NRC·CNRC

NRC Publications Archive
Archives des publications du CNRC

A Practical Approach To Creating Precise And Detailed 3d Models From Single And Multiple Views

El-Hakim, Sabry

This publication could be one of several versions: author’s original, accepted manuscript or the publisher’s version. / La version de cette publication peut être l’une des suivantes : la version prépublication de l’auteur, la version acceptée du manuscrit ou la version de l’éditeur.

NRC Publications Record / Notice d'Archives des publications de CNRC:
https://nrc-publications.canada.ca/eng/view/object/?id=654813a9-7747-406d-83f5-172410790be5
https://publications-cnrc.canada.ca/fra/voir/objet/?id=654813a9-7747-406d-83f5-172410790be5

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at
https://nrc-publications.canada.ca/eng/copyright
READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L’accès à ce site Web et l’utilisation de son contenu sont assujettis aux conditions présentées dans le site
https://publications-cnrc.canada.ca/fra/droits
LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D’UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n’arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

National Research Council Canada
Conseil national de recherches Canada

Canada

*A Practical Approach to Creating Precise and Detailed 3D Models from Single and Multiple Views**

S.F. El-Hakim
July 2000

***published in** Proceedings of the XIX ISPRS Congress, Amsterdam, Holland. July 16-23, 2000. Volume XXXIII, Book 5A, pp. 202-209. NRC 43655.

IAPRS, Vol. XXXIII, Amsterdam, Book 5A, pp 202-209, 2000

A PRACTICAL APPROACH TO CREATING PRECISE AND DETAILED 3D MODELS FROM SINGLE AND MULTIPLE VIEWS

Sabry EL-HAKIM
Visual Information Technology (VIT) Group
Institute For Information Technology, National Research Council Canada
Sabry.El-Hakim@nrc.ca

Working Group V/3

KEY WORDS: Image registration, Modeling, Multi-sensor, Scanner, Segmentation, Surface Reconstruction, VR

ABSTRACT

This paper describes a general system for creating geometrically correct and complete 3D models of objects and environments. The main process uses digital images and applies rigorous photogrammetric techniques. Once the images are registered using bundle adjustment, points in selected locations are measured in all the images where they appear. However, there are usually many parts of the scene that appear in only one image due to occlusions or lack of features. These parts can be reconstructed in the 3D space from the coordinates in a single image and the mathematical model of the surface determined by fitting a function to existing surface points. This produces a sampled geometry in the form of points in the three-dimensional space however the connectivity or the topology is not known and must be determined someway. The proposed approach relies on interactive point segmentation and automatic triangulation. Counting on images for modeling is limited because the features that can be extracted are usually fewer then the required level of details. To overcome this problem two options are implemented. First, large number of points can be automatically added to surfaces of known shape, such as spheres, cylinders and quadrics, using a polygon subdivision method. The second option integrates data from range sensors that can densely digitize the surface where filled out details are required. The presented approach will be assessed by several examples representing various types of objects and sites.

1 INTRODUCTION

The computer reconstruction of objects and environments has a wide range of applications. Manufacturing and virtual prototyping [Gomes de Sa et al], reverse engineering [Thompson et al, 1999], urban design and analysis [Teller, 1998], architecture [Liebowitz et al, 1999], and cultural heritage [Sablatnig et al, 1996], are just few examples. Methods to acquire the 3D data and reconstruct shapes have been progressing rapidly in recent years. However, many problems remain to be solved and no approach is suited for all applications and all types of object and environment. This paper presents an approach that aims to be limitless. It is semi-automated, can generate 3D data from multiple or single images, can integrate data from active range sensors, and can automatically add points on surfaces of known shapes. The resulting models are accurate, complete, easily created, and not restricted to specific application.

In this introduction the state of the technology is summarized followed by an overall description of our approach. In the second part of the paper details of the approach are given. In the third part an evaluation of the geometric accuracy and the ability to model details using various types of data is presented. Finally, some conclusions and future research directions are outlined.

1.1 The Current 3D Digitizing and Modeling Techniques

Techniques for 3D digitizing and modeling have been rapidly advancing over the past few years. The ability to capture details and the degree of automation vary widely from one approach to another. One can safely say that there is no single approach that works for all types of environment and at the same time is fully automated and satisfies the requirements of every application.

The process of creating 3D models from real scenes has a few well-known steps [figure 1]: data collection, data registration, and modeling (geometry, texture, and lighting). There are many variations within each step, some of which are listed in figure 1. Approaches that skip the geometric modeling step also exist. For example, panoramas [Szeliski et al, 1997 and Chen, 1995], light fields [Levoy et al, 1996], and image-based rendering (images plus depth) [Kang, 1999], are popular for applications that require only visualization or limited walkthrough. However, the lack of geometric model impedes the accuracy and the freedom to render the environment from arbitrary viewpoints.

