THE PHOTOGRAMMETRIC RECORD

The Photogrammetric Record 25(132): 437–453 (December 2010)

MULTIPLE CLOSE-RANGE IMAGE MATCHING BASED ON A SELF-ADAPTIVE TRIANGLE CONSTRAINT

Qing Zhu* (zhuqing@lmars.whu.edu.cn)

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

Yunsheng Zhang (sheng_650@163.com)

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

Bo Wu (lsbowhu@polyu.edu.hk)

Hong Kong Polytechnic University, Hong Kong

Yeting Zhang (zhangyeting@263.net)

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

*Now also at Research Center of Remote Sensing and Spatial Information Science, Shanghai University, China

Abstract

Reliable image matching is an essential and difficult task in digital photogrammetry and computer vision. Possible problems from geometric distortions, illumination changes, scale changes and difficult texture conditions will result in matching ambiguity, especially for close-range image matching. This paper presents a multiple close-range image matching method for surface reconstruction based on a self-adaptive triangle constraint. This method features two aspects. First, the triangles constructed from the previously matched interest points provide strong geometric constraints for the subsequent point matching combined with gradient orientation and disparity constraints. The dynamic update of the triangulation adapts automatically to the changes of image textures. Secondly, a consistency check in object space is performed to remove possible mismatches. Using three sets of actual triple overlapped close-range images for the experiment, the results revealed that the proposed method provides improved matching reliability.

Keywords: close-range images, consistency check, multiple image matching, self-adaptive triangle constraint

Introduction

Image-based three-dimensional (3D) modelling of objects and scenes is an important and long-term research task in the digital photogrammetry and computer vision communities

(Baltsavias, 2004; Mayer, 2008). Three-dimensional modelling from images can not only provide geometric reconstruction capability, it also provides photorealistic texture reconstruction. As multi-source images including satellite images, aerial images and ground images have become more easily available, image-based 3D modelling will serve as the most feasible, economic and independent method for some time to come (Remondino and El-Hakim, 2006; Zhu et al., 2007a; Remondino et al., 2008). The increasing application of 3D spatial data in different areas requires high automation, efficiency and reliability of the 3D modelling technologies. However, the most recent software systems for image-based modelling are still not fully automatic and require many human interactions because of the limitations of image matching.

Image matching is a challenging and often ill-posed problem, especially for images with textures difficult for matching, such as homogenous textures and repetitive or similar texture patterns. During the past decades, much effort has been devoted to improving the reliability, automation and efficiency of image matching (Heipke, 2001; Scharstein and Szeliski, 2002; Brown et al., 2003; Zhang, 2005; Zhu et al., 2005; 2007a). Zhu et al. (2005; 2007a) presented a novel image matching method based on a self-adaptive triangle constraint, which was used for stereoscopic aerial image matching and which proved able to produce reliable matching results (Wu, 2006; Zhu et al., 2007a). However, for close-range image matching, it is more difficult to obtain reliable matching results because problems such as occlusions and distinct surface discontinuities are often encountered in close-range images (Kanade and Okutomi, 1994; Hirschmüller and Scharstein, 2007; Hirschmüller, 2008). For example, when reconstructing façades from images, the bottleneck is to obtain corresponding points from façade images which are full of homogenous areas and repetitive texture. Multiple image matching proved to be an effective strategy to improve the matching reliability using the redundant information (Zhang, 2005). Therefore, this paper investigates the multiple image matching strategies based on the self-adaptive triangle constraint for close-range images.

The rest of this paper is organised as follows. After giving a literature review on how to improve the reliability of image matching, a new multiple close-range image matching method is presented. Then, actual close-range images are employed for experimental analysis. Finally, the concluding remarks are presented and future work is discussed.

Related Work

Image matching techniques can be generally classified into two groups according to the length of their stereo baselines. First are image matching methods for "short range motion", such as video frame (Pollefeys et al., 2008) and other short-baseline stereo (Scharstein and Szeliski, 2002); second are image matching methods for "long range motion" or "wide baselines" (Lowe, 2004; Mikolajczyk and Schmid, 2005; Mikolajczyk et al., 2005). A short baseline (usually dozens of centimetres, typically with stereo cameras mounted on a rigid camera bar) is good for image matching since it is relatively easy to find corresponding points, but it often results in large depth errors. For image matching with a long or wide baseline (usually from a few metres to dozens of metres, with images taken from different locations using the same camera), it is hard to find correct corresponding points, since the look angles may be significantly different owing to the longer baseline. However, it provides better accuracy in depth, which is important for 3D modelling from images. For close-range images in particular, long- or wide-baseline images will provide accurate measurements at longer ranges compared to short-baseline images. Most of the current image matching methods try to use a moderate baseline for good measurement accuracy in depth and improve the matching reliability using other strategies. They can be generally categorised into the following three groups:

(1) The first group includes the methods introducing available constraints to improve the matching reliability. Marr and Poggio (1979) proposed several compatibility constraints for stereoscopic image matching, such as: the uniqueness constraint, meaning that one point in an image can only have at most one corresponding point in another image; the continuity constraint, meaning that, in a local area, the depth is smooth; ordering constraints, where the order of neighbouring correspondences on the corresponding epipolar lines is always preserved for opaque surfaces. The epipolar constraint is a typical geometrical constraint that can narrow the corresponding point search from a two-dimensional search to a one-dimensional search. There are several constraints based on image segmentation that rely on some seed points or other segmentation methods. Tang et al. (2002) tried to reduce the search area of matching by making use of a Voronoi diagram. Rengarajan et al. (2004) used triangulations to help in matching the Mars Orbiter Camera narrow-angle (MOC NA) stereo-images for generation of digital elevation models (DEMs) of Mars. The triangulations are generated from a few seed points and the triangle pairs provide a geometric constraint to help match the points within the triangles. Zhu et al. (2005) proposed a novel constraint to improve the matching reliability which also uses triangulations. But in their method, the triangulations are dynamically updated using the successfully matched points and the dynamic updating of the triangulations is adaptive to the image texture changes so that better matching results can be obtained.

