THREE DIMENSIONAL OBJECT RECONSTRUCTION BY OBJECT SPACE MATCHING

Franz Rottensteiner*
Institute of Photogrammetry and Remote Sensing
TU Vienna
Gußhausstraße 27-29
A-1040 Vienna, Austria
frottens@fbgeo1.tuwien.ac.at

ISPRS Commission III, Working Group 3

KEY WORDS: Matching, Surface, Modeling, Reconstruction, Three-dimensional Object Reconstruction, Feature Adjacency Graphs, Object Space Matching

ABSTRACT:

One of the main current fields of research in photogrammetry is concerned with the reconstruction of object surfaces from digital images. Both feature based and area based matching algorithms have been described in literature. Most of these algorithms aim at a 2½D representation of the object surface which is to be derived from a digital stereo pair. They give good results when the surface to be reconstructed is smooth, and they face problems in areas with surface discontinuities. Feature based matching techniques appear to be more flexible with respect to surface discontinuities and requirements for approximate values; area based matching techniques offer a higher accuracy potential. In the course of the Austrian Research Program on Digital Image Processing we have developed a concept for an image matching algorithm considering topological relations which should work under quite general circumstances, e.g. in close range applications. A generalized framework based on the Förstner operator will be used for feature extraction. The generation of correspondence hypotheses will be based on region adjacency graphs; we are investigating more complex models in object space for the evaluation of these correspondence hypotheses in order overcome the problem of occlusions. The integration of a bundle block software should provide geometrical constraints for correspondence hypotheses and give us the possibility to use more than two images for surface reconstruction. Approximate values will be improved by hierarchical methods (image pyramids) starting from a very coarse level. The final result of our algorithm should be a fully three-dimensional representation of the surface to be reconstructed.

1. MOTIVATION

One of the main current fields of research in photogrammetry is concerned with the reconstruction of object surfaces from digital images. A high level of automation can be achieved for this task by applying matching techniques to digital images (Gülch, 1994). Several different approaches have been proposed in literature. Some systems have left the experimental state and can now be used in the production process of Digital Elevation Models (DEM), at least for small and medium image scales. These systems aim at a 2½D representation of the object surface which is to be derived from a pair of epipolar images (Krzystek, 1995); (Gülch, 1994). The matching algorithms used in such systems give good results when the surface to be reconstructed is smooth but face problems in areas with surface discontinuities, especially in urban areas, because most of them implicitly use models assuming the object surface to be smooth or even planar. If the object meets this smoothness assumption and if it can be modeled in 2½D, the above algorithms can also be used for close-range applications. However, if those conditions are hurt, they will fail to give good results.

In the course of the Austrian Research Program on Digital Image Processing we are investigating matching techniques for object reconstruction which can also be used in cases when the requirements mentioned above are hurt, as it happens with many close range applications and with large scale aerial images. Less emphasis is laid on the optimization of computational speed than on finding matching algorithms which give accurate and reliable results under quite general circumstances and render possible a high degree of automation. The demand for applicability of the algorithm to close-range images and for large aerial image scales leads to certain requirements:

- Many classes of objects can no longer be described by 2½D models. Thus, more sophisticated methods for fully 3D representation of the object to be reconstructed become necessary.
- In close-range photogrammetry, more than two images of a certain object region are usually available. The viewing directions might be convergent, thus the assumption of a near-normal-case configuration might have to be dropped.
- Matching will be heavily influenced by surface discontinuities and occlusions; the object surface can thus no longer be modeled to be smooth.
- In order to achieve a high degree of automation, very coarse approximations for the object surface should be sufficient for the algorithm.

In section 2 we want to give an overview about related work in the fields of image matching and object representation. In that section we will also discuss the applicability of the techniques described in literature to our problem. Our concept will then be described in section 3.

* This work is funded by the Austrian Research Program S7004-MAT on Theory and Applications of Digital Image Processing and Pattern Recognition, Project IV: Stereovideometry and Spatial Object Recognition

2. RELATED WORK

2.1 Matching Techniques

One possibility to characterize matching algorithms is given by the geometric and radiometric models they use for the mapping functions (figure 1).

