

Scale and rotation robust line-based matching for high resolution images

Min Chen^a, Zhenfeng Shao^{a,*}, Chong Liu^a, Jun Liu^b

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

^b Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, China

ARTICLE INFO

Article history:

Received 19 October 2012

Accepted 19 March 2013

Keywords:

Line-based matching

Saliency-line

General-line

Control network

ABSTRACT

A line-based matching method is proposed to overcome the low significant level of point feature and the shortage in the matching between weak texture images. Edges are extracted to generate line segments at first. Then, saliency-lines are detected and a control network is constructed in each image. Every remaining general-line is grouped into a saliency-line through a process of clustering. Finally, an iterative algorithm is applied to search for correspondences based on their position relative to the saliency-line. Sufficient spatial information is available to reduce the ambiguity and avoid false matches. Experimental results demonstrate that the proposed method is robust to image scale change and rotation, and the performance is better than point-based method in low texture area.

© 2013 Elsevier GmbH. All rights reserved.

1. Introduction

Image matching is a fundamental issue in computer vision being used in target tracking [1], image mosaicking [2], 3D reconstruction [3] and so on. Generally, image matching algorithms are mainly divided into two categories: intensity-based and feature-based. In intensity-based matching methods, image data is used in the form of a matrix of gray values. This kind of methods provides sub-pixel accuracy and even better. However, they are sensitive to image intensity changes and geometric deformations. These disadvantages can be overcome in feature-based methods by matching local features extracted from images, such as patches, corners, junctions and edges. Therefore, we mainly focus on feature-based method in this paper.

Some famous feature-based methods, like SIFT [4], PCA-SIFT [5], Shape context [6], Steerable filters [7], Differential Invariants [8], have been compared by Mikolajczyk and Schmid [9]. It is indicated that the SIFT method performs better than other methods in most instances. However, the space and time efficiency of SIFT is limited, and the visual significance of the blob features extracted by SIFT are not intuitive enough. Even more, these point-based methods usually fail to deal with low texture scenes that are common in man-made environments. Fortunately, high resolution images are used more widely with the development of high resolution technology. In high resolution remote sensing image, the details of objects are clearer than other images, showing great texture and structure information. Line segments are often abundant in these situations, and complex object boundaries can usually be

approximated with sets of line segments. Many line-based matching methods have been proposed based on this property. Some of them match segments by implementing a weak constraint that adjacent line matches have similar disparities [10], or by checking the consistency of segment relationships [11]. These methods require known epipolar geometry. Some others rely on intensity [12], or color distribution of pixels [13] around line segments to obtain matches. They are sensitive to illumination changes.

In this paper, a novel algorithm is proposed for lines matching of high resolution images. It can successfully match general images and low texture images. Firstly, line segments are extracted and grouped into saliency-line and general-line. Then, saliency-lines are matched and a control network is constructed using the matching result. Each general-line is assigned a position based on the control network. Finally, general-line correspondences are computed through an iterative algorithm according to the position. In our proposed method, one challenge to ensure feature repeatability under line segment extraction is handled by adopting multi-scale strategy. Another challenge to construct distinctive feature descriptors for line segments with different length is avoided by only using spatial information of lines.

2. Robust line-based image matching

2.1. Line segment extraction

In the proposed method, edge pixels of images are extracted by Canny detector [14]. Before they can be used in the image matching, they need to be linked into connected curves and be divided into line segments. For each edge curve, the following algorithm is employed:

* Corresponding author.

E-mail address: shaozhenfeng@whu.edu.cn (Z.F. Shao).

- (1) Calculate the vertical distances from every edge pixel to the line constructed by the two end-points of the curve. And find the maximum distance M and the corresponding edge pixel D ;
- (2) If the value M is less than a threshold, this curve is accepted to generate line segment by the least squares method; Otherwise, it is divide into two parts from the point D , and the original curve is replaced with two new curves. Then, go to the step (1) and until all curves processing are completed;
- (3) Two line segments formed from the above method are merged if both their gap is smaller than a threshold T_g and their line fitting error is smaller than another threshold T_a inversely related to the gap.

Generally, no single threshold can divide all curves in consistent ways when there is scale change between two images. Therefore, a multi-scale scheme is employed in which each curve is divided with a set of scales and all possible results are kept. Similarly, multi-scale scheme is applied in the procedure of merging and all possible line segments are kept.

