Semi-automatic 3D Reconstruction of Occluded and Unmarked Surfaces from Widely Separated Views

El-Hakim, Sabry

Publisher's version / Version de l'éditeur:

Close Range Visualization Techniques, 2002

National Research Council Canada
Institute for Information Technology

Conseil national de recherches Canada
Institut de technologie de l'information

NRC-CNRC

*Semi-automatic 3D Reconstruction of Occluded and Unmarked Surfaces from Widely Separated Views **

El-Hakim, S.F.
September 2002

* published in Proceedings of ISPRS Commission V Symposium, Close Range Visualization Techniques, Corfu, Greece, September 1-2, 2002. pp. 143-148. NRC 44944.

Canada

[1]SEMI-AUTOMATIC 3D RECONSTRUCTION OF OCCLUDED AND UNMARKED SURFACES FROM WIDELY SEPARATED VIEWS

Sabry F. El-Hakim

Visual Information Technology (VIT) Group
Institute For Information Technology, National Research Council Canada (NRC)
Ottawa, Ontario, Canada K1A 0R6
E-mail: Sabry.El-Hakim@nrc.ca

Commission V, Working Group V/2

KEY WORDS: Architecture, Automation, Three-dimensional, Reconstruction, Texture, Virtual Reality, Accuracy, Registration

ABSTRACT:

Three-dimensional modeling from images, when carried out entirely by a human, is time consuming and impractical for large-scale projects. On the other hand, full automation may still be unachievable for many applications. In addition, 3D modeling from images requires the extraction of features and needs them to appear in multiple images. However, in practical situations those features are not always available, sometimes not even in a single image, due to occlusions or lack of texture. Taking closely separated images or optimally designing view locations can preclude some occlusions. However, taking such images is often not practical and we are usually left with images that do not properly cover every detail. This paper argues that widely separated views and a semi-automated technique are the logical solutions to 3D construction of large and complex objects or environments. The proposed approach uses both interactive and automatic techniques, each where it is best suited, to accurately and completely model man-made structures and objects. It particularly focuses on automating the construction of unmarked surfaces such as columns, arches, steps and blocks from minimum seed points. It also extracts the occluded or invisible corners and lines from existing ones. Many examples, such as Arc de Triomphe in Paris and Florence's St. John baptistery, are completely modeled from a small number of images taken by tourists.

1. INTRODUCTION

A wide range of applications requires 3D reconstruction of real world objects and scenes. In general, most applications specify a number of requirements:

1. High geometric accuracy
2. Copturing all details
3. Photo-realism

In addition the following would ideally be desired in a system that creates such models:

4. Full automation
5. Low cost
6. Portability
7. Flexibility in applications
8. Efficiency in model size

The order of importance of these requirements depends on the application, but in many all are important. A single system that satisfies all requirements is still in the future. In particular, accurately covering all the details with a fully automated system for a wide range of objects and scene remains elusive. For small and medium sized objects, up to the size of human or a statue, range-based techniques such as laser scanners can provide accurate and complete details with high degree of automation [Beraldin et al, 1999], but being a relatively new technology that is not produced in large quantities, they remain costly. They are also not portable enough for a single person to carry around and use in a manner similar to a video or digital camera. Image based approaches entail widely available hardware and potentially the same system can be used for a wide range of objects and scenes. They are also capable of producing realistic looking models and those based on photogrammetry have high accuracy. The issues that remain are coverage of details on unmarked and sculpted surfaces and full automation. This paper focuses on image-based methods aiming at increasing the details level and automation for man-made objects.

Three-dimensional measurement from images naturally requires that interest points or edges be visible in the image. This is often not possible either because a region is hidden or occluded behind an object or a surface, or because there is no mark, edge, or visual feature to extract. This problem exists even with only one object in the scene and when we can take images from well-planned positions. In objects such as architectures and monuments in their normal settings we are also faced with the restrictions of limited locations from which the images can be taken as well as the existence of other objects and illumination variations and shadows. All this causes problems for automatic modeling techniques and also generates incomplete models. Our approach, which does not aim to be fully automatic nor completely rely on human operator, is especially designed to model occluded and unmarked surfaces for structure such as classical architectures. The approach provides enough level of automation to assist the operator without sacrificing accuracy. Image registration and scene segmentation into separate regions are done interactively. This is followed by automatics corner detection and correspondence. Most points on blocks, windows, doors, steps, cylinders or columns, arches, quadrics and planes are measured automatically even when they are occluded or unmarked. On average, 20% of the points are measured interactively and the remaining 80% are added automatically.

