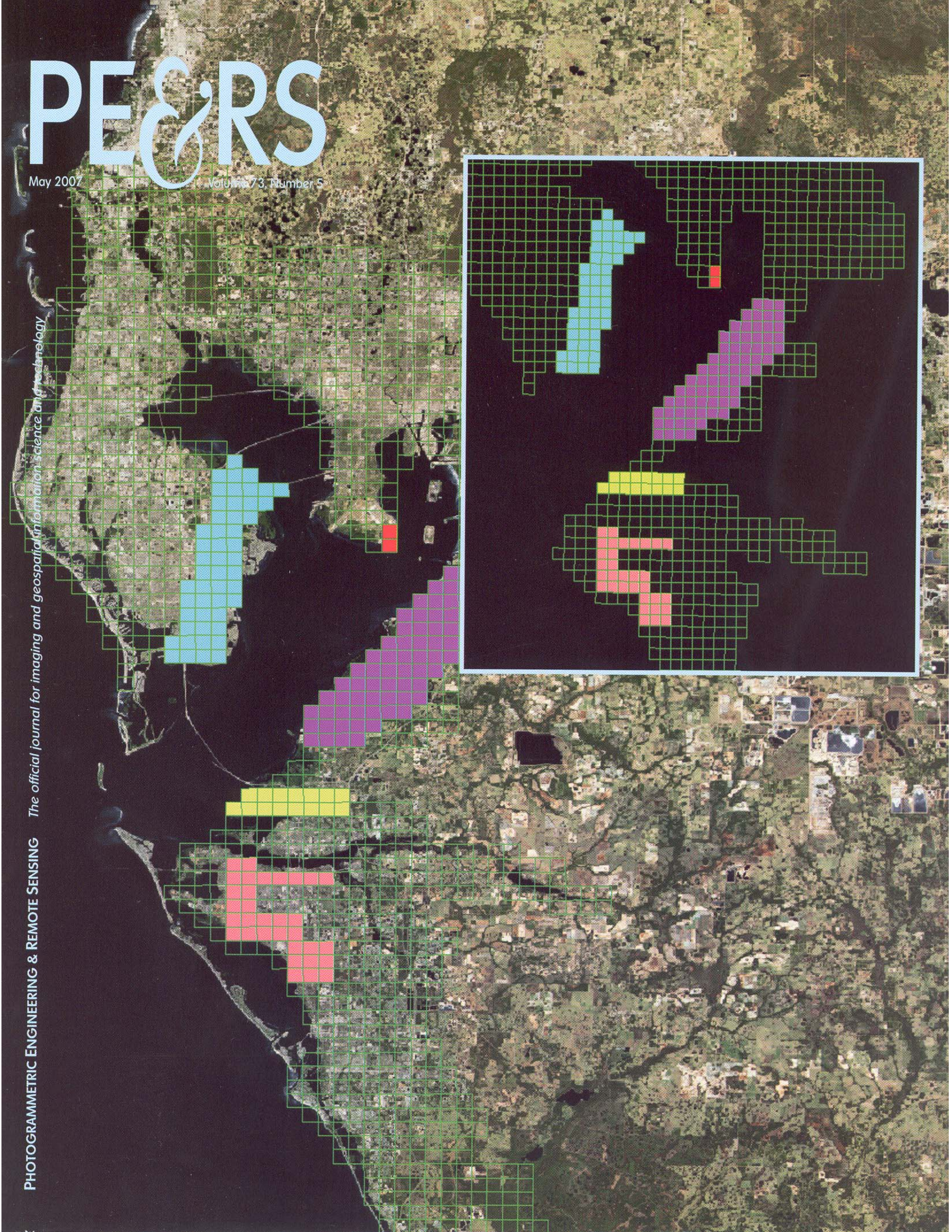
PE&RS

May 2007

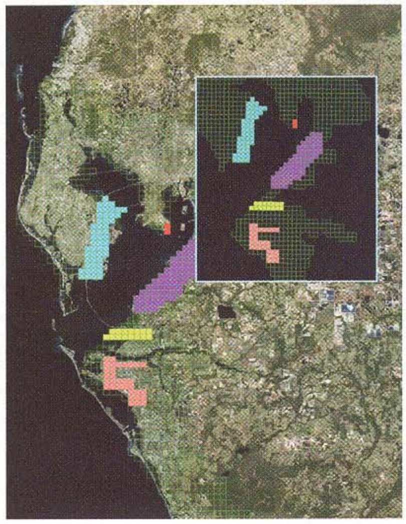
Volume 73, Number 5

The official journal for imaging and geospatial information science and technology

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING



This month's cover image depicts a Distributed Production Management System (DPMS) based on GeoCue Corporation's core project management system and GeoCue Web Server. This new system, developed in cooperation with URS Corporation's Geospatial Technologies Team of Gaithersburg, Maryland. The backdrop graphic depicts a simulated Master Project view over the Tampa Bay area of Florida. The tiles represent project components such as LIDAR or Orthophoto work units. The fill color, corresponding to the checklist depicted in the lower left of the cover, depict the status of various project tiles being worked by a number of geographically dispersed contractors. Superimposed over the Master Project is an Internet Explorer view through the Master Project Web Portal of the project, allowing project status viewing from any location in the world with a simple web browser. For additional information contact Bob Ryan (bob.ryan@urscorp.com) of URS Corporation or Lewis Graham (lgraham@geocue.com) of GeoCue Corporation. The backdrop NaturalVue™ images were provided courtesy of MDA Federal (formerly EarthSat) of Rockville, Maryland (contact Jon Dykstra, jon.dykstra@mdafederal.com).