(2) The second group improves the matching reliability by making use of local/global information or some propagation strategies. Lowe (1999) presented a scale invariant feature transform (SIFT operator), which combines a scale invariant region detector and a descriptor based on the gradient distribution in the detected local regions. In the SIFT descriptor, each interest point is characterised by 128 unsigned eight-bit numbers generated from a local region, which define the multi-scale gradient orientation histogram. The similarity is measured by comparing the two vectors with 128 components associated with the matching (Lowe, 2004). The SIFT operator provides robustness against localisation errors and small geometric distortions. Relational matching methods are global optimisation methods relying on the geometric relations, ratio relations and topological relations of primitives to improve the matching reliability. Relational matching can be implemented in many ways, such as energy minimisation (Geman and Geman, 1984), graph searching techniques (Boykov et al., 2001) or relaxation labelling techniques (Zhang et al., 2000). Wang (1998) proposed a structural image matching method for relative orientation or for tie point extraction which is a kind of graph searching method. Graph searching and energy minimisation methods to optimise the matching result are time consuming when the image resolution becomes higher. For the relaxation labelling method, the ordering constraint and uniqueness constraint are often used and can smooth the discontinuity. Hierarchical methods are used in many matching algorithms in order to reduce the ambiguity problem and to extend the pull-in range, resulting in a reliable matching. One familiar strategy of hierarchical methods is the pyramid image matching based on the "from coarse to fine" strategy. The matching results in the low-resolution layer image being used to guide the image matching at the higher resolution level. Lhuillier and Quan (2000) presented a quasi-dense matching algorithm based on "the best the first" matching propagation strategy. This can improve the matching reliability to some extent, but the matching propagation often fails when encountering repetitive patterns and discontinuities. Zhu et al. (2007a) proposed a novel self-adaptive propagation from some seed points used for aerial stereo-image matching. It can adapt to the changes in image textures automatically; nevertheless, it can suffer

from wrong matching, which would lead to wrong matching of sub-areas based on triangle constraints.

(3) The third group relies mainly on using redundancy to improve the reliability of matching. For close-range images and digital aerial images, an object point can often be projected to more than two images. Matching with multiple images provides a high redundancy of feature points and this will increase the reliability of the matching result. Zhang (2005) studied a multiple aerial image matching strategy based on the hypothesis that the surface constructed from the matched points should be relatively smooth in object space. The commercial software of Match-T (Lemaire, 2008) depends on selecting stereopairs to run stereoscopic image matching; it then merges several matching results by filtering in 3D space. Hirschmüller (2008) also merged the matching results from several stereopairs to improve the matching reliability. The methods which merge stereomatching results can suffer from the ill-posed problem of the stereomatching method. Seitz et al. (2006) performed a study to compare and evaluate multi-view stereo reconstruction algorithms. They compared the existing methods from the following aspects: the scene representation, photo-consistency measure, visibility model, "shape prior" (a priori shape), reconstruction algorithm and initialisation requirements. Several existing methods need a visual hull, a rough bounding box or volume, but these initial requirements are often very hard to satisfy for outdoor scenes; however, matching with multiple images is believed to be useful for reducing matching ambiguity resulting from occlusions, surface discontinuities and repetitive patterns or homogenous texture.

The proposed multiple image matching method is targeted at close-range images with moderate baseline lengths. It is based on the method described in Zhu et al. (2005; 2007a) which was originally developed for stereoscopic aerial image matching. The stereo-image matching method based on the self-adaptive triangle constraint needs to insert matched points in the triangulations from time to time. A possible mismatch may lead to a group of mismatches or absence of successful matches in the neighbouring areas. For close-range images, matching is more difficult because of the occlusions, distinct surface discontinuities, homogenous texture areas and repetitive texture frequently encountered in close-range images. This provoked the research presented in this paper.

Multiple Close-range Image Matching Methodology

Overview of the Method

The inputs for this method are the images and image orientation parameters. A few seed points manually defined or derived from the previous image orientation process are employed to generate the initial triangulations for image matching. The matching pipeline as shown in Fig. 1 includes the following steps:

(1) Image pre-processing using the Wallis filter (Pratt, 1991). This step enhances the image texture.
(2) Interest point extraction using the Harris–Laplace operator. This step is performed to provide the interest points for later multi-image matching.
(3) Interest point matching to obtain potential corresponding points for every stereopair. The point matching is performed on the basis of several constraints. The details are described in the following sections.
(4) Verifying the matching hypothesis by a consistency check. This step is used to obtain triple corresponding points to exclude possible mismatches.

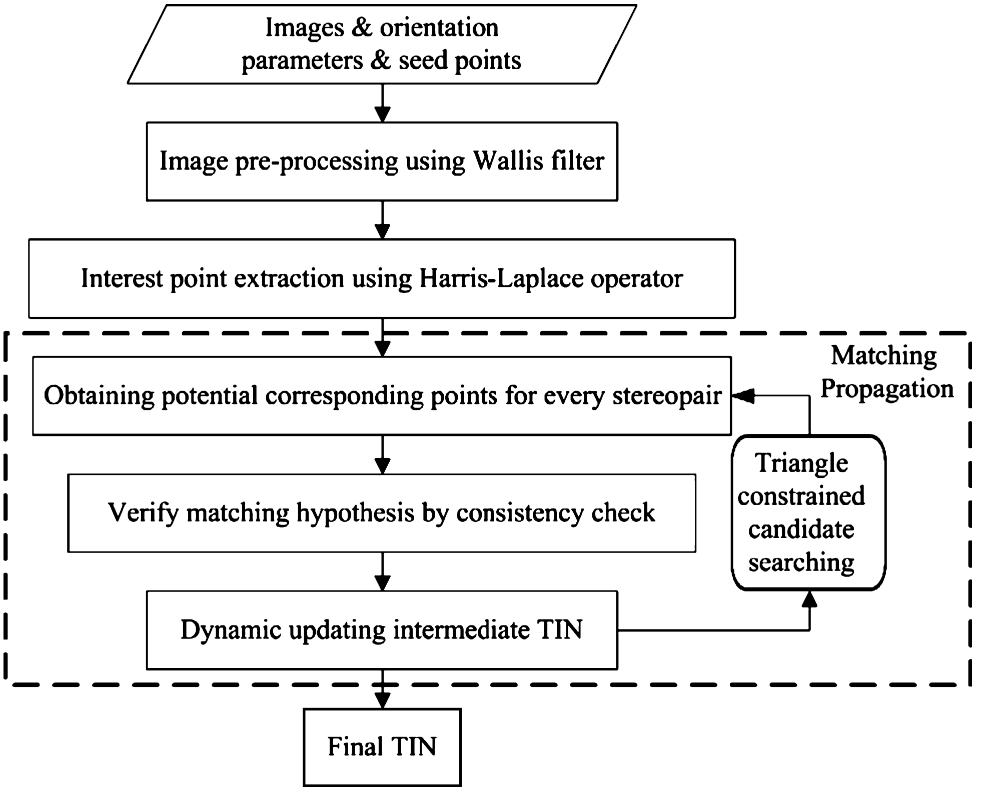


FIG. 1. Framework of the multiple close-range image matching method.

(5) Matching propagation. The successfully matched interest points provide constraints for the neighbouring unmatched points through a matching propagation process that is automatically adaptive to changes in the image texture. The details are described in the following sections.

Image Pre-processing using Wallis Filter

The quality of matching results depends strongly on the quality of the images. With the illumination changing with each change of viewpoint, the contrast may be lower or not equal, which would reduce the number of interest points extracted. The Wallis filter (Pratt, 1991) is employed to strongly enhance and sharpen the already existing texture patterns and increase the signal-to-noise ratio, which is good for feature extraction.