The remainder of the paper is organized as follows. In section 2, an overview of image-based 3D reconstruction techniques is presented. This will lead to a deduction in the following section that widely separated views and semi-automated techniques are currently the logical solution to 3D construction of large and complex objects or environments. A discussion on the effect of occlusion and lack of textures is given in section 4 followed by

[1] ISPRS Comm. V Symposium, Corfu, Greece, Sept. 2-6, 2002

details of the proposed approach. Section 6 illustrates many examples, such as Arc de Triomphe in Paris, Florence's St. John baptistery, San Giacomo dell'Orio church in Venice, and other structure from around the world that were fully modeled rapidly from a small number of images taken by tourists. The paper concludes with a short discussion in section 7.

2. SYNOPSIS OF 3D RECONSTRUCTION TECHNIQUES

The ultimate goal of all 3D reconstruction methods is to satisfy the eight requirements listed in the previous section. Since this is not an easy task, they focus on some of the tasks at the expense of the others. We will use this to distinguish between methods in order to depict a comparison. The methods may:

1- Focus on accuracy without any automation.
2- Focus on full automation.
3- Try to reach a balance between all requirements.

The most widely used method remains to be first method, which is the traditional approach. This is a labor-intensive endeavor where engineering plans or drawings plus surveying and/or standard photogrammetry techniques are employed followed by importing the measurements into a CAD system to create a 3D model. The results are often unsatisfactory in appearance and seem computer-generated. Efforts to increase the level of automation became essential in order to meet the increasing demand for 3D models. However, the efforts to completely automate the process from taking images to the output of a 3D model, while promising, are thus far not always successful. The automation of camera pose estimation, self-calibration, and computation of pixel 3D coordinates will be summarized. This procedure, which is now widely used in computer vision [e.g. Faugeras et al, 1998, Fitzgibbon et al, 1998, Pollefeys et al, 1999, Liebowitz, et al, 1999], starts with a sequence of images taken by un-calibrated camera. The system automatically extracts interest points, like corners, sequentially matches them across views, then computes camera parameters and 3D coordinates of the matched points using robust techniques. The key to the success of this fully automatic procedure is that successive images may not vary significantly, thus the images must be taken at short intervals. The first two images are usually used to initialize the sequence. It is important that the points are tracked over a long sequence or in every image where they appear to reduce the error propagation. This is all done in a projective geometry basis and is usually followed by a bundle adjustment, also in the projective space. Self-calibration to compute the intrinsic camera parameters, usually only the focal length, follows in order to obtain metric reconstruction, up to scale, from the projective one [Pollefeys et al, 1999]. Again, bundle adjustment is usually applied to the metric construction to optimize the solution. The next step, the creation of the 3D model, is more difficult to automate and is usually done interactively to define the topology and edit or post process the model. An output model based only on the measured points will usually consist of surface boundaries that are irregular and overlapping and need some assumption to be corrected using for example planes and plane intersections. For large structures and scenes, since the technique may require a large number of images, the creation of the model requires a significant human interaction regardless of the fact that image registration and a large number of 3D points were computed fully automatically. The degree of modeling automation increases when certain assumptions about the object, such as architectures, can be made. Since automated image-based methods rely on features that can be extracted from the scene, occlusions and un-textured surfaces are problematic. We often end up with areas with too many features that are not all needed for modeling, and areas without any or have minimum features that cannot produce a complete model.

