IMAGE-BASED 3D ACQUISITION TOOL FOR ARCHITECTURAL CONSERVATION

Joris Schouteden, Marc Pollefeys, Maarten Vergauwen, Luc Van Gool
Center for Processing of Speech and Images, K.U.Leuven,
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
Tel. +32 16 321064; Fax +32 16 321723
Marc.Pollefeys@esat.kuleuven.ac.be

KEY WORDS: photogrammetry, architecture, heritage conservation, image-based 3D reconstruction, dense stereo.

ABSTRACT
In this paper an image-based 3D acquisition tool is presented. The tool has been developed with architectural conservation in mind. The proposed tool is developed as part of a project that aims at developing a technology that enables the operator to build up an accurate three dimensional model - without too much repetitive work - starting from photos of the objects and measured reference co-ordinates. This model can in a later phase be used as a core for a multimedia database, to represent designs for interventions, or for distribution to the wider public. The presented 3D acquisition tool is largely automated. A limited amount of interaction is used to indicate the reference points in the images to allow global localization in world coordinates. These points are also used to simplify the matching process so that wide baselines can be coped with. Depending on the number of reference points the camera needs to be pre-calibrated or not. From this point on the approach automatically finds additional matches and refines the calibration. A globally consistent estimate can be obtained through bundle adjustment. The next step consists of dense depth estimation based on stereo matching. From the computed information a dense textured 3D surface reconstruction of the recorded structure is computed. The tool also allows extracting 2D slices or pointing measurements. The results can be used both for measurements and for visualization purposes.

1 INTRODUCTION

At this moment many architects involved in conservation still work in the traditional way. They use instrument based (theodolite, total station, photogrammetry) or even hand-measured (tapes, plumb-bobs, levels...) survey methods. As many architects are shifting towards computer-aided design for new buildings, they also try to apply these programs to renovation or conservation projects. However, the number of tools available to accomplish the task of 'getting the existing building in the CAD program' is limited, and mainly directed to 'translate' traditional methods to CAD (automatic import of full station co-ordinates, error-adjustment of triangulation...). Based on a limited number of actually measured points, 2D plans and sections or a 3D model can be constructed. This typically results in a very 'simplified' representation of the building, which is absolutely not in line with the high requirements for conservation purposes.

The presented 3D acquisition tool is largely automated. A limited amount of interaction is used to indicate the reference points in the images to allow global localization in world coordinates. These points are also used to simplify the matching process so that wide baselines can be coped with. Depending on the number of reference points the camera needs to be pre-calibrated or not. From this point on the approach automatically finds additional matches and refines the calibration. A globally consistent estimate can be obtained through bundle adjustment. The next step consists of dense depth estimation based on stereo matching. A complete model is obtained by combining the results of different disparity maps into a 3D reconstruction. The texture is extracted from the original photographs. From the computed information a dense textured 3D surface reconstruction of the recorded structure is computed. The tool also allows extracting 2D slices or pointing measurements. The results can be used both for measurements and for visualization purposes.

The proposed tool is developed as part of a project that aims at developing a technology that enables the operator to build up an accurate three dimensional model - without too much repetitive work - starting from photos of the objects and measured reference co-ordinates. This model can in a later phase be used as a core for a multimedia database, to represent designs for interventions, or for distribution to the wider public (Nuyts 2001).

2 CALIBRATION

The first module allows to calibrate the camera and to compute the pose in world coordinates for the different views. For this purpose the tool assumes the knowledge of 3d reference points measured by surveying methods as well as their projection in the imagery. As a part of this tool an interface is provided for identifying the 3d reference points in the images. A snapshot of this interface can be seen in Fig. 1. The reference points have been indicated in both images.