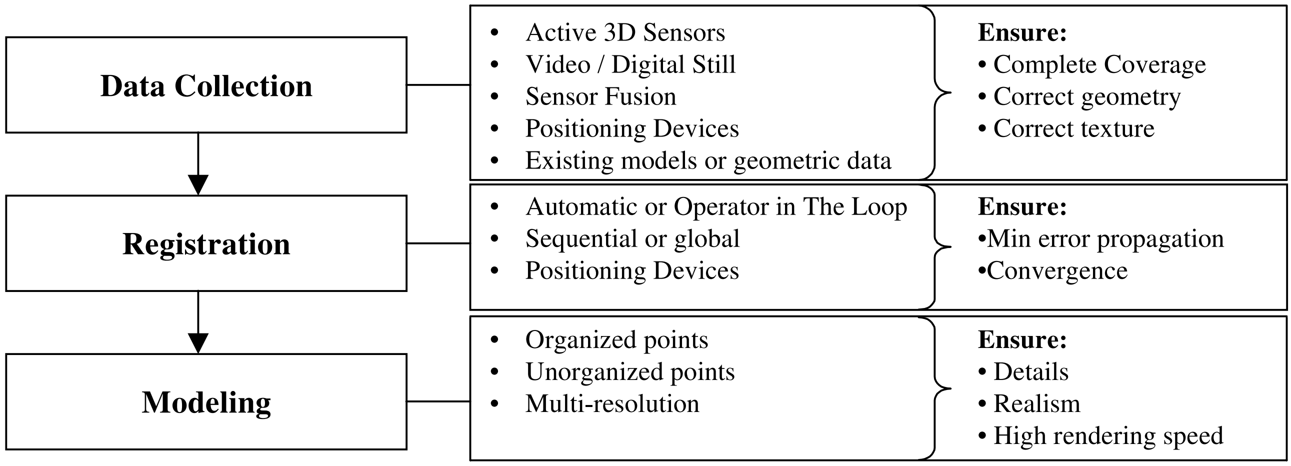


Figure 1: The main steps for creating 3D models from real scenes.

Passive image-based methods, mostly based on photogrammetry, have been developed for specific applications such as architecture [Debevec et al, 1996]. Those needing user interaction have matured to a level where commercial software is now available (e.g. Photomodeler [http://photomodeler.com] and ShapeCapture [http://www.shapequest.com]). Approaches that automatically acquire 3D points from a sequence of images at unknown locations, using projective reconstruction, are available [Faugeras, 1992 and Polleyfeys et al, 1999]. These methods require images taken close to each other (short baseline) in order for the automatic correspondence to work. This makes them more noise sensitive and numerically unstable. However, they can be useful for applications that do not require high geometric accuracy but need realistic looking and easy to create models. Techniques based on a single image plus constraints were developed for specific objects such as buildings [Liebowitz et al, 1999 and van den Heuvel, 1998]. The Main Advantage of all passive image-based methods is that the sensors are inexpensive and very portable. However, due to the need for features, incomplete models may result particularly on irregular geometry or sculptured surfaces.

Active sensors (e.g. laser scanners) have the advantage of acquiring dense 3D points automatically [Beraldin et al, 1999 and Sequeira et al, 1999]. They also produce organized points suitable for automatic modeling [Soucy et al, 1996]. However, the sensors can be costly, bulky, and affected by surface reflective properties. Also a range sensor is usually designed for a specific range, thus a sensor designed for a close range can not be used for a long range. Active methods can be combined with passive methods to take advantage of the strength of each [El-Hakim et al, 1998].

For relatively simple objects, structures, or environments, most existing methods will work at varying degrees of automation, level of details, effort, cost, and accuracy. Many researchers have presented examples of those types of model in the past 5 years. However, when it comes to complex environments the only proven methods so far are those using positioning devices, CAD or existing models and operator in the loop. Limited examples are available, most in urban or city modeling [Coorg et al, 1999 and Gruen et al, 1998] and as-built factory models [Chapman et al, 1998].

1.2 Summary Of The Approach

The proposed approach attempts to integrate several techniques, each of which individually may not be general enough, into a coherent system that has a wider range of applications. It is based on the following key ideas:

1. An easy to use interactive system can be more effective than the existing automatic methods. Some aspects of 3D modeling, such as connectivity between points, can not be done based on the information extracted from the image alone. Full automation often imposes restrictions on the type of object being reconstructed [Bailard et all, 1999], the way the images are taken [Pollefeys et al, 1999], and may limit accuracy. Human intervention, at the critical phases, can solve most those drawbacks and save the time required to edit the output of a fully automated process.
2. Correct geometry demand no compromises in the camera model, and registration is most accurate with a global optimization technique. Moreover, reliability can be assured with continuous accuracy checks in every step.
3. For reasons detailed later, the system must be able to obtain 3D information not only from multiple images but also from a single image.
4. The system should be able to integrate data from different types of sensor. With the existing technology, it is not possible to capture all details with one type of sensor. Laser scanning is ideal for featureless sculptured surfaces and geometrically complex shapes but may not be practical or feasible for large structures. On the other hand image-based methods can model large structures, however may not be able to acquire details on unmarked surfaces.