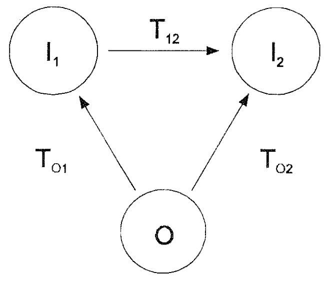


Figure 1: Object reconstruction and image matching with two images; taken from (Lang et al., 1995)

Matching algorithms which work in object space reconstruct the object O directly by inverting the perspective transformations T_{O1} and T_{O2} after having found initial correspondencies between homologous image features. Object space matching techniques have the advantage that they are closer to physical reality so that they may be capable of handling occlusions if more sophisticated object models are used for the evaluation of the initial correspondencies. On the other hand, the number of parameters to be estimated in the inversion process is very high (Lang et al., 1995). In order to avoid the computational complexity of object space matching, the algorithms used in most photogrammetrical systems apply image matching techniques (Gülch, 1994), (Krzystek, 1995) which relate the images I_1 and I_2 by a mapping function T_{12}. In this case, the object model is implicitly contained in the formulation of T_{12}, which can be approximated by an affine transformation if the object surface can be assumed to be locally flat. If this assumption is hurt, T_{12} becomes more complicated. This is the reason why image matching techniques using a flat terrain model give bad results in the presence of occlusions (Lang et al., 1995).

From another point of view, matching algorithms can be characterized by the image model they use (Gülch, 1994). Area based matching algorithms use a raster representation of the image, i.e. they try to find a mapping function between image patches by directly comparing the grey levels (Ackermann, 1984). Feature based matching techniques on the other hand first derive a symbolic description of the image by extracting salient features from the images using some feature extraction operator, e. g. the Förstner operator (Förstner, 1986), and then try to find corresponding features under certain assumptions regarding the local geometry of the object to be described and the mapping geometry (Krzystek, 1995). Feature based algorithms appear to be more flexible with respect to surface discontinuities and requirements for approximate values whereas area based matching techniques offer a higher potential of accuracy (Gülch, 1994).

Both area based and feature based methods as described in the works cited above rely on similar grey level distributions in different images. This is an assumption which holds true for stereo images but may be hurt for images taken from convergent viewing directions. The topology of features which can be stored in relational descriptions is an image property which is invariant under perspective transformation (Vosselman, 1995). The topological relations between neighbouring features can be extracted together with the features themselves. Matching of relational descriptions or relational matching thus is a very powerful concept which might work in rather general cases. However, its computational complexity is very high because it leads to rather complex search trees (Vosselman, 1995). Whereas common area and feature based techiques have already been applied simultaneously to more than two images, no multi-image relational matching algorithm is known (Förstner, 1995).

2.2 Three - Dimensional Object Representation

Common systems for surface reconstruction from digital images often use 2½D surface representations which is convenient for many applications, especially if the object in consideration is the earth and if the image scale is small (Gülch,1994). More complex objects cannot be described in that way. Modeling of arbitrary surfaces requires the surface description to be independent from the coordinate system. This can be achieved by decomposing the surface into basic geometric entities such as nodes, edges and triangles. The structure of the surface is then described by the topological relations between these elements; its geometry is determined by assigning the nodes uniquely to the measured surface points. Originally, only these points (and possibly, the surface normals in these points) are available. The topological relations can be found by triangulation. (Heitzinger, 1996) describes an incremental method for that purpose which is to form the basic surface description model for the new implementation of the program system SCOP. As triangulation of a given point set is not unique, this method aims at a triangulation which reproduces the surface as correctly as possible. Thus after inserting a node into the triangulation, the triangulation has to be optimized according to some criterion, e.g. smoothness. Triangulation should be able to establish constraints in order to render possible the consideration of break lines (Heitzinger, 1996).

2.3 Discussion

Having in mind requirements stated in section 1, we think that feature based matching in object space is best suited for a solution to our problem because such an algorithm seems to be capable of overcoming the problem of occlusions by introducing more complex models for the evaluation of correspondence hypotheses in object space. It also can benefit from the possibilities of bundle block geometry with regard to geometrical constraints and the usage of more than two images. As topological relations between neighbouring features provide valuable information for the generation of correspondence hypotheses, they should be considered by the algorithm, too.

