On the detection of systematic errors in terrestrial laser scanning data

Jin Wang,[1,*] **Hansjoerg Kutterer**[2] **and Xing Fang**[3]

[1] Leibniz Universität, Hannover, Germany

[2] Bundesamt für Kartographie und Geodäsie, Frankfurt am Main, Germany

[3] Wuhan University, Wuhan, P. R. China

Abstract. Quality descriptions are parts of the key tasks of geodetic data processing. Systematic errors should be detected and avoided in order to insure the high quality standards required by structural monitoring. In this study, the iterative closest point (ICP) method was invested to detect systematic errors in two overlapping data sets. There are three steps to process the systematic errors: firstly, one of the data sets was transformed to a reference system by the introduction of the Gauss–Helmert (GH) model. Secondly, quadratic form estimation and segmentation methods are proposed to guarantee the overlapping data sets. Thirdly, the ICP method was employed for a finer registration and detecting the systematic errors. A case study was casted in which a dam surface in Germany was scanned by terrestrial laser scanning (TLS) technology. The results indicated that with the conjugation of ICP algorithm the accuracy of the data sets was improved approximately by 1.6 mm.

Keywords. Terrestrial laser scanning, systematic errors, iterative closest point, structural monitoring, quadratic form estimation, Gauss–Helmert model.

1 Introduction

1.1 Background

Terrestrial laser scanning (TLS), which is an efficient approach to ascertain near real time 3D points, has frequently been used in structural monitoring. However, the quality of the data from the TLS is often impaired by various error sources. Although the magnitude of the errors may be not significant in some applications, even systematic errors which only measure a few millimetres should be

Corresponding author: Jin Wang, Geodätisches Institut, Leibniz Universität Hannover, Nienburger Str. 1, 30167 Hannover, Germany. E-mail: wang@gih.uni-hannover.de.

detected to guarantee the accuracy of data sets in structural deformation monitoring.

It is possible to render quality of 3D point clouds with accuracy and precision, as can be identified by such standard organisations as ANSI, ASTM and ISO [1, 9]. One indicator of the data quality is accuracy, which is defined, according to [9], "as the closeness of agreement between a measured quantity value and a true quantity value of a measurand". Another indicator of data quality is precision, which is given by [9], "as the closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions". It is also noted that "precision" does not relate to the true value and a measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results. Less precision is reflected by a larger standard deviation. "Accuracy" is a qualitative concept and the term "precision" should not be used in place of "accuracy" [5]. Systematic errors indicate the degree of accuracy and may not be reduced by averaging vast amounts of data. It is assumed that if there are no systematic errors, the distance between corresponding points should be zero after the transformation of point clouds into a common coordinate system.

A fine registration of two range images was casted by first aligning feature points, followed by feature surface elements, which are sets of points corresponding to the image area determined by the feature [3]. A modified ICP method [10] was generated to explore the systematic errors in overlapping strips of airborne laser scanning data sets for the detection of systematic errors. The errors usually come from vegetation or other non-permanent features. Based on a triangulated irregular network (TIN) structure, Maas [13] solved the problem that applying matching methods causes systematic errors of the shift parameters with partial occlusions in laser scanner data. Yoon et al. [20] carried out an error compensation process for detecting systematic errors at depth discontinuities. For relevant researches, we refer to Vosselman [17] and Schreier and Sutton [15].

Various matching algorithms, such as Scale Invariant Feature Transform (SIFT) [12] and Speeded Up Robust Features (SURF) [2], are used in many disciplines. However, for objects without distinctive features, it is hard to find enough available identical patches, especially for structures with similar curvature and colour. In this case, many matching methods may fail. The iterative closest point (ICP) algorithm was investigated by Besl and McKay [4],

Yang and Medioni [18], and Zhang [21]. The pros and cons of the ICP approach have been illustrated by Gruen and Akca [7]. The basic version of the ICP method is based on the search for pairs of nearest points in the two sets, and estimates a rigid transformation. One of the data sets is considered as a reference surface. The other data set is called the matching surface. Proper initial values are necessary to guarantee correct matching. This means a rough transformation needs to be performed before the ICP method is launched. We introduce the ICP algorithm to adjust for systematic errors. Strategies were employed to speed up computation time and improve data quality: reduction of the number of matching points through quadratic form estimation, and speeding up the accuracy of data sets with providing of proper initial values.

