matching to obtain the exact matching feature points. Then according to the integer level result, the interpolation samples can be calculated by multiple bi-cubic interpolations. The interpolation samples are used to obtain the sub-pixel disparity by the proposed Lagrange interpolation method.

*Manuscript Received Dec. 01, 2016; Accepted Jan. 11, 2017. This work is supported by the National Natural Science Foundation of China (No.61672181).

© 2017 Chinese Institute of Electronics. DOI:10.1049/cje.2017.03.002

Finally, the elevation of remote sensing image can be calculated by the final sub-pixel disparity. The basic idea of the algorithm is illustrated in Fig.1 as follows.

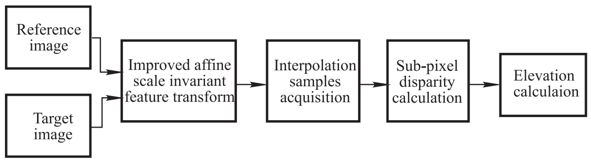


Fig. 1. The flowchart of SDRL algorithm

2. Integer level pixel matching

In the process of integer level pixel matching, the steps of the improved ASIFT method are as follows: 1) Simulation of multi affine distortion to obtain the analog image; 2) The analog image matching by an improved Scale invariant feature transform (SIFT) method.

1) Simulation of multi affine distortion

The space plane scene captured by the camera can be described by an affine model as

$$u = S_1 G_1 A T u_0 \tag{1}$$

where u denotes a digital image, u_0 denotes the projection of the plane scenery on the infinite distance, S_1 denotes standard sampling operation, G_1 denotes the Gaussian visual fuzzy model, T denotes image transformation caused by the camera movement, A denotes a translation transformation simplified as an affine transformation. Therefore, every local visual projection effect in the image can be simulated by the local affine transformation shown as

$$u(x,y) = (ax + by + e, cx + dy + f) \tag{2}$$

where (a,b,c,d,e,f) denotes the image transform parameter collection.

The affine transformation matrix $\boldsymbol{A} = \begin{bmatrix} a, b \\ c, d \end{bmatrix}$ can be decomposed into

$$\begin{aligned} \boldsymbol{A} = \lambda \boldsymbol{R}_1(\psi) \boldsymbol{T}_t \boldsymbol{R}_2(\varphi) = \lambda \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} \\ \cdot \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \end{aligned} \tag{3}$$

where λ denotes the scaling parameter ($\lambda > 0$), λt denotes the determinant of $\boldsymbol{A}$, $\boldsymbol{R}_1(\psi)$ and $\boldsymbol{R}_2(\varphi)$ denotes the latitude transformation matrix and rotation transformation matrix respectively, $\boldsymbol{T}_t$ denotes the tilt transformation matrix. The first eigenvalue of the diagonal matrix $\boldsymbol{T}_t$ is $t = |1/\cos\theta|$, and the second one is 1. θ denotes the camera longitude angle, φ denotes the camera latitude angle, and ψ denotes the camera rotation angle. The affine deformation caused by the optical axis of the camera can be simulated through the collection of θ and φ to achieve the result of image transformation. The sampling of θ follows the geometric sequence sampling principle as $1, a, a^2, \cdots, a^n$ ($n \geq 5$). When $a = \sqrt{2}$, the result can achieve the best accuracy. The sampling of φ follows the arithmetic sequence sampling principle as $0, b/t, \cdots, kb/t$ ($b = 72°$[9]), where k denotes maximal integer value which meets the condition $kb/t < \pi$. The analog image can be obtained by parameter series θ and φ by Eq.(1).