The method for surface representation by (Heitzinger, 1996) appears to be well-suited for describing quite general surfaces. As triangulation is rather complex a step, it will not be done during the matching phase, but will be applied to the point set which is the result of matching. If surface discontinuities have been detected by the matching algorithm, this information should be considered in triangulation.

3. A CONCEPT FOR OBJECT RECONSTRUCTION

3.1 General Strategy

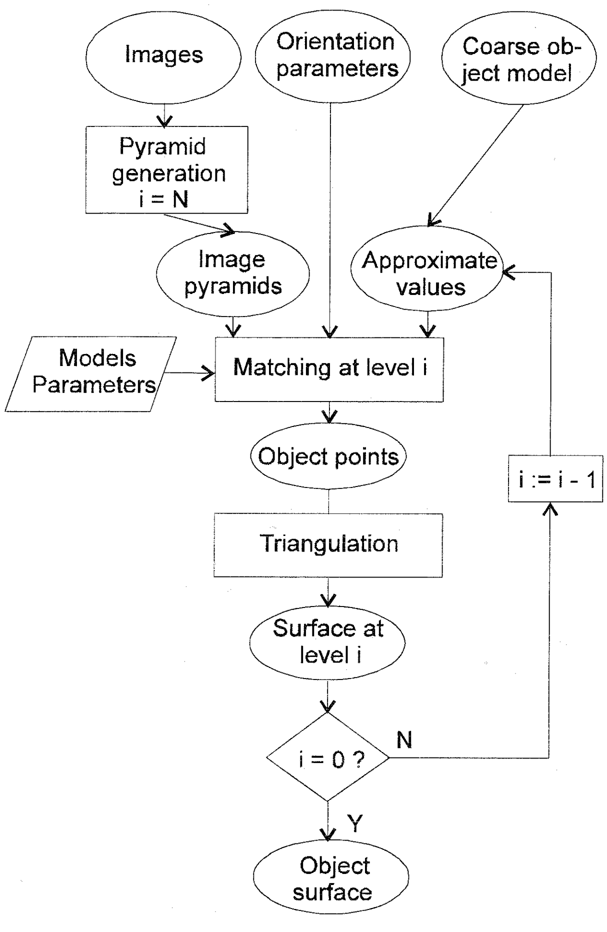


Figure 2: General matching strategy

The general strategy of our concept for object reconstruction is shown in figure 2. Two or more digital images, their orientation parameters and a coarse model of the object surface (e.g. a tilted plane or a cube) provide the input for the algorithm. In a first step, image pyramids have to be generated from the original images. The matching algorithm which is controlled by a small number of parameters and by pre-defined object models, is first applied to the upper level (level N) of the image pyramids with approximate values derived from the coarse object model. The matching result is a cluster of points supposed to be on the object surface, possibly together with some information about surface discontinuities. A triangulation of these points delivers the description of the object surface at the upper level of the image pyramids which is now used as an approximation for the next lower level.

Matching and triangulation are now iteratively applied to each level i of the image pyramids; the results of level i provide the approximate values for level i-1. The process is stopped as soon as the lowest level of the image pyramids (i.e. the level with the highest spatial resolution; i = 0) has been reached.

The matching algorithm at a certain level i of the image pyramids is described in more detail in figure 3. Matching starts with the extraction of features and their mutual relations under the assumption of a certain image model (section 3.2). The topological relations between neighbouring features are described by a feature adjacency graph. Basically, feature extraction delivers both point and line features as well as homogeneous image regions. In our concept, we only use points to represent surfaces for the time being. However, the topological relations between points, lines and homogeneous regions will be used in the course of matching. Additionally, the concept can be extended to matching of line segments in the future.