Once line segments have been extracted, they are grouped into saliency-line and general-line. The value of gradient magnitude is not the only criterion to evaluate the saliency of line segment, its spatial information is also important to judge whether a line is noise or not. The saliency value of a line is computed with:

$$s = l^\alpha \cdot g^\beta, \tag{1}$$

where l is the length of the line, g is the average gradient magnitude of the edge pixels on the line, α and β are two parameters to control the relative importance between the length and the gradient magnitude. In each image, the H line segments which have the greatest saliency value are detected as saliency-lines, and the remaining lines are general-lines. In the proposed method, it is required that all the saliency-lines should not be collinear. If several saliency-line candidates are collinear, the line with the maximum saliency value is kept. The following matching framework includes two steps: saliency-line matching and general-line matching.

2.2. Saliency-line matching

The goal of saliency-line matching is to construct a control network which is the basis of the general-line matching. Meanwhile, the scale and rotational change between images can be regularized.

The saliency-line matching procedure is as follows:

- (1) A saliency-line from each saliency-line set is randomly selected as a root, and other saliency-lines are leaves. Then, a control network is constructed for each saliency-line set;
- (2) In one control network, a leaf is selected randomly as the starting line. The angle γ and center distance d between the root and the starting line are computed. In another control network, the leaf with the closest angle to γ and distance to d is selected as the corresponding starting saliency-line;
- (3) The relative distance and angles are assigned to each saliency-line in both control networks as Fig. 1.

In Fig. 1, S_{Pr} is the root, S_{Ps} is the starting line, S_{Pi} is a leaf selected randomly, P is the intersection of the root and the starting line, α_{Pi1} is the angle between the leaf and the root, α_{Pi2} is the angle between the leaf and the starting line, d_{PMi} is the distance from the point P to the leaf, *axis1* and *axis2* are two axis which divide the plane into four quadrants. The quadrant in which the root exists is the first quadrant, and other three quadrants are determined counterclockwise. θ_{P1} must be equal to θ_{P2} when the plane is divided.

- (4) After each saliency-line in both control networks has been assigned relative distance and angles, a corresponding saliency-line in one control network may be found for every

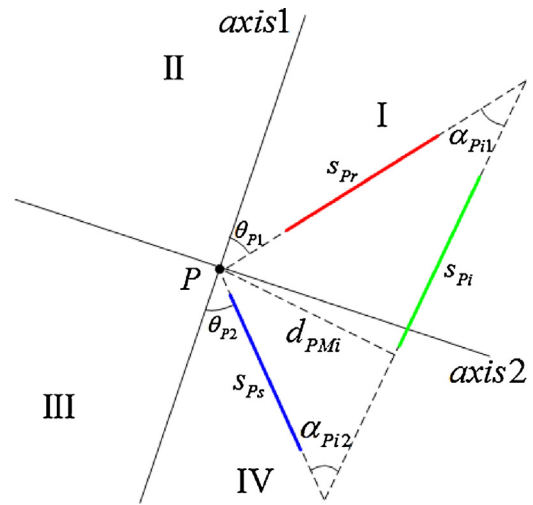


Fig. 1. Relative distance and angles assignment.

saliency-line in another control network based on the following function:

$$\begin{aligned} & \text{correspondence} \\ & = \text{Min} \left\{ \sum_{i=1}^{m-1} \sum_{j=1}^{m-1} \frac{1}{2} \cdot [abs(\alpha_{Pi1} - \alpha_{Qj1}) + abs(\alpha_{Pi2} - \alpha_{Qj2})] \right\} \\ & \times \text{Min} \left[\sum_{i=1}^{m-1} \sum_{j=1}^{m-1} abs(d_{PMi} - k \cdot d_{QNj}) \right]_{M=N}, \tag{2} \end{aligned}$$

where m denotes the number of leaves in the control network, α_{Pi1} , α_{Pi2} and α_{Qj1} , α_{Qj2} are relative angles in the two control networks, d_{PMi} and d_{QNj} are relative distances in the two control networks, M and N are the quadrants in which the leaves exist in the two control networks, respectively. k is a scale factor between the two images. In each iteration, all the possible values of d_{PMi}/d_{QNj} which are consistent with the angle constraint are kept. Then, a histogram for the values of d_{PMi}/d_{QNj} is formed from the minimum to the maximum by the step of 0.1. Finally, the average of values in the bin corresponding to the highest peak is regarded as k . The distances d_{PMi} and d_{QNj} are computed from the intersection point of the root and the starting line to the current leaf. Therefore, the two parameters and the scale factor k will not be affected by the position of the start and the end point on a line.