5. Even when using different types of sensor, some parts of a complex environment will not be captured due to, for example, occlusions. Methods to interpolate or extrapolate to fill the gaps should be employed.
6. The system should be able to integrate models constructed from independent sets of images (e.g. a set of images capturing a front of a building and another set capturing the back, the top, or the inside).
7. For large complex environments, the system must be able to take input from positioning devices (e.g. GPS).

This paper argues that implementing the above ideas will result in:

- Flexibility and less restriction on the type of object or environment in terms of complexity and size.
- High accuracy and reliability.
- Complete scene description.

This approach is suited for a variety of applications including accurate documentation of complex sites.

2 DETAILS OF THE APPROACH

To completely describe a complex scene we propose to practically combine several techniques. The basic approach uses multiple overlapped images and photogrammetric bundle adjustment to find the global shape of the environment or object. However, we also need 3D points from single images because the following cases are prevalent:

- Most images do not have enough features to describe all surfaces. Without feature, multi-images can not be used.
- Parts of the surface appear in one image only due to occlusions or incomplete coverage.
- In complex scenes, finding feature in multiple images can be time consuming and error prone.

We will also need additional techniques to automatically generate sufficient 3D points on non-flat surfaces that have only few features. If the surface is of known shape, for example, a sphere, cylinder, or quadric, we can determine the parameters of the surface from available features. Once this is determined, additional points can be added automatically and projected into the images using the known image parameters. Another approach, which can be applied to smooth a triangulated part of the surface, is polygon subdivision (Zorin, 1997). The initial set of triangles is split into smaller triangles by adding, for example, a point in the middle of each side and fit B-splines to determine 3D coordinates of new points. The process is repeated as many times as needed. However, the existing features may not be sufficient to start this process. In this case an active range sensor such as a laser scanner is best suited for the local details

2.1 Main Features

Based on the ideas presented in section 1.2 and the above discussion, we designed a system with the following features:

- Complete camera calibration with all distortion parameters to provide solid basis for all subsequent computations
- Sub-pixel target and corner extraction
- Manual selection and labeling of points
- Photogrammetric bundle adjustment with or without control points (free network)
- Least squares surface fitting (planes, cylinders, spheres, quadrics, circles)
- Computes 3D from a single image, with fitted surface model
- Automatic addition of randomly distributed points on known surfaces
- Automatic polygon subdivision to smooth surface between existing triangles
- Automatic extraction and matching of targets, if available, after registration
- Interactively register and integrate data from laser scanners with main data sets from digital images
- Interactively register and integrate models created by independent sets of images
- Generates accuracy numbers for calibration, 3D points, registration, and fitting
- Accepts input from positioning devices
- Human assisted determination of point connectivity followed by automatic modeling

2.2 Main Steps

Figure 2 summarizes the main procedure steps. In the data collection step, overlapped images, usually from digital still camera, are taken at wide base line and made sure to cover the intended object or site. If local details are needed on some parts, those can be acquired with a range sensor. Placement of sensors dependents entirely on the site and the application requirements and is beyond the scope of this paper.