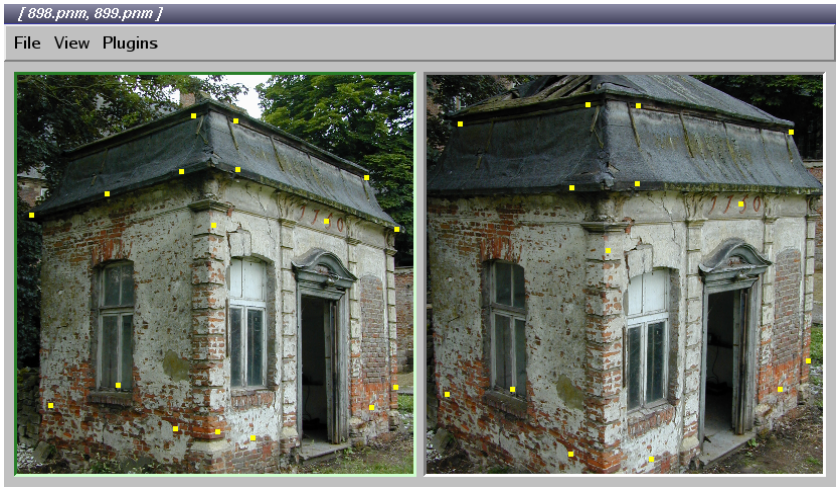

Fig. 1: Measured 3D reference points are indicated in the images

Several different approaches for camera calibration have been implemented. Experiments have shown these different approaches to be complementary. The camera calibration not only has to compute the parameters of perspective projection, but also has to deal with radial distortion. Both tools for a priory as for on-line calibration are provided.

2.1 Pre-calibration of internal camera parameters

In many cases the calibration can be simplified by pre-calibrating the camera off-line. Note that this is only feasible when the camera has fixed settings or, at least, the settings can reliably be reproduced (e.g. minimal or maximal settings). In this case the parameters describing the internal structure of the camera can be determined in a separate procedure, so that they can be assumed known once the pose of the camera has to be determined. This can be especially helpful when only a restricted number of reference points are available.

To achieve this a photograph of a pattern of circles is used. The circles are automatically detected in the images and their centres accurately computed up to sub-pixel accuracy. Then Tsai camera calibration for coplanar points is performed to determine the cameras internal parameters including the parameter of radial distortion. Fig. 2 shows an example of the pattern and a pattern detected from a real photograph. Note that the internal parameters can change from image to image due to change of focus or even auto-focus. In these cases it can be necessary to refine the calibration results based on the images used for reconstruction.

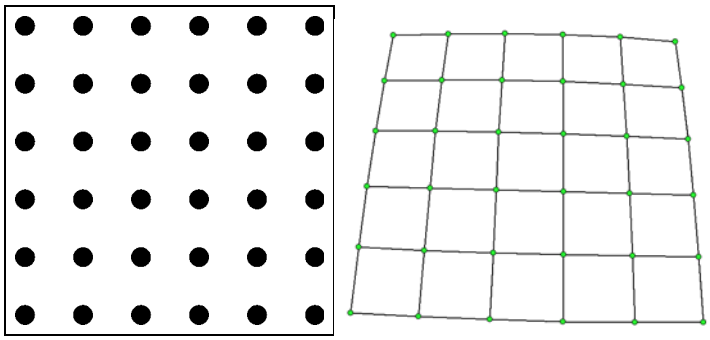

Fig. 2: calibration pattern (left) and pattern detected from a real photograph (right).

2.2 Calibration from images and reference points

The calibration approach that is used for this tool consists of a combination of a Tsai´s calibration (Tsai 1986) and Grunert´s pose estimation algorithm (see Haralick 1994). In this case the version of Tsai´s algorithm for non-coplanar points is used on the reference points. Although Tsai´s algorithm can also be used to determine the pose of the camera, more robust results are obtained by using Grunert´s algorithm once the internal calibration has been determined.

Grunert´s pose determination algorithm only requires 3 points. When more points are available the pose is computed for different combinations and the one yielding the smallest global reprojection error is selected. Finally, a global minimization can be carried out that takes all observations and constraints into account (Slama 1980).