The most impressive results remain to be those achieved with interactive approaches. Rather than full automation, a hybrid easy to use system named Façade has been developed [Debevec et al, 1996]. The method's main goal is the realistic creation of 3D models of architectures from small number of photographs. The basic geometric shape of a structure is first recovered using models of polyhedral elements. In this interactive step, the actual size of the elements and camera pose are captured assuming that the camera intrinsic parameters are known. The second step is an automated matching procedure, constrained by the now known basic model to add geometric details. The approach proved to be effective in creating geometrically accurate and realistic models. The drawback is the high level of interaction and the restrictions to certain shapes. Also since assumed shapes determine all 3D points and camera poses, the results are as accurate as the assumption that the structure elements match those shapes. Our method, although similar in philosophy, replaces basic shapes with a small number of seed points in multiple images to achieve more flexibility and levels of detail. In addition, the camera poses and 3D coordinates are determined without any assumption of the shapes but instead by a full bundle adjustment, with or without self-calibration depending on the given configuration. This achieves higher geometric accuracy independent from the shape of the object.

The Façade approach has inspired several research activities to automate it. Werner and Zisserman, 2002, proposed a fully automated Façade-like approach. Instead of the basic shapes, the principal planes of the scene are created automatically to assemble a coarse model. These are three dominating directions that are assumed to be perpendicular to each other. Like Façade, the coarse model guides a more refined polyhedral model of details such as windows, doors, and wedge blocks. Since this is a fully automated approach, it requires feature detection and closely spaced images for the automatic matching and camera pose estimation using projective geometry. Dick et al, 2001, proposed another automated Façade-like approach. It employs model-based recognition technique to extract high-level models in a single image then use their projection into other images for verification. The method requires parameterized building blocks with a priori distribution defined by the building style. The scene is modeled as a set of base planes corresponding to walls or roofs, each of which may contain offset 3D shapes that model common architecture elements such as windows and columns. Again, the full automation necessitates feature detection and projective geometry approach, however the technique used here also employs planner constraints and perpendicularity between planes to improve the matching process. Another approach [Tao et al, 2001] to improve the automatic matching and scene segmentation for modeling, after image registration, applies depth smoothness constraints on surfaces combined with color similarity constraints.

The presence of noise, which result from extracting features from images, will make the choice of camera positions, or more precisely motion versus object distance, critical for correct construction. This has been studied widely in photogrammetry [e.g. Fraser, 1994]. It has been lately recognized in computer vision that photogrammetric bundle adjustment provides the optimum solution to image-based modeling [Triggs et al, 2000]. This has resulted in the inclusion of bundle adjustment following the sequential techniques as mentioned above.

Critical analyses of automated techniques that use projective geometry were undertaken [Oliensis, 2000, Bougnoux, 1998]. Configurations that lead to ambiguous projective construction have also been identified [Hartley, 2000, Kahl et al, 2001]. For metric construction, certain sequences will cause self-calibration to fail or not give a unique construction. These critical motions have been studied when camera intrinsic parameters do not vary [Sturm, 1997] and when they do vary from image to image [Kahl et al, 2000].

This paper focuses on the 3D construction, which is the least automated of all the steps, rather than correspondence, pose estimation or calibration.

3. AUTOMATION AND WIDELY-SEPARATED VIEWS

From the above overview of current techniques, the following points can be made:

- In practical situations taking the sequences required for a fully automated techniques might not be feasible.
- Large objects or complex scenes require a large number of images to carry matching and pose estimation automatically.
- Since the modeling process still requires human interaction to define the topology, assign constraints, and post-process the results, large number of images makes this difficult.

It is therefore important to develop an approach that requires only a small number of widely separated views and at the same time offers a high level of automation, and be able to deal with occluded and unmarked surfaces.

4. ON OCCLUSION AND LACK OF TEXTURE

Occlusions and lack of texture hinder image-based modeling since the methods require features that can be seen either by a computer or human. The scene in figure 1 has occlusions and most of the columns surfaces have no texture. Both fully automated and fully manual methods will have difficulty here. Yet, this is typical in much classical architecture. In our approach, with less than 30 manually measured points, the full scene [figure 2] with automatically added 300 points can be completed without further human intervention.