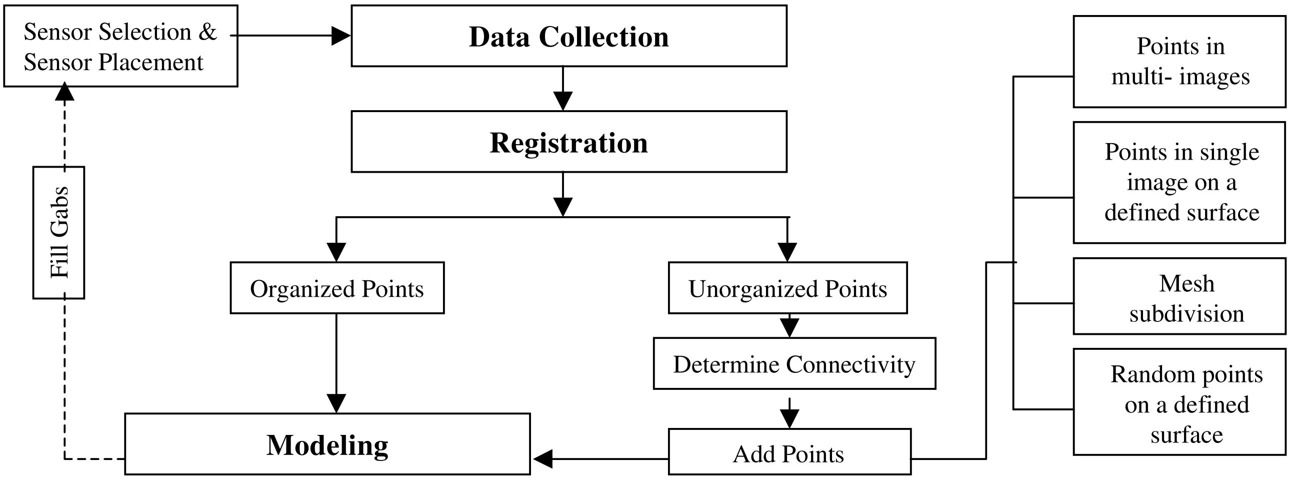


Figure 2: Summary of the overall procedure

2.2.1 Registration of Main Images. The images are displayed in the proper order and common points are extracted and labeled interactively. If correct scale is required, some distances in the scene are also measured. Control points, if available, or data from positioning devices (e.g. GPS), may also be utilized in this step. A bundle adjustment is carried out to register the images. In addition to registered images, we also have a number of unorganized scattered 3D points.

2.2.2 Segmentation, Fitting, and Automatic Point Densification. The 3D points generated so far are not sufficient for modeling. They are also unorganized, thus the connectivity, or the topology, is unknown. Three interrelated operations are needed in order to add sufficient points and organize them to create a complete 3D model. Segmenting or grouping 3D points into sets each belonging to a different surface is the first step. Most existing automatic modeling methods were developed for organized 3D points, such as the range images obtained from a laser scanner [Soucy et al, 1996], or unorganized points belonging to specific types of object [Hoppe et al, 1992]. Unorganized points obtained from features on various surfaces on different objects are almost impossible to model automatically since they are subject to many possible interpretations. In our approach, the scene is visually divided into surface patches, each is triangulated and texture mapped separately. Although this is specified manually by a human operator, it is easy to do since all that is required is to draw, with the mouse, a window around the points belonging to the same surface set. Once this is done, the modeling will be carried out fully automatic. Each set may be on a different surface, or the same surface may be divided into several sets depending on the complexity of its shape.

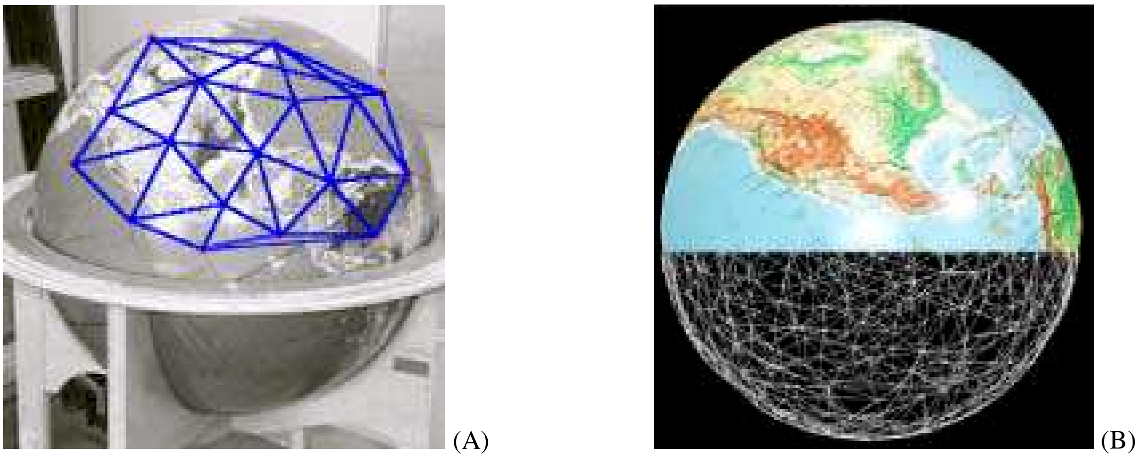

(A) (B)

Figure 3: Using existing features to fit a known shape (sphere) then automatically adding any number of points.

Using any existing features on the surface set, 3D-point computation is first done interactively [Figure 3.A]. These 3D-points are then used to compute the function defining the surface, using least squares fitting. The function is in turn used to automatically generate new points on the surface [Figure 3.B, half with texture]. On more complex surfaces, we can only interpolate between the existing triangles by a subdivision technique [Zorin, 1997]. For partially occluded surfaces, a single image can be used to extend the surface. For example, in figure 4, we can extend the floor (4.A) or the side of the cabinet (4.B) by fitting a plane using the corners of the surface. We then use a single image and pick any point on the surface. The 3D coordinates of this point can be computed from the image coordinates and the plane parameters.