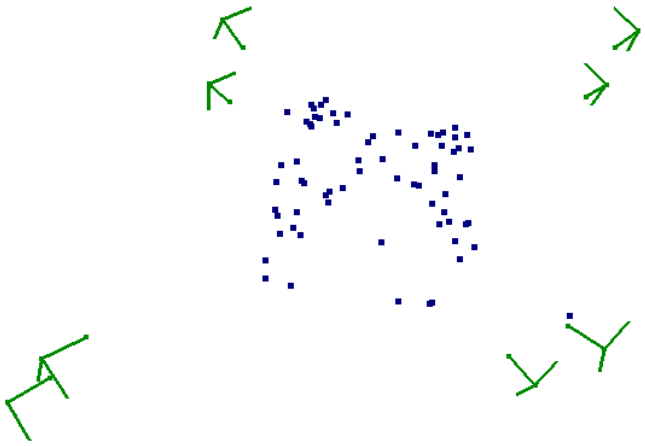

Fig. 3: Cameras poses in world reference frame

2.3 Self-calibration

Although this was not yet integrated within this tool, some experiments with automatic matching and self-calibration have been carried out. The purpose was to calibrate the camera without any manual interaction. A more complete description of this approach can be found in (Pollefeys et al. 1999a, Pollefeys et al. 2000). This method works very well when the images are not too far apart. Of course, it is still necessary to indicate a number of reference points in the images if the reconstruction has to be located in world. However, in this case much less reference points are required. If the reference points where accurately reconstructed, it is sufficient to indicate them in one image or alternatively on the 3D model. If this is not the case, then it is necessary to indicate them in two images. This interesting not only to reduce interaction but also to enable reconstruction of a close-up sequence where reference points are few or sites where surveying is hard.

3 DENSE SURFACE ESTIMATION

Once the calibration has been obtained, dense stereo matching can be used to recover the detailed surface geometry from the images. This is done using dense stereo matching. The output is a depth map or a point cloud for the requested surface areas together with a map containing an estimate of the depth error for each point. A number of additional features where implemented to allow a flexible and efficient use of this module.

Regions of interest can be defined to speed up the matching process, as well as various regions of disinterest to eliminate unwanted sections or sections that are matched badly (such as trees in front of the scene). These regions can be interactively indicated with the user interface developed in this project.

Using the calibration obtained as described in the previous section, the images can be rectified towards the standard stereo configuration. In this case corresponding epipolar lines are warped to corresponding scanlines in the images. This allows important gains in efficiency. In the case of this tool a standard homography-based stereo rectification algorithm was used. An example is shown in Fig. 4. This approach works well in the case where the camera motion is mostly sideways (i.e. the epipoles are located far away from the image), but would fail otherwise. If one intends to reconstruct from image pairs where the camera mostly moves in the direction of the optical axis, then a more advanced rectification approach (Pollefeys et al. 1999) should be used.

Fig. 4: Rectified image pair

In addition to the epipolar geometry other constraints like preserving the order of neighbouring pixels, bi-directional uniqueness of the match, and detection of occlusions can be exploited. These constraints are used to guide the correspondence towards the most probable scan-line match using a dynamic programming scheme (Cox et al. 1996). The matcher searches at each pixel in one image for maximum normalized cross correlation in the other image by shifting a small measurement window along the corresponding scan line. Matching ambiguities are resolved by exploiting the ordering constraint in the dynamic programming approach (Koch, 1996). The algorithm was further adapted to employ extended neighbourhood relationships and a pyramidal estimation scheme to reliably deal with very large disparity ranges of over 50% of image size (Falkenhagen, 1997). Using the common reference points seen in the two images, the disparity search range is limited to the proximity of the scene. This results in much lower computation times and, in addition, improves the robustness by reducing the possibility for mismatches.

From the disparity estimates the 3D points can be reconstructed through triangulation and the depth can be computed as the distance between the point and the center of projection of the reference view. These depth values are stored in a depth map (an image where at each pixel the depth value is stored. On the left of Fig. 5 a depth map obtained from the image pair of Fig. 4 is shown. In this case darker means closer.