Figure 1: Corner of a courtyard

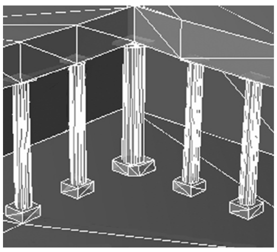

Figure 2: The constructed solid model

5. DETAILS OF THE APPROACH

The approach is designed mainly for man-made objects. A good example is classical architectures, which are designed within constraints of proportion and configurations. Classical buildings are divided into architectural elements. These elements are logically organized in space to produce a coherent work. There is a logical hierarchical relation among building parts and between parts and whole. The most common scheme divides the building into two sets of lines forming a rectangular grid [Tzonis and Lefaivre, 1986]. The distance between the grid lines are often equal or when they vary, they alter regularly. The grid lines are then turned into planes that partition the space and control the placement of the architectural elements. The automation of 3D reconstruction is better achieved when such understanding is taken into account. We will reconstruct the architecture elements from minimum number of points and put them together using the planes of a regular grid. Other schemes, such as a polar grid, also exist but the basic idea can be applied there too. Classical architecture can be reconstructed, knowing its components, even if only a fragment survives or seen in the images. For example, a columnar element consists of: 1) The capital, a horizontal member on top, 2) the column itself, a long vertical tapered cylinder, 3) a pedestal or a base on which the column rests. Each of those can be further divided into smaller elements. In addition to columns, other elements include pillars, pilasters, banisters, windows, doors, arches, and niches. Each can be reconstructed with a few seed points from which the rest of the element is built.

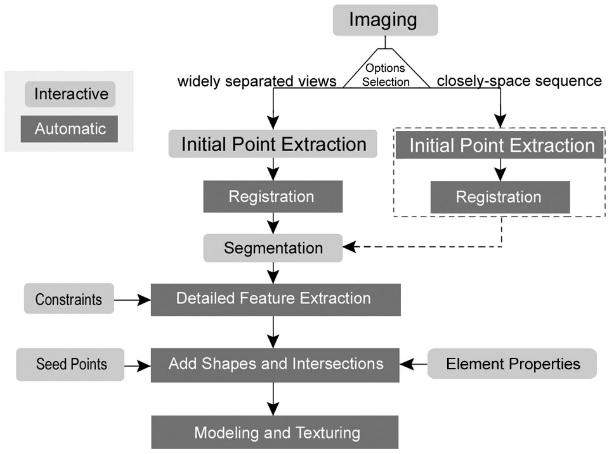

Figure 3. Simplified diagram of the general procedure that shows which is automatic.

Our approach is photogrammetry-based. In order to increase the level of automation, the process takes advantage of properties like those mentioned above for man-made objects and structures. The approach provides an adequate amount of automation to assist an operator to provide high level of details with excellent geometric accuracy. Figure 3 summarizes the procedure and indicates which step is interactive and which is automatic (interactive operations are light gray). The figure also shows an option of taking a closely-spaces sequence of images, if conditions allow, and increase the level of automation. In the remainder of the paper, we will discuss only the option of widely separated views. Images are taken, all with the same camera set up, from positions where the object is suitably showing. There should be a reasonable distance, or baseline, between the images. Several features appearing in multiple images are interactively extracted, usually 12-15 per image. The user points to a corner and label it with a unique number and the system will accurately extract the corner point. Harris operator is used [Harris, 1998] for its simplicity and efficiency. Image registration and 3D coordinate computation are based on the photogrammetric bundle adjustment approach for its accuracy, flexibility, and effectiveness compared to other structure from motion techniques [Triggs et al, 2000]. Advances in bundle adjustment eliminated the need for control points or physically entering initial approximate coordinates. Many other aspects

required for high accuracy such as camera calibration with full distortion corrections have long been solved problems in Photogrammetry and will not be discussed in this paper.
We now have all camera coordinates and orientations and the 3D coordinates of the set of initial points, all registered in the same global coordinates system. Unless a known distance is used, the coordinates are up to scale factor. The next interactive operation is to divide the scene into connected segments to define the surface topology. This is followed by an automatic corner extractor, again the Harris operator, and matching procedure across the images to add more points into each of the segmented regions. The matching is constrained, within a segment, by the epipolar condition and disparity range setup from the 3D coordinates of the initial points. The bundle adjustment is repeated with the newly added points to improve on previous results and re-compute 3D coordinate of all points.