2) Improved scale invariant feature transform

SIFT is used to calculate image scale space by Gaussian kernel and image convolution, the two-dimensional analog image scale space obtained in Section II.2.1) can be defined as

$$L(x,y,\sigma) = G(x,y,\sigma) \times I(x,y) \tag{4}$$

where $I(x,y)$ denotes the original image, $G(x,y,\sigma)$ denotes the Gaussian kernel with scale variability which can be calculated as $G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} \mathrm{e}^{\frac{-(x^2+y^2)}{2\sigma^2}}$, where σ denotes the scale, (x,y) denotes the spatial coordinate of pixels. Lowe has proposed a Gaussian differential operator Difference of Gaussian (DoG) which approaches the Laplacion of Gaussian (LoG) scale invariant operator to conduct the Gaussian smooth differential operation on different scale images[10]. The operation can be described as

$$\begin{aligned} D(x,y,\sigma) &= [G(x,y,k\sigma) - G(x,y,\sigma)] \times I(x,y) \\ &= L(x,y,k\sigma) - L(x,y,\sigma) \end{aligned} \tag{5}$$

The image feature points consist of the local extreme points in DoG space. The feature point can be obtained by the comparison of gray value with its 8-neighbor points on the same scale space and 18 points on adjacent scale spaces. If a feature point is the maximum or minimum among its 26 neighbor points in present DoG space and its adjacent scale spaces, the point can be identified as a feature point. Every feature point can be set an orientation parameter by using the gradient direction distribution of the feature points and their neighbor pixels to reduce the image rotation effect.

The pixels gradient magnitude $m(x,y)$ and direction $\theta(x,y)$ can be described as

$$\begin{aligned} m(x,y) = [(L(x+1,y) - L(x-1,y))^2 \\ + (L(x,y+1) - L(x,y-1))^2]^{\frac{1}{2}} \end{aligned} \tag{6}$$

$$\theta(x,y) = \tan^{-1}[\frac{L(x,y+1) - L(x,y-1)}{(L(x+1,y) - L(x-1,y))}] \tag{7}$$

where L denotes the scale space containing the feature points to calculate the corresponding 128-dimensional space description vector for each feature point to obtain the feature point matching result. The feature point detection has finished so far, every feature point has position, scale and direction information, but the mismatching of feature points also exist.

In order to confirm the final matching feature point pairs, the gradient angle variation based on neighbor feature points method is proposed. The feature point can be described as $\boldsymbol{S}_i = [X_i, r_i, \theta_i, f(X_i, r_i, \theta_i)]^{\mathrm{T}}$ where X_i denotes the feature point coordinate, r_i denotes the feature point scale, θ_i denotes the gradient direction of feature point, $f(X_i, r_i, \theta_i)$ denotes the descriptor of corresponding matching feature point which has a similar description. We assume S_1 is a feature point and its neighbor feature points S_3 and S_2 which have the first and the second nearest distance from S_1. At the same time, the corresponding matching feature points $f(S_1)$, $f(S_3)$ and $f(S_2)$ can be extracted in the target image by the connection of S_1S_2, S_1S_3, $f(S_1)f(S_2)$ and $f(S_1)f(S_3)$. The gradient relation of neighbor feature matching points is illustrated in Fig.2. The difference of gradient direction $\Delta\theta_x$ in the reference image and $\Delta\theta'_x$ in the target image can be calculated by the gradient direction of S_1, $f(S_1)$ and their connection directions. They can be obtained by the formula $\Delta\theta_x = |\theta_1 - \theta_2|$ and $\Delta\theta'_x = |\theta'_1 - \theta'_2|$. In consideration of the error caused by different shooting angles, when $|\Delta\theta_x - \Delta\theta'_x| < 10°$, the matching feature point pairs will be retained. Otherwise, it is considered to be a wrong matching point and will be eliminated from the feature point group. This method will traverse all the feature points to determine the final matching feature point pairs.

