Potential and limitation for the 3D documentation of cultural heritage from a single image

André Streilein and Frank A. van den Heuvel

Faculty of Civil Engineering and Geosciences
Section of Photogrammetry and Remote Sensing
Delft University of Technology
Thijsseweg 11, 2629JA Delft
The Netherlands
Phone: +31-15-2781737 / 7609
Fax: +31-15-27842745
E-mail: {a.streilein,f.a.vandenheuvel}@geo.tudelft.nl

Abstract:

Photogrammetry has many advantages as a technique for the acquisition of three-dimensional information of the cultural heritage. However, the photogrammetric process to extract 3D geometric information from multiple images is often considered to be too labor-intensive and complicated. As a result of this consideration, existing images are often not used for a photogrammetric documentation of the object or the accomplishment of a photogrammetric project is not taken into account at all.

In order to re-emphasize the value of single images for the documentation of the cultural heritage this paper shows on some practical examples the potential and limitations of different methods and techniques for the geometric reconstruction of objects from a single image. A general overview of such methods is given. Attention is paid especially to the 3D reconstruction from a single image and the role of a priori object information. Different software packages that are able to handle such tasks are mentioned exemplary.

Keywords:

Digital photogrammetry, line photogrammetry, object reconstruction, single image

1. Introduction

Photogrammetry has many advantages as a technique for the acquisition of three-dimensional information of the cultural heritage. However, the photogrammetric process to extract 3D geometric information from multiple images is often considered to be too labour-intensive and complicated. As a result existing images are often not used for a photogrammetric documentation of the object or the accomplishment of a photogrammetric project is not taken into account at all.

As a discipline, architectural photogrammetry is currently undergoing profound changes. New technologies and techniques for data acquisition (CCD cameras, Photo-CD, photoscanners), data processing (computer vision), structuring and representation (CAD; simulation, animation, visualization) and archiving, retrieval and analysis (spatial information systems) are leading to novel systems, processing methods and results.

In order to re-emphasise the value of single images for the documentation of the cultural heritage this paper shows the potential and limitations of different methods and techniques for the geometric reconstruction of objects from single images. The paper gives a general overview of the methods applicable to derive metric information from a single image and concentrates on some exemplary methods for the reconstruction of 2D and 3D information from a single image. For the reconstruction of 3D information emphasis is placed on the role of a priori object information

In photogrammetry it is common to derive 3D information from measurements in multiple images. For many applications a stereo approach is chosen. Archaeology and architecture are examples of such applications. In that case the collected information is better characterized as 2½D instead of 3D, because the object is mapped relative to a reference plane approximately perpendicular to the two optical axes of the stereo pair.

For a full 3D all-around object model multiple images or multiple stereo pairs are required (Waldhäusl and Ogleby, 1994). For this conventional approach some object information is needed beforehand. The first reason is the establishment of an object co-ordinate system. Improvement of the quality of the model is the second reason. Usually this object information consists of 3D co-ordinates of some reference points and distance measures between two points. 3D reconstruction from a single image is only possible if additional object information is available. The main type of a priori information needed for 3D reconstruction is parallelism of (straight) object edges. Straight and parallel object edges are frequently present in man-made structures, and buildings in particular.

2. General methods

In principle the reconstruction of a 3D object from a single 2D image is an ill-posed problem. A monocular image alone does not contain sufficient information to uniquely retrieve 3D information. For example, a point in object space is determined by three cartesian co-ordinates (X, Y, Z), but from a point in the image just two observations (image co-ordinates x, y) can be derived.

The third dimension however can be recovered from monocular images in conjunction with certain visual cues (e.g. size, shade, distortion, vanishing points) or prior knowledge of certain geometric properties of the object. For example, a 3D point can be reconstruct if a constraint is available, e.g. the point is located on a known plane.

The techniques to recover the object in three dimensions from a single image are often referred to as "shape from X", such as shape from shading, shape from texture, shape from focus or shape from geometry (Jain et al., 1995).

The **shape-from-shading** technique infers the depth of an image pixel based on its shade (or intensity). The underlying theory of the technique is that the intensity of a pixel is determined in part by the angle between the surface normal and the illumination direction. Limitations of this technique are the assumption of a point source, of a Lambertian surface and of weak perspective projection.