An approach to obtain 3D coordinates from a single image is essential to cope with occlusions and lack of features. Several approaches are available [e.g. van den Heuvel, 1998, Liebowitz et al, 1999]. Our approach uses several types of constraints for surface shapes such as planes and quadrics, and surface relationships such as perpendicularity and symmetry. The equations of some of the planes can be determined from seed points previously measured. The equations of the remaining plane are determined using the knowledge that they are either perpendicular or parallel to the planes already determined. With little effort, the equations of the main planes on the structure, particularly those to which structural elements are attached, can be computed. From these equations and the known camera parameters for each image, we can determine 3D coordinates of any point or pixel from a single image even if there was no marking on the surface. When some plane boundaries are not visible, they can be computed by plane intersections. This can also be applied to surfaces like quadrics or cylinders whose equations can be computed from existing points. Other constraints, such as symmetry and points with the same depth or same height are also used. The general rule for adding points on structural elements and for generating points in occluded or symmetrical parts is to do the work in the 3D space to find the new points then project them on the images using the known camera parameters. The texture images are edited afterwards to remove the occluding objects and replace them with texture from current or other images. The main steps are shown in figure 4.

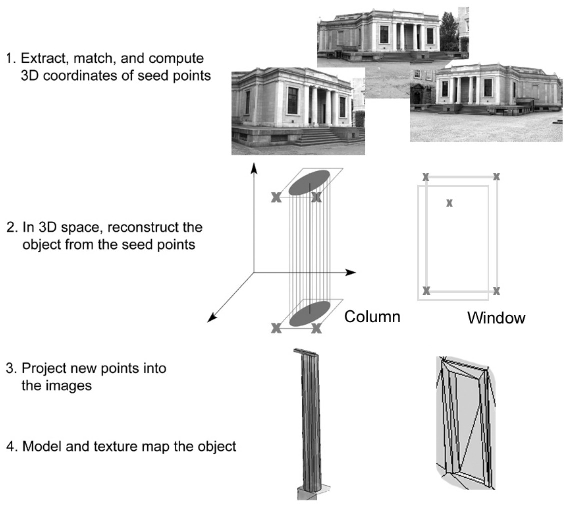

Figure 4. Main steps of constructing architectural elements semi-automatically (column and window examples)

We will now give more details on the use of seed points. A cylinder is constructed after its direction and approximate radius and position have been automatically determined from four seed points (figure 5-a) using quadric formulation [Zwillinger, 1996]. The ratio between the upper and the lower circle can be set in advance. It is set to less than 1.0 (about 0.85) to create a tapered column. From this information, points on the top and bottom circle of the column (figure 5-b) can be automatically generated in 3D resulting in a complete model.

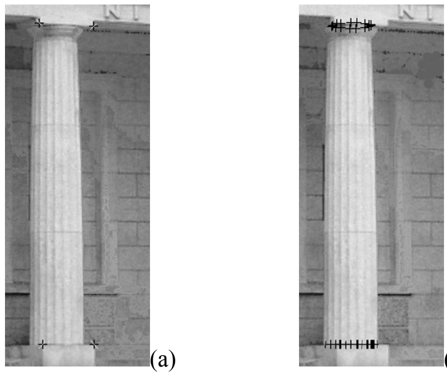
(a) (b)

Figure 5. (a) Four seed points are extracted on the base and crown, (b) column points are added automatically.

Reconstructing arches is similar to the approach used in Façade except that our approach uses seed points instead of blocks and the arch points are extracted automatically. First a plane is fitted to seed points on the wall (figure 6-a). An edge detector (a morphological operator, revision to [Lee et al, 1987]) is applied to the region (figure 6-b) and points at constant interval along the arch are automatically sampled. Using image coordinates of these points (in one image), the known image parameters, and the equation of the plane, the 3D coordinates are computed and projected on the images (figure 7). A procedure for constructing blocks, even when partially invisible, is developed. For example in figure 8 the part of the middle block where it meets the base is not visible and needs to be measured in order to reconstruct the whole block. To solve this problem, we first extract the visible corners on all blocks from several images and compute their 3D coordinates. We then fit a plan to the top of the base block, using the gray points in figure 8, then project a normal to this plane from each of the corners of the block attached to it (the white points). The intersections of each normal will produce a new point (a black point in figure 8) automatically. Using symmetry, we can fully construct the block.