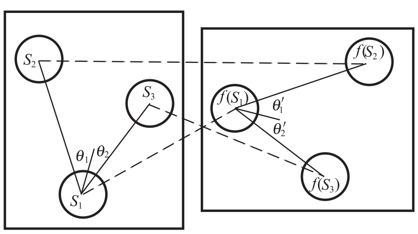


Fig. 2. Gradient relation of neighbor feature matching points

The exact integer level matching feature points can be obtained through the improved ASIFT method. In the image with corrective epipolar line, the integer level pixel disparity can be calculated as

$$D = x_1 - x'_1 \tag{8}$$

where D denotes the integer level pixel disparity, x_1 and x'_1 denotes the horizontal coordinate value of matching feature point pairs. The integer level pixel disparity as an input parameter is used by the following sub-pixel disparity estimation.

3. Sub-pixel disparity solution

In the case of epipolar line correction, the elevation of the target construction can be calculated as

$$H = \frac{f \times D}{b/h} \tag{9}$$

where H denotes the elevation of the target construction, f denotes the resolution of the remote sensing image, D denotes the disparity of the target construction, b/h denotes the base-height ratio which is usually less than 0.5. The integer level disparities are discrete values, so the calculated elevations will have relatively large errors. The sub-pixel disparity solution is necessary for the remote sensing image to improve the elevation accuracy.

The image gray interpolation has an advantage of high accuracy in sub-pixel disparity solution. In Lagrange interpolation theorem, the Lagrange function meets the uniqueness and continuity conditions which can be described as

$$f(x) = \sum_{k=0}^{n} l_k(x) y_k = \sum_{k=0}^{n} \Big(\prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{x - x_j}{x_k - x_j}\Big) y_k \tag{10}$$

where $l_k(x)$ denotes the interpolation node polynomial, and it can be calculated by polynomial factorization as

$$l_k(x) = \frac{\pi_{j+1}(x)}{(x - x_k)\pi'_{j+1}(x_k)} \tag{11}$$

$$\pi_{j+1}(x) = (x - x_0)(x - x_1) \cdots (x - x_j) \tag{12}$$

$$\begin{aligned} \pi'_{j+1}(x_k) = {} & (x_k - x_0) \cdots (x_k - x_{k-1}) \\ & \cdot (x_k - x_{k+1}) \cdots (x_k - x_j) \end{aligned} \tag{13}$$

According to Marr visual theory, the uniqueness and continuity constraints are designed similarly in the same object in image pairs. The Lagrange function can be determined on the target construction to obtain the sub-pixel disparity with the integer level matching pixels. The number of the samples of Lagrange interpolation determines the accuracy of the interpolation function. In order to get more interpolation samples, the obtained matching feature points in section II.2.2) will be conducted three times magnification by using multiple bi-cubic interpolations method. In traditional bi-cubic interpolation method[11], the points which are not on the axis of integer level pixels can be obtained by its 16 peripheral integer level pixels. But the points on the axis of integer level pixels have to be determined by multiple bi-cubic interpolations as shown in Fig.3. The point a_{1k} not only lies on 16 pixels around a_{jk}, but also the below 16 pixels around a_{jl}. The final gray value $g(a_{1k})$ is calculated as $g(a_{1k}) = \overline{g(a_{t1k}) + g(a_{b1k})}$ where $g(a_{t1k})$ and $g(a_{b1k})$ is gray value of a_{1k} through multiple bi-cubic interpolations respectively.

Because the feature points obtained by improved ASIFT have texture property, the Lagrange interpolation can be confirmed to obtain smooth interpolation curve by the matching feature point with its two compact neighbor points and four auxiliary sub-pixel points calculated by multiple bi-cubic interpolations on horizontal axis direction. These seven pixels are considered to be the interpolation samples. According to the possible distortion of remote sensing images, the interpolation samples need to be filtered to obtain accurate gray value with their adjacent area pixels. In Eq.(10), x_k and x_j denotes the different position of the samples along the horizontal axis respectively, y_k denotes the post filtered gray value of the pixel with the horizontal coordinate x_k. y_k is calculated by x_k and its adjacent area pixels to reduce gray value distortion as follows

