The Photogrammetric Record 22(120): 321–335 (December 2007)

A SUB-PIXEL LOCATION METHOD FOR INTEREST POINTS BY MEANS OF THE HARRIS INTEREST STRENGTH

QING ZHU (zhuqing@lmars.whu.edu.cn)

BO WU* (wu.573@osu.edu)

NENG WAN (wanneng99@gmail.com)

State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, China

Abstract

The sub-pixel location of interest points is one of the most important tasks in refined image-based 3D reconstruction in digital photogrammetry. The interest point detectors based on the Harris principles are generally used for stereoscopic image matching and subsequent 3D reconstruction. However, the locations of the interest points detected in this way can only be obtained to 1 pixel accuracy. The Harris detector has the following characteristics: (1) the Harris interest strength, which denotes the distinctiveness of an interest point, is a grey scale descriptor which computes the gradient at each sample point in a region around the point, and (2) the Harris interest strengths of the pixels in a template window centred on the interest point exhibit an approximately paraboloid distribution. This paper proposes a precise location method to improve the precision of the interest points on the basis of these characteristics of the Harris interest strength. Firstly, a least squares fit of a paraboloid function to the image grey scale surface using the Harris interest strength is designed in a template window and a Gaussian-distance algorithm is employed to determine the weight. Then, the precise coordinates of this interest point are obtained by calculating the extremities of the fitting surface. The location accuracy of this method is studied both from the theoretical and the practical point of view. Experimental analysis is illustrated with synthetic images as well as actual images, which yielded a location accuracy of 0·15 pixels. Furthermore, experimental results also indicate that this method has the desired anti-image-noise and efficiency characteristics.

KEYWORDS: Gaussian-distance weight determination, Harris interest strength, least squares fitting, sub-pixel location

INTRODUCTION

IMAGE FEATURE EXTRACTION plays important roles in the fields of digital photogrammetry and computer vision. The essential elements of image features are interest points, which include

*Now at The Ohio State University, USA.

© 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.
Blackwell Publishing Ltd. 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street Malden, MA 02148, USA.

corners, junction points, dominant points and all other distinctive points at which the signals change two-dimensionally. Interest point detection should satisfy the following criteria: (1) all the true interest points should be detected; (2) the interest points should be located precisely (Sojka, 2002). This paper focuses on the precise location of the detected interest points for the purpose of stereoscopic image matching and refined three-dimensional (3D) reconstruction (Zhu et al., 2005).

Many interest point detectors are discussed in the literature (Deriche and Giraudon, 1993; Alexandrov, 2002). In the field of image matching and 3D reconstruction, the Moravec detector (Moravec, 1977), the Förstner detector (Förstner, 1994), the SUSAN detector (Smith and Brady, 1997) and the Harris detector (Harris and Stephens, 1988) are generally used. Schmid et al. (2000) experimented with several interest point detectors in several ordinary images, and noted that the Harris detector obtains the best results in terms of repeatability and information content, which are the two key issues for image matching and 3D reconstruction. Zhu et al. (2007) also experimented with generally used interest point detectors in actual aerial images, and their conclusions are mainly in agreement with those of Schmid et al. (2000) that the Harris detector is preferred for building reconstruction from large-scale aerial images. It has to be noted, however, that the standard Harris detector is sensitive to changes in image scale, and also to rotations out of the camera plane. Recent work has shown that scale-space interest points perform better at finding correspondences between images in which there are large changes in viewpoint, scale and illumination (Brown and Lowe, 2002). Lowe (1999, 2004) proposed a new method for image feature extraction called the scale invariant feature transform (SIFT) operator, which is invariant over a wider set of transformations, especially scale change. Recent experiments show that the SIFT operator performs better than the standard Harris detector and other detectors (Mikolajczk and Schmid, 2003), and it is more appropriate for object identification and tracking. However, the SIFT operator detects mainly blob-like interest points (Mikolajczk and Schmid, 2004), while significant points such as the corners of buildings and road edge intersections or corners could not be successfully extracted, and this disadvantage is critical to the subsequent 3D reconstruction of such objects. In contrast to the SIFT method, which detects mainly blobs, the Harris detector responds to corners and highly textured points. Mikolajczk and Schmid (2004) proposed an improved Harris method for detecting interest points which is also invariant to image scale and affine change, and they also presented a comparative evaluation of different detectors and showed that their improved Harris method provided better results than the existing methods.

The Harris detector and the improved Harris method are suitable for image matching and 3D reconstruction. However, they have the drawback that the location accuracies of the detected interest points can only be obtained to 1 pixel, whereas the sub-pixel location of points is a very important task in refined image-based 3D reconstruction, and greatly affects the accuracy and stability of 3D machine vision (Schmid et al., 2000; Chen and Zhang, 2005). This paper therefore investigates the sub-pixel location method for interest points by means of the Harris interest strength.