Interest Point Extraction using Harris–Laplace Operator

In the field of image matching and 3D reconstruction, according to the studies of Schmid et al. (2000), Zhu et al. (2007b) and Tuytelaars and Mikolajczyk (2008), the Harris–Laplace detector has good repeatability for interest point extraction, which is a critical factor for image matching. Therefore, the Harris–Laplace detector is used in this research to extract the interest points from all images. The implementation details can be found in Schmid et al. (2000). All the interest points are precisely located to sub-pixel level by means of the Harris strength (Zhu et al., 2007c). After extraction and precise location, all the interest points are assigned a gradient orientation as given in Lowe (2004).

Matching Process

(1) Constraints for Image Matching. The following constraints are considered in the present method for image matching:

(a) Disparity gradient constraint: each triangle is considered as a local continuous area, the disparity of the corresponding points in the triangle having some relation to the triangle vertex (Pollard et al., 1985; Zhu et al., 2005).
(b) Point gradient orientation constraint: for corresponding points in local areas from images with baselines which are not very large, the gradient orientation differences should be similar. In the proposed algorithm, the gradient orientation difference is approximated to the seed points. Every interest point is assigned a main point gradient orientation, which can reduce the search time to some extent.
(c) Epipolar geometrical constraint: epipolar geometry constraint can narrow the search only to the epipolar line, which results in fewer ambiguity problems and lower time costs. As there is often some error in the image orientation parameters, the correct match will sometimes not lie exactly along the epipolar line, so all the interest points in the search image near to the corresponding epipolar line are considered to be matching candidates.
(d) Triangle constraint: each triangle is considered as a local continuous area, and then corresponding points should lie in the corresponding triangle. The triangle constraint implicitly contains the ordering constraint and continuity constraint in the local area. This can reduce the search distance along the epipolar line. During the matching propagation, the corresponding triangles are updated dynamically, so the search distance along the epipolar line is also determined dynamically.

After potential matching candidates are selected, the next task is to measure the similarities of the corresponding points in the image pairs, simulating human eyes by means of similarity descriptors. The similarity descriptor will be the matching score to judge the corresponding points or will be used for a global strategy to judge the corresponding points. There have been many similarity descriptors for image matching presented in the past decades, such as the normalised cross-correlation (*NCC*) (Helava, 1978), sum of squares difference, normalised mutual information (Knops et al., 2006). These are simple computationally, but not distinct enough, and are sensitive to geometrical distortion and discontinuity problems. When using normalised cross-correlation to measure the similarity between the features on stereo-images, a key problem is the selection of an appropriate window size to calculate the correlation values. To avoid or minimise the geometrical distortion and discontinuity problems, Kanade and Okutomi (1994) presented an adaptive method to select the proper window size for their resolution. So the *NCC* combined with an adaptively selected window is chosen for one constraint for the proposed matching method. The *NCC* was used as the matching score to measure the similarity between corresponding points of a stereopair. For corresponding points, the normalised cross-correlation should be higher than a given threshold, which is set to be 0·7 in the present experiment.

(2) Consistency Check using Multiple Images. First, a cross check in image space was performed. The uniqueness constraint is often imposed by cross-consistency checks by correlating an interest point of the first to the second image and inversely by correlating those of the second to the first image; only the best matches consistent in both directions are retained. This check will result in a less dense matching result, and the error rate after cross-consistency checking would reduce. But with half-occlusions and repetitive patterns found in close-range images, the error rate can be high, even after cross checking in image space. Therefore, the consistency check in object space is focused on mainly as described in the following section.

Under the uniqueness constraint, disparity gradient constraint, point gradient orientation constraint, epipolar geometry constraint and triangle constraint, the search range is constrained to a small range, but some interest points may have a wrong match owing to the occlusion or repetitive patterns, while still achieving a high matching score. The stereomatching process cannot deal easily with this problem, so a third image is added to improve the reliability of the matching method based on the self-adaptive triangle constraint.

For each interest point on the reference image I_0 depicted in Fig. 2, in every stereopair several potential corresponding points are selected under the above constraints. For every stereopair, a winner-take-all strategy is adopted and the corresponding point with maximum *NCC* is thought to be the matching hypothesis of the selected interest point. As in Fig. 2, the maximum correlation between x_1 and x is NCC_1 and between x and x_2 is NCC_2; then $x_1 - x - x_2$ is thought to be a triple matching hypothesis. For every corresponding point projected from the same object point, the forward intersection results of every stereopair should be consistent; however, as there is often some error in the orientation parameters, all the rays will not exactly pass through the same object point. It is hard to define a threshold to force all the rays to intersect at the same point in object space, so an alternative is adopted: the forward intersection result of one stereopair is projected to the other image to verify the matching hypothesis. As in Fig. 2, the forward intersection result of $x_1 - x$ is projected to the search image I_2; if the projected point is in the circle with centre at x_2 and radius ϵ, the matching hypothesis can be classed as a successful match. The mean of NCC_1 and NCC_2 forms a reliability indicator *MNCC* (mean of *NCC*) for each triple corresponding point. The reliability indicator *MNCC* is defined as follows:

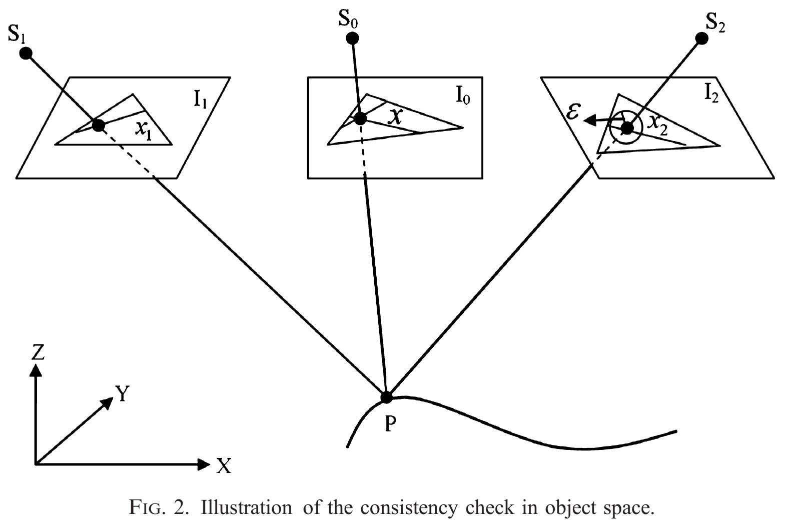


Fig. 2. Illustration of the consistency check in object space.

$$MNCC = \frac{1}{2}(NCC_1 + NCC_2). \tag{1}$$

From equation (1), it can be seen that the larger the *MNCC* value, the more reliable is the match. The reliability indicator defined above will essentially serve as an important indicator for the triangle feature descriptor.

The consistency check in object space will also result in a less dense matching result, but it can ensure that more reliable corresponding points are inserted into the triangle list to constrain later matching, allowing the whole matching result to be more reliable.