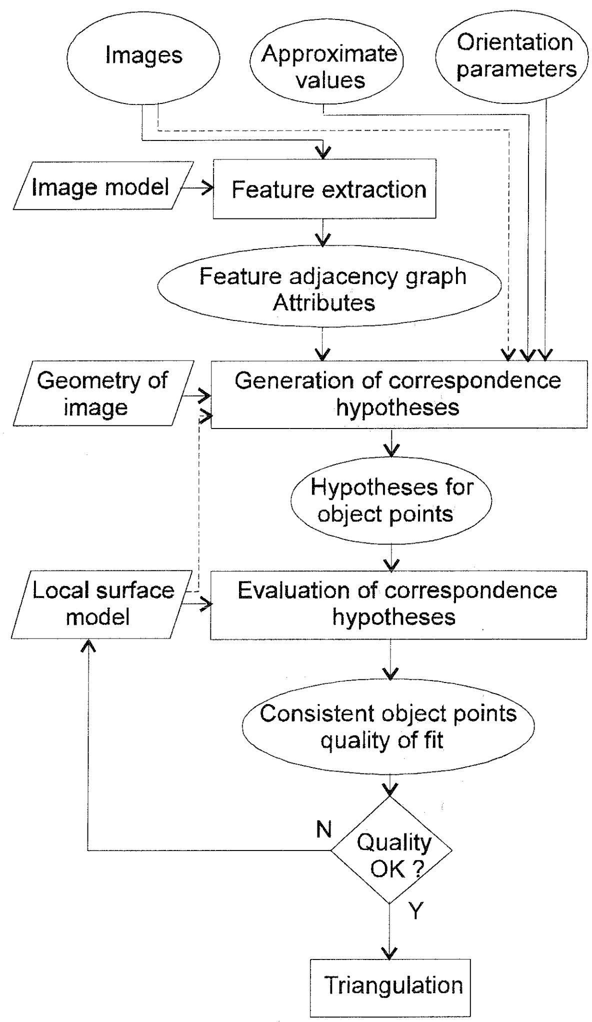


Figure 3: Matching algorithm

Having detected point features in two or more images, correspondencies between homologous features from different images have to be found (section 3.4). Finding such correspondencies comprises two steps (Gülch, 1994):

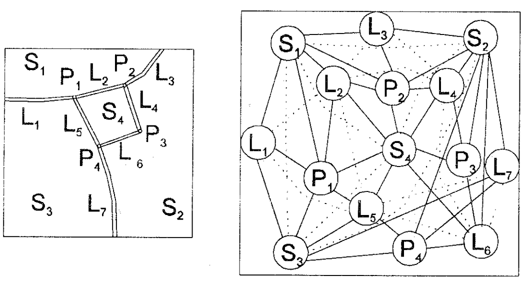
- The generation of correspondence hypotheses which makes use of approximate values and the orientation parameters under the assumption of a model of image geometry.
- The evaluation of these hypotheses under the assumption of some (pre-defined) local surface model in object space. Only hypotheses consistent with the model will be accepted.

The result of matching is a list of points which are consistent with the object model and quality measures of the fit between data and model. A final quality check shall make sure that an adequate model has been used. A bad fit of the measured points to the object model might indicate that a wrong object model was used. In this case, another model should be selected. Both generation and evaluation of correspondence hypotheses are the main topics of research work in our concept.

3.2 Feature Extraction

Many feature based matching algorithms for photogrammetric surface reconstruction use the Förstner operator to extract distinct points from digital images (Krzystek, 1995). On a symbolic level, the image is then described by an unstructured cluster of such point features. Evidently, a considerable amount of information contained in the images is thrown away. We think that this information, e.g. line information, but also information about the mutual relations between the extracted features, should be used in order to increase the reliability of a matching algorithm.

We thus want to use the more complex image model proposed in (Fuchs et al., 1995) which was originally designed for automatic building extraction. In this model, the ideal image is assumed to be composed of homogeneous segments, piecewise smooth boundary lines of these segments and points. The digital image is a blurred and sampled version of the ideal image which is additionally afflicted by noise. Thus one can no longer speak of finding lines and points in the image, but more reasonably of regions containing line segments or points (figure 4a); (Fuchs et al., 1995).



a b

Figure 4: Image regions (a) and region adjacency graph with direct (full lines) and indirect (dotted lines) neighbourhood relations (b); adopted from (Fuchs et al, 1995)

Each pixel can be classified as belonging either to a homogeneous region S, to a region P containig a point or to a region L containing a line (figure 4a) using a measure for homogeneity and a measure for isotropy of texture, both of which can be derived from a local function of the grey levels. From a segmentation of the classified image, all regions are extracted and a region adjacency graph is created which describes the topological relations between neighbouring regions (figure 4b); (Fuchs et al., 1995).