RELATED WORK

Many methods for the sub-pixel location of interest points have been reported. They may be divided into the following three groups:

(1) The first group contains the methods that locate the interest points to sub-pixel accuracies through image matching (Yang et al., 1993; Bobick and Intille, 1999; Shimizu and Okutomi, 2002). Assuming (i_1, i_2) is a pair of best matches, the sub-pixel

 © 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

accuracy is determined through a correspondence $i_2 + \bar{t}_2$ for each i_1. To determine the optimal value of the displacement $\bar{t}_2$, a continuous constant i_1 is reconstructed from the matching cost function values neighbouring (i_1, i_2) and $\bar{t}_2$ is chosen to bring the reconstruction to its extreme value. Gleason et al. (1991) developed an algorithm that is based on a least squares fitting of a paraboloid function to the surface generated by correlating a reference image feature against a test image search area. Sub-pixel accuracies of better than 1/16 of a pixel have been achieved through calculating the position of the surface maximum. These sub-pixel refinement methods attempt to determine the precise location of points in the secondary image, which correspond to discrete positions in the reference image. Nehab et al. (2005) observed that these strategies could lead to a systematic bias associated with the violation of the general symmetry of matching cost functions, which produces random or coherent noise in the final reconstruction. They presented a sub-pixel refinement strategy by refining both image coordinates simultaneously in a symmetric way. Although these methods can obtain location accuracy of better than 0·1 pixels, they mix the point location and matching together, which leads to a limited application, and also produce high complexity due to the concomitant image matching.

(2) The second group includes the methods that locate the points to sub-pixel accuracies during the detection process directly. Brand and Mohr (1994) made use of a model-based corner detector, matching a part of the image containing a corner against a predefined corner model. Once the fitting is accomplished, the precise position of the corner in the image can be deduced by the knowledge of the corner position in the image. This method can obtain a more accurate location within 0·1 pixels; however, it requires an initial position of the corner and relies principally on the corner model. Zhang et al. (1994) presented a high-precision location operator for straight lines and corner points on digital images. It locates the corner point precisely through the accurate positioning of its two edges. The results of experiments show that the accuracy of corner point location is about 0·02 pixels in ideal conditions; however, the efficiency of this method is lower due to its complicated computing process.

(3) The last group sets out first to detect interest points and then to locate the points to sub-pixel accuracy afterwards, according to certain fitting procedures within a neighbouring area. Chen and Zhang (2005) proposed a sub-pixel X-corner detector for camera calibration, which firstly detects the pixel position of the X-corner, and then a second-order Taylor polynomial is formulated to describe the local intensity profile around the corner. The sub-pixel position of the X-corner can be determined directly by calculating the saddle point of this intensity profile. This method has been tested on computer-simulated data, and the results show that location accuracy of about 0·1 pixels can be obtained, but it only performs well with X-corners. Malik et al. (2002) first use the Harris detector to extract interest points, and then a linear fitting function is employed to compute the sub-pixel position of each point based on weighting the Harris interest strengths of the four connected neighbours. Their algorithm is simple and efficient, but the location accuracy is relatively low.

Compared with the traditional sub-pixel location methods mentioned above, the method proposed in this paper highlights the following three innovations:

(1) The Harris interest strengths of the pixels surrounding an interest point, which denote a certain characteristic in grey-value space in a neighbour area centred on the interest point, are used to construct a local paraboloid.

© 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

(2) A least squares fitting method is employed to refine the paraboloid and then to implement the sub-pixel location of the interest point.
(3) This method is capable of sub-pixel location for general interest points, not only including the corners and junction points, but also the blob-like points and all other types of distinctive points with important two-dimensional texture features.

The method proposed here falls into the last group and refers to the basic idea of Malik et al. (2002) but uses a least squares fitting of a paraboloid function instead of the linear fitting; the eight neighbour pixels are employed instead of the four connected neighbours and will certainly obtain more accurate results. This method can be directly applied to the interest point locations detected by the Harris detector (as well as the improved Harris method) by means of the calculated Harris interest strength of the pixels during the foregoing detection process. When using other detectors to detect interest points, the calculation of the Harris interest strength of the pixels is necessary.

THE HARRIS INTEREST STRENGTH

The basic idea of the Harris detector (Harris and Stephens, 1988) is to use the auto-correlation function in order to determine locations where the signal changes in two directions. A matrix

$$\mathbf{A} = G(\sigma) \otimes \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix}$$

related to the auto-correlation function is computed, which takes into account the first derivatives of the signal on a window, where g_x and g_y are the gradient values in the X and Y directions, respectively, and $G(\sigma)$ is a Gaussian filter to smooth the images. The eigenvectors of this matrix are the principal curvatures of the auto-correlation function, and if at a certain point the two eigenvalues of the matrix are large, then this point is indicated as being an interest point. The response function to calculate the interest strength (called interest value) of every pixel is given by

$$I = \det(\mathbf{A}) - \alpha Trace(\mathbf{A})^2 \tag{1}$$

where α is a parameter set to 0·04 (a suggestion of Harris), and the second term is used to eliminate contour points with one strong eigenvalue.

The improved Harris method (Mikolajczk and Schmid, 2004) considers another matrix

$$\tilde{\mathbf{A}} = \sigma_D^2 G(\sigma_I) \otimes \begin{bmatrix} g_x^2(x, \sigma_D) & g_x g_y(x, \sigma_D) \\ g_x g_y(x, \sigma_D) & g_y^2(x, \sigma_D) \end{bmatrix},$$

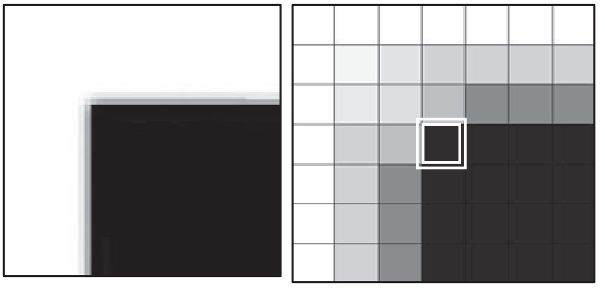
and calculates the interest strength of every pixel through the following response formulation:

$$\tilde{I} = \det(\tilde{\mathbf{A}}) - \alpha Trace(\tilde{\mathbf{A}})^2. \tag{2}$$

Detailed information on this method can be found in Mikolajczk and Schmid (2004).