$$y_k = \frac{\sum_{k=0}^{n} g_{cn}(S_n) + g_c}{\sum_{k=0}^{n} S_n + 1} \tag{14}$$

$$S_n = \begin{cases} 1, \text{ if } t > |g_c - g_{cn}| \\ 0, \text{ if } t \le |g_c - g_{cn}| \end{cases} \tag{15}$$

$$g_{cn} \in G = \{g_{c1}, g_{c2}, \cdots, g_{c8}\} \tag{16}$$

where G denotes the gray value collection of the adjacent area pixels, g_c and g_{cn} denotes the gray value of the sample and the gray value of its 8-neighbor pixels respectively, t denotes a threshold which is evaluated as 10 in the proposed algorithm. In this paper, the area of sample is defined as a 3×3 window, and every sample has a post filtered gray value y_k. According to the samples and the post filtered gray values, the only Lagrange fitting result can be obtained to find the sub-pixel position of feature points in the image pair. In the reference image, the gray value of feature point as a reference value compares with the vertical coordinate value on estimated Lagrange fitting function in the target image to find the most suitable new position. The feature points will be relocated its new sub-pixel position by comparing corresponding gray value relationship.

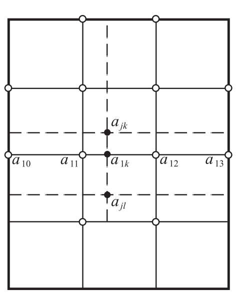


Fig. 3. Multiple bi-cubic interpolations to obtain pixels on the axis of integer level

During the reposition process, we should consider not only the gray value relationship, but also the gradient relationship to obtain the sub-pixel position of feature points ultimately. When the effect of absolute gray value is lower, the gradient direction is the only criterion to determine whether the deviation result is correct or not. The actual matching point must have the same trend under gradient condition with the reference value. The gradient condition of sub-pixel position can be described as

$$[f(x_r^+) - f(x_r^-)] \times [f(x_t^+) - f(x_t^-)] > 0 \tag{17}$$

where x_r denotes the horizontal coordinate value of the matching feature point in the reference image, x_t denotes the horizontal coordinate value of the corresponding point in the target image, $f(x^+)$ and $f(x^-)$ denote the right and left neighbor limit function value of feature point respectively. The simulation of the deviation results under ideal sine Lagrange interpolation function conditions is shown in Fig.4. In Fig.4(b), the new location has multiple results within the sample range after artificial 0.25 pixels shifting. The searching results of x_5 are x_{5a} and x_{5b}. Due to $[f(x_5^+) - f(x_5^-)] \times [f(x_{5a}^+) - f(x_{5a}^-)] < 0$ and $[f(x_5^+) - f(x_5^-)] \times [f(x_{5b}^+) - f(x_{5b}^-)] > 0$, the actual matching point of x_5 is x_{5b}. The final sub-pixel disparity is $d' = |x_5 - x_{5b}|$. If there is more than one matching points with the same gradient direction, the nearest matching point from the feature point will be selected to obtain the sub-pixel deviation result.

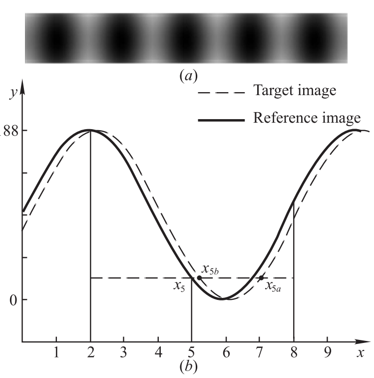


Fig. 4. The simulation of the deviation results under ideal sine Lagrange interpolation function conditions. (a) A part of simulated ideal Lagrange interpolation function image; (b) Simulation of the migration results under ideal Lagrange interpolation conditions. x denotes the integer level pixel, y denotes the gray value range

4. Description of the SDRL algorithm

The detailed description of the proposed algorithm is as follows:

Step 1: Through the input reference image and target image, the analog image pair parameters φ and θ can be obtained by the latitude and longitude sampling. The analog images u can be calculated as Eq.(1).