(3) Matching Propagation for Multiple Image Matching. The propagation of image matching is the process of matching new points using prior knowledge related to the previously matched corresponding points (Zhu et al., 2007a). The proposed multiple image matching propagation adopts a strategy of matching propagation taking the texture conditions into consideration, starting the matching from the triangles with high matching reliability and good textures, and then propagating the matching to the other triangles with poorer textures. The texture feature of triangle coverage can be approximated by the matching reliability and distinctiveness of the three triangle vertices. The distinctiveness of points is defined using the Harris–Laplace interest strength calculated through the response formulation given in Mikolajczyk and Schmid (2004), and the matching reliability is calculated when carrying out the image matching under the triangle constraint. A triangle feature descriptor I_t for each triangle is designed as

$$I_t = \frac{1}{3}\sum_{i=0}^{2} MNCC_i \times HL_i \tag{2}$$

where HL_i is the Harris–Laplace interest strength and $MNCC_i$, computed by equation (1), is the matching reliability of the corresponding vertex numbered i.

The process of multiple image matching propagation is illustrated in Fig. 3, and is accomplished by the following steps.

The first step of propagation is to triangulate the seed points to form a coarse triple corresponding Delaunay triangulation and calculate the feature descriptor for all the triangles using equation (2). Then the triple corresponding triangles are ranked according to the calculated feature descriptor. Each triangle in these triangulations is considered as a local continuous area.

The second step is to obtain matching hypotheses under available constraints and cross checks for every stereopair. Assuming the middle image to be the reference image, the other images are search images. The search image and reference image can form two stereopairs. Choose one interest point with maximum interest strength in the selected triangle on the reference image. After that, for every stereopair, potential corresponding points are obtained under the triangle constraint, point gradient constraint, disparity gradient limit constraint, epipolar constraint and similarity constraint. If there is no potential corresponding point under the constraints, turn to the next interest point. If multiple potential corresponding points are obtained, the normalised cross-correlation is used as the matching score for the stereopair matching. The corresponding point with the highest matching score is chosen as the matching hypothesis for the stereopair and then a cross check is carried out. If the matching from the first to the second image is not consistent with the matching from the second image to the first, then turn to the next interest point. After obtaining the matching hypothesis in the search images for the selected interest point in the reference image, a triple corresponding point hypothesis is obtained.

The third step is to verify the triple corresponding point hypothesis through a consistency check in object space. If the triple hypothesis can pass the consistency check, the triangulations

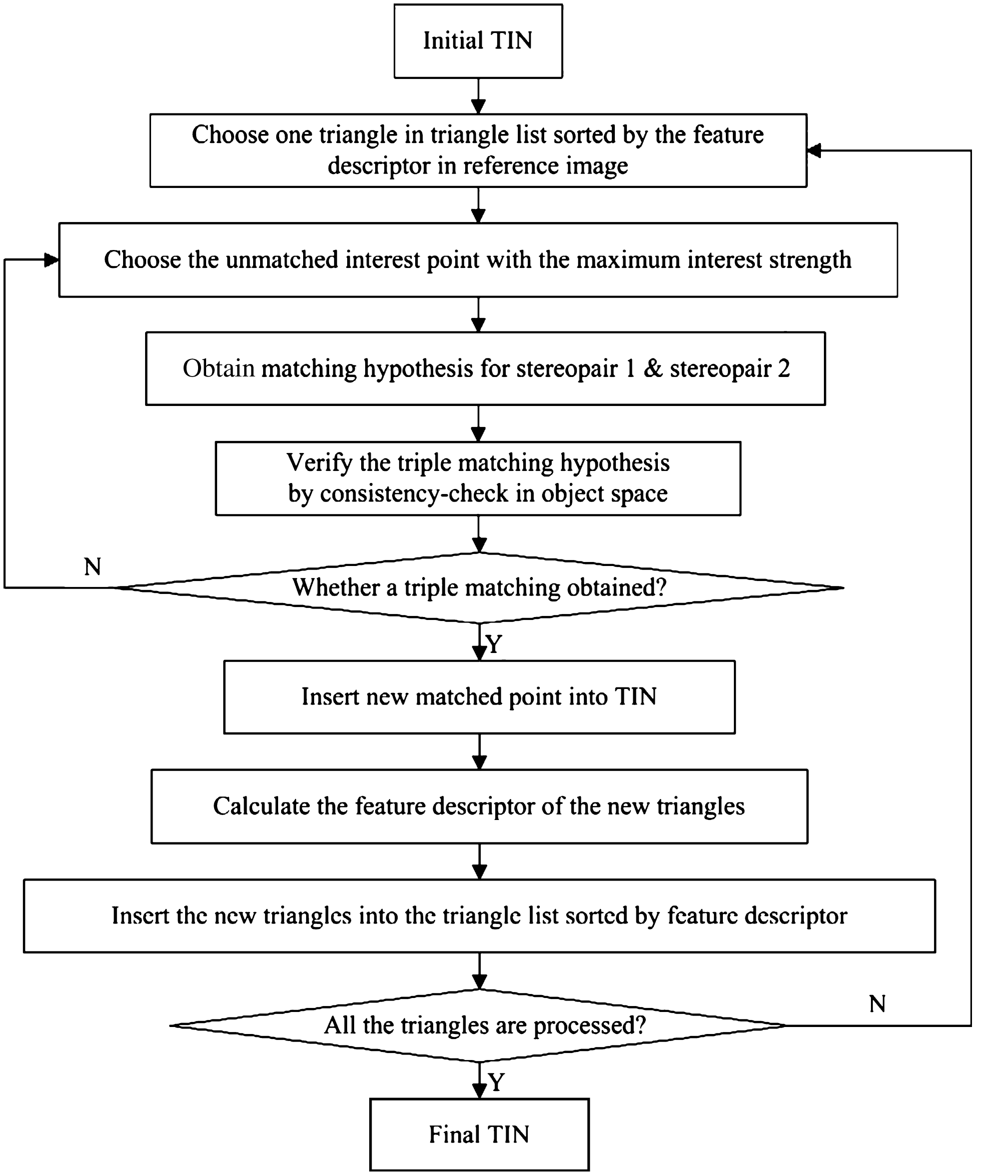


FIG. 3. Flowchart of the matching propagation algorithm.

are updated dynamically to insert the triple corresponding point, unless the propagation turns to the next interest point. After the triangulations are updated, the feature descriptor is calculated for the new triangles, and then the new triangles are inserted into the triangle list in the order of the feature descriptor. The new inserted triangles are also used to constrain the subsequent multiple image matching propagation. The unprocessed pair of triangles is selected iteratively, and the same matching process is repeated until the termination conditions of the propagation are met.

The dynamic updating of triangulation is exactly the process of self-adaptive matching propagation. Along with the matching propagation, the local geometry constraint of the dynamically updated triangles can adapt to the changes in image textures automatically, and will finally produce more reliable matching results.