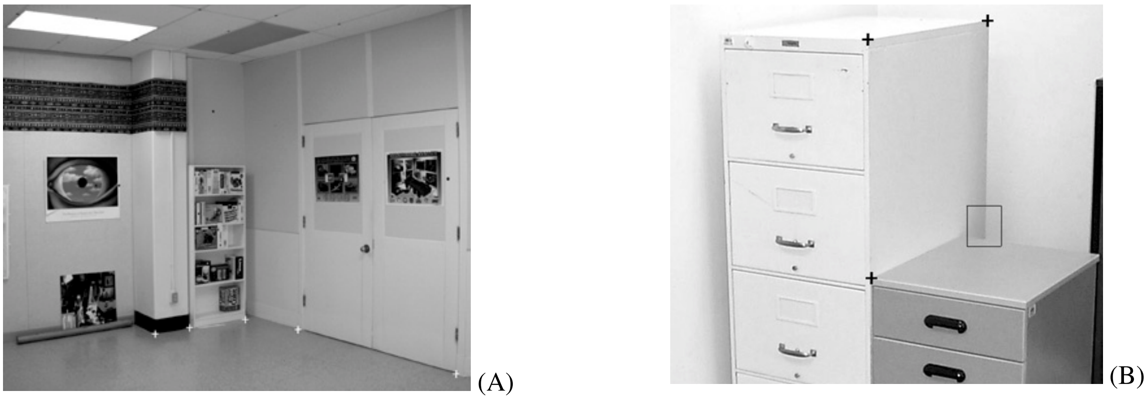
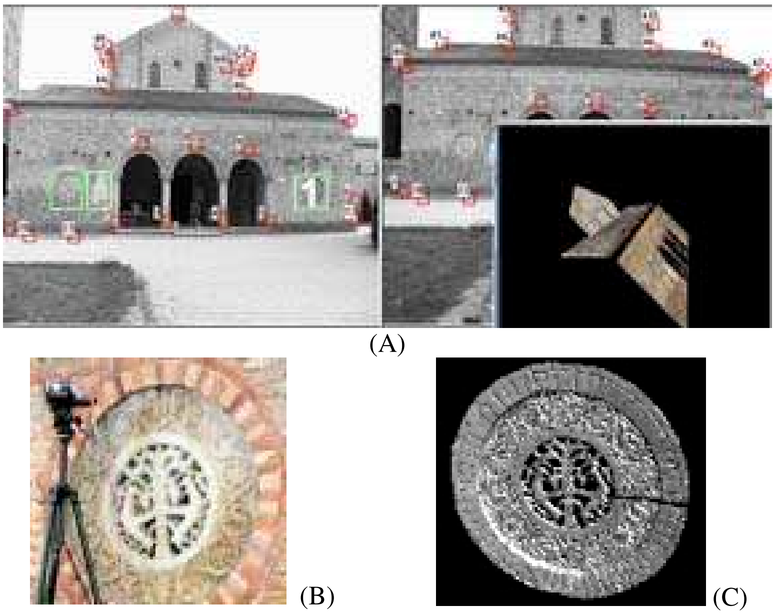

(A) (B)

Figure 4: Using available features to fit planes and then extrapolate featureless points from single images.

2.2.3 Adding Range Data. This involves matching and integrating local detailed points obtained by a laser scanner to the global model obtained by the image-based method. This is best described by an example. In figure 5, nearly all the structure is easy to model with images taken by digital camera. However, parts of the surface contain fine geometric details that will be very difficult or impractical to model from images, such as the one shown in figure 5.B. Those parts are best acquired by a laser scanner and added to the global model created from the images. To register the detailed model shown in figure 5.C we measure several features, usually 6, using the images then extract the 3D coordinates of the same features from the scanned data. This is done interactively using intensity images generated by the laser scanner. A similarity transformation is then used to register the two coordinate systems of the two data sets.



(A) (B) (C)

Figure 5: (A) Images for global modeling, green boxes are the areas to be scanned. (B) The range sensor. (C) The detailed model of area 1.

The same idea is applied to integrate models reconstructed from different sets of images (item 6 in section 1.2 above).