The **shape-from-focus** technique infers the depth from a focal gradient. Due to the finite depth of field of optical systems, only objects which are at a proper distance are focussed in the image whereas those at other depths are blurred in proportion to their distance. The depth is then recovered by estimating the amount of blur in the image. The technique is limited in precision, and cameras with large depth of field yield less accurate depth estimation.

The **shape-from-texture** technique uses cues from the image plane variations in the texture properties such as density, size and orientation. For example, the texture gradient, defined as the magnitude and direction of maximum change in the primitive size of the texture elements, determines the orientation of the surface. From images of surfaces with textures made up of regular grids of lines, possibly due to structured lighting, orientation may be uniquely determined by finding the vanishing points. These methods suffer from difficulties in accurately locating and quantifying texture primitives and their properties.

The **shape-from-geometry** techniques, also referred to as "shape from inverse perspective projection", reconstruct a 3D geometric entity from a single image in conjunction with geometric constraints on the 3D geometric entity. The geometric constraints employed include euclidian distance constraints and orientation constraints. These techniques are typically able to produce a model of the object with high accuracy.

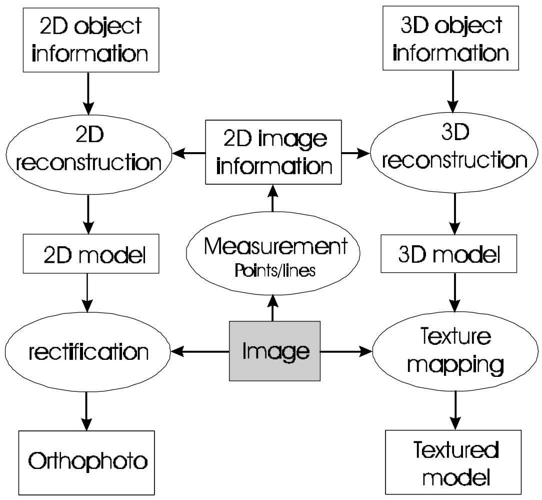

Figure 1: Overview of methods to derive 2D and 3D information from a single image

The general possibilities given for the reconstruction of 2D and 3D object information from a single image by shape-from-geometry techniques are outlined in Figure 1.

Starting with a single image, two-dimensional image information can be retrieved by measurements of image points and/or image lines either manually, semi-automatically or automatically.

In collaboration with known two-dimensional object information (e.g. co-ordinates of points, assumptions of plane surfaces, etc.) the 2D reconstruction leads to a 2D model of the object. This is shown on the left side of Figure 1. The same holds in three dimensions with given 3D object information (e.g. point co-ordinates in 3D, parallelism of object edges) for the three-dimensional reconstruction, as shown on the right side of Figure 1. It should be mentioned here that in most cases of the 2D reconstruction the camera parameters (parameters of interior and exterior orientation) are not needed explicitly as they are calculated implicitly during the reconstruction process. For the 3D reconstruction process from one image the camera parameters can be estimated if sufficient object information (parallelism and perpendicularity constraints) is available. The orientation of the image relative to the object can be a limiting factor also (van den Heuvel, 1999). The availability of the parameters of the interior orientation of the camera improves the possibilities, and the quality of the 3D reconstruction from a single image.

From the 2D model of the object an orthophoto can be derived by applying a resampling procedure onto the original image with the transformation parameters of the 2D reconstruction. Often the 2D reconstruction and the resampling procedure are combined in one step, which is known as digital rectification.

The same procedure for rectification can be applied to each face of a 3D model. Here the resampling process is usually known as texture mapping and results in a textured 3D model.

3. 2D information from a single image

The reconstruction of 2D information from a single image usually requests the knowledge of 2D metric information about the object. The methods described in this section have in common that they mainly intend to produce different types of ortho-photographies in order to provide the image content with a unique scale. Depending on the type of object information available (i.e. description or assumption of the object itself, known co-ordinates of reference points) different methods can be distinguished (see Table 1).

Object information	*Method*
Planar object 4 control points	Projective transformation
Piecewise planar object 4 control points per plane	Combination of projective transformations
Any object Dense grid of control points	Non-parametric rectification
Mathematically definable object	Parametric rectification
Digital Surface Model (DSM)	Differential rectification

Table 1: Methods for the reconstruction of 2D information from a single image.