An application in the Harz dam monitoring task was performed for data collection of the dam surface by Trimble GX 3D scanner. Eling [6] demonstrated the existence of systematic errors between points from different scanner stations, even when the calibration of vertical angles and distances was undertaken before the scanning task. Because this precalibration may be not sufficient, we propose a feature matching method which alleviates the systematic errors in TLS data. The experiments showed that the systematic errors are allayed approximately by 1.6 mm with the implementation of the ICP algorithm.

The paper is organised as follows. The proposed method will be illustrated in Section 2. In Section 3, an application of dam monitoring will be presented in detail. Conclusions are drawn in Section 4.

2   Proposed method

The proposed method is illustrated in the following three steps: First, the Gauss–Helmert model (GH model) represents the transformation between a scanner station and a local coordinate system, and this model is performed in order to estimate the seven transformation parameters by using identical points. Second, quadratic form parameters representing the object surface are estimated in order to build a uniform framework, which is then used to segment the point clouds into the corresponding small blocks. The point clouds in one block are estimated as one representing point. The representing point stands on the points in the block. Third, the ICP algorithm is then applied as a finer registration strategy for reducing the discrepancies between two representing point sets. The two representing point sets come from two scanner stations.

2.1   Registration with the Gauss–Helmert model

The 3-D similarity transformation of coordinates is a computational procedure that maps one set of coordinates in a given system onto another. In this scenario, an estimation of the transformation parameters is achieved by identical points which provide best fitting of the conversion between the scanner system and the local system. The transformation functions can be expressed as follows:

$$\mathbf{y}_j^i = \mathbf{\Delta x}^i + \mathbf{s}^i \mathbf{R}(w^i, \phi^i, \kappa^i)\mathbf{l}_j^i, \tag{1}$$

where $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, q$; p and q equal the total number of scanner stations and identical points, respectively; $\mathbf{y}$ is the vector of coordinates in the local system; $\mathbf{l}$ is the vector of the coordinates in the scanner system; $\mathbf{\Delta x}$ is the vector of translations of scanner station; $\mathbf{s}$ is the scale; and $\mathbf{R}$ is the rotation matrix with three rotation angles (ω, ϕ, κ).

Linearising the above formula by a first-order Taylor approximation, the GH model is formed as follows:

$$\mathbf{Bv} + \mathbf{A\delta x} = \mathbf{w}, \tag{2}$$

where $\mathbf{B}$ and $\mathbf{A}$ are Jacobian matrices with respect to observation vector and parameter vector; $\mathbf{v}$ and $\mathbf{\delta x}$ are the corrections to the approximate values of observations and parameters; $\mathbf{w}$ is the misclosure vector.

The solution of unknown parameters in the established GH model is illustrated by [14, pp. 175–180]:

$$\mathbf{\delta \hat{x}} = -(\mathbf{A}^T(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^T)^{-1}\mathbf{A})^{-1}\mathbf{A}^T(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^T)^{-1}\mathbf{w}, \tag{3}$$

where $\mathbf{Q}_{\mathrm{ll}}$ is the weight coefficient matrix of the observations. The solution can be obtained iteratively, provided that the Jacobian matrices are correctly updated.

The effects of gross errors to the estimator are eliminated through a properly built equivalent weight matrix [19]. Furthermore, variance components estimation is proposed for modifying the precision of observations from three types [16]. In this study, we classified the observations according to ranges, horizontal and vertical angles.

2.2   Quadratic form estimation and segmentation

For structure monitoring, a huge data volume and automated mathematical model need to be considered and handled. In many cases, the shape of the object surface could be described through quadratic form estimation, where several parameters stand for curvature of the object [8].