Step 2: The analog images u are used to calculate the local extreme points in DoG space through Eq.(5) to obtain the feature point pairs $\boldsymbol{S}_i = [X_i, r_i, \theta_i, f(X_i, r_i, \theta_i)]^{\mathrm{T}}$. According to the gradient direction θ_i, the final retained matching feature point pairs can be extracted by the condition $|\Delta\theta_x - \Delta\theta'_x| < 10°$. The integer level disparity D of matching feature points can be calculated by Eq.(8).

Step 3: According to the matching feature point and its close neighbor points, the samples can be calculated by multiple bi-cubic interpolations method. x_k is one of the positions of the samples, and its gray value y_k is optimized by Eq.(14). Likewise, the basic information of other samples can also be obtained.

Step 4: The optimized samples are used for fitting function $f(x)$ through Lagrange interpolation method, the final sub-pixel disparity d' will be obtained by using the gray value and the difference of gradient direction.

III. Experimental Results and Analysis

In order to confirm the accuracy of the SDRL algorithm under the epipolar line constraint, we use the city analog remote sensing images with building elevations. The analog remote sensing image created by SE-SCENARIO under OKTAL is shown in Fig.5.

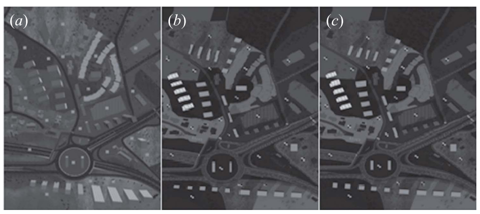


Fig. 5. The analog experimental images. (*a*) The analog remote sensing image created by SE-SCENARIO under OKTAL; (*b*) The analog reference remote sensing image; (*c*) The analog target remote sensing image

The resolution ratio f is calculated as

$$f = \frac{2 \times h \times \tan\beta}{p} \tag{18}$$

where h denotes the set camera height, β denotes the field of view, p denotes the number of pixels in unit area. In this experiment, $h = 5000$, $\beta = 7.1527/2$ and $p = 2084$. There are 40 buildings in the image as the target constructions to conduct the experiment to calculate the elevations which compared with the true elevation information. The elevation calculation formula of the epipolar line correction image should be the Eq.(9). However, there is a systematic error between the calculated elevation and the real elevation in the image. Therefore, Eq.(9) needs to be adjusted by a benchmark as

$$H = d \times \frac{f}{b/h} + \Delta H \tag{19}$$

where d denotes the disparity with sub-pixel level, ΔH denotes the benchmark difference, and the determination of the benchmark difference requires a reference building with the known elevation information. According to Eq.(18) and the baseline setting, the value of $\frac{f}{b/h}$ is 5.7. In this experiment, a small difference in elevation will be brought by the choice of different benchmark building. The difference cannot bring the loss of accuracy, so it will be ignored in this paper.

In Fig.6, five target buildings are chosen to conduct the experiment. There are a plenty of feature points which are obtained by the improved ASIFT on each target building. Because the number of matching feature points is so large that they cannot be calculated one by one, we filter out a part of the matching point pairs. The partly chosen feature point pairs with relative stable gray value on the No.3 building is obtained to calculate the mean value of sub-pixel deviation result as shown in Fig.6(*b*). In order to obtain a more accurate sub-pixel disparity, the chosen feature point should have a certain gradient change parameter. In this paper, the parameter is identified as 10. It means that there are 10 gray value differences from the chosen feature point and its neighbor pixels no matter the row and the column direction. This method partly improves the texture properties of the feature points and their neighborhoods to obtain a more accurate interpolation function. The final results and error of target buildings are shown as Table 1. In Table 1, the mean value of error is 0.1372m. Therefore, the sub-pixel accuracy of the obtained results is 0.024. Table 2 is the comparison of sub-pixel accuracy including Ref.[7], Ref.[8] and SDRL algorithm. The calculated elevation is obtained by mean sub-pixel deviation result of five target buildings. It can be seen that our algorithm has a better matching accuracy.