It is worthwhile to mention that, despite of the last two methods, the methods for the reconstruction of 2D information from a single image in general do not depend on available information of the camera parameters. This makes them usually easy to handle and applicable to many ‘unusual’ tasks in heritage documentation. In the following some examples for the methods shown in Table 1 are explained in more detail.

The **projective transformation** is probably the most known and established technique to produce in an easy and fast way a map of a (planar) object or terrain. Planar in this context means that the object can still have a small variation in depth compared to the object distance from the camera. For a projective transformation of a building facade deviations from a plane of about 30cm can be considered as acceptable (Jänsch, 1976). The main advantage of this method is that neither the camera parameters nor the camera or the type of camera have to be known. Only four control points on the facade or terrain have to be known in two dimensions. Therefore this method is very convenient in producing a map from a single image where no other information is available, especially not about the camera. However, it should be mentioned that for obvious reasons this method is not able to compensate any lens distortions of the used camera. Hence, while applying this technique to an image, the user should avoid using a camera with large lens distortions (e.g. with wide angle or fisheye lenses). The method of projective transformation is well known and distributed and found already its way into various software packages for image processing and Computer Aided Design.

A practical example for the projective transformation of an image with an unknown camera is given in Figure 2. The task was to generate a map of the front facade of the Basilica di San Salvatore in Spoleto, which shows the situation in the year 1920. This was intended for a comparative study of the situation in the year 1920 and the actual situation. Hence, a projective transformation of the actual facade and an image showing the facade in the year 1920 was performed. Both transformations were based on the two-dimensional co-ordinates of four corner stones on the facade. Figure 2 shows the original scanned image from 1920 on the left and the rectified ortho-photo on the right.

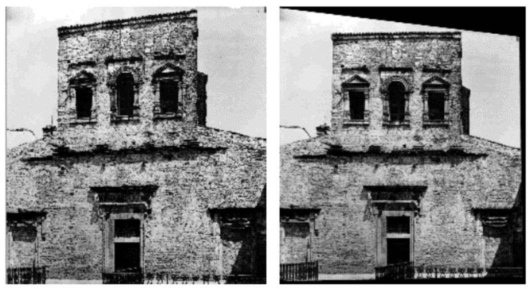

Figure 2: Basilica di San Salvatore, Spoleto, Italy. Original image (left) and rectified image (right)

In many cases an object or a complex facade, which can not be considered as one plane, can be subdivided into several planar facades. These single planar facades can then be transformed separately and combined later into one ortho-photo. This **combination of projective transformations** is a very valuable tool and often used in order to derive an accurate and fast result for the reconstruction of a more complex facade. An example for this procedure is given in Figure 4, where the images

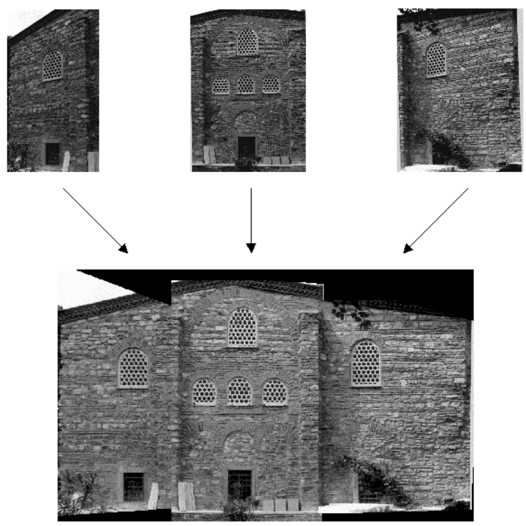

Figure 3: Attik Mustafa Camii (Istabul); example for the

of three adjacent facades are combined into one orthophoto.

The **non-parametric transformation** might be not of high practical relevance for heritage documentation as this method requires a dense set of known control points in order to describe the polynomial transformation between the object and the image. And in heritage documentation the geometry of an object has usually to be considered as unknown. However, due to the dense set of control points and the type of transformation, this method does not require any camera parameters. And depending on the known data of the object, this method might be of use in some projects. An example for the application of this technique is given by (Marten et al., 1994).