In the context of heritage conservation it is important to also have an estimate of the error that is present on the measurements. Therefore, error estimation was also implemented. Assuming the dense stereo matching to be accurate within e pixels, one can calculate the uncertainty margin on the depth estimate by calculating a new depth for a corresponding point displaced by a distance e. For camera viewpoints that are located close to each other, a small difference in the image yields a large error on the estimate. On the other hand, with cameras more orthogonal to each other the error is much more contained. This error estimate also depends on the distance of the scene to the cameras. On the right side of Fig. 5 a map of error estimates is shown. One can observe that indeed both maps look very similar (up to scale). This is because points further away have more uncertainty than points closer to the cameras. The average error estimate is around 3cm in this case.

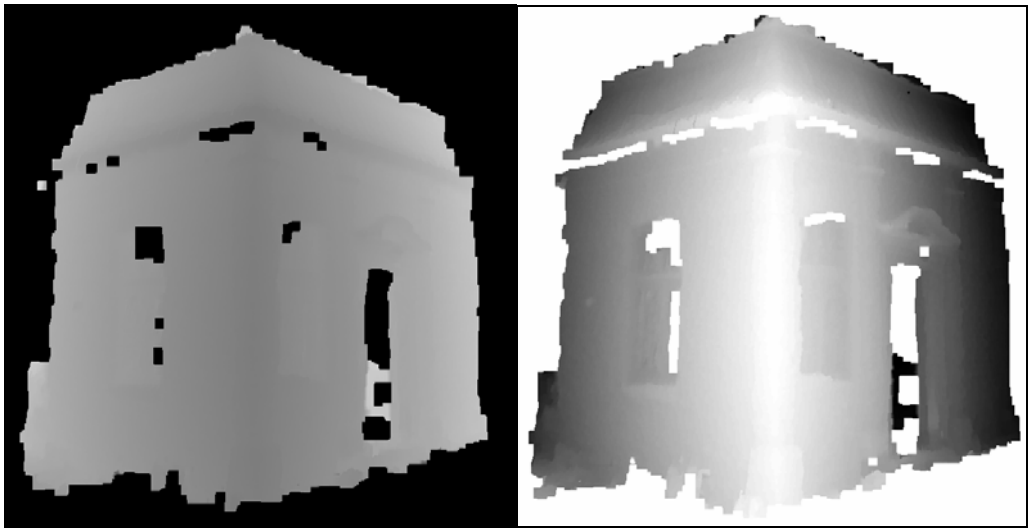

Fig. 5: depth map (left) and map containing error estimate (right)

5 3D MODELLING AND ANALYSIS

In the previous sections both the camera calibration and dense depth maps where computed. This yields all the necessary information to build 3D models. On the left of Fig. 6 a 3D reconstruction consisting of the reprojection of a sub-sampled depth image is shown. In the system developed in the context of our project, this data is then passed on to another tool developed during this project (Lauwers and Li 2001) for integration of the results of different depth maps into a single 3D model. On the right side of Fig. 6 a textured 3D model is shown that was obtained by combining several depth maps and using the original photographs as texture.