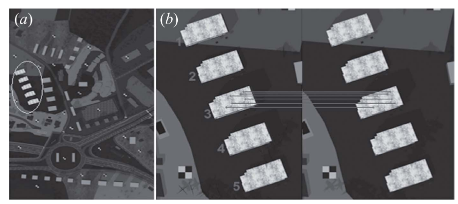


Fig. 6. The result of experiment in the analog remote sensing image. (*a*) The image area for experiment; (*b*) The construction numbers and the matching connection between the feature point pairs

Table 1. The result of five target buildings

Building	Disparity	True elevation	Calculated elevation	Error
1	–8.86	74.58m	74.468m	0.112m
2	–7.02	85.14m	84.956m	0.184m
3	–5.75	92.28m	92.195m	0.085m
4	–9.36	71.51m	71.618m	0.108m
5	–7.31	83.50m	83.303m	0.197m

In general, the traditional sub-pixel algorithm can be simply applied to the image with the epipolar line constraint. However, the SDRL algorithm not only has an excellent accuracy on the image with the epipolar line constraint, but also can act on the remote sensing images with non-corrective epipolar. We conduct artificial affine transformation of the No.2 building in Fig.6(*b*) as shown in Fig.7. The affine transformation includes rotation, shearing and translation. In our experiment, the angle of rotation is 45°, the horizontal and vertical shearing angle are both 20°, and the translation of Y axis direction is 2 unit pixels. In Fig.7, the initial reference image and the affine transformed target image conduct the proposed sub-pixel disparity solution on row direction and column direction respectively to calculate the coordinate values of feature points. Then we conduct inverse transformation on the affine transformed target image. The final experiment results are shown in Table 3. In Table 3, the coordinate values after inverse transformation compare with the original positions of reference image to obtain the errors. The true disparity on the row direction of No.2 building is -6.99, so the errors of horizontal and vertical coordinates are 0.42 and 0.1275 respectively which are less than half unit pixel. The results show that our algorithm can achieve a sub-pixel accuracy with non-corrective epipolar line.

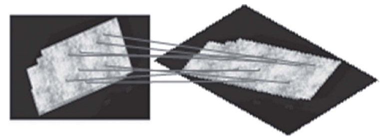


Fig. 7. The No.2 building matching results by improved ASIFT after the artificial affine transformation

Table 2. The comparison of sub-pixel accuracy in five target buildings

Algorithm	Building 1		Building 2		Building 3		Building 4		Building 5		Mean value of error
	Calculated	Error	Calculated	Error	Calculated	Error	Calculated	Error	Calculated	Error	
SDRL	74.468m	0.112m	84.956m	0.184m	92.195m	0.085m	71.618m	0.108m	83.303m	0.197m	0.1372m
Ref.[7]	74.672m	0.092m	85.328m	0.188m	92.387m	0.107m	71.659m	0.149m	83.688m	0.188m	0.1448m
Ref.[8]	74.452m	0.128m	84.945m	0.195m	92.142m	0.138m	71.289m	0.221m	83.259m	0.241m	0.1846m

Table 3. The sub-pixel deviation of eight chosen feature points in No.2 building

Point	Position of reference image	Sub-pixel deviation result	Position deviation	Point	Position of reference image	Sub-pixel deviation result	Position deviation
1	(492, 318)	(485.66, 318.25)	$(-6.34, 0.25)$	5	(494, 342)	(487.29, 342.17)	$(-6.71, 0.17)$
2	(486, 325)	(479.53, 326.13)	$(-6.47, 1.13)$	6	(470, 344)	(463.31, 344.34)	$(-6.69, 0.34)$
3	(471, 331)	(464.69, 330.21)	$(-6.31, -0.79)$	7	(459, 352)	(452.28, 352.47)	$(-6.72, 0.47)$
4	(461, 333)	(454.48, 333.26)	$(-6.52, 0.26)$	8	(476, 354)	(469.20, 353.19)	$(-6.80, -0.81)$