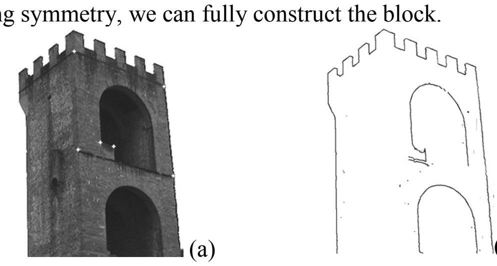
(a) (b)

Figure 6. (a) Seed points (b) detected edge.

Figure 7. Automatic point extraction on arches.

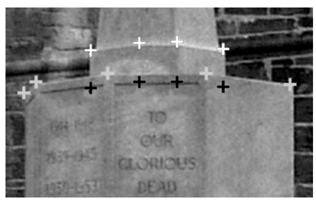

Figure 8. Constructing blocks.

For windows and doors we need three (preferably four) corner points and one point on the main surface (figure 4 above). By fitting a plane to the corner points, and a plane parallel to it at the surface point, the complete window or door can be reconstructed. For steps, sufficient seed points to compute the two side planes, plus one point on either side of each step are needed. Table 1 summarizes seed point requirements.

Element	Seed points
Plane	3: non-linear
Column	4: two on base, two on crown
Window or door	5: 4 corners and one inside
Block attached to a plane	Top corner points
Arch	3: on front plus arch edge
Steps	A corner in either side

Table 1: Seed points for some structural elements

6. EXAMPLES

Over the past year, members of our group took images of various interesting monuments in cities all over the world. The images were taken during routine tours without any advanced planning of where to take the images. We took the images just like any typical tourist, by walking around the monument and getting the best view under real conditions such as presence of other tourists, vehicles, and other buildings and objects. Several types of digital cameras and regular film cameras (digitized later) were used. The results were indeed very encouraging. Over 100 models were created in 6 months, each one usually in 1-2 days of work by one person. The number of points and level of interaction and automation obviously varied significantly from one model to another. At least 80% of the points were generated automatically. Seven examples [more at El-Hakim, 2002] are presented here, each to illustrate specific feature. They are presented in wire-frame, solid model without texture, and solid model with texture. In some of the structures, we found dimensional information available in travel or history books. This information was not used or needed in the model construction, but was valuable in evaluating the accuracy.

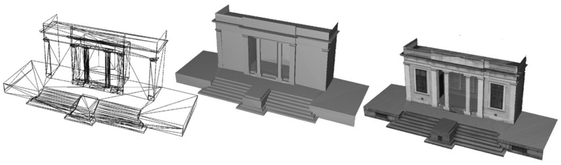

Figure 9: A Trinity Collage building with columns and steps

Figure 9 shows a building on Trinity Collage campus in Dublin. This typical classical architecture includes columns, steps, and several other blocks. From only two images, the whole entrance is constructed from about 400 points of which only about 50 seed points were measured interactively. Figure 10-A shows the Arc de Triomphe in Paris. The Olympus C3030 digital camera (3.1 Mega-pixels) was used (14 images). The arc measures 45 meter x 22 meter, as indicated in some tourist guides (accurate height was not available). We used one distance (the 22 meter width) to scale our model. From the model, the dimensions on the four sides were: 22 meter (fixed for scale), 22.06 meter, 44.85 meter, and 44.89 meter, an error of 0.28%. One should point out that the given dimensions are probably rounded off and the sides are not perfectly identical.

The next example is the St. John baptistery in Florence (figure 10-B). The Olympus E-10 camera (4 Mega pixels) was used to take eight images. The baptistery has eight sides. The actual dimensions were obtained from a plan in a book. The sides average about 13 meter in length. Again we assigned 13 meter to one side and used it to scale the whole model.