Only in the condition that the two leaves exist in the same quadrant, the search is performed. The closest leaves with the smallest angle differences and the shortest distance differences will be selected as initial correspondences, where both differences are less than their corresponding thresholds. After each iteration, the number of initial correspondences is recorded as T .

In the method, every saliency-line can be either the root or the starting line. When the whole iteration algorithm is completed, the maximum possible number of initial correspondences is obtained. The correspondences under the iteration with the maximum of T are the final result of saliency-line matching, and the corresponding root and starting line are the final root and starting line of the control network. Because the saliency-line matching is an iterative and exhaustive search process, the number H of saliency-lines should not be too large in order to improve the time efficiency of the iteration.

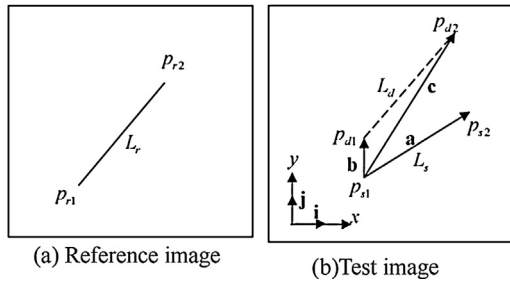


Fig. 2. Transformation between line correspondence. (a) Reference image (b) test image.

The result of saliency-line matching is the basis of the general-line matching. Therefore, wrong matches in the saliency-line correspondences will affect the matching performance of the general-line matching. Although the Eq. (2) includes angle constraint and distance constraint, some outliers may exist in the saliency-line matching result. We use a method based on the RANSAC algorithm to eliminate wrong matches. In this procedure, the transformation parameters between two images are calculated as follows.

As shown in Fig. 2, L_r and L_s is a pair of correspondence. The start point and end point of L_r are P_{r1} and P_{r2} . The start point and end point of L_s are P_{s1} and P_{s2} . L_d is the transformed line of L_s . The start point and end point of L_d are P_{d1} and P_{d2} . Obviously, L_d and L_r are collinear. L_d and L_s are collinear.

We suppose \mathbf{a} is the vector from P_{s1} to P_{s2} , \mathbf{b} is the vector from P_{s1} to P_{d1} , \mathbf{c} is the vector from P_{s1} to P_{d2} . Then, they can be expressed as

$$\begin{cases} \mathbf{a} = (x_{s2} - x_{s1})\mathbf{i} + (y_{s2} - y_{s1})\mathbf{j} \\ \mathbf{b} = (x_{d1} - x_{s1})\mathbf{i} + (y_{d1} - y_{s1})\mathbf{j} \\ \mathbf{c} = (x_{d2} - x_{s1})\mathbf{i} + (y_{d2} - y_{s1})\mathbf{j} \end{cases} \quad (3)$$

From the properties of collineation, it can be described as

$$\begin{cases} \mathbf{a} \times \mathbf{b} = 0 \\ \mathbf{a} \times \mathbf{c} = 0 \end{cases} \quad (4)$$

Then, the error equation can be obtained,

$$\begin{cases} V_{ab} = (y_{s2} - y_{s1})(x_{d1} - x_{s1}) - (x_{s2} - x_{s1})(y_{d1} - y_{s1}) \\ V_{ac} = (y_{s2} - y_{s1})(x_{d2} - x_{s1}) - (x_{s2} - x_{s1})(y_{d2} - y_{s1}) \end{cases} \quad (5)$$

Affine transformation

$$\begin{cases} x_d = a_0 + a_1x_r + a_2y_r \\ y_d = b_0 + b_1x_r + b_2y_r \end{cases} \quad (6)$$

is selected as the mapping function between the two images. Given three pairs of line correspondence, the transformation parameters can be computed based on the Eqs. (5) and (6).

2.3. General-line matching

After the saliency-line matching, two control networks are generated in the two images. The following general-line matching in our method is based on the two control networks. Firstly, each general-line is grouped with the closest leaf in the control network through a process of clustering. Then, a subcontrol network is constructed for each leaf of the control network as the example shown in Fig. 3.