EXPERIMENTAL RESULTS AND ANALYSIS

Test Data Description

In order to evaluate the performance of the proposed multiple close-range image matching method, three sets of typical images are employed for experiments as shown in Fig. 4. The experimental image sets were selected from data-sets known as "castle-P30", "Herz-Jesu-P8" and "fountain-P11", provided via http://cvlab.epfl.ch/~strecha/multiview/denseMVS.html (Strecha et al., 2008). Table I gives a brief description of these image sets. All the images have been corrected for radial distortion. The internal orientation (IO) and external orientation (EO) parameters of the images were known. From these parameters, the epipolar line can be derived.

As illustrated in Fig. 4, the image sets are typical façade images with serious repetitive patterns.

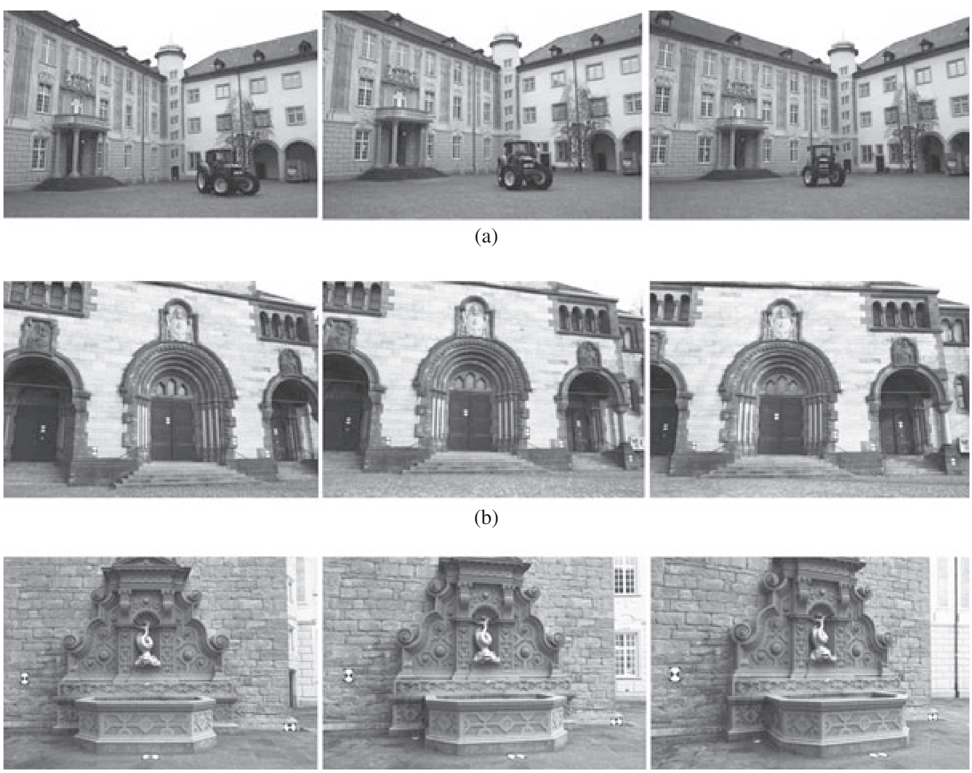


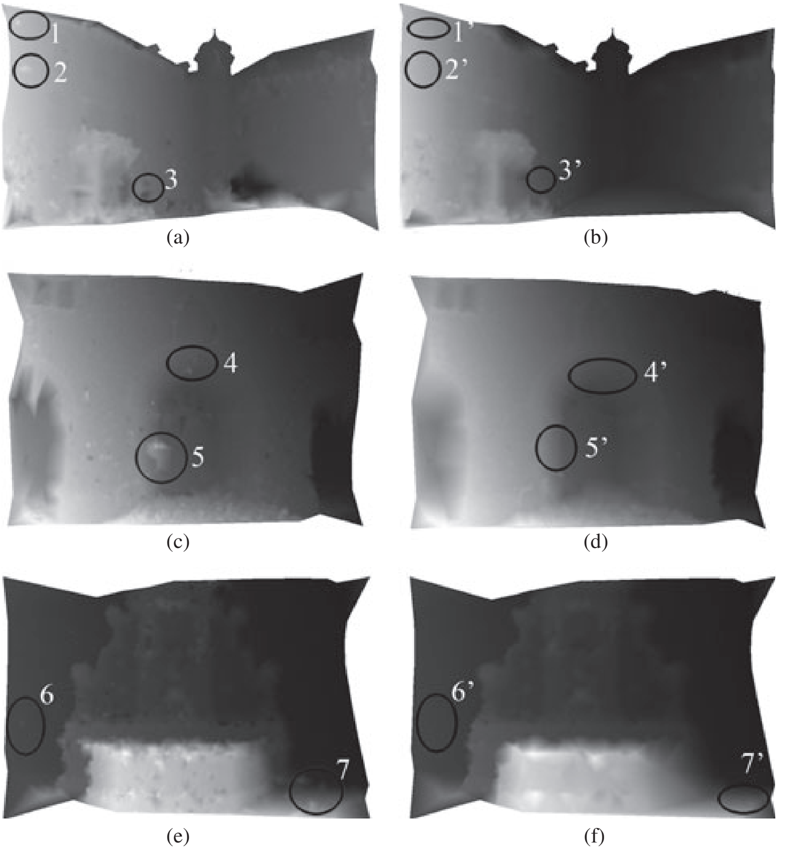
FIG. 4. The test data-sets: (a) castle-P30; (b) Herz-Jesu-P8; (c) fountain-P11.

TABLE I. Description of the test data-sets.

Data-set	*Size (pixels)*	*Focal length (mm)*	*Maximum intersection angle (degrees)*
Castle-P30	3072 × 2048	13·81	8·73
Herz-Jesu-P8	3072 × 2048	13·81	11·83
Fountain-P11	3072 × 2048	13·81	11·21

Experimental Results and Analysis

After the matching process, a final triangulated irregular network (TIN) was obtained, from which dense disparity maps were interpolated to give an intuitive overall description of the matching result. These are shown in Fig. 5. For the images in Fig. 4, the available area in the image derived from the final TIN was divided into a regular grid with size of $u \times u$ pixels (values of 2 were used for u in all the experiments), then every grid was assigned a disparity interpolated from the derived corresponding triangle. For Fig. 4(a), the background of sky was