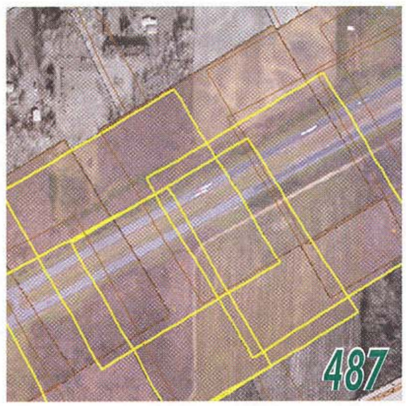


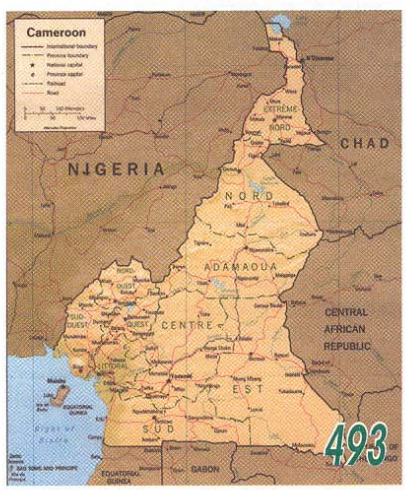
Highlight Article

Columns & Updates

Announcements

Departments





PE&RS

May 2007 Volume 73, Number 5

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING

The official journal for imaging and geospatial information science and technology

asprs

Immediate electronic access to all peer-reviewed articles in this issue is available to ASPRS members at www.asprs.org. Just log in to the ASPRS web site with your membership ID and password and download the articles you need.

Peer-Reviewed Articles

Correction: In last months' Region of the Month column, the Potomac Region was erroneously cited as the winner in the text accompanying the column. The correct winner was the Columbia River Region.

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING is [the of]ficial journal of the American Society for Photogrammetry and [Remote] Sensing. It is devoted to the exchange of ideas and informatio[n about] the applications of photogrammetry, remote sensing, and geo[graphic] information systems.

The technical activities of the Society are conducted thro[ugh the] following Technical Divisions: Geographic Information System[s, Pho]togrammetric Applications, Primary Data Acquisition, Profe[ssional] Practice, and Remote Sensing Applications. Additional informa[tion on] the functioning of the Technical Divisions and the Society can b[e found] in the Yearbook issue of *PE&RS*.

Correspondence relating to all business and editorial [matters] pertaining to this and other Society publications should be d[irected] to the American Society for Photogrammetry and Remote S[ensing], 5410 Grosvenor Lane, Suite 210, Bethesda, Maryland 2081[4-2160,] including inquiries, memberships, subscriptions, changes in a[ddress,] manuscripts for publication, advertising, back issues, and publi[cations.] The telephone number of the Society Headquarters is 301-49[3-0290;] the fax number is 301-493-0208; email address is asprs@aspr[s.org.]

PE&RS. *PE&RS* (ISSN0099-1112) is published monthly by the A[merican] Society for Photogrammetry and Remote Sensing, 5410 Gr[osvenor] Lane, Suite 210, Bethesda, Maryland 20814-2160. Periodicals [postage] paid at Bethesda, Maryland and at additional mailing offices.

SUBSCRIPTION. Effective January 1, 2007, the Subscription [rate for] non-members per calendar year (companies, libraries) is $31[0;] $373 for **Canada Airmail** (includes **6% for Canada's Goods a[nd Ser]vice Tax (GST#135123065)**; $370 for all other foreign.

POSTMASTER. Send address changes to *PE&RS*, ASPRS Headq[uarters,] 5410 Grosvenor Lane, Suite 210, Bethesda, Maryland 2081[4-2160.] CDN CPM #(40020812)

MEMBERSHIP. Membership is open to any person actively eng[aged in] the practice of photogrammetry, photointerpretation, remote [sensing] and geographic information systems; or who by means of ed[ucation] or profession is interested in the application or development [of these] arts and sciences. Membership is for one year, with renewal b[ased on] the anniversary date of the month joined. Membership Dues [include] a 12-month subscription to *PE&RS* valued at $60. Subscriptio[n is part] of membership benefits and cannot be deducted from annu[al dues.] Annual dues for Regular members (Active Member) is $120; for [Student] members it is $45; for Associate Members it is $80 (see des[cription] on application in the back of this Journal). An additional [postage] surcharge is applied to all International memberships: Add [... for] **Canada Airmail**, and **6% for Canada's Goods and Service T[ax (GST] #135123065)**; all other foreign add $60.00.

A Filtering Strategy for Interest Point Detecting to Improve Repeatability and Information Content

Qing Zhu, Bo Wu, and Neng Wan

Abstract

This paper compares several stereo image interest point detectors with respect to their repeatability and information content through experimental analysis. The Harris-Laplace detector gives better results than other detectors in areas of good texture; however, in areas of poor texture, the Harris-Laplace detector may be not the best choice. A feature-related filtering strategy is designed for the Harris-Laplace detector (as well as the standard Harris detector) to improve the repeatability and information content for imagery with both good and poor texture: (a) the local information entropy is computed to describe the local feature of the image; and (b) the redundant interest points are filtered according to the interest strength and the local information entropy. After the filtering process, the repeatability and information content of the final interest points are improved, and the mismatching then can be reduced. This conclusion is supported by experimental analysis with actual stereo images.

Introduction

Image feature extraction plays an important role in the field of image matching, object description, movement estimating, and object tracking. Interest points are the essential elements of an image feature, where an interest point simply means any distinctive point in the image for which the signal changes two-dimensionally. In image analysis and stereovision, the choice of an interest point detector to deal with different application requirements is especially significant (Schmid *et al.*, 2000; Sebe *et al.*, 2003).