The Harris interest strength calculated through equation (1) or (2) takes into account both the gradient of grey values in two reciprocally orthogonal directions in a region around the point location, and responds positively to all points with distinctive texture features. If an

 © 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.



(a) (b)

0	0	0	0	0	0	0
0	0	0·02	0·4	0·6	0·7	0·7
0	0·02	2	0·4	8	9	9
0	0·4	0·9	15	4	3	3
0	0·5	8	4	0·8	0·03	0·03
0	0·6	9	3	0·03	0	0
0	0·6	9	3	0	0	0

(c)

FIG. 1. The distribution of the Harris interest strength: (a) simulated image; (b) selected local area (7×7); (c) Harris interest strength ($\times 10^5$).

interest point is detected, then this point has the largest Harris interest strength in a neighbour area, and a small motion in any direction will cause an important change of the Harris interest strength. The Harris interest strength of an interest point denotes a descriptor in the grey-value space, which describes the grey magnitude and distribution information of a local area centred on the interest point. Considering a local area in the image, the Harris interest strength of all its pixels composes a surface in the grey-value space, which describes the grey magnitude and distribution in this local area.

Fig. 1(a) shows a simulated image with a typical corner generated by convolving a 2D Gaussian function from an ideal corner. Its grey distribution includes corner, edges and smooth area. The Harris detector is used to detect interest points in the image. Fig. 1(b) shows a local 7×7 window centred on the corner, and Fig. 1(c) illustrates the Harris interest strength of all the pixels in this window. Fig. 1(c) indicates that the distribution of the pixels' Harris interest strength presents an approximately paraboloid distribution. Assuming the paraboloid surface is continuous the extreme value of the paraboloid should then be the precise location of the corner.

THE SUB-PIXEL LOCATION METHOD

Least Squares Fitting

For a group of observed data x_1, y_1, z_1, x_2, y_2, z_2, …, x_n, y_n, z_n, its "best" fitting function $f(x, y)$ should fulfil the following standard least squares equation:

$$I = \sum_{i=1}^{n} [z_i - f(x_i, y_i)]^2 \tag{3}$$

where x, y are the independent variables and z is the dependent variable, $n \geq k$, k is the number of independent parameters in the function $f(x, y)$, and it is also the least number of samples required (Yuan, 1996).

The fitting technique can be easily generalised from a best-fit line to a best-fit polynomial. If a low-order polynomial is employed, the fitting accuracy must be bad, while high-order polynomials may lead to unstable fitting results. So, this paper selects a quadratic polynomial function ($k = 6$, which is a_0, a_1 … a_5) as equation (4) to fit a paraboloid surface, and then obtains the parameters of this function through least squares adjustment:

© 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

$$f(x,y) = a_0x^2 + a_1y^2 + a_2xy + a_3x + a_4y + a_5. \tag{4}$$

Because the least squares adjustment method is well known, this paper omits the detail of the process and only describes the key equations. This adjustment model is expressed as follows:

$$\underset{n,1}{\hat{L}} = \underset{n,k}{\mathbf{B}}\,\underset{k,1}{\hat{X}} + \underset{n,1}{d} \tag{5}$$

where $\underset{k,1}{\hat{X}}$ denotes the k independent parameters, $\underset{n,1}{\hat{L}}$ is the n samples and $\underset{n,1}{d}$ is the constant item in the expression.

Assume

$$\hat{X} = X^0 + \hat{x} \tag{6}$$

where X^0 denotes the approximation of the parameters. Substituting equation (6) into (5) it can be assumed that

$$l = L - (\mathbf{B}X^0 + d) = L - L^0 \tag{7}$$

where $L^0 = \mathbf{B}X^0 + d$. Because L^0 are the approximations of the sample values, l is the difference between the samples and their approximations, so that the error equation can be stated as follows:

$$V = \mathbf{B}\hat{x} - l. \tag{8}$$

Assuming $\mathbf{P}$ as the weight matrix of the samples, and taking $V^{\mathrm{T}}\mathbf{P}V = \min$ as the adjustment criteria, the undetermined coefficient $\hat{x}$ in the error equation can be calculated. The following equation can be reached from the solution of the function extreme value:

$$\frac{\partial V^{\mathrm{T}}\mathbf{P}V}{\partial x} = 2V^{\mathrm{T}}\mathbf{P}\frac{\partial V}{\partial X} = V^{\mathrm{T}}\mathbf{PB}.$$

The solution of this equation is

$$\hat{x} = (\mathbf{B}^{\mathrm{T}}\mathbf{PB})^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P}l. \tag{9}$$

Substituting equation (9) in (8), the correction V can be calculated and the results of the least squares adjustment are consequently obtained as follows:

$$\hat{X} = X^0 + \hat{x}, \quad \hat{L} = L + V.$$

Gaussian-Distance Weight Determination

In the least squares adjustment method, the weight values in the weight matrix $\mathbf{P}$ denote the correlation between the interest point and its surrounding pixels. Equation (3) is just a simple square summation calculation of the difference between the fitting function value and the sample value. That is to say, the weight values of all the sampling pixels are the same. However, the contributions of different pixels to the fitting function are actually different because the distances from different pixels to the interest point are different and therefore their weight values must also be different. For a pixel numbered i, its weight value P_i is correlated with the distance d_i from this pixel to the interest point, and the greater the value of d_i, the less

 © 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

the correlation between them, so the lower the value of P_i. Otherwise, the pixel i has a larger value of P_i.