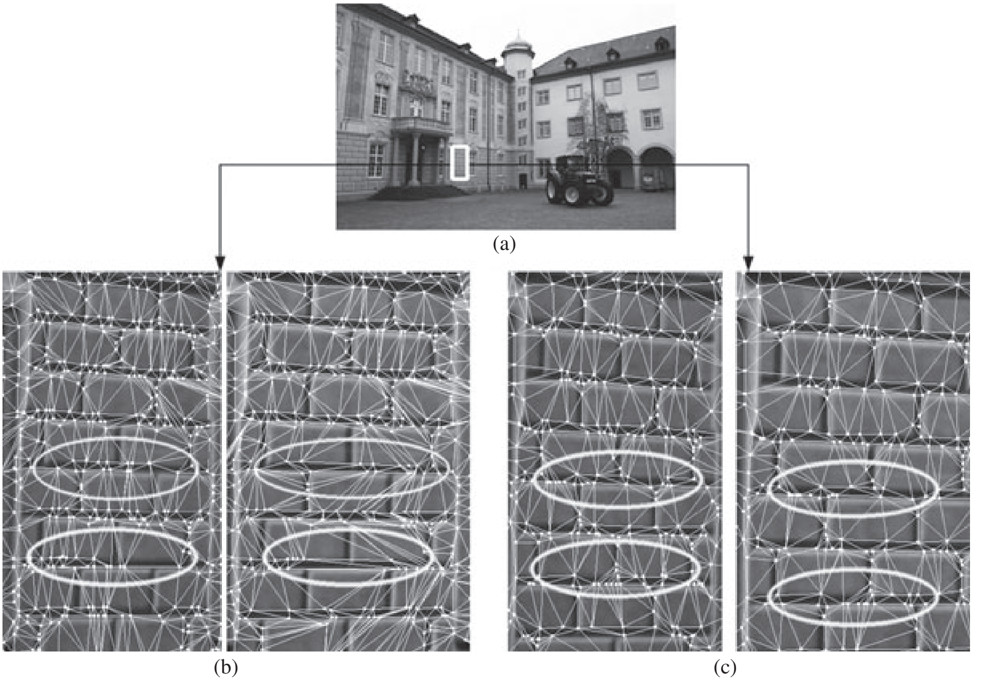
(a) (b) (c) (d) (e) (f)

FIG. 5. Interpolated dense disparity map: (a), (c), (e) disparity maps interpolated from the matching results only using two images; (b), (d), (f) disparity maps interpolated from the matching results using three images and the proposed method.

marked manually, and used as a mask to prevent the pixels in the background from being interpolated. For the experiment image sets shown in Fig. 4(a) and (c), there were some wrong matching results when only two images were used for matching, which resulted in the unsmooth disparity regions indicated by ellipses numbered 1, 2, 3, 6 and 7 in Fig. 5(a) and (e). When using the triple overlapped images, the results were significantly improved as the disparity map became smooth in the corresponding areas 1′, 2′, 3′, 6′ and 7′ in Fig. 5(b) and (f). For the image set in Fig. 4(b), when only two images were used there are large brightness changes in the disparity map, as shown in areas 4 and 5 in Fig. 5(c); this brightness discontinuity implies wrong corresponding points in the derived final TIN. When an additional image was used for the consistency check, the same areas become smooth as shown at 4′ and 5′ in Fig. 5(d).

In the small study areas shown in Figs. 6(a) and 7(a), there are several interest points on the edge that can confuse the matching result significantly, as shown in the areas indicated by white ellipses in Figs. 6(b) and 7(b). When an additional image was added for the consistency check in object space, the wrong matching was avoided in the final TIN, as shown in Figs. 6(c) and 7(c).

When there is a lack of texture as shown in the study area of Fig. 8(a), the corresponding points in the areas indicated by black ellipses in Fig. 8(b) are clearly wrong, while there are no wrong matches in the same area indicated by the black ellipse in Fig. 8(c). It is obvious that the reliability improvement of the proposed multiple close-range image matching method is notable when dealing with repetitive patterns and areas lacking texture.



(a)

(b) (c)

FIG. 6. Consistency check in object space for points on edges: (a) original image with the labelled study area; (b) matching results using only two images; (c) matching results using three images and the proposed method.

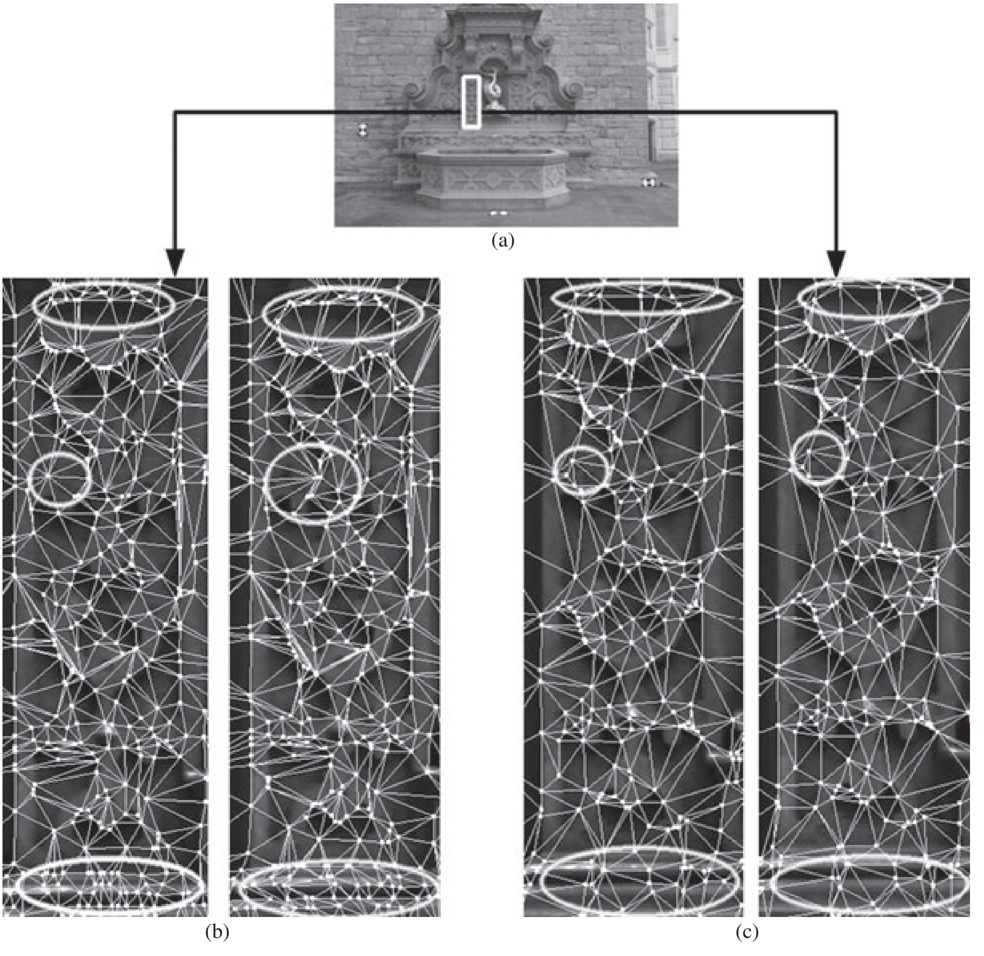


FIG. 7. Consistency check in object space for points on edges: (a) original image with the labelled study area; (b) matching results using only two images; (c) matching results using three images and the proposed method.