For the purpose of stereo image matching and the subsequent three-dimensional (3D) reconstruction, only the detecting efficiency and location accuracy were considered in detail in the past when choosing an interest point detector (Brand and Mohr, 1994; Baker and Nayar, 1999). Schmid *et al.* (2000) pointed out that the choice of an interest point detector should be based on its repeatability and information content. The repeatability of interest points determines the matching reliability, while the information content indicates the significance of such interest points to the 3D object reconstruction.

Qing Zhu and Neng Wan are with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, 129 LuoYu Road, Wuhan, 430079, P.R. China.

Bo Wu was with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, and is currently with the Mapping and GIS Laboratory, Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University, 470 Hitchcock Hall, 2070 Neil Avenue, Columbus, OH 43210 (wu.573@osu.edu).

In experiments with general images, the Harris detector gives better repeatability and information content than other detectors (Schmid *et al.*, 2000). However, for large-scale aerial images or satellite images, many poor interest points would be detected by the Harris detector in imagery with poor texture, and some interest points with small interest strength (also called interest-values, calculated through the response formulation given in Harris (1988)) decrease the repeatability and information content, increase the probability of mismatch, and lower the efficiency of the subsequent image matching. Although some improved methods to the Harris detector such as Mikolajczk and Schmid (2004) proposed to strengthen its invariance to scale and affine transformations, the interest points detection with good repeatability and information content for the stereo image matching and the 3D reconstruction is not studied thoroughly.

This paper proposes a filtering method related to the local image texture features to pick out those points that will decrease the overall repeatability and information content. This method takes into account both the interest strength and the local feature of interest point at the same time. A filtering formulation is presented to calculate the response of every pixel, and then a threshold is designed for selecting the interest points.

This paper makes two contributions. First, in the next section the concepts of repeatability and information content as applied to stereo image interest point detection are outlined, and the following section compares several typical detectors, which include the traditional and the up-to-date methods, and presents the comparison results in repeatability and information content using standard test images. The second contribution is by making use of the image feature analysis based on information entropy; a filtering method to select interest points related to the local image feature is introduced. The final section describes the detail of this method, and illustrates the improvement of its repeatability and information content through experimental analysis.

Repeatability and Information Content

Repeatability

Given a 3D point P and two projection matrices M_1 and M_2, the projections of P into two images I_1 and I_2 are $p_1 = M_1P$ and $p_2 = M_2P$. The point p_1 detected in image I_1 is repeated in image I_2 if the corresponding point p_2 is detected in image I_2. To measure the repeatability, a unique relationship

Photogrammetric Engineering & Remote Sensing
Vol. 73, No. 5, May 2007, pp. 547–553.

0099-1112/07/7305–0547/$3.00/0

between p_1 and p_2 has to be established. In the case of a planar scene, this relation is defined by a homography: $p_2 = H_{21}\, p_1$, $H_{21} = M_1\, M_2^{-1}$. The percentage of detected points that are repeated is the repeatability rate, which explicitly describes the geometrical stability of the interest point detector between different images of a given scene taken under varying viewing conditions (Schmid *et al.*, 2000).

However, for large-scale aerial images or satellite images, it is difficult to establish the relationship between one interest point and its corresponding point in a stereo pair, so the image matching method is utilized to determine the "relative" repeatability rate. The traditional image matching method in digital photogrammetry is based on the epipolar geometric constraint and the grey value correlation constraint, and the relative repeatability rate is then defined as:

$$R\,(\Psi) = \{(x_1 x_k) \mid \Psi\,(\rho, dx_1, dx_k) < \Psi\}$$

$$r_k(\Psi) = \frac{|\, R(\Psi)\,|}{\min\,(n_1, n_k)} \tag{1}$$

where n_1 and n_k are the number of points detected in the overlapping part of images I_1 and I_k, respectively, and $R(\psi)$ is the number of points repeated in two images in the condition of ψ. $\psi(\rho, dx_1, dx_k)$ describes the matching reliability of one point x_1 in an image and the corresponding point x_k in another image. The matching reliability can be defined as (Zhu *et al.*, 2005; Zhong and Zhang, 2002):

$$\Psi(\rho,\, dx_1,\, dx_k) = \rho \times f\left(\sqrt{dx_1^2 + dx_k^2}\right) \tag{2}$$

$$\rho = Cov\,(x_1, x_k)$$

$$f\,(x) = \begin{cases} 1 - \dfrac{x}{\sigma} & if\;\; 0 \le x \le \sigma \\ 0 & else \end{cases}$$

where ρ is the grey value correlation coefficient of the windows (e.g., 5*5) in two images which x_1 and x_k are the centric pixels, respectively, $|\rho| \le 1$, ranging from -1 to 1. dx_1, dx_k provide the distance between x_1 and x_k to the same epipolar, and σ is the epipolar allowed distance (e.g., approximately 0.5 to 1.5 pixel). ψ ranges from -1 to 1, and the matching reliability is best when $\psi = 1$, and $\psi = 0.8$ is generally used. $r_k\,(\psi)$ is the relative repeatability rate in the specific condition of ψ.

Information Content

Information content is a measure of the distinctiveness of an interest point. Distinctiveness is based on the likelihood of a gray value descriptor computed at the point within the population of all observed interest point descriptors. If all descriptors are spread out, information content is high, and the matching is likely to succeed. On the other hand, if all descriptors are close to each other, the information content is low, and matching can easily fail as any point can be matched to any other.