This paper proposed a Gaussian-distance weight determination method as follows:

$$P_i = \mathrm{e}^{-d_i^2/k^2} \tag{10}$$

where d_i is the distance between the pixel i to the interest point, and k is an optional constant (suggest: $k = 0{\cdot}2$).

Operation

Considering a 3×3 template window as shown in Fig. 2, equation (4) is used to carry out the least squares fitting. The sample values f_i ($i = 0, 1, \ldots, n$) are the Harris interest strengths calculated through equation (1) or (2) and $n = 9$ is the number of sample values.

Set $\underset{n,1}{\hat{L}} = [f_0, f_1 \ldots f_n]^{\mathrm{T}}$ and $\underset{6,1}{\hat{X}} = [a_0, a_1 \ldots a_5]^{\mathrm{T}}$, where a_i ($i = 0, 1, \ldots, 5$) are the six independent parameters. Select any six pixels at reasonably small distances from the centre point for initial sampling, for example, f_0, f_1, f_2, f_3, f_5 and f_7. Because the fitting function is invariant to parallel movement, the coordinates of the centre point in the template window (initial interest point) can be assumed to be (0, 0), and the width of each pixel can be assumed to be 1 unit. The initial values of the parameters in the fitting function can then be calculated as $\underset{6,1}{X^0} = [a_0^0, a_1^0, \ldots, a_5^0]^{\mathrm{T}}$, which in detail is

$$\underset{6,1}{X^0} = \mathbf{T} \underset{6,1}{F} \tag{11}$$

where

$$\mathbf{T} = \begin{pmatrix} 0 & 0 & 0{\cdot}5 & -1 & 0{\cdot}5 & 0 \\ 0{\cdot}5 & 0 & 0 & -1 & 0 & 0{\cdot}5 \\ -1 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & -0{\cdot}5 & 0 & 0{\cdot}5 & 0 \\ 0{\cdot}5 & 0 & 0 & 0 & 0 & 0{\cdot}5 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{pmatrix}.$$

f_0 (−1,1)	f_1 (0,1)	f_2 (1,1)
f_3 (−1,0)	f_4 (0,0)	f_5 (1,0)
f_6 (−1,−1)	f_7 (0,−1)	f_8 (1,−1)

FIG. 2. The template window for least squares fitting.

© 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

While

$$\underset{n,6}{\mathbf{B}} = \begin{pmatrix} x_1^2 & y_1^2 & x_1y_1 & x_1 & y_1 & 1 \\ x_2^2 & y_2^2 & x_2y_2 & x_2 & y_2 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_n^2 & y_n^2 & x_ny_n & x_n & y_n & 1 \end{pmatrix}.$$

Using the Gaussian-distance weight determination method as equation (10) showed that to calculate the weight matrix **P**, substituting $\hat{L}, \hat{X}, X^0, \mathbf{B}, \mathbf{P}$ in the adjustment equation (5), and the final correction $\hat{X}$ to the initial independent parameters X^0 can be obtained through equations (8) and (9), and finally the adjustment values for all the independent parameters can also be obtained.

After calculating the position of the fitted paraboloid surface maximum, the distance from the accurate coordinate of the interest point to the initial coordinate can be calculated as

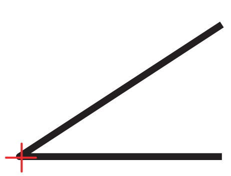
$$\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} \dfrac{2a_1a_3 - a_2a_4}{a_2^2 - 4a_1a_0} \\ \dfrac{2a_0a_4 - a_2a_3}{a_2^2 - 4a_1a_0} \end{pmatrix}. \tag{12}$$

So the accurate coordinate of the interest point is

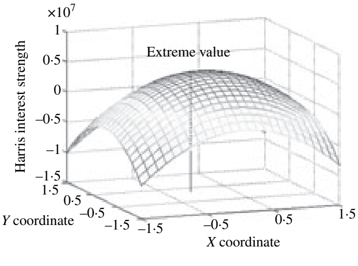
$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \dfrac{2a_1a_3 - a_2a_4}{a_2^2 - 4a_1a_0} + x_{\mathrm{ini}} \\ \dfrac{2a_0a_4 - a_2a_3}{a_2^2 - 4a_1a_0} + y_{\mathrm{ini}} \end{pmatrix} \tag{13}$$

where (x_{ini}, y_{ini}) is the initial pixel coordinate when they are detected by the Harris detector or the improved Harris method.

Fig. 3 shows the sub-pixel location process using a computer-simulated image. Firstly, interest points are detected as shown in Fig. 3(a), and then the corner, of which the actual pixel



216244	696907	1495135
397338	4043951	1095007
167681	1211299	1246681



(a) (b) (c)

FIG. 3. The precise location process with a computer-simulated image: (a) reference image with an interest point; (b) Harris interest strengths in a 3×3 template window; (c) the fitted paraboloid surface.