The **parametric rectification** of a mathematically definable object is well known in the photogrammetric community. It is always of interest if an object can be described in a mathematical form, which allows an simple projection on a defined plane (e.g. cylinder or dome). This holds especially for mathematical settlements, which are often used in heritage documentation (Karras et al., 1996). However, for each mathematical form a specific solution has to be developed. Hence this method is only applicable for special cases.

The **differential rectification** or classical orthophoto production is a well-known standard procedure in aerial photogrammetry. In order to apply this method to a single image the parameters of the interior and exterior orientation of the camera as well as the underlying digital terrain model (DTM) have to be known. As this data usually comes with the triangulation process of an aerial photogrammetric project, this method is especially of interest for the field of archaeological studies. Examples for the use of low-altitude aerial orthophoto production in for archaeological purposes are given in (Doneus, 1996). But not only for aerial photogrammetry also for close-range applications, like in architectural photogrammetry, this technique is a valuable source for information (Wiedemann, 1997).

4. 3D information from a single image

The basics in three steps

In this section a short overview is given of how three-dimensional information is inferred from a single image and a priori object information. The three steps of the 3D reconstruction from a single image are:

- Find orientations of object edges (and faces) using parallelism and perpendicularity.
- Model construction: position edges and faces (oriented in step 1) relative to each other by using common points.
- Position and orient the model by using one or more known point positions, and/or known orientations of edges and faces ("absolute orientation").

The second and third step of the procedure can sometimes not be separated clearly. The first step is crucial and will be explained in more detail. The central projection from object to image (all light rays travel through the optical centre of the lens), is known to preserve straightness of lines, if lens distortion is absent. Furthermore, the projections of edges that are parallel are known to intersect in one point in the image plane. This point is called the vanishing point (see Figure 4). Several procedures have been developed especially for automatic vanishing point detection in images of buildings (van den Heuvel, 1998a) (Shufelt, 1996). The line through the vanishing point and the projection

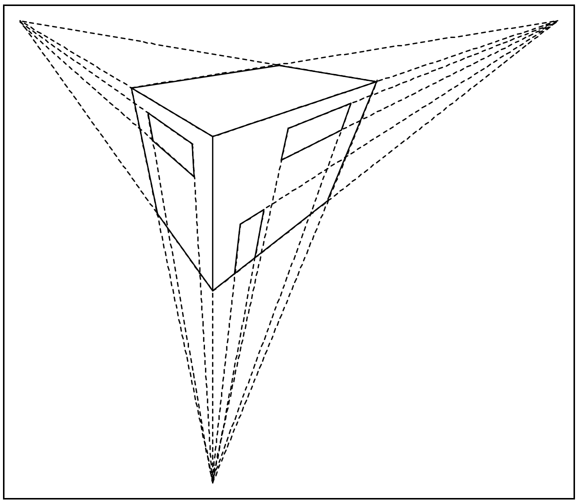

Figure 4: Parallel edges and related vanishing points

centre has the same orientations as the related parallel edges of the object. The edges are in (and border) object faces and thus the orientation of a face can be derived from at least two of its edges with different orientation. Perpendicularity information allows finding orientations in case of insufficient parallel edges, and it improves the accuracy of the derived orientations.

Example

Figure 5 shows the image and the line measurements that were performed manually. For the example presented here, parallelism and perpendicularity was specified for the edges of the three major object orientations. The resulting geometric constraints were processed in a least-squares adjustment, resulting in adjusted image line observations. These adjusted image lines intersect at their vanishing point. From each vanishing point an object orientation of the related edges was computed. Then step 2 of the procedure was executed. Two views of the texture mapped model are depicted in Figure 6. The model is available in VRML-format on the Internet (www.geo.tudelft.nl/frs/architec/single.html).