Koenderink and van Doorn (1987) gave a rotation invariant descriptor that was a combination of derivatives of local gray values of interest points. In this paper, invariants up to second order are used:

$$\overrightarrow{V}[0..3] = \begin{bmatrix} L_x L_x + L_y L_y \\ L_{xx} L_x L_x + 2L_{xy} L_x L_y + L_{yy} L_y L_y \\ L_{xx} + L_{yy} \\ L_{xx} L_{xx} + 2L_{xy} L_{xy} + L_{yy} L_{yy} \end{bmatrix} \tag{3}$$

where the first component of $\overrightarrow{V}$ is the square of the gradient magnitude, and the third is the Laplacian.

The entropy of these descriptors measures the information content of a set of interest points. The computation of entropy requires a partitioning of the space $\overrightarrow{V}$. Partitioning is dependent on the distance measure between descriptors (Schmid *et al.*, 2000). The Mahalanobis distance is utilized to measure the descriptors and normalize the descriptors. This normalization allows the use of distance cells of equal size in all dimensions. This is important since the entropy is directly dependent on the partition used. The probability of each cell of this partition is used to compute the entropy of a set of vectors $\overrightarrow{V}$.

Interest Point Detectors Used in Stereo Image Matching

State of the Art

In the field of photogrammetry, the Moravec detector (Moravec, 1977), the Förstner detector (Förstner, 1994), and the Harris detector (Harris, 1988) are generally used in stereo image matching. As to the two novel criteria of repeatability and information content, Schmid *et al.* (2000) experimented with several interest point detectors under image rotation, illumination variation, and viewpoint change, and pointed out that the Harris detector obtains the best results for these two aspects. Note however, that the Harris detector is sensitive to changes in scale, and also to rotations out of the camera plane.

Kadir and Brady (2001) proposed a salient region detector by making use of local complexity as a measure of saliency. The salient scale is selected at the entropy extreme of the local descriptors. The method searches for scale localized features with high entropy, with the constraint that the scale is isotropic. This method is robust to image scale change and perturbation, and is repeatable under intra-class variability (Kadir *et al.*, 2004). The salient region detector is preferable to deal with small images for object recognition under background perturbations.

Lowe (1999) proposed a method for image feature generation called the Scale Invariant Feature Transform (SIFT) operator based on local 3D extrema in the scale-space pyramid built with difference-of-Gaussian (DOG) filters, which is invariant over a wider set of transformations especially scale change. The input image is successively smoothed with a Gaussian kernel and sampled. The DOG representation is obtained by subtracting two successive smoothed images. Thus, combined smoothing and sub-sampling construct all the DOG levels. The local 3D extrema in the pyramid representation determines the localization and the scale of the interest points. The SIFT operator finds at multiple scales, and describes the region around each interest point by a histogram of gradient orientations (Lowe, 2004). This description provides robustness against localization errors and small geometric distortions. Recently, experiments have been done to test several descriptors computed for local interest regions (Mikolajczk and Schmid, 2003). The results of this test show that SIFT operator performs best.

However, the SIFT operator has two drawbacks in the case of stereo image matching in photogrammetry. First of all, the DOG detects mainly blob-like interest points (Mikolajczk and Schmid, 2004), while the significant points, such as the corners of buildings and the saddle points near the edge of roads, could not be successfully extracted, and this disadvantage is critical to the subsequent 3D reconstruction of such objects. Secondly, the interest points DOG detected may be not dense enough to fulfill the generation of Digital Surface Model (DSM) through image matching and the later exterior orientation. Each SIFT point is characterized by 128 unsigned eight-bit numbers, which define the multi-scale gradient orientation histogram. To match SIFT points it is necessary to compare these descriptors, and this become difficult when dealing with large scale aerial or satellite images which possess large numbers of interest points.

Differing with the DOG detects mainly blobs, the Harris detector responds to corners and highly textured points. Mikolajczk and Schmid (2004) proposed a Harris-Laplace method for detecting interest points also invariant to scale, which computes a multi-scale representation for the Harris interest point detector and then selects points at which a local measure (the Laplacian) is maximal over scales. The scale-adapted matrix $\tilde{A}$ is used instead of the auto-related matrix in the standard Harris detector:

$$\tilde{A} = \sigma_D^2 \; G(\sigma_1) \otimes \begin{bmatrix} g_x^2(x, \sigma_D) & g_x g_y(x, \sigma_D) \\ g_x g_y(x, \sigma_D) & g_y^2(x, \sigma_D) \end{bmatrix} \quad (4)$$

where σ_I is the integration scale, and σ_D is the differentiation scale. The strength of an interest point is measured by: $\tilde{I} = \det(\tilde{A}) - \alpha Trace(\tilde{A})^2$. Furthermore, they extend this scale invariant detector to affine invariance by estimating the affine shape of a point neighborhood. An iterative algorithm modifies location, scale, and neighborhood of each point and converges to affine invariant points. The characteristic scale and the affine shape of neighborhood determine an affine invariant region for each point. The authors also present a comparative evaluation of different detectors, and show that this method provides better results than DOG, the standard Harris and other existing methods.

Comparison in Repeatability and Information Content

This paper chooses three actual stereo pairs, of which some properties are the same, the size is 1,000*1,000 pixels, the scale is 1:10 000, and the image overlap is 65 percent, but they are different in texture complexity (Figure 1). The experiment is illustrated with these stereo pairs to compare the Moravec detector, the Förstner detector, the standard Harris detector, the DOG method and the Harris-Laplace detector, for analyzing the repeatability and information content of the extracted interest points.

In our experiment, the parameter thresholds of the Moravec detector and the Förstner detector were chosen according to the values recommended by Moravec (1977) and Förstner (1994): the parameter σ of Gaussian weight template in Harris of 0.5, α equal to 0.04, and the filtrate mask of 5*5. As for the DOG method, two octaves are engaged, with four sampled scales in each octave. The simplified algorithm (Mikolajczk and Schmid, 2004) is used to fulfill the Harris-Laplace detector.