All general-lines in the ellipse are grouped with the saliency-line S_4 . S_4 is selected as the root of the subcontrol network, and

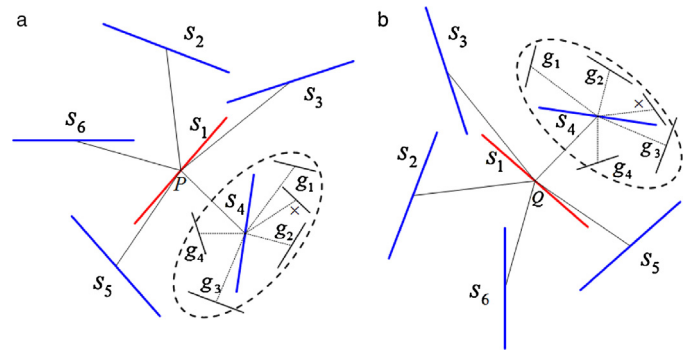


Fig. 3. Subcontrol network construction and general-line correspondence search. (a) Is the control network P . (b) is the control network Q . In each control network, the subcontrol network is constructed with the father root, the corresponding saliency-line and all general-lines grouped with the saliency-line in the ellipse area. In the two corresponding subcontrol networks, four general-line matches are obtained, and “x” denotes an outlier.

the father root S_1 is selected as the starting line in each subcontrol network. Next, every general-line in the subcontrol network is assigned a relative distance and angles as the method in the saliency-line matching. Finally, general-line matching is performed between the two subcontrol networks whose root nodes are correspondences. A corresponding general-line in one subcontrol network may be found for a general-line in another subcontrol network based on the Eq. (2). Correspondences are those general-lines that with the smallest angle differences and smallest distance differences, where both differences are less than their corresponding thresholds. The scale factor k in the general-line matching is determined by the distances D_1 and D_2 between the intersections of the matched saliency-lines in the two control networks, i.e. $k = D_1/D_2$. After general-line matching, we use the RANSAC algorithm again to eliminate wrong matches.

Our method holds ability to handle matching of line segments mainly depend on the direction and intercept of features instead of accurate positioning endpoints. In the matching result, the match is correct if the two line segments are collinear. It breaks through those traditional methods within rigid limit that the two features of a pair of match should be matched accurately.

3. Experiments and results

3.1. Dataset

In our experiments, the areas of city and farmland of GeoEye-1 image are selected to evaluate the proposed method. Image pairs are obtained by applying known transformations to one reference image. The applied transformations are 0.25 times scale change and 30 degrees rotation, respectively. The urban area includes buildings, roads and other facilities as shown in Fig. 4. The farmland area includes farmland and some agricultural infrastructures as shown in Fig. 5.

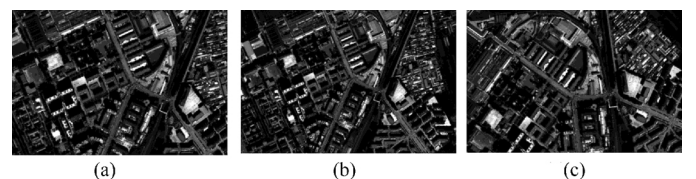


Fig. 4. Experimental dataset 1. Image in (a) is the reference image. Images in (b) and (c) are the test images with scale change and rotation, respectively.

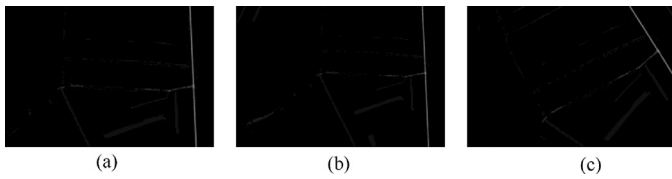


Fig. 5. Experimental dataset 2. Image in (a) is the reference image. Images in (b) and (c) are the test images with scale change and rotation, respectively.

3.2. Image matching results

3.2.1. Parameters

In order to determine the parameters in our approach, 25 pairs of images are selected and the parameters are tuned. After that, a group of suitable parameters are fixed: in line segments extraction, multi-resolution edge split threshold M is set as $\{0.5, 1.5, 2.5\}$, multi-resolution edge merge threshold (T_g, T_a) is set as $\{(3, 0.02), (5, 0.03), (8, 0.04)\}$; in line segments matching, the angle difference threshold is set as 0.03, the distance difference threshold is set as 3.