The general equation for the quadratic form estimation can be read as

$$\mathbf{x}_k^T \mathbf{M} \mathbf{x}_k + \mathbf{m}^T \mathbf{x}_k + \alpha = 0, \tag{4}$$

where $\mathbf{x}_k$ is the coordinate vector of a single point; $k = 1, 2, \ldots, n$; n denotes the total number of points scanned by TLS; $\mathbf{M}$ is a 3×3 symmetric coefficient matrix; $\mathbf{m}$ is the

coefficient vector; and α is the scalar. The parameters are expressed as

$$\mathbf{x}_k = \begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix}, \quad \mathbf{M} = \begin{pmatrix} a_1 & a_4 & a_5 \\ a_4 & a_2 & a_6 \\ a_5 & a_6 & a_3 \end{pmatrix},$$

$$\mathbf{m} = \begin{bmatrix} a_7 \\ a_8 \\ a_9 \end{bmatrix}, \quad \alpha = a_{10}. \tag{5}$$

Equation (4) can be consequently represented by

$$a_1 x_k^2 + a_2 y_k^2 + a_3 z^2 + 2a_4 x_k y_k + 2a_5 x_k z_k$$
$$+ 2a_6 y_k z_k + a_7 x_k + a_8 y_k + a_9 z_k + a_{10} = 0. \tag{6}$$

The parameters from a_1 to a_{10} could be estimated by a GH model with several steps of iteration for convergence. The determinant method, in which an extended form matrix $\mathbf{M}^*$ is constructed, was launched to estimate the four motion invariant parameters δ, Δ, B and J (see [11]).

The extended form matrix is identified as

$$\mathbf{M}^* = \begin{pmatrix} \mathbf{M} & \mathbf{m} \\ \mathbf{m}^T & \alpha \end{pmatrix}. \tag{7}$$

For the functions of the four parameters of δ, Δ, B and J refer to Eling [6]. The shape of the object was determined by looking for the test tree of automatic form recognition [8].

A uniform framework of the object surface is established in a spherical coordinate system with the parameters calculated by quadratic form estimation. Based on the framework, the object surface is segmented into small blocks and the point clouds from the TLS are divided into corresponding blocks. Each block is represented by a representing point, which is estimated by a least squares method. The representing points are situated in the centres of gravities of the blocks when the blocks have enough point clouds for the least squares adjustment. This reduces the volume of data, and builds a uniform framework for point clouds from both different scanner stations and different epochs. The object surface is then demonstrated and analysed by the representing points. In other words, the representing points stand for the 3D representations of the object surface, which we also call a surface or a feature in this paper.

The GH model processes the random errors both from observations and from configuration matrix. For the detection and reduction of systematic errors, the ICP algorithm is introduced to correct the discrepancies between two matching surfaces.

2.3 Implementation of the ICP method

The idea of the ICP algorithm is to find the closest distance between two matching surfaces without considering the scale index. This computation is performed for each representing point in the matching surface against all representing points in the reference surface. It iteratively revises the transformation (translation and rotation) in order to minimise the distance between the points of two data sets.
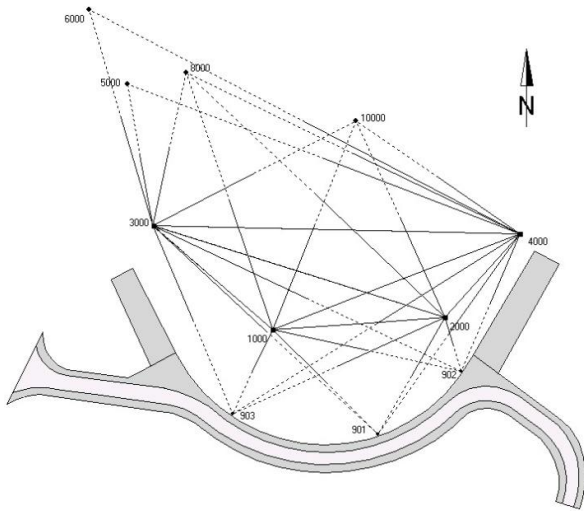
Thus the workflow of the ICP algorithm could be simply involved as follows:

1. Input: two point sets from a reference model and a matching model.

2. Initialisation: proper initial values which are benefit for fast and correct convergence.

3. Iteration: calculate the closest distance and then the transformation matrix, thus utilising the rigid transformation to the matching surface. The iteration stops if the points in the matching model are close enough to the points in the reference model.