When determining the repeatability using Equation 1, different thresholds of the matching reliability ψ were chosen ($\psi \in [0.75, 0.95]$). When calculating the information content, in order to obtain a statistically significant measure, all interest points have been considered, and the cell sizes are set to 30 in all dimensions when partitioning the set of normalized descriptors. The result is shown in Figure 2 and Table 1.

From Figure 2, the Harris-Laplace detector performs better than the other detectors in repeatability on the whole. However, the repeatability of Harris-Laplace in stereo pair 1 does not have significant superiority to other detectors, and is still not always the best in stereo pair 3, this may be the reason of disturbance of similar and complex textures in stereo pair 1, and the smoothness of textures in stereo pair 3. As for information content, Table 1 shows that the DOG method gives the worst results in all test images. As shown in Figure 1a and Figure 1b, the Harris-Standard detector produces the higher information content than Moravec detector and Förstner detector, while the Harris-Laplace detector is better than the Harris-Standard detector. But the Förstner detector is better than the Harris-Laplace detector in the image of Figure 1c, which is due to the Förstner detector being more sensitive to the linear texture as Figure 1c shows.

More intensive experiments give the similar results by making use of other stereo images. Taking into account the repeatability and the information content at the same time, the Harris-Laplace detector gives better results than the other detectors in moderately texture images, but in the images of poor textures, the Harris-Laplace detector is not the best because of some poor points selected, and at the same time, these poor points increase the probability of mismatch and lowers the information content. Therefore, this paper proposes a method of further improving the Harris-Laplace detector (as well as the standard Harris detector) through a filtering method that is related to the local image texture feature. Those points with less interest strength will be omitted, which will result in better repeatability and better information content in all image features.

Feature-related Filtering Strategy

The interest points detected by the Harris-Laplace detector can be filtered using the magnitude of their interest strength; the interest points are ranked according to their interest strength, and those points below a chosen value are filtered out (Schmid *et al.*, 2000; Sebe *et al.*, 2003). However, the number of interest points to be omitted and the method of defining good interest points are not studied thoroughly. This becomes important in the case of large-scale aerial images or satellite image matching. The following sections presents a feature-related filtering strategy for the Harris-Laplace detectors, giving an heuristic formulation taking into

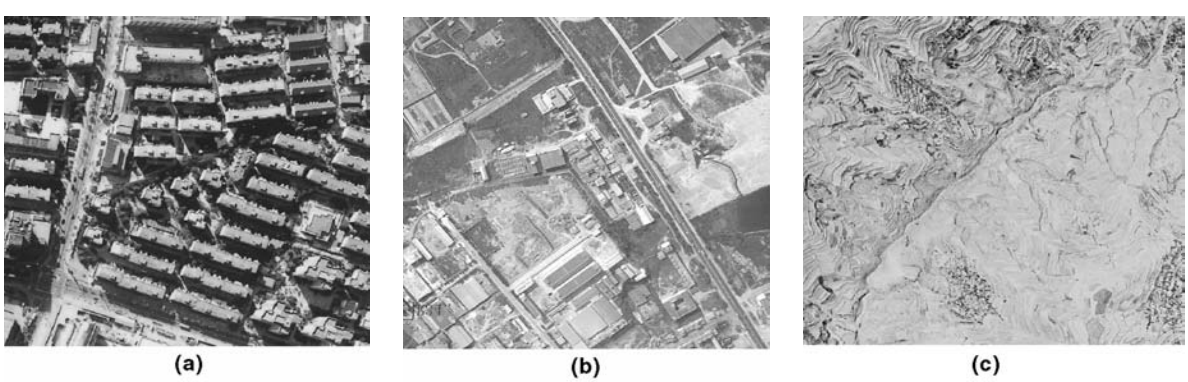


Figure 1. Reference images for the comparison analysis: (a) Stereo pair 1, (b) Stereo pair 2, and (c) Stereo pair 3.