3.2.2. Evaluation criteria

The matching quality is evaluated by the Recall and RMSE defined as follows:

$$\text{Recall} = \frac{\text{true_matches}}{\text{total_matches}},$$

where *true_matches* is the number of correctly matched line pairs in the matching result, *total_matches* is the total number of existing matched line pairs in the correspondences set.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n [(x'_i - Hx_i)^2 + (y'_i - Hy_i)^2]},$$

where n is the number of selected feature correspondence $(x_i, y_i) \sim (x'_i, y'_i)$, H is the transformation from the reference image to the test image. For the line segments matching result, we select the intersection point of lines to compute RMSE.

3.2.3. Matching results

Firstly, the robustness of the proposed method for scale change and rotation in urban area is evaluated based on the dataset in Fig. 4. The statistical results are shown in Table 1.

The results in Table 1 show that 150 pairs of correct matches are obtained and the recall value achieves 96.77% when scale change exists between images. It is benefited from that the line segments extraction is scale robust since multi-scale schemes are used in the process of split and merging, and the procedure of matching is also scale robust by exploiting the scale factor k . When there is rotation existed between images, the recall value also exceeds 95%. This is because the relative angles are used in the matching. Besides the invariance of the method for image scale change and rotation, the main two reasons to the high recall value are the strong discrimination of line features and the elimination method based on RANSAC. The few wrong matches still existed in the final correspondences can be eliminated by using a group of stronger thresholds. Moreover, in the two groups of test, the RMSE values are less than 1 pixel.

Table 1
Statistical results based on the dataset in Fig. 4.

Variation between images	<i>total_matches</i>	<i>true_matches</i>	Recall	RMSE
Scale	155	150	96.77%	0.513
Rotation	143	136	95.10%	0.499

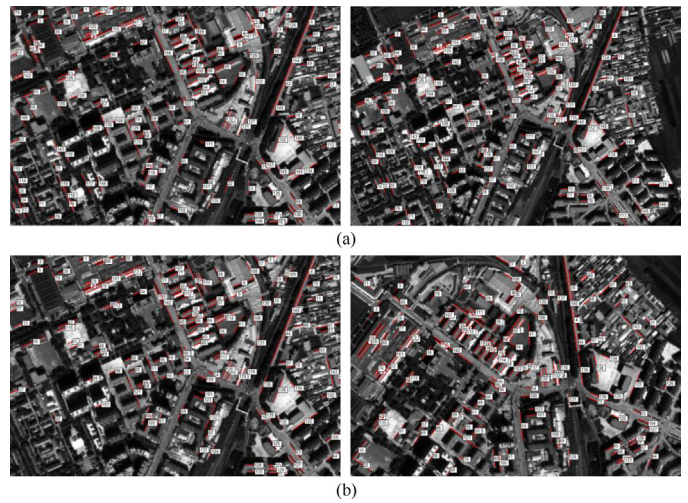


Fig. 6. Experimental results based on dataset 1. Images in (a) and (b) are the matching results of the proposed method under scale and rotational change, respectively.

It indicates the proposed method can achieve a satisfactory accuracy. All the experimental results mentioned above demonstrate our method is scale and rotation robust for the matching of images in urban area. The matching results are shown in Fig. 6.

Another group of experiments are performed to evaluate the performance of the proposed method in low texture images based on the dataset 2 in Fig. 5. Furthermore, the proposed method is compared with the point-based method in low texture area. Considering the outstanding performance of SIFT in point-based methods, we select it as the reference method. The statistical results of experiments are shown in Table 2.

The results in Table 2 show that the proposed method can achieve a high recall value and accuracy when scale change and rotation exist between images in low texture area. It is consistent with the results in Table 1. However, the number of correct matches and recall value of SIFT are much less than those of our method under the two image changes. Furthermore, since the position accuracy of SIFT point is limited in weak texture area, the RMSE is higher than that of our proposed method. The matching results of our method are shown in Fig. 7.

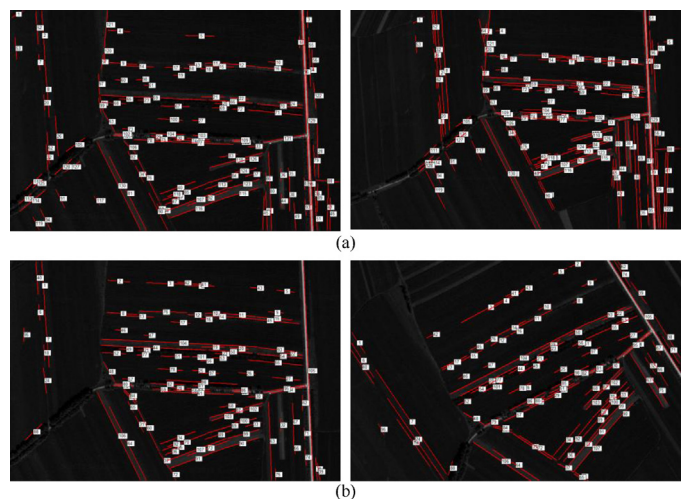


Fig. 7. Experimental results based on dataset 2. Images in (a) and (b) are the matching results of the proposed method under scale change and rotation, respectively.