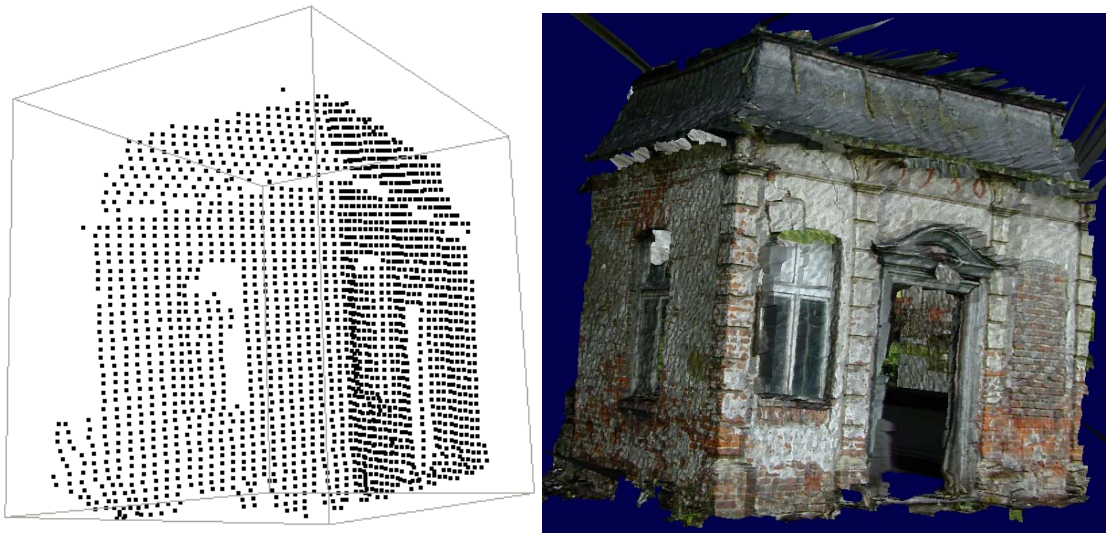

Fig. 6: 3D point reconstruction (left) and textured 3D reconstruction (right)

In the context of conservation it is important that the computed 3D model can also be used to obtain measurements from it. As an example Fig.7 shows a number of isolines automatically computed by the system. These can for example be exported to a CAD package to derive plans from it. More details on this and other applications can be found in (Nuyts et al. 2001) where the third tool of the system is described.

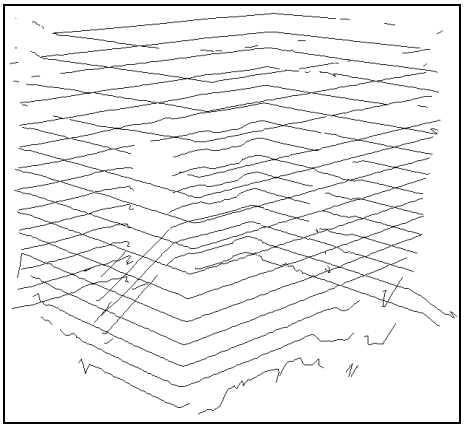

Fig. 7: Isolines computed from the 3D model

5 EXAMPLES

In this section some additional examples are shown. On the left of Fig.8 an image of the reconstruction of a painted vault is shown. The vault is located in the Arenberg castle in Leuven. On this reconstruction isolines are also indicated. In the middle of Fig.8 a part of the inner court of the Arenberg castle is shown. On the right of Fig.8 a 3D model of ancient tower in St-Truiden can be seen.

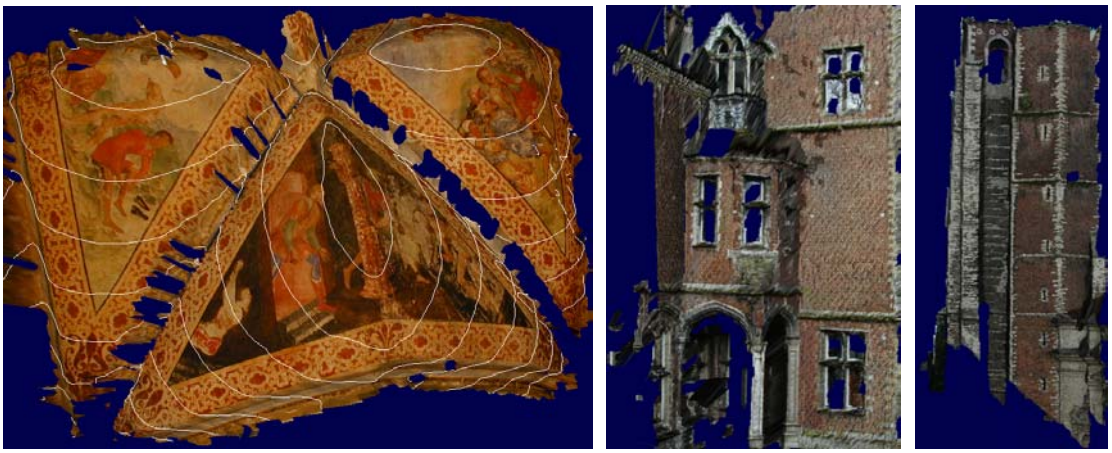

Fig. 8: Some examples of 3D reconstructions obtained by the system

In Fig. 9 a 3D model is shown that was obtained fully automatically from images, without using reference points for calibration. This approach is not yet integrated into the presented tool. Our goal is to also integrate this approach in the 3D modelling tool for conservation that was presented in this paper. This would especially be useful for including high-resolution models of decorated building parts as seen in the figure below. More details on this approach can be found in (Pollefeys et al. 1999a, Pollefeys et al 2000).