CONCLUSIONS

This paper presented a multi-image matching method for close-range images based on a self-adaptive triangle constraint. The experimental analyses conveyed the following conclusions:

(1) The self-adaptive triangle constraint combined with several other constraints can provide a strong constraint for reliable image matching.
(2) The consistency check strategy using multiple images can remove possible mismatches.
(3) "The best the first" matching propagation strategy can adapt to the image texture changes automatically and provide reliable final matching results.

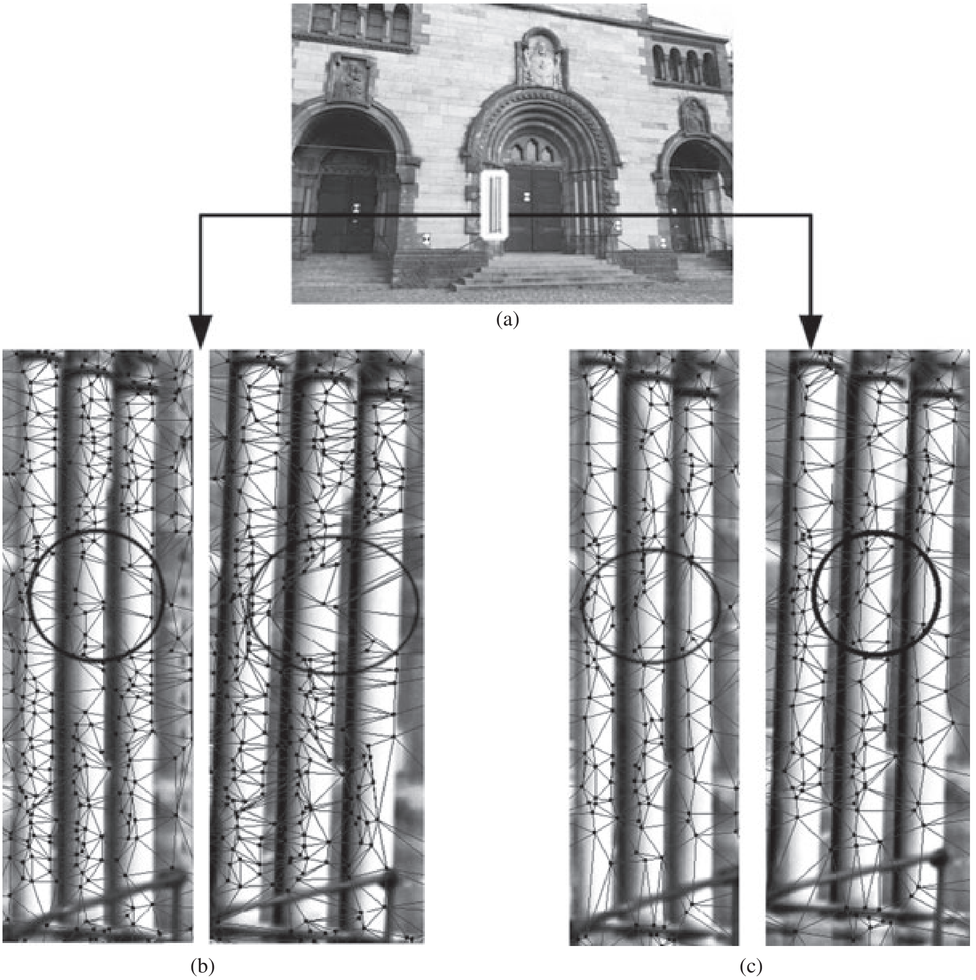


FIG. 8. Consistency check in object space for repetitive patterns or areas lacking texture: (a) original image with labelled study area; (b) matching results using only two images; (c) matching results using three images and the proposed method.

Future work will be area-based matching under the constraints of the interest point matching results (the final TIN) from this proposed method to obtain dense matching points, which can be used for subsequent 3D object modelling and reconstruction.

ACKNOWLEDGEMENTS

The work described in this paper is supported by the National Basic Research Program of China (973 Program, No. 2010CB731800), Research Fund for the Doctoral Program of Higher Education of China (20090141110030), the National Natural Science Foundation of China (40871212) and the Leading Academic Discipline Project of Shanghai Educational Committee

(No. J50104). The work is also supported by a grant from the Research Grants Council of Hong Kong (Project No: PolyU 2312/10E) and a grant from the Start-up Research Fund Program of the Hong Kong Polytechnic University.

REFERENCES

Baltsavias, E. P., 2004. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3–4): 129–151.

Boykov, Y., Veksler, O. and Zabih, R., 2001. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11): 1222–1239.

Brown, M. Z., Burschka, D. and Hager, G. D., 2003. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8): 993–1008.

Geman, S. and Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6): 721–741.

Heipke, C., 2001. Digital photogrammetric workstations—a review of the state-of-the-art for topographic applications. *GIM International*, 15(4): 35–37.

Helava, U. V., 1978. Digital correlation in photogrammetric instruments. *Photogrammetria*, 34(1): 19–41.

Hirschmüller, H., 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2): 328–341.

Hirschmüller, H. and Scharstein, D., 2007. *Evaluation of cost functions for stereo matching*. IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, Minnesota. 8 pages.

Kanade, T. and Okutomi, M. O., 1994. A stereo matching algorithm with an adaptive window: theory and experiment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(9): 920–932.

Knops, Z. F., Maintz, J. B. A., Viergever, M. A. and Pluim, J. P. W., 2006. Normalized mutual information based registration using *k*-means clustering and shading correction. *Medical Image Analysis*, 10(3): 432–439.

Lemaire, C., 2008. Aspects of the DSM production with high resolution images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B4): 1143–1146.

Lhuillier, M. and Quan, L., 2000. Match propagation for image-based modeling and rendering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8): 1140–1146.

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.

Marr, D. and Poggio, T., 1979. A computational theory of human stereo vision. *Proceedings of the Royal Society of London, Series B (Biological Sciences)*, 204(1156): 301–328.

Mayer, H., 2008. Object extraction in photogrammetric computer vision. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(2): 213–222.

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

Mikolajczyk, K. and Schmid, C., 2005. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10): 1615–1630.

Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T. and Van Gool, L., 2005. A comparison of affine region detectors. *International Journal of Computer Vision*, 65(1–2): 43–72.

Pollard, S. B., Porrill, J., Mayhew, J. E. W. and Frisby, J. P., 1985. Disparity gradient, Lipschitz continuity, and computing binocular correspondences. *Proceedings of the Third International Symposium of Robotics Research*, 19–26.

Pollefeys, M., Nistér, D., Frahm, J.-M., Akbarzadeh, A., Mordohai, P., Clipp, B., Engels, C., Gallup, D., Kim, S.-J., Merrell, P., Salmi, C., Sinha, S., Talton, B., Wang, L., Yang, Q., Stewénius, H., Yang, R., Welch, G. and Towles, H., 2008. Detailed real-time urban 3D reconstruction from video. *International Journal of Computer Vision*, 78(2–3): 143–167.