 © 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

coordinate is (248·5, 248·5), is selected to carry out precise location. Fig. 3(b) shows the Harris interest strength of the pixels in a 3×3 template window, in which the selected interest point is located in the centre. Fig. 3(c) is a paraboloid surface fitted with the least squares adjustment using the Harris interest strength of those pixels. After calculating the position of this surface maximum, the interest point can be located to a position (248·553, 248·459), and a sub-pixel accuracy of 0·067 pixels is obtained.

Precision Analysis

Theoretical Precision

The theoretical precision of this method can be determined error propagation using covariance propagation law. From $\underset{6,1}{X^0} = \mathbf{T} \underset{6,1}{F}$, the covariance matrix $\mathbf{Q_{XX}}$ of the six independent parameters $\underset{6,1}{\hat{X}} = [a_0, a_1 \ldots a_5]^{\mathrm{T}}$ can be calculated as

$$\mathbf{Q_{XX}} = \mathbf{T Q_{FF} T}^{\mathrm{T}} \tag{14}$$

where $\mathbf{Q_{FF}}$ is the covariance matrix of the samples F. Then, from equation (13), the position variance of the interest point in the X and Y directions can be calculated respectively according to the covariance propagation law of non-linear functions, which are shown as follows:

$$\sigma_x^2 = K_1^{\mathrm{T}} \mathbf{Q_{XX}} K_1 \tag{15}$$

$$\sigma_y^2 = K_2^{\mathrm{T}} \mathbf{Q_{XX}} K_2 \tag{16}$$

where K_1 and K_2 are the partial derivative matrices of the non-linear function in equation (13) to a_i. They are described in detail as follows:

$$\begin{aligned} K_1^{\mathrm{T}} &= \left(\frac{\partial x}{\partial a_0} \; \frac{\partial x}{\partial a_1} \; \frac{\partial x}{\partial a_2} \; \frac{\partial x}{\partial a_3} \; \frac{\partial x}{\partial a_4} \; \frac{\partial x}{\partial a_5} \right) \\ &= \left(\frac{4a_1(2a_1a_3 - a_2a_4)}{(a_2^2 - 4a_1a_0)^2} \; \frac{2a_2^2a_3 - 4a_0a_2a_4}{(a_2^2 - 4a_1a_0)^2} \; \frac{a_2^2a_4 - 4a_1a_2a_3 + 4a_0a_1a_4}{(a_2^2 - 4a_1a_0)^2} \right. \\ &\qquad \left. \frac{2a_1}{a_2^2 - 4a_1a_0} \; \frac{-a_2}{a_2^2 - 4a_1a_0} \quad 0 \right) \end{aligned} \tag{17}$$

$$\begin{aligned} K_2^{\mathrm{T}} &= \left(\frac{\partial y}{\partial a_0} \; \frac{\partial y}{\partial a_1} \; \frac{\partial y}{\partial a_2} \; \frac{\partial y}{\partial a_3} \; \frac{\partial y}{\partial a_4} \; \frac{\partial y}{\partial a_5} \right) \\ &= \left(\frac{2a_2^2a_4 - 4a_1a_2a_3}{(a_2^2 - 4a_1a_0)^2} \; \frac{8a_0^2a_4 - 4a_0a_2a_3}{(a_2^2 - 4a_1a_0)^2} \; \frac{a_2^2a_3 + 4a_0a_1a_3 - 4a_0a_2a_4}{(a_2^2 - 4a_1a_0)^2} \right. \\ &\qquad \left. \frac{-a_2}{a_2^2 - 4a_1a_0} \; \frac{2a_0}{a_2^2 - 4a_1a_0} \quad 0 \right). \end{aligned} \tag{18}$$

© 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

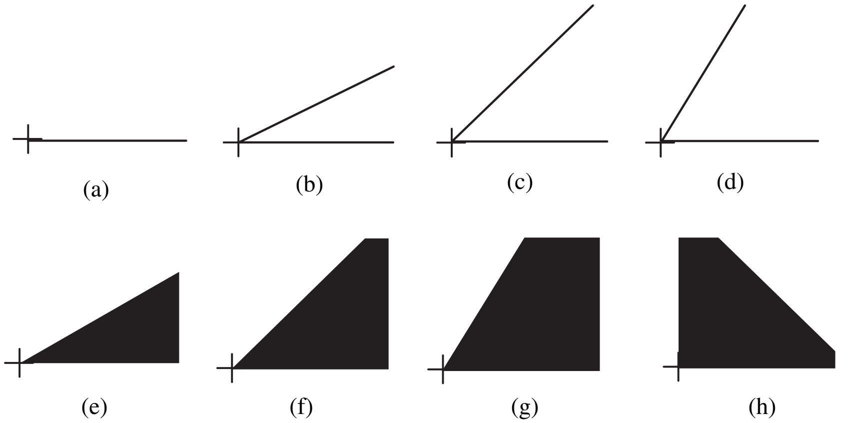


FIG. 4. Computer-simulated images with interest points marked as crosses: (a) end point; (b) 30° corner; (c) 45° corner; (d) 60° corner; (e) 30° solid corner; (f) 45° solid corner; (g) 60° solid corner; (h) 90° solid corner.

TABLE I. Sub-pixel location of interest points in simulated images.