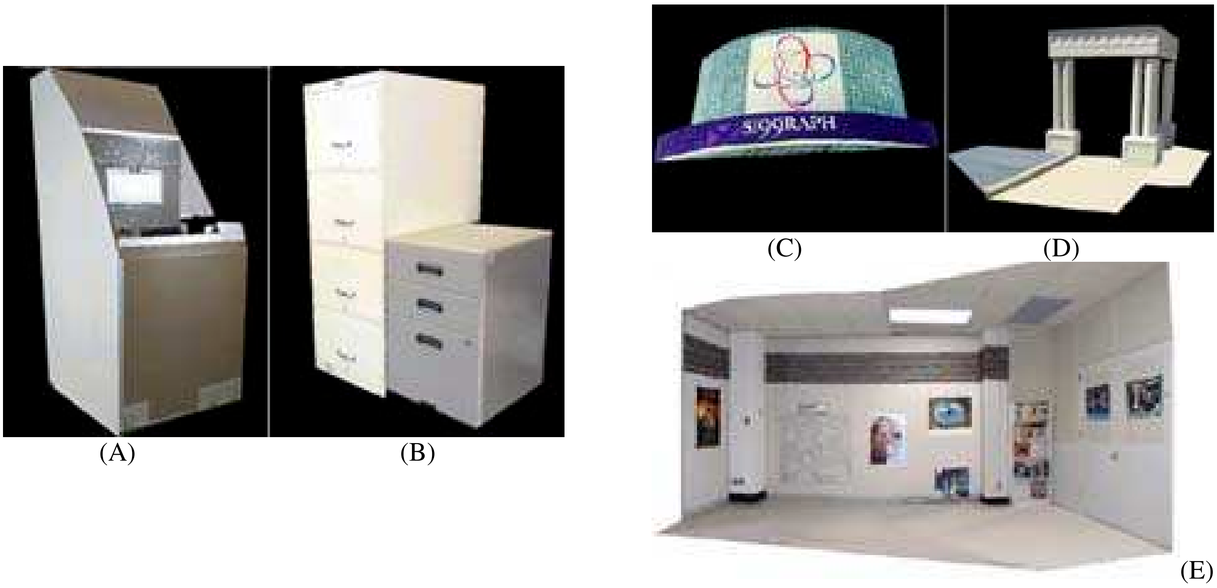
3 EXPERIMENTAL RESULTS AND ACCURACY ANALYSIS

We have extensively tested the performance of our system. Different types of object and environment have been tested to demonstrate the concept of automatic point densification, measuring details that show in one image only, and integrating different sensor data. Rarely any of the publications describing 3D modeling techniques address the issue of geometric accuracy of the model. Here, we quantitatively evaluate the accuracy using directly measured dimensions between features or known geometric parameters of objects (for example a radius of a sphere). We used simple objects for the accuracy tests, however, this should be sufficient to represent the accuracy for any other object.

3.1 Example Models

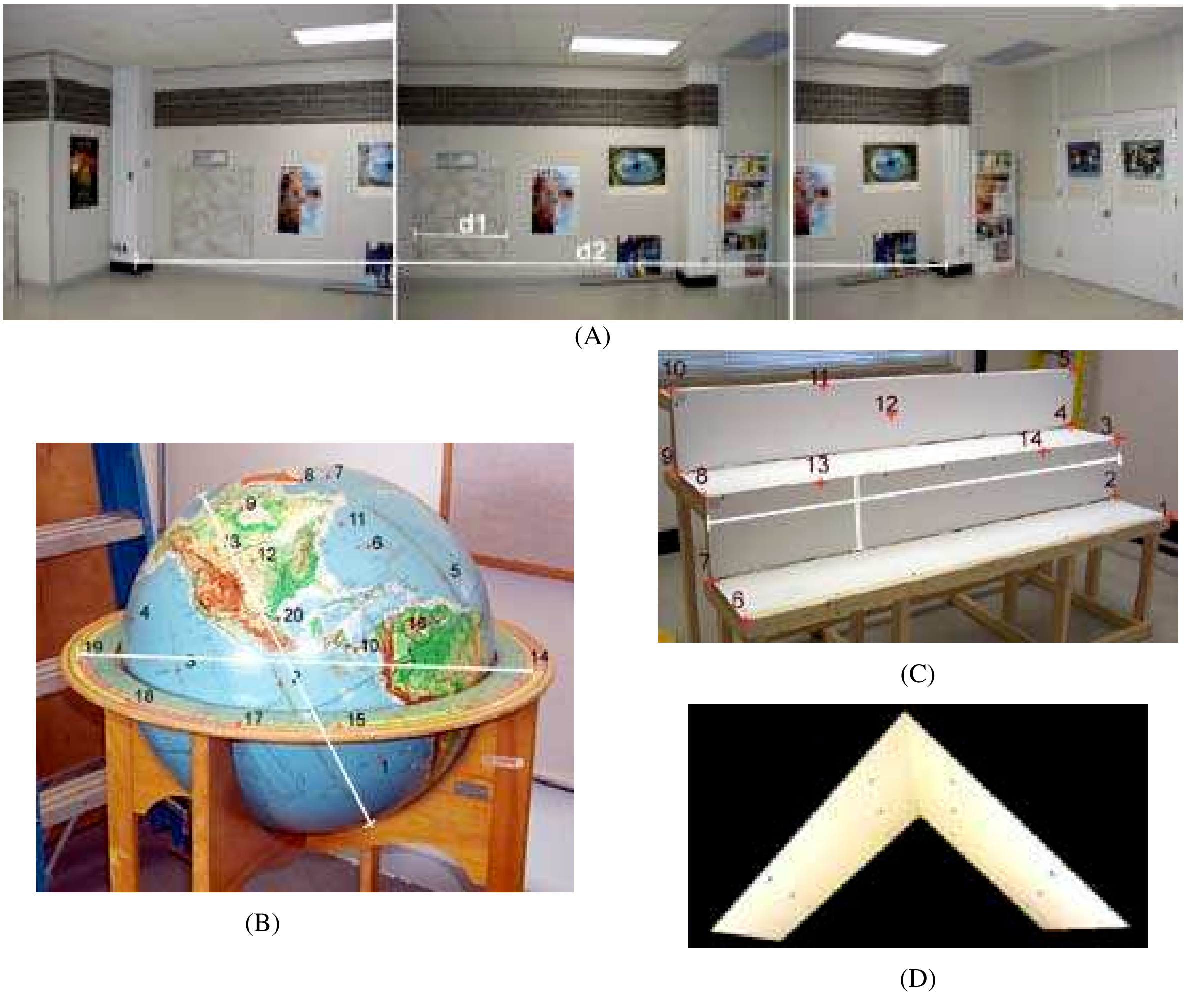
Figure 6.A, B, and E show examples of models where most of the surfaces are planes. The use of plane fitting and 3D from a single image was a must to complete those models. Many of the surfaces had no features except when they join other surfaces. This made it impossible to obtain 3D from multiple images. Figure 6.C shows a structure where the surfaces are quadrics. To get smooth surfaces, thousands of points were added automatically after fitting quadrics to manually extracted points. This model took only 20 minutes to complete. Figure 6.D shows a structure consisting of