Table 2
Statistical results based on the dataset in Fig.5.

		<i>total_matches</i>	<i>true_matches</i>	Recall	RMSE
Scale change	Proposed method	131	127	96.95%	0.469
	SIFT	84	58	69.05%	0.809
Rotation	Proposed method	107	102	95.33%	0.397
	SIFT	106	82	77.36%	0.782

4. Conclusion

In this paper, we propose a novel line-based matching method for high resolution images. The line segments extraction is scale robust by using multi-scale scheme. Before line segments matching, they are grouped into saliency-line and general-line and matched separately. The hierarchical strategy employed in the method avoids a mass of iteration. In addition, it is only necessary to search for correspondences within the corresponding subcontrol network. It makes the matching procedure more efficient. In the proposed method, the difficulty to find appropriate feature descriptor of line segment is solved by adopting the spatial information. Because each line has unique linear equation, the discrimination is very strong. Therefore, the recall value in the matching is very high. Of course, like other methods, the proposed method cannot solve every matching problem. Because only scale and rotational change are considered in the method, we think that this method can only be used for high resolution images that were captured with a narrow field-of-view camera or other images that were captured with a short baseline.

Acknowledgements

Our research was supported by National Basic Research Program of China (973 Program, No. 2010CB731800), National Natural Science Foundation program (No. 61172174), the Fundamental Research Funds for the Central Universities (No. 201161902020014), National Science & Technology Specific Projects (No. 2012YQ16018505 & No. 2013BAH42F03), Program of New Century Excellent Talents in University (No. NCET-12-0426)

and the Chongqing Major Programs for Science and Technology Development (No. cstc2011ggC40008).

References

- [1] S.N. Sinha, J.-M. Frahm, M. Pollefeys, Yakup Genc, Feature tracking and matching in video using programmable graphics hardware, *Mach. Vision Appl.* 22 (1) (2011) 207–217.
- [2] D.J. Holtkamp, A.A. Goshtasby, Precision registration and mosaicking of multicamera images, *IEEE Trans. Geosci. Remote Sens.* 47 (10) (2009) 3446–3455.
- [3] E. Kratochvil, L.X. Dong, L. Zhang, B.J. Nelson, Image-based 3D reconstruction using helical nanobelts for localized rotations, *J. Microsc.* 237 (2) (2010) 122–135.
- [4] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.* 60 (2) (2004) 91–110.
- [5] Y. Ke, R. Sukthankar, PCA-SIFT: a more distinctive representation for local image descriptors, in: *Proc. Comput. Vis. Pattern Recogn.*, 2004, pp. 511–517.
- [6] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (4) (2002) 509–522.
- [7] W. Freeman, E. Adelson, The design and use of steerable filters, *IEEE Trans. Pattern Anal. Mach. Intell.* 13 (9) (1991) 891–906.
- [8] P. Montesinos, V. Gouet, R. Deriche, Differential invariants for color images, in: *Proc. Int. Conf. Pattern Recogn.*, 1998, pp. 838–840.
- [9] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (10) (2005) 1615–1630.
- [10] G. Medioni, R. Nevatia, Segment-based stereo matching, *Comput. Vis. Graph. Image Process.* 31 (1) (1985) 2–18.
- [11] Venkateswar, R. Chellappa, Hierarchical stereo and motion correspondence using feature groupings, *Int. J. Comput. Vis.* 15 (3) (1995) 245–269.
- [12] A. Schmid, Zisserman, Automatic line matching cross views, in: *Proc. Comput. Vis. Pattern Recogn.*, 1997, pp. 666–671.
- [13] H. Bay, V. Ferrari, L.V. Gool, Wide-baseline stereo matching with line segments, in: *Proc. Comput. Vis. Pattern Recogn.*, 2005, pp. 329–336.
- [14] J. Canny, A computational approach to edge detection, *IEEE Trans. Pattern Anal. Mach. Intell.* 8 (1986) 679–698.