Interest points	*Actual coordinates (x, y)*	*Sub-pixel located coordinates (x, y)*	v_x	v_y	v
(a)	248·500, 248·500	248·499, 248·500	−0·001	0	0·001
(b)	248·500, 248·500	248·594, 248·518	0·094	0·018	0·094
(c)	248·500, 248·500	248·552, 248·459	0·053	−0·041	0·067
(d)	248·500, 248·500	248·558, 248·394	0·059	−0·106	0·121
(e)	248·500, 248·500	248·728, 248·506	0·228	0·006	0·228
(f)	248·500, 248·500	248·758, 248·485	0·258	−0·015	0·259
(g)	248·500, 248·500	248·694, 248·482	0·194	−0·018	0·194
(h)	248·500, 248·500	248·546, 248·462	0·046	−0·038	0·059
Location precision m: $(m = \sqrt{\sum V^2/n})$				0·15	

Finally, the theoretical precision of the interest point can be determined through $\sigma = \sqrt{\sigma_x^2 + \sigma_y^2}$. Making use of computer-simulated data (shown in Fig. 4 and Table I) proves that this method can obtain the theoretical precision of 0·14 pixels.

Actual Precision

Several computer-simulated images (Fig. 4) have been employed to calculate the actual precision of this method. Firstly, some identified interest points (corners and end points) were detected using the standard Harris detector, and then the sub-pixel location method was used to calculate the sub-pixel position of these interest points. The equation used to calculate the actual precision is

$$m = \sqrt{\frac{\sum V^2}{n}}$$

where $V = \sqrt{v_x^2 + v_y^2}$, and n is the number of interest points.

 © 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

The experimental results are shown in Table I. The results show that the actual precision of 0·15 pixels is obtained on condition that there is no noise in the data, and this agrees with the theoretical precision approximately.

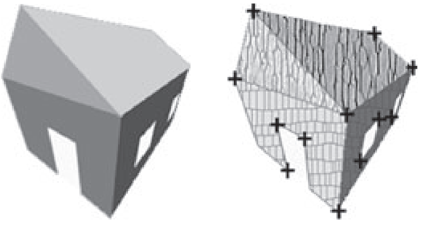
Experimental Analysis

Experiment with a Synthetic Image

A more complicated synthetic image with respect to the images shown in Fig. 4 was tested. In Fig. 5(a), the coordinates of all the junction points can be precisely calculated. Fig. 5(b) is the textured image of Fig. 5(a). After detecting interest points using the Harris detector in Fig. 5(b), the sub-pixel location method was conducted on the detected interest points. The 12 points marked with crosses were used to calculate the location precision, and a precision of 0·17 pixels was obtained.

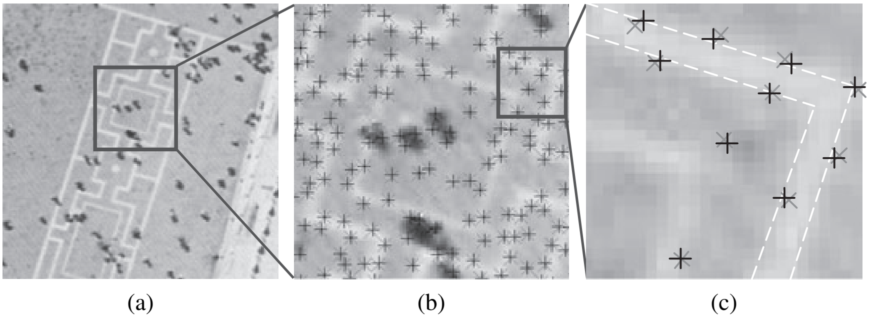
Experiment with an Actual Aerial Image

Another actual aerial image shown in Fig. 6(a) was also employed to test the proposed sub-pixel location method. The Harris detector was firstly used to detect interest points in this



(a) (b)

FIG. 5. The precise location of points with a textured image: (a) a house image; (b) the textured house image with interest points.



(a) (b) (c)

FIG. 6. Sub-pixel location of interest points in actual aerial image: (a) the reference image; (b) zoomed image section (4×) with interest points detected by Harris detector (grey × crosses) and their corresponding precise location (black + crosses); (c) zoomed image (15×) section, in which the white dashed lines indicate the actual edge of the road in the square.

© 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

TABLE II. The location precision (pixel) with noise.

Noise level σ	0·01	0·05	0·10	0·15	0·20
Location precision m	0·22	0·23	0·24	0·22	0·23

aerial image, and then the sub-pixel location method was carried out for the detected interest points. Fig. 6(b) shows the zoomed image section, in which the grey × crosses denote the original interest points detected by the Harris detector and the black + crosses denote their position after precise location. In Fig. 6(c), the actual edges of the road in the square were projected onto the aerial image, which derived from the corresponding 2D cadastral map. The differences between the initial position and the precision location of these interest points are illustrated in Fig. 6(c), which shows that the latter are closer to the actual position.

Image Noise

To study the anti-image-noise ability in the presence of image noise, a Gaussian noise ($f(x) = (1/\sqrt{2\pi}\sigma)\exp(-x^2/2\sigma^2)$, where σ is the variance of the Gaussian function) was added to the images in Fig. 4. The results of this experiment are displayed in Table II.

The results show that the location precision is better than 0·25 pixels even when the σ of the Gaussian noise reaches 0·20. This proves that this method is capable of obtaining good results even where there is a significant level of image noise.

Efficiency

The proposed method was implemented on a 1·5 GHz Pentium 4 machine, which processes 24 000 to 30 000 points per second (corresponding image size range from 5681 × 8560 pixels to 1982 × 3043 pixels) using this sub-pixel location procedure. The high efficiency confirms the benefits of the application of this method.