cylinders and planes. The planes forming the grounds were extended from single image measurements, while the columns were smoothed with several hundreds of automatically added points using cylinder fitting. Figure 5 above shows an example of a model where digital camera and laser scanner data were combined to increase the details.



(A) (B) (C) (D) (E)

Figure 6: Example of created models: (A) and (B) are objects with planar surfaces, (C) a large structure with quadric surfaces, (D) is a structure with cylindrical surfaces, and (E) is an indoor site.



(A) (B) (C) (D)

Figure 7: Accuracy test dimensions: (A) Indoor site, (B) Globe, (C) Test benches, (D) Planer walls.

3.2 Accuracy Verification

In the tests shown in figure 7.A and C, several distances were measured between 3D points. The differences between the computed distances and the directly measured distances were computed. Also radius computed from surface fitting of the globe and a circle on the rim in figure 7.B are compared to directly measured values. Deviations from plane in the walls showing in figure 7.D were also computed. Table 1 displays the results. Accuracy is estimated to be from 1: 2300 to 1: 6000, which is good considering that natural features were used for image registration and point measurement.

	Description	Actual value	Difference	Relative to site size
Distances	d1, figure 6-A (5 m view)	1215	-1.5	1: 3700
	d2, figure 6-A	3810	-1.9	1: 2900
	length of bench, figure 6-C (3 m view)	2001	-0.5	1: 6000
	width of bench, figure 6-C	300	1.3	1: 2300
	between targets 13-14, figure 6-C	802	0.5	1: 6000
Fitted surfaces	Sphere radius in figure 6-B (2 m view)	300.5	0.35	1: 5700
	Circle (rim) radius in figure 6-B	351	0.85	1: 2300
	Planes in figure 6-D (2.5 m view)	plane	0.60 off plane	1: 4200

Table 1: Results of accuracy tests. The values and differences are in mm.

4 CONCLUSIONS AND FUTURE WORK

The described system exhibits notable improvement in flexibility, accuracy, and completeness over existing approaches. The system is mostly interactive with easy to use interface. Depending on the type of surface and the environment, certain components are automatic. The main advantage of the approach is its flexibility in that it can use image-based modeling from multiple or single images, combine multiple image sets, use data from positioning devices, and integrates data from range sensors such as laser scanners. The accuracy achieved by applying a complete camera model and simultaneous photogrammetric global adjustment of bundles is sufficient for most applications.

Although this interactive system can be used to model a wide spectrum of objects and sites, it is still desirable to reduce human intervention, particularly when using a large number of images. Automation is particularly needed for:

- image acquisition and view planning (incremental on-site modeling may be needed),
- point extraction and matching before registration, especially for widely spaced camera positions, and
- determining point connectivity by segmentation of 3D points into groups.