4. Output: a transformation matrix which transforms the matching surface to the reference model.

For step 3, as the two data sets converge, the corrections of the resulting rigid transformation become smaller, and the solution iteratively approaches a global minimum. Once the minimum distances between the matching surface and the reference surface have been achieved, a covariance matrix was calculated with the mean values of coordinates from both surfaces. A quaternion matrix was utilised for the estimation of a transformation matrix, where arrays of features are estimated in a least square sense in closed form. Therefore, the points in the matching surface are transformed into the reference surface.

One issue for the application of the ICP algorithm is about initial approximations. This means a satisfaction of good initial approximations in the beginning of iteration. With the introduction of registering by the GH model, good initial values could be reached, and correct convergence is also guaranteed. Another issue is about the overlapping regions of two surfaces. The adoptions of quadratic form estimation and segmentation methods throw out the points that are unique to each surface. Therefore, the two surfaces iterated in the ICP algorithm are entirely overlapping regions. A finer registration is obtained, and systematic errors existing in the first step of transformation are decreased by the implementation of the ICP algorithm.

Because the deformation of structures is quite small, the interference of systematic errors lets the analysis of deformation be inaccurate. Thus, the detection method of systematic errors is very reasonable for monitoring tasks. The corrected coordinates of the representing points from different epochs are compared for analysing the possible variations of objects.

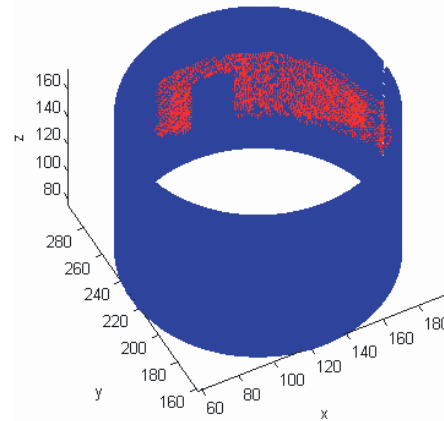**Figure 1.** Geodetic monitoring network of the dam [6].

Figure 2. Elliptic cylinder fitting (the red section is the registered points. The blue section is the fitted elliptic cylinder).

3   Application and discussion

A case study was examined in a monitoring task for a dam located in the Harz Mountains, Germany. Four scanner stations (station 1000, 2000, 3000 and 4000) were established to scan the dam surface for data collection (Figure 1). Pre-calibration to range and vertical angle was tested before scanning [6]. The geodetic network with the original stations 5000 and 6000 was built using a survey of total stations. Without loss of generality, we issued data sets only on epoch 4.

In order to get a full representation of the object surface, the similarity transformation was implemented to transform point clouds into the local coordinate system. At the beginning of data processing, gross errors were eliminated because points contaminated by the outliers could influence the accuracy of registration and stop the ICP algorithm during the iteration process. In order to get the resulting weight matrix of observations and remove gross errors, the equivalent weight matrix and the variance components estimation were introduced for gross errors detection and robust estimation.

With quadratic form estimation, the shape of the dam surface was estimated as an elliptic cylinder (Figure 2). The dam surface was divided into small rectangular blocks with the criterion of horizontal and vertical angles in the spherical coordinate system. A representing point, which should be situated in the gravities of the blocks, was obtained in the block by the least squares method. The accuracy estimation was achieved by the analysis of the representing points. Approximately 12,000 blocks were plotted in the dam surface. Unique series numbers were attached with the character information of each block. The series number of the block started from the right bottom of the dam surface.

A proper block size warrants the accuracy of estimated representing points. If the definition of block size is too large, there may be a significant variation of the block curvature. Smaller deformation of object surface could be hard to detect. If a block size is too small, not enough point clouds can be divided into the block. These would cause deviations of the estimation of representing points.

After registration with artificial targets in the GH model, approximately 5 mm systematic errors were detected in point clouds between two scanner stations. Our application then employed the ICP algorithm to the representing points to check and reduce the discrepancies between the two scanner stations.

The representing points which were unique to one surface model are not included in the transformation computation. We confirmed the correspondences of representing points with the unique series number during the segmentation process. The representing points that are not a portion of the overlapping surface were removed during the ICP calculation. This guarantees the exact overlapping areas.