IV. Conclusions

The SDRL algorithm is proposed to estimate the accurate sub-pixel disparity and calculate the elevation of the target construction on remote sensing images. The optimized samples improve the Lagrange interpolation result on the accuracy of the sub-pixel disparity. The calculated elevation is more accurate than other sub-pixel interpolation algorithms on the remote sensing image with corrective epipolar line. The SDRL algorithm can apply to the remote sensing image with non-corrective epipolar line to obtain the sub-pixel deviation on vertical and horizontal direction respectively. In the future, we need to focus on the accuracy of sub-pixel level on the remote sensing image with non-corrective epipolar line.

References

[1] G. Peng, C. Lan and H. Wang, "An improvement of image registration based on phase correlation", *Optik-International Journal for Light and Electron Optics*, Vol.125, No.22, pp.6709–6712, 2014.

[2] O. Arandjelović, D.S. Pham and S. Venkatesh, "Efficient and accurate set-based registration of time-separated aerial images", *Pattern Recognition*, Vol.48, No.11, pp.3466–3476, 2015.

[3] A. Das and M. Bhattacharya, "Multiresolution framework based global optimization technique for multimodal image registration", *Communications in Computer & Information Science* , Vol.276, pp.336–347, 2013.

[4] C.T. Dang and Z. Wu, "Adaptive multi-layer particle swarm optimization with neighborhood search", *Chinese Journal of Electronics*, Vol.25, No.6, pp.1079–1088, 2016.

[5] W. Huan, L. Li and J. Yang, "A modified self-correction subpixel registration algorithm for passive millimeter wave images", *International Conference on Computational Problem-Solving*, Sichuan, China, pp.151–154, 2013.

[6] I. Haller and S. Nedevschi, "Design of interpolation functions for subpixel-accuracy stereo-vision systems", *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, Vol.21, No.2, pp.889–898, 2012.

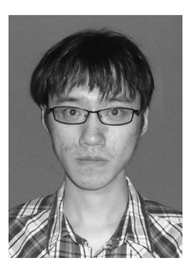
[7] V.C. Miclea, C.C. Vancea and S. Nedevschi, "New sub-pixel interpolation functions for accurate real-time stereo-matching algorithms", *IEEE International Conference on Intelligent Computer Communication and Processing*, Cluj-Napoca, Romania, pp.173–178, 2015.

[8] C.Y. Nan and X.H. Yang, "Spline interpolation sub pixel edge detection method based on improved morphological gradient", *Advanced Materials Research*, Vol.971–973, pp.1756–1759, 2014.

[9] J.M. Morel and G. Yu, "ASIFT: A new framework for fully affine invariant image comparison", *Siam Journal on Imaging Sciences*, Vol.2, No.2, pp.438–469, 2009.

[10] D.G. Lowe, "Distinctive image features from scale-invariant key points", *International Journal of Computer Vision*, Vol.60, No.2, pp.91–110, 2004.

[11] C.R. Vogel and M.E. Oman, "Iterative methods for total variation denoising", *Siam Journal on Scientific Computing*, Vol.17, No.1, pp.227–238, 1997.



MEN Yubo was born in 1989. He received B.S. degree from Harbin Institute of Technology in 2011, and he got direct admission from master to Ph.D. in Harbin Engineering University in 2014. He is now a Ph.D. candidate of Harbin Engineering University. His research interests are stereo matching, image processing. (Email: menyubo@hrbeu.edu.cn)