At the same time as the region adjacency graph is created, attributes can be assigned to the extracted regions such as the subpixel position of points, the average grey level of homogeneous regions, curve parameters, e.g. spline coefficients, 'for lines, etc. These attributes will become very important for the creation of correspondence hypotheses. As the subpixel estimates for point coordinates (and, perhaps, in a future step, curve parameters of lines) will be essential for the description of the object, we refer to the region adjacency graph as 'feature adjacency graph' although it also contains homogeneous regions.

3.3 Image Geometry

Many matching algorithms use epipolar images, e.g. (Gülch, 1994); (Krzystek, 1995). In case epipolar images are used, the matching problem can be reduced to a one-dimensional problem. This strategy considerably reduces the complexity of the matching algorithm. However, in our concept we do not use epipolar images for mainly three reasons:

- By using epipolar images, one is restricted to using stereo image pairs for matching. We eventually want to use more than two images for that purpose.
- Small errors in the orientation parameters of the images might deteriorate the matching result, especially when lines which are almost parallel to the epipolar lines are used.
- Epipolar images are derived from the original images by resampling. Feature extraction may be influenced by the lowpass characteristics of resampling methods.

Instead of using epipolar images, we will rely on bundle block geometry. The bundle block adjustment system ORIENT (Kager, 1989) will be integrated into the matching software to be developed. Bundle block geometry will be used to establish geometrical constraints as well as for the formulation of models for the local object surface.

3.4 Generation and Evaluation of Correspondence Hypotheses

The generation of hypotheses for the correspondence of features from different images is based on some measure of similarity between these features. This similarity measure is based on the comparison of the feature attributes which have been extracted. If the viewing directions are nearly parallel, the correlation coefficient of the grey levels in a small region surrounding the point can additionally be used as a similarity measure. We also use the feature adjacency graph for that purpose because we assume that a correspondence between features from different images is more likely if the neighbouring image regions also show similar attributes (Zhang et al, 1992). Up to now, we have not yet decided which feature attributes will be used and how the similarity measure shall be composed. These questions are among the main topics of our research.

By just using similarity as a criterion for the generation of hypotheses, one would get too many wrong hypotheses. The number of initial hypotheses is reduced in two ways:

- Introduction of geometrical constraints: Only features with image residuals smaller than a certain threshold may correspond to the same object point.
- Reduction of search space by approximate values. The algorithm is successively applied to relatively small homologous image patches which are extracted according to the approximate values. Additionally, thresholds for local height differences can be introduced.

Matching of relational descriptions leads to search trees (Vosselman, 1995). Finding the optimum matching result corresponds to finding an optimal path through that search tree, assuming that a cost function is associated to each leaf of that search tree. The trees involved in image matching can become very extensive even for two images. Things even get worse when more images are used. This is the reason why the number of possible correspondencies has to be reduced considerably. Still, no complete search tree using topology in more than two images shall be generated due to the high computational cost of that method (Vosselman, 1995).

In a first step, hypotheses of correspondence between features of all possible image pairs are generated. After that, the hypotheses of all image pairs have to be combined in order to consider all images at one time. (Tsingas, 1992) gives a graph - theoretical approach for the detection of hypotheses in more than two images which makes use of heuristic search tree methods. This approach is designed for aerial triangulation, where no orientation parameters are available. The method might become easier and faster if these parameters are assumed to be known, as it happens with our application.