Occlusions and variations in illumination between images affect existing automatic methods for correspondence and image registration. Therefore, they need images taken at close intervals, which result in too many images as well as reduced geometric accuracy. In addition, the resulting 3D points are not likely to be suitable for modeling. Therefore, improved automated methods that do not suffer from these shortcomings are the subject of future research.

ACKNOWLEDGMENTS

I would like to thank my colleagues Angelo Beraldin and Luc Cornouyer for providing the range sensor data.

REFERNENCES

Baillard, C., Zisserman, A., 1999. “Automatic reconstruction of piecewise planar models from multiple views”, IEEE Conference on Computer Vision and Pattern Recognition, pages 559-565.

Beraldin, J.-A., Blais, F., Cornouyer, L., Rioux, M., El-Hakim, S.F., 1999. “3D imaging for rapid response on remote sites”, SIGGRAPH’99 Technical Sketches, pp. 225.

Chapman, D., Deacon, A., 1998. "Panoramic imaging and virtual reality - filling the gaps between the lines", ISPRS Journal for Photogrammetry and Remote Sensing, 53(6), pp. 311-319, December.

Chen, E., 1995. “QuickTime VR–an image-based approach to virtual environment navigation”, SIGGRAPH’95, pp. 29-38.

Coorg, S., Teller, S., 1999. “Extracting textured vertical facades from controlled close-range imagery”, IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp. 625-632.

Debevec, P.E., Taylor, C.J., Malik, J., 1996. “Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach”, SIGGRAPH’96, pp.11–20.

El-Hakim, S.F., Brenner, C., Roth, G., 1998. "A multi-sensor approach to creating accurate virtual environments", ISPRS Journal for Photogrammetry and Remote Sensing, 53(6), pp. 379-391, December.

Faugeras, O., 1992. “What can be seen in three dimensions with an uncalibrated stereo rig”, ECCV’92, pp. 583-578.

Gruen, A. Wang, X., 1998. “CC-Modeler: a topology generator for 3-D city models”, ISPRS Journal for Photogrammetry and Remote Sensing, 53(5), pp. 286-295, October.

Gomes de Sa, A., Zachmann, G., 1999. “Virtual reality as a tool for verification of assembly and maintenance processes”, Computers & Graphics, Vol. 23, pp. 389-403.

Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., Stuetzle, W., 1992 “Surface reconstruction from unorganized points”, Computer Graphics (SIGGRAPH ’92 Proceedings), 26(2), pp. 71–78.

Kang, S.B., 1999. “Survey of image-based rendering techniques”, SPIE Vol. 3641, Videometrics VI, pp. 2-16.

Levoy, M., Hanrahan, P., 1996. “Light field rendering”, SIGGRAPH’96, pp. 31-42.

Liebowitz, D., Criminisi, A., Zisserman, A., 1999. ”Creating Architectural Models from Images”, EUROGRAPHICS ’99, 18(3).

Pollefeys, M., Koch, R., Vergauwen, M., Van Gool, L., 1999. “Hand-held acquisition of 3D models with a video camera”, IEEE proceedings of 2nd. Int. Conf. On 3D Digital Imaging and Modeling (3DIM’99), pp. 14- 23.

Sablatnig, R., Menard, C., 1996. “Computer based acquisition of archaeological finds: the first step towards automatic classification”, 3rd International Symposium on Computing and Archaeology, Rome, Vol. 1, pp. 429-446.

Sequeira, V., Ng, K., Wolfart, E., Goncalves, J.G.M., Hogg, D., 1999. “Automated reconstruction of 3D models from real environments”, ISPRS Journal for Photogrammetry and Remote Sensing, 54(1), pp. 1-22, January.

Soucy, M., Laurendeau, D., 1996. "Multiresolution surface modeling based on hierarchical triangulation", Computer Vision and Image Understanding 63(1), pp. 1–14.

Szeliski, R., Shum, H.-Y., 1997. "Creating full view panoramic image mosaics and environment maps", SIGGRAPH’97, pp. 251-258.

Szeliski, R., Torr, P., 1998. “Geometrically constrained structure from motion: points on planes”, Microsoft Research Tech. Report MSR-TR-98-64.

Teller, S., 1998. “Towards urban acquisition from geo-located images”, 6th Pacific Conference on Computer Graphics and Applications, Pacific Graphics’98, IEEE Press, pp. 45-51, Oct. 26-29.

Thompson, W.B., Owen, J.C., de St. Germain, H.J., Stark, S.R., Henderson, T.C., 1999. “Feature-based reverse engineering for mechanical parts”, IEEE Trans. Robotics and Automation, 15 (1), pp. 57-66, February.

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

Zorin, D.N., 1997. Subdivision and Multiresolution Surface Representation. Ph.D. Thesis, Caltech, California.