Choosing feature from one scanner station as a reference model, the ICP algorithm was executed so that the matching model was transformed into the reference system. An initial alignment, which takes the overlapping surface close enough, is indispensable for a fast convergence of the ICP method.

In order to test the validation of the ICP method, we compared the results from the ICP method and the GH model. We show results from epoch 4. The mean value of the standard deviation (STD) represents the average distance between the corresponding representing points (Table 1). If no systematic errors exist, the distance between representing points from different stations should be zero. In epoch 4, the mean STD with the ICP approach was 3.4 mm, while the mean STD with the GH model was 5 mm. In the procedure of the GH model, the mean value of 1000 vs. 4000 was 6.0 mm, while the mean value using the ICP approach was 2.8 mm (Table 1).
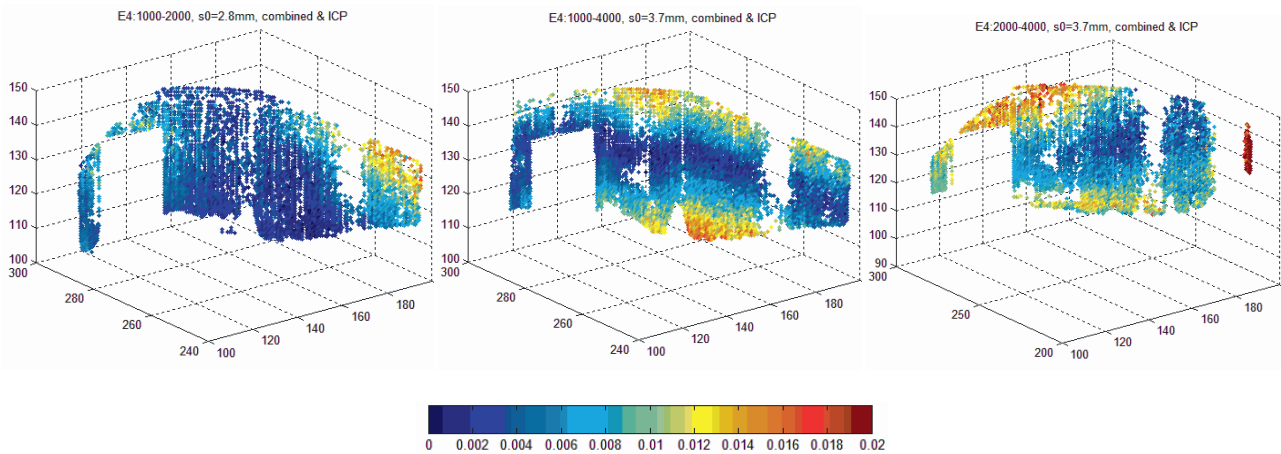
Figure 3. Distance difference of the representing points between two scanner stations (the ICP method).

STD (mm)	ICP method	GH model
1000 vs. 2000	2.8	6.0
1000 vs. 4000	3.7	4.6
2000 vs. 4000	3.7	4.1

Table 1. Mean values of distance between the representing points.

Figure 3 illustrates the distance differences between the corresponding representing points in two of the scanner stations. Comparisons (station 1000 vs. 2000, station 1000 vs. 4000, and station 2000 vs. 4000) are shown from left to right. The middle parts of the dam are quite stable; meanwhile the largest biases happen in the boundary of the dam. For station 1000 vs. station 2000, distance difference varies from 0 mm to 12 mm, while the largest variation appears on the top right part of the dam. The most obvious difference of distance occurs at the bottom of the dam with station 1000 vs. station 4000. It is possibly because station 4000 is farthest away from the dam and the distance has a direct effect on the quality of observations.

For structural deformation analysis, the representing points were compared in order to test if the corresponding points had significant variations in different epochs. Deformation of the structure is usually small. Thus, the confirmation and elimination of systematic errors are indispensable in the task of structure monitoring.

4 Conclusions

We have presented an ICP approach to adjust systematic errors within TLS data. This is based on identifying surface in a data set and matching it to the overlapping surface in a reference model. The problems associated with the ICP algorithm are categorised in two ways. The first is by describing the dam surface with the representing points. This method decreased the data volume and assured the overlapping area. The other way is by choosing proper initial values to guarantee correct convergence and save running time. In this research, the application of the ICP algorithm reduced the systematic errors by about 1.6 mm.