A correspondence hypothesis between features in image space leads to a hypothesis for a surface point in object space. Many matching algorithms evaluate correspondence hypotheses by assuming the object surface to be smooth and eliminate hypotheses which contradict to that model by a robust estimation technique , e.g. (Krzystek, 1995). We will also assume the object surface to be smooth in a first step. If the image data do not fit that model, another surface model should be assumed. This means that a knowledge base of different object models which can be formulated in object space has to be developed. However, the assumption of another object model might allow different possibilities of correspondence between features of different images so that the generation step might have to be repeated (indicated by the broken line in figure 3). Again, we want to use topology to generate more complex object models. Up to now, first tests regarding the formulation of rather simple object models have been made using the program system ORIENT. However, this point remains the probably most important one for research in our concept.

4. CONCLUSION

Our concept for 3D object reconstruction aims at developing a feature based matching using topology. More complex object models should be formulated in 3D object space, which might give us the possibility to overcome the problem of occlusions. The geometrical constraints necessary to reduce the computational complexity of matching are provided by the integration of the bundle block adjustment system ORIENT. Basic modules, e.g. the data interface to ORIENT and the creation and handling of image pyramids have already been implemented in C++. Feature extraction is in the implementation phase, and first tests regarding the formulation of object models have been run. The whole development is closely connected to the new SCOP environment (Molnar et al., 1996).

5. REFERENCES

Ackermann, F., 1994: High Precision Digital Image Correlation. In: 39th Photogrammetric Week. Schriftenreihe des Instituts für Photogrammetrie, Heft 9, Universität Stuttgart, 1984.

Förstner, W., 1986: A Feature Based Correspondence Algorithm for Image Matching. Int. Arch. of P. & RS., Vol. XXVI, P. 3/3, Com. III, Rovaniemi, Finland, 1986.

Förstner, W., 1995: Matching Strategies for Point Transfer. In: Fritsch, Hobbie, (Eds.): Photogrammetric Week '95. Herbert Wichmann Verlag, Heidelberg, 1995.

Fuchs, C., Lang, F., Förstner, W.,1994: On the Noise and Scale Behaviour of Relational Descriptions. Int. Arch. of P. & RS., Vol. XXX, P. 3/1, Com. III, München, 1994.

Fuchs, C., Förstner, W., 1995: Polymorphic Grouping for Image Segmentation. In: Grimson, E. (ed): Proc. 5th ICCV '95, IEEE Computer Society Press, 1995.

Gülch, E., 1994: Erzeugung digitaler Geländemodelle durch automatische Bildzuordnung. Dissertation at Stuttgart University. Deutsche Geodätische Kommission, Heft Nr. 418, München, 1994.

Heitzinger, D., 1996: 3D-Oberflächenmodellierung mit topologischen Grundelementen. Diploma Thesis at the Institute of Photogrammetry and Remote Sensing at TU Vienna, 1996.

Kager, H., 1989: ORIENT: A Universal Photogrammetric Adjustment System. In: Gruen, Kahmen (eds.): Optical 3-D Measurement Techniques. Herbert Wichmann Verlag, Karlsruhe, 1989

Krzystek, P., 1995: Generation of Digital Elevation Systems. In: Second Course in Digital Photogrammetry. Notes of that course held at the Institute of Photogrammetry at Bonn University, Feb. 6 - 10, 1995

Lang, F., Förstner, W., 1995: Matching Techniques. In: Second Course in Digital Photogrammetry. Notes of that course held at the Institute of Photogrammetry at Bonn University, February 6 - 10, 1995

Molnar, L., Wintner, J., Wöhrer, B.: DTM System SCOP in a New Technological Generation. Int. Arch. of P. & RS., Vol. XXXI, Com. IV, WG 4, Vienna, 1996

Tsingas, V., 1992: Automatisierung der Punktübertragung in der Aerotriangulation durch mehrfache digitale Bildzuordnung. Dissertation at Stuttgart University. Deutsche Geodätische Kommission, Heft Nr. 392, München, 1992

Vosselman, G., 1995: Applications of tree search methods in digital photogrammetry. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 50, Number 4, pp 29 - 37, Aug. 1995.

Zhang, Z., Zhang, J., Wu, X., Zhang, H., 1992: Global Image Matching with Relaxation Method. In: Proceedings of the International Colloquium on Photogrammetry, Remote Sensing and GIS, Wuhan, China, 1992, pp. 175-188.