CONCLUSION

In this paper, a sub-pixel location method for interest points has been investigated. The theoretical and experimental analyses led to the following conclusions:

(1) Based on the distribution characteristics of the Harris interest strengths of the pixels in a template window, where the interest point is located at the centre, least squares fitting of a paraboloid function to the image grey scale surface by means of the Harris interest strength can be used to locate the interest points to sub-pixel accuracy.
(2) The proposed method can be used to locate interest points including corners, junction points, blob-like points and all other types of distinctive points with important two-dimensional texture features.
(3) It is shown that an accuracy of 0·15 pixels can be achieved with this method in ideal conditions.
(4) This method has the desired performance in terms of efficiency and of anti-image-noise characteristics.

The proposed sub-pixel location method can be widely applied in the fields of digital photogrammetry, computer vision and image understanding. The principles of this method will be further studied and extended to the accurate location of edge and surface features.

 © 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

Acknowledgements

The work described in this paper was supported by National Natural Science Foundation of China (40671158), National Basic Research Program of China (2002CB312101) and also by the National High Technology Research and Development Program of China (2006AA12Z224). All the reviewers' comments on this paper are appreciated.

References

Alexandrov, A., 2002. Corner detection overview and comparison. http://alex.bmwquebec.ca/cs558/CornerDet.pdf [Accessed: 6th July 2007].

Bobick, A. F. and Intille, S. S., 1999. Large occlusion stereo. *International Journal of Computer Vision*, 33(3): 181–200.

Brand, P. and Mohr, R., 1994. Accuracy in image measure. *SPIE*, 2350: 218–228.

Brown, M. and Lowe, D., 2002. Invariant features from interest point groups. *Proceedings of the 13th British Machine Vision Conference*, Cardiff. Pages 656–665.

Chen, D. and Zhang, G., 2005. A new sub-pixel detector for X-corners in camera calibration targets. *Proceedings of the WSCG International Conference in Central Europe on Computer Graphics and Visualization*. Pages 97–100.

Deriche, R. and Giraudon, G., 1993. A computational approach for corner and vertex detection. *International Journal of Computer Vision*, 10(2): 101–124.

Förstner, W., 1994. A framework for low level feature extraction. *Proceedings of the 3rd European Conference on Computer Vision*, Stockholm, Sweden, II: 383–394.

Gleason, S. S., Hunt, M. A. and Jatko, W. B., 1991. Sub-pixel measurement of image features based on paraboloid surface fit. *SPIE*, 1386: 135–144.

Harris, C. and Stephens, M., 1988. A combined corner and edge detector. *Proceedings of 4th Alvey Vision Conference*, Manchester. Pages 147–151.

Lowe, D. G., 1999. Object recognition from local scale-invariant features. *Proceedings of the International Conference on Computer Vision*, Corfu, Greece, II: 1150–1157.

Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2): 91–110.

Malik, S., Roth, G. and McDonald, C., 2002. Robust 2D tracking for real-time augmented reality. *Proceedings of Vision Interface (VI)*, Calgary, Alberta, Canada. Pages 399–406.

Mikolajczk, K. and Schmid, C., 2003. A performance evaluation of local descriptors. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Madison, Wisconsin. Pages 257–264.

Mikolajczk, K. and Schmid, C., 2004. Scale and affine invariant interest point detectors. *International Journal of Computer Vision*, 60(1): 63–86.

Moravec, H. P., 1977. Towards automatic visual obstacle avoidance. *Proceedings of the 5th International Joint Conference on Artificial Intelligence*, Cambridge, Massachusetts, 2: 584.

Nehab, D., Rusinkiewicz, S. and Davis, J., 2005. Improved sub-pixel stereo correspondences through symmetric refinement. *10th IEEE International Conference on Computer Vision (ICCV)*, 1: 557–563.

Schmid, C., Mohr, R. and Bauckhage, C., 2000. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2): 151–172.

Shimizu, M. and Okutomi, M., 2002. Precise sub-pixel estimation on area-based matching. *Systems and Computers in Japan*, 23(7): 1–10.

Smith, S. M. and Brady, J. M., 1997. SUSAN—a new approach to low level image processing. *International Journal of Computer Vision*, 23(1): 45–78.

Sojka, E., 2002. A new and efficient algorithm for detecting the corners in digital images. *Proceedings of the 24th DAGM Symposium on Pattern Recognition*. Pages 125–132.

Yang, Y., Yuille, A. and Lu, J., 1993. Local, global, and multilevel stereo matching. *Proceedings of CVPR'93, IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Pages 274–279.

Yuan, J. Y., 1996. Numerical methods for generalized least squares problems. *Journal of Computational and Applied Mathematics*, 66(1/2): 571–584.

Zhang, J. Q., Zhang, Z. X. and Wang, Z. H., 1994. High-precision location operator of straight line and corner point on digital images. *Acta Geodaetica et Cartographica Sinica*, 23(2): 98–106.

Zhu, Q., Zhao, J., Lin, H. and Gong, J., 2005. Triangulation of well-defined points as a constraint for reliable image matching. *Photogrammetric Engineering & Remote Sensing*, 71(9): 1063–1069.

Zhu, Q., Wu, B. and Wan, N., 2007. A filtering strategy for interest point detecting to improve repeatability and information content. *Photogrammetric Engineering & Remote Sensing*, 73(5): 547–553.