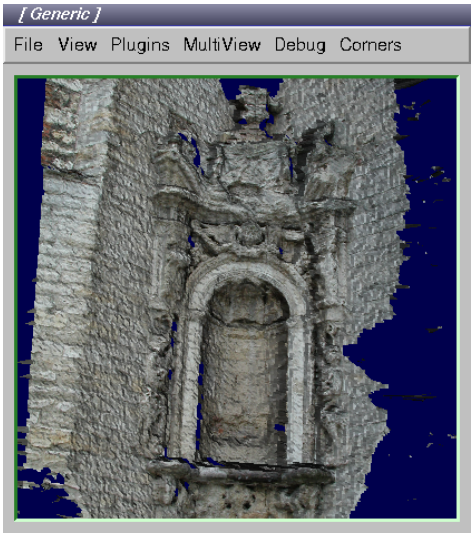

Fig. 9: 3D reconstruction obtained using a fully automatic reconstruction approach

6 CONCLUSION

In this paper a semi-automatic tool for acquiring 3D models of architectural monuments was presented. The tool requires a number of reference points to be measured and the corresponding points to be indicated in the images. The further processing is then fully automatic. First the pose and calibration of the camera for the different views is computed. Then using dense stereo matching techniques and estimate of the object surface is obtained. From this a 3D model is reconstructed. This 3D model can then be analyse and several types of measurements can be extracted from it. Our approach has been validated on a number of different test cases, ranging from architectural artefacts to complete monuments.

ACKNOWLEDGEMENTS

This work was partially funded by the STWW VirtErf and the ITEA BEYOND project of the IWT. Marc Pollefeys is a post-doctoral researcher of the fund for scientific research-Flanders.

REFERENCES

I. Cox, S. Hingorani and S. Rao, 1996. A Maximum Likelihood Stereo Algorithm, Computer Vision and Image Understanding, Vol. 63, No. 3.

L. Falkenhagen, 1997. Hierarchical Block-Based Disparity Estimation Considering Neighbourhood Constraints, Proc. International Workshop on SNHC and 3D Imaging, pp.115-122, Rhodes, Greece.

R. Haralick, C. Lee, K. Ottenberg, M. Nolle, 1994, Review and Analysis of Solutions of the Three Point Perspective Pose Estimation Problem, International Journal Of Computer Vision 13, pp. 331-356.

R. Koch, 1996. Automatische Oberflachenmodellierung starrer dreidimensionaler Objekte aus stereoskopischen Rundum-Ansichten, PhD thesis, Univ. of Hannover, Germany.

B. Lauwers, Q. Li, 2001, The application of reverse engineering technology in model reconstruction for conservation, Proc. CIPA, International Archive of Photogrammetry and Remote Sensing.

K. Nuyts, P. Smars, K. Van Balen, 2001, Layered Geometric Information System, Proc. CIPA, International Archive of Photogrammetry and Remote Sensing.

M. Pollefeys, R. Koch and L. Van Gool, 1999a. Self-Calibration and Metric Reconstruction in spite of Varying and Unknown Internal Camera Parameters, International Journal of Computer Vision, 32(1), pp. 7-25.

M. Pollefeys, R. Koch and L. Van Gool, 1999b. A simple and efficient rectification method for general motion, Proc. International Conference on Computer Vision, pp.496-501.

M. Pollefeys, R. Koch, M. Vergauwen, L. Van Gool, 2000. Automated reconstruction of 3D scenes from sequences of images, ISPRS Journal Of Photogrammetry And Remote Sensing (55) 4, pp. 251-267.

C. Slama, 1980. Manual of Photogrammetry, American Society of Photogrammetry and Remote Sensing, Falls Church, Virgninia, USA.

R. Tsai, 1986. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision, Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 364-374.