RETURN

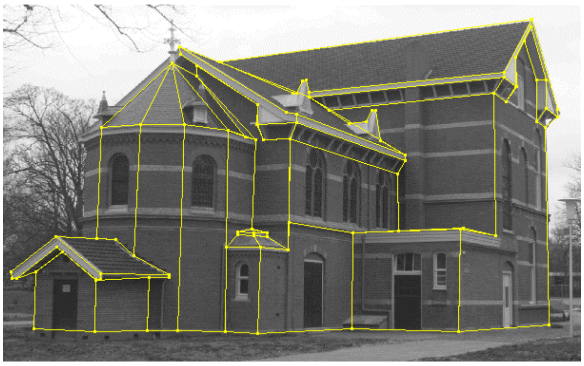

Figure 5: The line mesurements used for 3D reconstruction

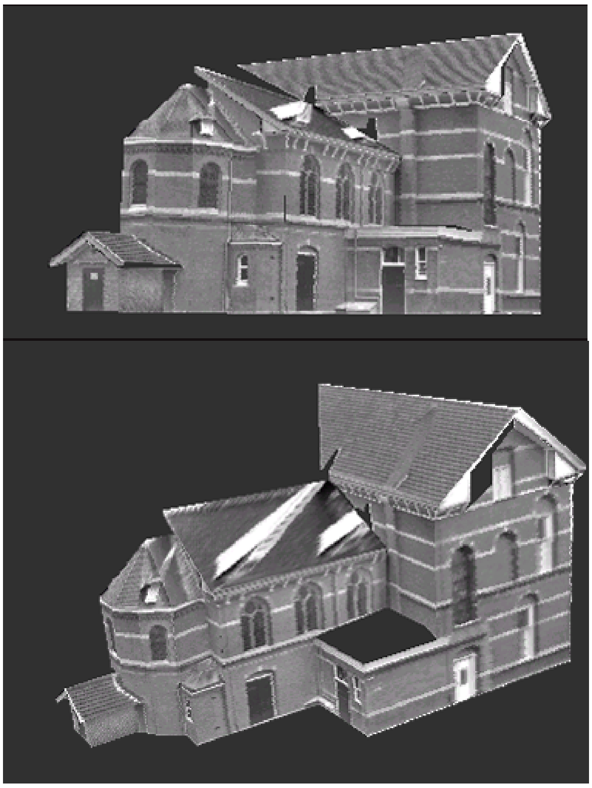

Figure 6: Two views of the 3D reconstruction

Reconstruction from a single image: not a common task

In contrast to 3D reconstruction from multiple images (and in contrast to 2D reconstruction from a single image discussed previously), there is not a big variety of software packages available for the 3D reconstruction from a single image. It has been a subject of research for many decades, and it still is (Braun, 1994) (van den Heuvel, 1998b).

Since a few years, commercial software is available in which techniques for monocular reconstruction are applied.

3D Builder is such a package that is primarily designed for processing multiple images, but can work with a single image also (3D Builder, 1999). It allows the use of parallelism and perpendicularity of object edges. The lines in the image are measured by marking building corner points in the image. Four lines can be specified to form a rectangle. In this way parallelism and perpendicularity information is supplied. Furthermore, lines can be specified to have a principle object orientation (X, Y, or Z), and the same holds for faces (XY, YZ, XZ).

The PhotoModeler software supports reconstruction from a single image only if the object plane is known relative to the camera (PhotoModeler, 1999). To achieve this, a number of control point co-ordinates is required. Then the rays, defined by point measurements in the image, are intersected with the object plane to calculate 3D co-ordinates.

Relatively new is the package called Canoma (Canoma, 1999). It is different from the other packages in the sense that it works with so-called primitives. Primitives are simple shapes that are defined by a limited number of parameters. A box is such a primitive that is defined by three parameters (the lengths of the three sides), and contains a lot of parallelism and perpendicularity relations between the edges. Instead of directly measuring points in the image, the operator drags corner points of a back-projection of a primitive to corresponding points in the image. The model is constructed from multiple primitives of whom the relative position and orientation can be restricted.

All these commercial software packages allow reconstructing a 3D model from one or more images. It has to be noted that they generally act as “black boxes”, and do not supply detailed information on the accuracy of the resulting model. Furthermore, statistical testing and blunder detection is not very advanced, if present at all.

A non-commercial packages for 3D reconstruction from a single image that uses the same principles as 3D Builder is SolidFit (SolidFit, 1999).

What about camera calibration?