Pratt, W. K., 1991. *Digital Image Processing*. Second edition. John Wiley, New York. 503 pages.

Remondino, F. and El-Hakim, S., 2006. Image-based 3D modelling: a review. *Photogrammetric Record*, 21(115): 269–291.

Remondino, F., El-Hakim, S. F., Gruen, A. and Zhang, L., 2008. Turning images into 3-D models. *IEEE Signal Processing Magazine*, 25(4): 55–65.

RENGARAJAN, R., YOON, J.-S. and SHAN, J., 2004. *Triangulation based hierarchical image matching for Mars DEM generation using MOC NA stereo images*. Proceedings of the ASPRS Annual Conference, Denver, Colorado. 12 pages (on CD-ROM).

SCHARSTEIN, D. and SZELISKI, R., 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1–3): 7–42.

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

SEITZ, S. M., CURLESS, B., DIEBEL, J., SCHARSTEIN, D. and SZELISKI, R., 2006. A comparison and evaluation of multi-view stereo reconstruction algorithms. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, I: 519–528.

STRECHA, C., VON HANSEN, W., VAN GOOL, L., FUA, P. and THOENNESSEN, U., 2008. *On benchmarking camera calibration and multi-view stereo for high resolution imagery*. IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, Alaska. 8 pages. http://cvlab.epfl.ch/~strecha/multiview/denseMVS.html [Accessed: 31st July 2010].

TANG, L., TSUI, H. T. and WU, C. K., 2002. *Dense stereo matching based on propagation with a Voronoi diagram*. Proceedings of 3rd Indian Conference on Computer Vision, Graphics & Image Processing, Ahmadabad, India. 6 pages.

TUYTELAARS, T. and MIKOLAJCZYK, K. 2008. Local invariant feature detectors: a survey. *Foundations and Trends in Computer Graphics and Vision*, 3(3): 177–280.

WANG, Y. 1998. Principles and applications of structural image matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(3): 154–165.

WU, B., 2006. *A reliable stereo image matching method based on the self-adaptive triangle constraint*. Ph.D. dissertation, Wuhan University, China. 110 pages.

ZHANG, L., 2005. *Automatic digital surface model (DSM) generation from linear array images*. Ph.D. thesis, No. 16078, Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology (ETH), Zurich. 199 pages.

ZHANG, Z., ZHANG, J., LIAO, M. and ZHANG, L., 2000. Automatic registration of multisource imagery based on global image matching. *Photogrammetric Engineering & Remote Sensing*, 66(5): 625–629.

ZHU, Q., ZHAO, J., LIN, H. and GONG, J. Y., 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 TIAN, Y., 2007a. Propagation strategies for stereo image matching based on the dynamic triangle constraint. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(4): 295–308.

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

ZHU, Q., WU, B. and WAN, N., 2007c. A sub-pixel location method for interest points by means of the Harris interest strength. *Photogrammetric Record*, 22(120): 321–335.

Résumé

L'appariement fiable d'images est une tâche essentielle et délicate dans le domaine de la photogrammétrie numérique et de la vision par ordinateur. Des problèmes de distorsions géométriques, de changements d'illumination, de changement d'échelle et de configurations texturales complexes peuvent entraîner des ambiguïtés d'appariement, notamment pour l'appariement d'images rapprochées. Cet article présente une méthode d'appariement d'images rapprochées multiples pour la reconstruction de surfaces basée sur une contrainte auto-adaptative sur des triangles. Cette méthode a deux particularités. Tout d'abord, les triangles construits à partir de points précédemment appariés définissent de fortes contraintes pour les appariements ultérieurs, qui sont combinées aux contraintes d'orientation et de disparité des gradients. La mise à jour dynamique de l'appariement s'adapte automatiquement aux changements de textures. Par ailleurs, un contrôle de consistance est effectué dans l'espace objet pour éliminer d'éventuelles erreurs d'appariement. Les résultats expérimentaux obtenus sur trois jeux de triplets stéréoscopiques d'images rapprochées montrent que la méthode proposée améliore la fiabilité de l'appariement pour des images rapprochées.

Zusammenfassung

Eine zuverlässige Bildzuordnung ist eine entscheidende und schwierige Aufgabe in der Digitalen Photogrammetrie und im Computer Vision Bereich. Geometrische Verzerrungen, Beleuchtungsänderungen, Maßstabsänderungen und schwierige Texturbedingungen können in Mehrdeutigkeiten der Bildzuordnung resultieren, die insbesondere im Nahbereich auftreten. Für eine Mehrbildzuordnung zur Oberflächenrekonstruktion im Nahbereich wird hier ein Beitrag auf der Basis selbstadaptierender Dreiecksbedingungen vorgestellt. Entscheidend sind dabei zwei Aspekte. Erstens stellen die Dreiecke, die aus den zunächst zugeordneten Interestpunkten erzeugt werden, starke geometrische Bedingungen für die nachfolgende Punktzuordnung, kombiniert mit Gradientenrichtung und Disparitätsbedingungen, dar. Dabei passt sich die dynamische Erneuerung der Triangulation automatisch den Änderungen der Bildtexturen an. Zweitens wird eine Konsistenzprüfung im Objektraum eingesetzt, um mögliche Fehlzuordnungen zu eliminieren. Mit drei Datensätzen von je dreifach überlappenden Nahbereichsaufnahmen wird experimentell gezeigt, dass die vorgeschlagene Methode eine verbesserte Zuverlässigkeit für die Bildzuordnung im Nahbereich bietet.

Resumen

Lograr una correspondencia de imágenes fiable es una operación tan esencial como difícil en la fotogrametría digital y en la visión por ordenador. Los posibles problemas derivados de las distorsiones geométricas, los cambios de iluminación, los cambios de escala y texturas complicadas resultan en correspondencias ambigüas, especialmente en la de imágenes de objeto cercano. Este artículo describe un método de correspondencia de múltiples imágenes de objeto cercano para realizar una reconstrucción de superficies, basado en un constreñimiento triangular autoadaptativo. Este método presenta dos características principales. En primer lugar, los triángulos construidos a partir de los puntos de interés identificados previamente proporcionan unos constreñimientos geométricos sólidos para la subsiguiente correspondencia de puntos combinada con constreñimientos de orientación y disparidad del gradiente. La actualización dinámica de la triangulación se adapta automáticamente a los cambios en la textura de las imágenes. En segundo lugar, se realiza una comprobación en el espacio objeto cuyo fin es eliminar posibles desajustes o correspondencias erróneas. Los resultados del experimento, utilizando tres conjuntos de imágenes de objeto cercano con solape triple, indican que el método propuesto proporciona una fiabilidad mejorada en la correspondencia de imágenes de objeto cercano.