As explained, a suitable block size is beneficial for the estimation of the representing points, because the ICP algorithm was operated with respect to the representing points, and the accuracy of the estimated representing points affected the results of the ICP method as well. Large deviations of the initial values would also lead to ICP method failure.

Although the implementation of the ICP algorithm does reduce the systematic errors, 3.6 mm systematic errors still remain within the data sets. Future work needs to include configuration and other information in order to detect and eliminate the systematic errors.

Acknowledgments

References

[1] ANSI/NCSLZ540-2-1997, American National Standard for Expressing Uncertainty, U.S. Guide to the expression of Uncertainty in Measurement, National Conference of Standards Laboratories, 1997.

[2] H. Bay, A. Ess, T. Tuytelaars and V. L. Gool, SURF: Speeded up robust features, *Comput. Vision Image Understanding* **110** (2006), 346–359.

[3] G. H. Bendels, P. Degener, R. Wahl, M. Körtgen and R. Klein, Image-based registration of 3d-range data using feature surface elements, in: *5th Int. Symp. on Virtual Reality,*

Archaeology and Cultural Heritage (VAST'04) (2004), 115–124.

[4] P. J. Besl and N. D. McKay, A method for registration of 3-D shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* **14** (1992), 239–256.

[5] G. S. Cheok (ed.), *Proceeding of the 2nd NIST LADAR Performance Evaluation Workshop*, NISTIR 7266, National Institute of Standards and Technology, Gaithersburg, 2005.

[6] D. Eling, *Terrestrisches Laserscanning für die Bauwerks-überwachung*, Fachrichtung Geodäsie und Geoinformatik, Leibniz Universität., Hannover, 2009.

[7] A. Gruen and D. Akca, Least squares 3D surface and curve matching, *ISPRS J. Photogramm. Remote Sens.* **59** (2005), 151–174.

[8] C. Hesse and H. Kutterer, Automated form recognition of laser scanned deformable objects, in: *Geodetic Deformation Monitoring: From Geophysical to Engineering Roles*, IAG Symposia 131, Springer (2006), 103–111.

[9] ISO VIM, International vocabulary of basic and general terms in metrology (VIM), International Organization for Standardization, 2004.

[10] P. Kumari, W. E. Carter and R. L. Shrestha, Adjustment of systematic errors in ALS data through surface matching, *Adv. Space Res.* **47** (2010), 1851–1864.

[11] H. Kutterer and S. Schoen, Statistische Analyse quadratischer Formen, *AVN* **10** (1999), 322–330.

[12] D. G. Lowe, Object recognition from local scale-invariant features, in: *The Proceedings of the Seventh IEEE International Conference on Computer Vision* 2 (1999), 1150–1157.

[13] H. G. Maas, Least-squares matching with airborne laserscanning data in a TIN structure, *Int. Arch. Photogramm. Remote Sens.* **33** (2000), 548–555.

[14] W. Niemeier, *Ausgleichungsrechnung*, 2nd ed., De Gruyter, Berlin, 2008.

[15] H. W. Schreier and M. A. Sutton, Systematic errors in digital image correlation due to undermatched subset shape functions, *Exp. Mech.* **42** (2002), 303–310.

[16] L. E. Sjoeberg, Non-negative variance component estimation in the Gauss-Helmert adjustment model, *Manuscr. Geodaet.* **9** (1984), 247–280.

[17] G. Vosselman, On the estimation of planimetric offsets in laser altimetry data, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **34** (2002), 375–380.

[18] C. Yang and G. Medioni, Object modelling by registration of multiple range images, *Image Vision Comput.* **10** (1992), 145–155.

[19] Y. Yang, L. Song and T. Xu, Robust estimator for correlated observations based on bifactor equivalent weights, *J. Geod.* **76** (2002), 353–358.

[20] S. Yoon, S. K. Park, S. Kang and Y. K. Kwak, Fast correlation-based stereo matching with the reduction of systematic errors, *Pattern Recognit. Lett.* **26** (2005), 2221–2231.

[21] Z. Zhang, Iterative point matching for registration of free-form curves and surfaces, *Int. J. Comput. Vision* **13** (1992), 119–152.