If sufficient object information is available, and the image orientation relative to the object is not unfavourable, all software packages mentioned can derive the focal length using a single image. PhotoModeler and 3D Builder allow the use of the results of a separate (multi-image) camera calibration procedure that results in the estimation of the focal length, the location of the principal point, and lens distortion parameters. In principle, all these parameters can also be estimated from a single image if sufficient lines and constraints are available (van den Heuvel, 1999). However, a number of limiting factors can reduce the precision of the estimated parameters considerably.

5. Conclusions

This paper has shown on a few examples the potential of the 3D reconstruction for cultural heritage documentation from a single image. A variety of different tools and methods is available for use. This is especially of interest if just one image of an object and no information about

RETURN

the camera is available. A priori knowledge about the object, which is needed in different ways for the different methods, is often easy to gain (assumptions, simple measurements). However, it has to be mentioned very clearly that there are also limitations in using just one single image.

A major limitation of reconstruction from a single image is the incompleteness of the 3D object model. Obviously, because there is no information about the backside of the object. Furthermore, often object parts (other than the backside) are occluded. To arrive at a more complete model, one can combine reconstructions of several images.

A second limitation for monocular 3D reconstruction is the need for additional object information. This information is usually available in the application area of architecture. In archaeology however, it is not likely that 3D reconstruction from a single image can be applied. In spite of these clear limitation there is a sustained interest in 3D reconstruction from a single image. In case of demolished historic buildings for instance, sometimes not more than a single image is available. Furthermore, techniques developed for monocular reconstruction reduce the number of images required to obtain a complete model.

6. References

3D Builder, 1999, URL: www.3dconstruction.com/

Braun,C., 1994. Interpretation von Einzelbildern zur Gebäuderekonstruktion, Druckerei Schwarzbold, Witterschlick, Germany

Canoma, 1999, URL: www.metacreations.com/products/canoma/

Doneus, M.: Photogrammetrical Applications to Aerial Archaeology at the Institute for Prehistory of the University of Vienna, Austria. IAPRS Volume XXXI, Part B5, 1996, 124-131.

Jain, R., Kasturi, R., Schunck, B.G., 1995. Machine Vision. McGraw-Hill. Inc. New York, 1995. 542 pages.

Jansch, R.D., 1976. Genauigkeitsansprüche bei der photogrammetrischen Bauaufnahme. Arbeitsheft 17, Landeskonservator Rheinland, Reihnland-Verlag Köln, 1976, pp. 103-106.

Karras, G., Patias, P., Petsa, E., 1996. Digital Monoplotting and Photo-Unwarping of developable surface in Architectural Photogrammetry. IAPRS, Vol. XXXI, Part B5, 1996, pp. 290-294.

Marten, W., Mauelshagen, L., Pallaske, R., 1994. Digital ortho-image system for architecture representation. IAPRS Vol. 49 Part 5, 1994. Pp. 16-22.

PhotoModeler, 1999, URL: www.photomodeler.com/

Shufelt, J.A., 1996, Performance Evaluation and Analysis of Vanishing Point Detection Techniques, ARPA Image Understanding Workshop, Morgan Kaufmann Publishers, Palm Springs, pp.1113-1132

SolidFit, 1999, URL: www.cssip.edu.au/~vision/solidFit.html

van den Heuvel, F.A., 1998a, Vanishing point detection for architectural photogrammetry, IAPRS, Vol. 32 part 5, pp. 652-659

van den Heuvel, F.A., 1998b, 3D reconstruction from a single image using geometric constraints, ISPRS Journal of Photogrammetry and Remote Sensing, Elsevier, Vol. 53 (6), pp. 354-368

van den Heuvel, F.A., 1999, Estimation of interior orientation parameters from constraints on line measurements in a single image, IAPRS, Vol. 32, Part 5W11, pp. 81-88

Waldhäusl, P.; Ogleby, C.L.,1994. 3 x 3 rules for simple photogrammetric documentation of architecture,Proceedings. ISPRS Commission V Intercongress Symposium Melbourne, pp.426-429

Wiedemann, A., 1997. Orthophototechnik in der Architekturphotogrammetrie – Möglichkeiten und Grenzen. Architekturphotogrammetrie gestern-morgen-heute. TU Berlin, 1997, pp.. 79-94.