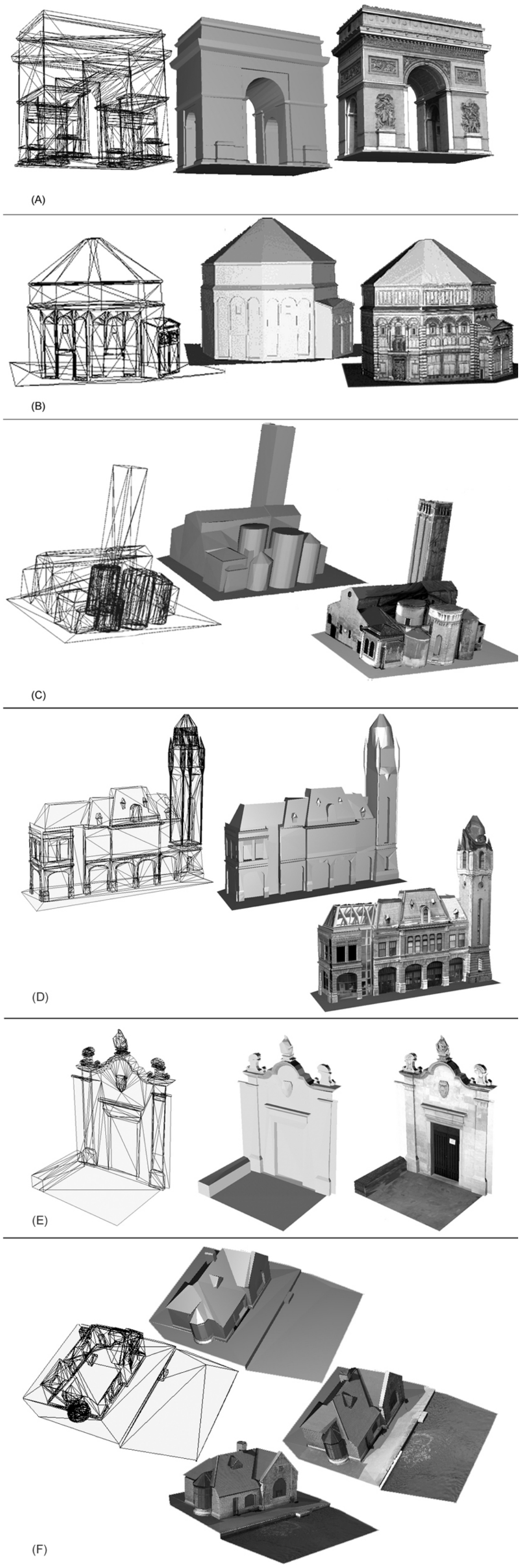

Figure 10: Examples of constructed structures

The average difference between the model sides and the actual sides was less than 1 cm, or 0.07%. This is significantly better than the accuracy of the Arc de Triomphe. This is due to the better quality camera and smaller size object with high contrast feature definition. A compo section in Venice, with the San Giacomo dell'Orio church, is shown in figure 10-C. It illustrates cylindrical shaped structures and great deal of occlusions. Figure 10-D shows a large historic building, la Caserne Dalhausie, in Quebec City. Ten images were taken to cover three sides of the building (the back side was attached to other buildings). The model, which has over 3000 points of which only 150 were measured interactively, captured all the details of the building including the tower. The model in figure 10-E depicts details, including irregular elements, of the entrance to Saint Cristina Crypt in Apulia, Italy. A historic building on the Rideau Canal in Ottawa, figure 10-F, is completely modeled from 9 images.

7. CONCLUDING REMARKS

A semi-automatic approach for constructing medium and large-scale man-made objects, such as classical architecture, was presented. Several representative examples from a small number of images taken by tourists were given. Parts of the process that can straightforwardly be performed by humans, such as registration, extracting seed points, and topological segmentation, remain interactive. Numerous details plus the occluded and the un-textured parts are added automatically by taking advantage of the object characteristics and making some realistic assumptions about the elements shapes and the relations between them. Efforts to automate the whole procedure are continuing. When conditions allow, the steps of initial point extraction and image registration can be fully automated, although this still requires numerous closely-spaced images. In the mean time, to achieve immediate and useful results, parts of the process necessitate human interaction.

Acknowledgements

My colleagues François Blais and Angelo Beraldin took many of the images. Emily Whiting constructed some models.