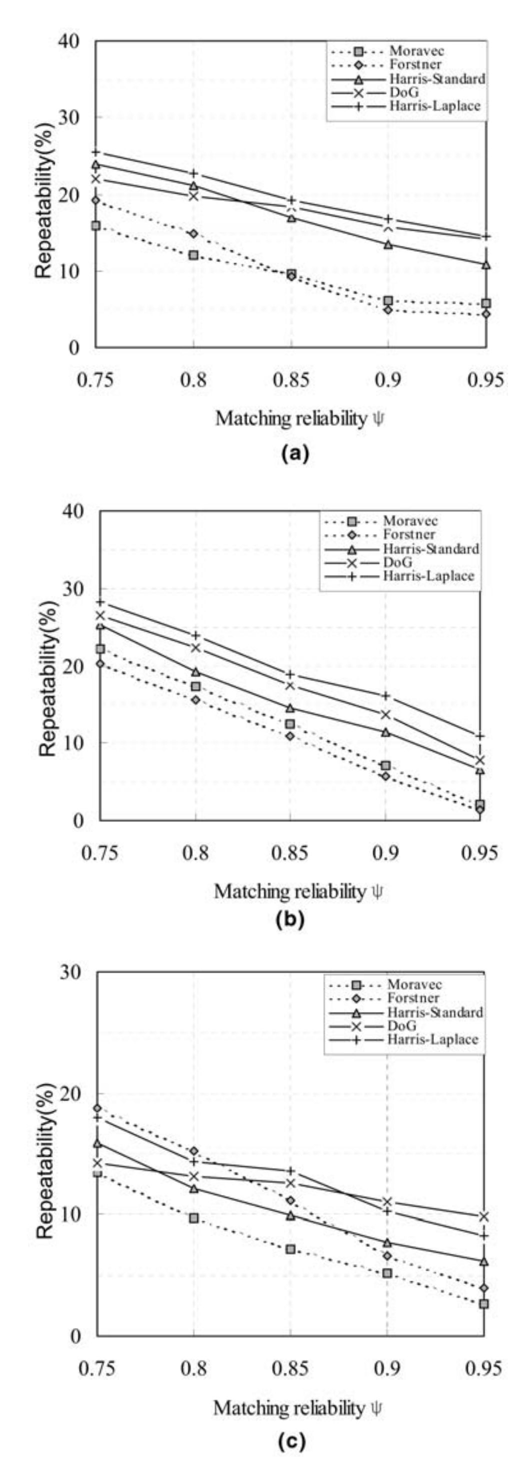


Figure 2. The repeatability comparison: (a) Stereo pair 1, (b) Stereo pair 2, and (c) Stereo pair 3.

account the strength and the local feature of interest point synthetically for the purpose of stereo image matching and 3D object reconstruction.

Local Feature of Interest Points

The information entropy is generally used in the field of image analysis and understanding. Kadir and Brady (2001) used information entropy to get a salient region, while this paper uses it to describe the local feature of image. The distinctiveness of an interest point is not only related to its strength in the whole image, but to the local features of the point. A grid partitioning strategy to the image is used for the analysis of this kind of local feature.

To project a grid with a fixed resolution to an image, the information entropy of each grid cell describes the feature of such cell. The concept of cell entropy is introduced as follows. If $H[k]$ is the cell entropy in the image which has the serial number k, then $H[k]$ can be defined as:

$$H[k] = -\sum_{j=1}^{N} p_j \log p_j \tag{5}$$

where N is the number of pixels in cell k. The grey value probability p_j is considered as its frequency:

$$p_i = \frac{f_i}{\sum_{j=1}^{N} f_j}$$

where f_i is the number of grey value i. When using the cell entropy to describe the feature of cells in an image, the whole image can be considered as an entropy matrix consisting of different cell entropy. This matrix expresses not only the whole feature of the image but also the local feature (Sun *et al.*, 2004).

The image (Figure 3a) is projected to a grid with a resolution of 30*40 (Figure 3b), and the entropy of each cell is then calculated. Different gray values are used to denote the cell entropy, where the white color represents high entropy, and the black color indicates small entropy, giving a feature-classified image (Figure 3c). From Figure 3c, the cell entropy can be used to analyze the feature of images appropriately. When using an entropy matrix to analyze the image features, the precision is obviously related to the grid resolution. The higher the grid resolution, the more precise the image feature, the entropy matrix processing therefore would be more time-consuming. Our test proves that it is appropriate to use 16*16 grid cells for a 256*256 image.

After analyzing the image feature by making use of the local cell entropy, we can consider all the points (pixels) in a cell as the same local feature. The local feature of a point $F[i]$ is defined as:

$$F[i] = \frac{H[k]}{\frac{1}{N}\sum_{j=1}^{N} H[j]} \tag{6}$$

where N is the number of cells, and i is the serial number of one pixel in the cell k. $F[i]$ stands for the ratio of the current cell entropy to the average cell entropy of the whole image.

The Feature-related Filtering Method

In the Harris-Laplace detector, the principal criteria to detect an interest point is that the strength of this point must be

Table 1. The Information Content of Detectors for Different Images

Detector	Information Content		
	Figure 1a	Figure 1b	Figure 1c
Moravec	1.613	2.815	3.100
Förstner	1.590	2.908	3.294
Harris-Standard	1.651	3.284	2.871
DOG	0.987	1.215	1.825
Harris-Laplace	1.719	3.542	2.918

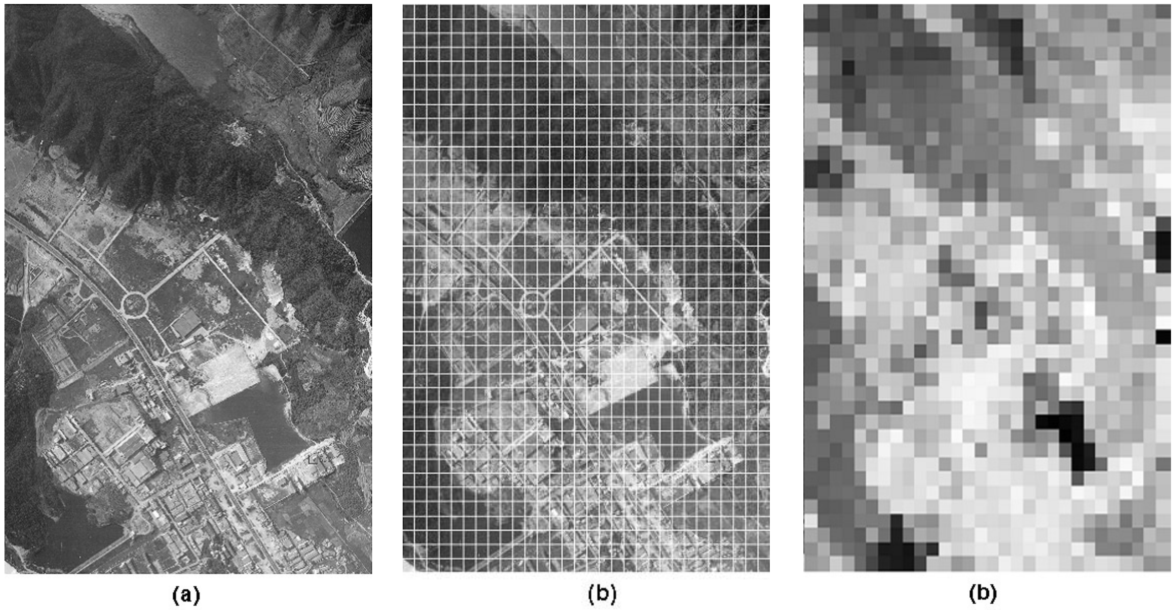


Figure 3. Image feature analysis: (a) Reference image, (b) Gridded (30*40) image, and (c) Feature classified image.