© 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

Résumé

La détermination sub-pixellaire de la position des points caractéristiques est l'une des tâches les plus importantes dans la reconstitution 3 D de qualité à base d'images en photogrammétrie numérique. On utilise généralement des identifiants de points caractéristiques basés sur les principes de Harris pour réaliser l'appariement des images stéréoscopiques et obtenir la reconstitution 3 D recherchée. Toutefois ces principes présentent des défauts de « délocalisation » et négligent certains éléments dans les formes. L'identifiant de Harris a les caractéristiques suivantes: 1) Le classement par intérêt de Harris qui signale les éléments distinctifs d'un point caractéristique, est un descripteur par échelle des gris qui calcule le gradient en chaque point échantillonné entourant le point caractéristique, et 2) les classements de Harris des pixels situés à l'intérieur d'une fenêtre-gabarit centrée sur le point caractéristique font apparaître une distribution approximativement en forme de paraboloïde. On propose dans cet article une méthode de positionnement précis permettant d'améliorer la précision des points caractéristiques tout en gardant les principes de Harris. On détermine d'abord les équations d'un paraboloïde correspondant à cette surface des échelles de gris définie par la méthode de Harris dans cette fenêtre-gabarit, en procédant à une compensation par moindres carrés où les poids sont obtenus par un algorithme d'après la distance gaussienne. Ensuite on calcule les coordonnées précises de ce point caractéristique en se reportant aux extrémités de la surface de compensation précédente. On a étudié la précision de ce positionnement d'un point de vue à la fois théorique et pratique. L'expérience a porté sur des images de synthèse ainsi que sur des images réelles qui ont fourni une précision de positionnement de 0,15 pixels. De plus les résultats de ces essais ont montré également que le bruit sur l'image était réduit et que l'efficacité était accrue.

Zusammenfassung

Bei der hochpräzisen photogrammetrischen 3D-Rekonstruktion mit Hilfe digitaler Bilddaten ist die Subpixel-Lokalisierung markanter Bildpunkte von entscheidender Bedeutung. Üblicherweise werden Interestoperatoren nach dem Harris Prinzip zur stereoskopischen Bildzuordnung und nachfolgender 3D-Rekonstruktion eingesetzt. Damit können die markanten Punkte jedoch nur mit einem Pixel Genauigkeit lokalisiert werden. Der Harris Operator hat die folgenden Eigenschaften: 1) der Harris Interessenwert beschreibt die Eindeutigkeit eines interessanten Punktes, und wird aus Grauwertgradienten in einem Gebiet um den Punkt berechnet und 2) die Harris Interessenwerte der Bildpunkte in einem Musterfenster um den markanten Punkt haben eine genähert paraboloide Verteilung. Auf der Basis dieser beiden Eigenschaften wird eine Methode zur präziseren Lokalisierung markanter Punkte vorgeschlagen. Zunächst wird eine Kleinste Quadrate Anpassung einer Paraboloid-Funktion an die Grauwertoberfläche mit Hilfe des Harris Interessenwerts in einem Musterfenster durchgeführt. Zur Ermittlung des Gewichts wird ein Algorithmus zur Bestimmung der Gauß-Distanz eingesetzt. Durch die Bestimmung der Extrema der angepassten Oberfläche werden präzise Koordinaten des markanten Punktes abgeleitet. Die Genauigkeit der Lokalisierung mit dieser Methode wird sowohl theoretisch, als auch empirisch untersucht. Die experimentelle Analyse wird dabei mit synthetischen und realen Bilddaten durchgeführt. Es konnte eine Lokalisiergenauigkeit von 0·15 Pixel erreicht werden. Zudem gibt es Hinweise, dass

 © 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.

diese Methode auch die erhofften Eigenschaften hinsichtlich Rauschunabhängigkeit und Effizienz aufweist.

Resumen

La localización subpíxel de puntos de interés es una de las operaciones más importantes de la fotogrametría para la reconstrucción fina tridimensional a partir de imágenes. Los detectores de puntos de interés basados en el detector de Harris se utilizan habitualmente para la correspondencia de imágenes estereoscópicas y la posterior reconstrucción tridimensional. Sin embargo, las localizaciones de los puntos de interés identificadas de este modo sólo pueden alcanzar exactitudes de un píxel. El detector de Harris tiene las siguientes características: 1) la intensidad de interés de Harris, que indica la singularidad de un punto de interés, es un descriptor en escala de grises que calcula el gradiente en cada punto de muestreo en una región alrededor de cada punto, y 2) las intensidades de interés de Harris de los píxeles dentro de una ventana centrada en el punto de interés presentan una distribución aproximadamente parabólica. Este artículo propone un método de localización precisa para mejorar la precisión de los puntos de interés basándose en esas características de la intensidad de interés de Harris. En primer lugar, se aplicó un ajuste por mínimos cuadrados de una función parabólica a la imagen en escala de grises, utilizando la intensidad de interés de Harris en una ventana y un algoritmo de distancia gaussiana que permite determinar los pesos. A continuación se obtienen las coordenadas precisas de este punto de interés calculando los extremos de la superficie de ajuste. La exactitud de localización proporcionada por este método se estudia desde los puntos de vista teórico y práctico. El análisis experimental se ilustra con imágenes sintéticas así como con imágenes reales, que alcanzaron una exactitud de localización de 0,15 píxeles. Además, los resultados experimentales también indican que este método posee las características esperadas de eficiencia y ante la presencia de ruido en la imagen.

© 2007 The Authors. Journal Compilation © 2007 The Remote Sensing and Photogrammetry Society and Blackwell Publishing Ltd.