References

Beraldin, J.-A., F. Blais, L. Cornouyer., M. Rioux, S.F. El-Hakim, R. Rodella, F. Bernier, N. Harrison, 1999. 3D imaging system for rapid response on remote sites. IEEE proc. of 2nd. Int. Conf. 3D Digital Imaging and Modeling, pp. 34-43.

Bougnoux, S., 1998. From projective to Euclidean space under any practical situation, a criticism of self-calibration. Proc. Intl. Conf. Computer Vision (ICCV98), pp. 790-796.

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

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

Dick, A.R., P.H.Torr, S.J. Ruffle, R.Cipolla, 2001. Combining single view recognition and multiple view stereo for architectural scenes. Proc. 8th IEEE International Conference on Computer Vision (ICCV'01), pp. 268-274, July.

El-Hakim, S. 2002. http://www.vit.iit.nrc.ca/elhakim/home.html

Faugeras, O., L. Robert, S. Laveau, G. Csurka, C. Zeller, C. Gauclin, I. Zoghlami, 1998. 3-D reconstruction of urban scenes from image sequences. *Computer Vision and Image Understanding*, 69(3), pp 292-309.

Fitzgibbon, A.W., Zisserman, A., 1998. Automatic camera recovery for closed or open image sequences. European Conf. On Computer Vision (ECCV'98), vol. 1, pp. 311-326.

Fraser, C.S., 1994. Network design considerations for non-topographic photogrammetry. *Photogrammetric Engineering & Remote Sensing*, 50(8), pp. 1115-1126.

Harris C., M. Stephens, 1998. A combined corner and edge detector. Proc. 4th Alvey Vision Conf., pp. 147-151.

Hartley, R., 2000. Ambiguous configurations for 3-view projective reconstruction. In European Conf. Computer Vision, pp. 922-935, Dublin, Ireland, Springer-Verlag.

Kahl, F., Triggs, B., Åström, K., 2000. Critical motions for auto-calibration when some intrinsic parameters can vary. *Journal of Mathematical Imaging and Vision* 13, pp. 131-146.

Kahl, F., Hartley, H., Åström, K., 2001. Critical configurations for N-views projective reconstruction. IEEE Computer Vision and Pattern Recognition (CVPR'01), vol. 2, pp. 158-163.

Lee, J., R. Haralick, L. Shapiro, 1987. Morphologic edge detection. *IEEE J. Robotics and Automation*, 3(2), pp 142-156.

Oliensis, J., 2000. A critique of structure-from-motion algorithms. *Computer Vision and Image Understanding*, 80(2), pp 172-214.

Pollefeys, M., R. Koch, L. Van Gool, 1999. Self-calibration and metric reconstruction in spite of varying and unknown intrinsic camera parameters. *International J. of Computer Vision*, 32(1), pp. 7-25.

Shum, H.Y., R. Szeliski, S. Baker, M. Han, P. Anandan, , 1998. Interactive 3D modeling from multiple images using scene regularities. Eur. Workshop 3D Structure from Multiple Images of Large-scale Environments – SMILE, pp 236-252.

Sturm, P. 1997. Critical Motion Sequences for Monocular Self-Calibration and Uncalibrated Euclidean Reconstruction. Proc. Int. Conf. Computer Vision and Pattern Recognition, San Juan, Puerto Rico, pp. 1100-1105.

Tao, H., Sawhney, H.S., Kumar, R., 2001. A global matching framework for stereo computation. Proc. IEEE 8th Int. Conf. On Computer Vision, ICCV'2001, Vol. 1, pp. 532-539.

Triggs, W., P. McLauchlan, R. Hartley, A. Fitzgibbon, 2000. Bundle Adjustment for Structure from Motion. in *Vision Algorithms: Theory and Practice*, Springer-Verlag.

Tzonis, A., Lefaivre, L., 1986. *Classical Architecture – The Poetics of Order*. MIT Press, Cambridge, MA.

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

Werner, T., Zisserman, A., 2002. New technique for automated architectural reconstruction from photographs. Proc. 7th European Conf. Computer Vision, Copenhagen, May.

Zwillinger, D. (ed.), 1996. *Standard Mathematical Tables*. 30th Edition, CRC Press, Inc., Florida, pp 311-316.