the maximum of all the points in a 5*5(or 3*3) mask with the point at its center. So that the feature-related filter method could be defined as follows: a detected point must not only fulfill this condition, but also fulfill the formulation described as below:

$$\frac{S[i]}{\frac{1}{N}\sum_{j=1}^{N} S[j]} \times F[i]^{e} > T \tag{7}$$

where N is the number of points in the mask window, and $S[i]$ is the strength of point i calculated using Equation 4. $F[i]$ is the local feature of point i, and e and T are constants, which determine the variety of image feature and the number of interest points selected, respectively.

Constant T is related to the size of the filtering mask. When the size of the filtering mask is 5*5, T can be set to [2, 5], while with the size of 3*3, T can be set to [1, 4], and the reduction number of the interest points is then about 25 percent to 50 percent. Constant e ranges from 0 to 3, and e is related to the local image texture features; the more complicated texture, the higher e value. The values of e and T here are experiential values based on extensive experiments with different images, and they also can be determined by users according to their actual requirements.

Figure 4 shows the different results of interest point detected in different strategies. Figure 4a is a reference

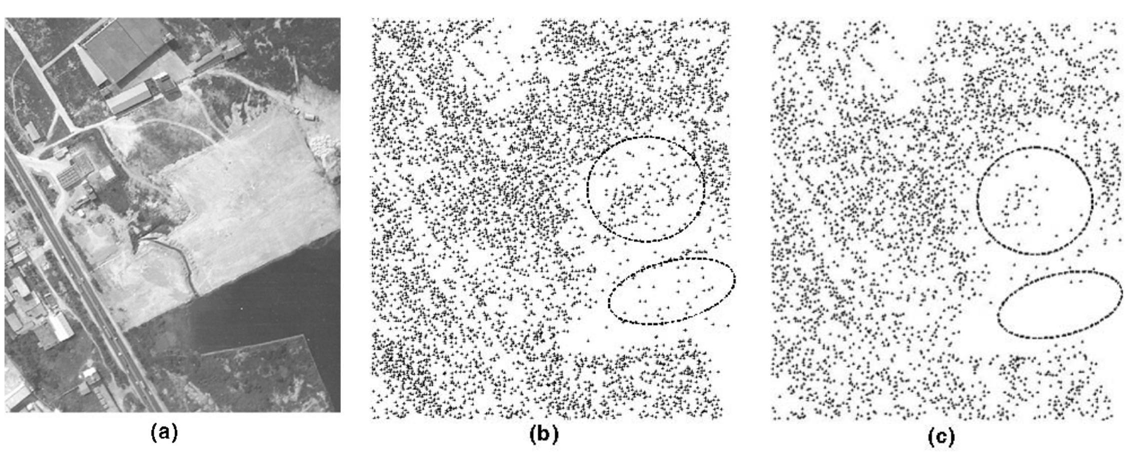


Figure 4. Images with interest points: (a) Reference image, (b) Interest points detected by Harris-Laplace, and (c) Interest points after filtering (T = 3.5, e = 1.5).

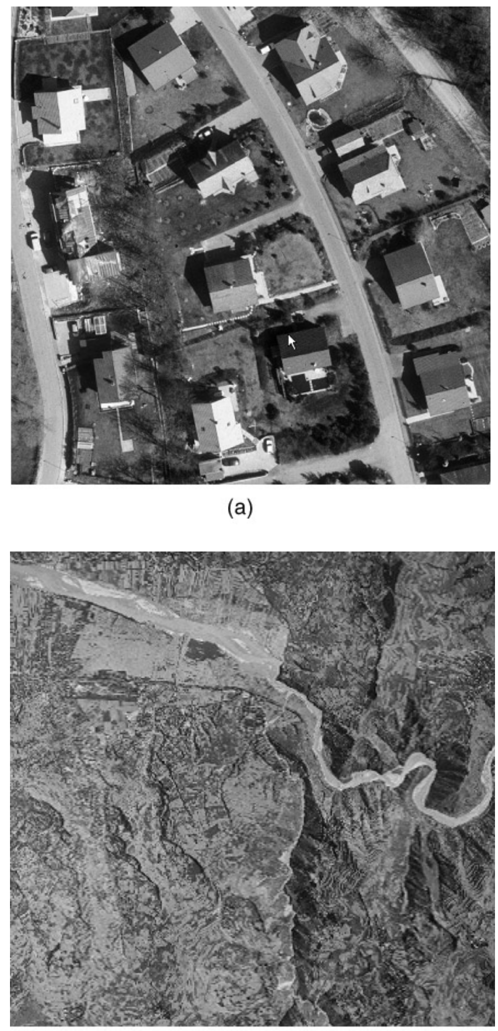
TABLE 2. DESCRIPTION OF THE TEST STEREO PAIRS

Stereo Pairs	Reference Image	Size (pixels)	Camera Focal Length	Scale	Image Overlap	Scan Resolution
Stereo pair 4	Figure 3a	1982*3043	213.734	1:10 000	65%	25 μm/pixel
Stereo pair 5	Figure 5a	1800*1800	152.850	1:5 000	60%	15 μm/pixel
Stereo pair 6	Figure 5b	6468*7002	153.710	1:10 000	65%	50 μm/pixel

image; Figure 4b shows the interest points detected by the Harris-Laplace detector without filtering (the number of the interest points is 4,993). Figure 4c shows the interest points after filtering when taking into account the local feature of the image (T = 3.5, e = 1.5). The number of interest points is 3,637, i.e., almost one third of the original points with lower strength are filtered out. The number of the remaining interest points in areas of poor textures (as in the playground and pond area marked with a circle and an ellipse, respectively, in Figure 4b) is correspondingly small, but in areas of good texture (near the roads and buildings), there is more interest points remained after filtering.

Experimental Analysis
Based on this filtering strategy with the Harris-Laplace detector, several stereo image pairs (Table 2 and Figure 5) are analysed, the stereo pair 5 was downloaded from the ISPRS official website (**http://www.isprs.org/data/avenches/**). The threshold of matching reliability is set to 0.8 to ensure a good matching quality, and the cell size is 100 for calculating the information content. After all the corresponding points are successfully matched, the DSMs then can be obtained, and 46 checkpoints for each stereo pair sampled from a digital photogrammetric workstation, most of which located on the ground, are utilized to compute the root mean square error (RMSE) by interpolating the elevation values of these checkpoints from the derived DSM. The results of Harris-Standard detector, DOG method and Harris-Laplace detector are compared, and the results are listed in Table 3.

From the test, the Harris-Laplace detector with a feature-related filtering strategy gives better results than other detectors in terms of repeatability and information content, such as in stereo pair 4, which covers a large area of poor texture forestry and water bodies, the repeatability is increased from 18.63 percent of the Harris-Laplace detector to 23.93 percent, and to the indistinctive and homogeneous texture of stereo pair 6, the repeatability and information content are increased from 22.30 percent and 2.318 of the Harris-Laplace detector to 25.79 percent and 3.108, respectively. The density of matched points after filtering is preferable to generate DSMS, because the RMSE of checkpoints of the Harris-Laplace detector with feature-related filtering are also better than that of Harris-Laplace detector, as well as more better than others.



(a)

(b)

Figure 5. Reference images of the stereo pairs for the test of the filtering strategy: (a) Stereo pair 5, and (b) Stereo pair 6.

Conclusions
This paper conveys the following conclusions:

1. For the purpose of stereo image matching and the 3D object reconstruction, the repeatability and information content are the two most important criteria when appraising interest point detectors;
2. The Harris-Laplace detector performs better in the aspects of repeatability and information content than other detectors in areas of good texture;
3. The Harris-Laplace detector with feature-related filtering stategy presented in this paper not only gives better repeatability and information content than other detectors, but also is more appropriate for stereo image matching and the subsequent 3D reconstruction.

Acknowledgments
The authors are indebted to Miss Tian Yixiang and Mr. Xu Zhixiang for their great effort during the experimental analysis of this paper. All the reviewers' comments and remarks to this paper are appreciated very much. The work described in this paper is supported by National Basic

Table 3. The Experimental Results of Different Detectors

Stereo Pairs	Detector	Interest Points	Repeatability (%) (ψ = 0.8)	Information Content	RMSE of DSM (m)
Stereo pair 4	Harris-standard	161 923	14.40	2.332	1.38
	DOG	8 618	15.91	1.761	2.32
	Harris-Laplace	134 270	18.63	2.513	1.29
	Harris-Laplace after filtering	90 136	23.93	2.816	1.22
Stereo pair 5	Harris-standard	35 194	15.80	2.771	1.78
	DOG	1 876	14.20	1.103	2.75
	Harris-Laplace	27 061	17.25	2.820	1.53
	Harris-Laplace after filtering	17 812	18.63	3.362	1.59
Stereo pair 6	Harris-standard	1 167 219	20.55	2.106	2.78
	DOG	48 196	22.89	0.869	4.29
	Harris-Laplace	861 029	22.30	2.318	2.56
	Harris-Laplace after filtering	542 813	25.79	2.525	2.47

Research Program of China (2002CB312101), Hubei Outstanding Young Researchers Foundation (2004ABB018), the National Natural Science Foundation of China (40671158), and the National High Technology Research and Development Program of China (2006AA12Z224).

References

Baker, S., and S.K. Nayar, 1999. Global measures of coherence for edge detector evaluation, *Proceedings of the 1999 IEEE Conference on Computer Vision and Pattern Recognition*, June, Fort Collins, Colorado, pp. 373–379.

Brand, P., and R. Mohr, 1994. Accuracy in image measure, *Proceedings of the SPIE Conference on Videometrics III*, November, Boston, Massachusetts, Vol. 2350, pp. 218–228.

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

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

Kadir, T., and M. Brady, 2001. Scale, saliency and image description, *International Journal of Computer Vision*, 45(2):83–105.

Kadir, T., A. Zisserman, and M. Brady, 2004. An affine invariant salient region detector, *ECCV*, (1):228–241.

Koenderink, J.J., and A.J. van Doorn, 1987. Representation of local geometry in the visual system, *Biological Cybernetics*, 55: 367–375.

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

Lowe, D.G., 2004. Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, 60(2):91–110.

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

Mikolajczk, K., and C. Schmid, 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*, August, Cambridge, Massachusetts, pp. 584.

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

Sebe, N., and M.S. Lew, 2003. Comparing salient point detectors, *Pattern Recognition Letters*, 24(1–3):89–96

Sun, J.D., X.X. Wu, and L.H. Zhou, 2004. The image index based on information entropy, *Journal of Xian Electronic Science and Technology University*, 31(2):223–229.

Zhong, Y., and H.F. Zhang, 2002. Control points based semi-dense matching, *Proceedings of the 5th Asian Conference on Computer Vision*, Melbourne, Australia, pp. 137–144.

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

(Received 10 October 2005; accepted 29 November 2005; revised 07 December 2005)