
MODELLING THE WORLD IN REAL TIME

Rui Liu, Michael Suppa, Gerd Hirzinger, Darius Burschka
Institute of Robotics and Mechatronics, German Aerospace Center (DLR)
Oberpfaffenhofen, 82234 Wessling, Germany.
E-mail: Rui.Liu@dlr.de

Abstract: This paper presents the automatic landscape modeling system of German Aerospace Center. It focuses on the real time 3D modeling process, which is performed at three scales: the first is large scale, in which the outdoor landscape point clouds generated by stereo-methods with a high resolution stereo-camera on a flying plane will be tiled and processed to a group of simplified meshes with compressed textures automatically; the second is medium scale, in which the indoor scans obtained with a laser range scanner are registered to each other without any marker, then the resulted point-clouds are meshed and simplified; and the third is small scale, in which a hand-guided multi-sensor device captures the detailed information of small objects, which will be merged into the whole 3D models.

1. Introduction

With the rapid development of computer technology and widely spreading of internet, digital landscape becomes a high topic in scientific area recently. Man uses airplane, airship, even satellite to gain the colored 3D information outdoors and utilizes medium- and small-sized devices to capture the scene indoors. Of course, the obtained data amounts are very large. Therefore, how to handle these data, to simplify them, to transform them to 3D models fast and automatically for further applications, becomes a challenging task.

This paper presents a real time automatic modeling process to generate 3D models at three layers (See Figure 1.).

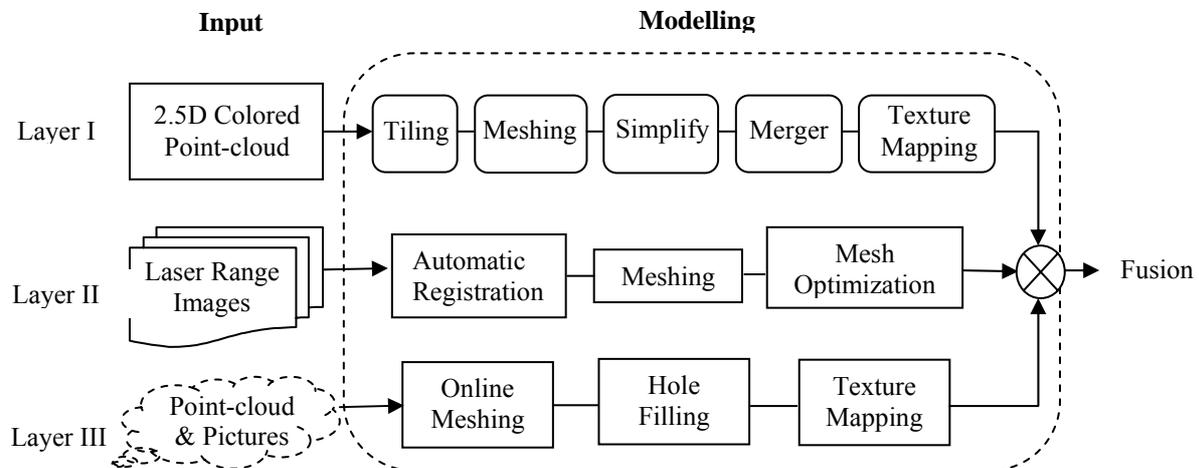


Figure 1: Three-layer real-time world modelling system

The first layer is the at the largest scale, which converts the huge amount of 2.5D colored point clouds of landscape to simplified meshes mapped with compressed pictures. The input data are resulted from stereo matching methods [1] with resolution of 0.2 m. The hardware device on this layer is a HRSC (High Resolution Stereo Camera) [2] assembled on an airplane (See Figure 2a) with sensing distance of 2~50 m. The second layer is at medium scale with sensing distance of 2~50 m and resolution of 0.5~10 cm. It combines the automatic matching, meshing, mesh optimization in one work flow to generate 3D models for the interior of buildings. The 3D point clouds are obtained with an advanced laser range scanner – Z+F Imager [3] (See Figure 2b). And the last layer is at the smallest scale. With a multi-sensory 3D-Modeler [4, 5] (See Figure 2c) developed by robotics institute of Germany Aerospace Center, it digitalizes the small objects in distance from 5~200 cm and resolution of 1~5 mm. 3D-modeling and visualization of these data is performed parallel to the scanning process.

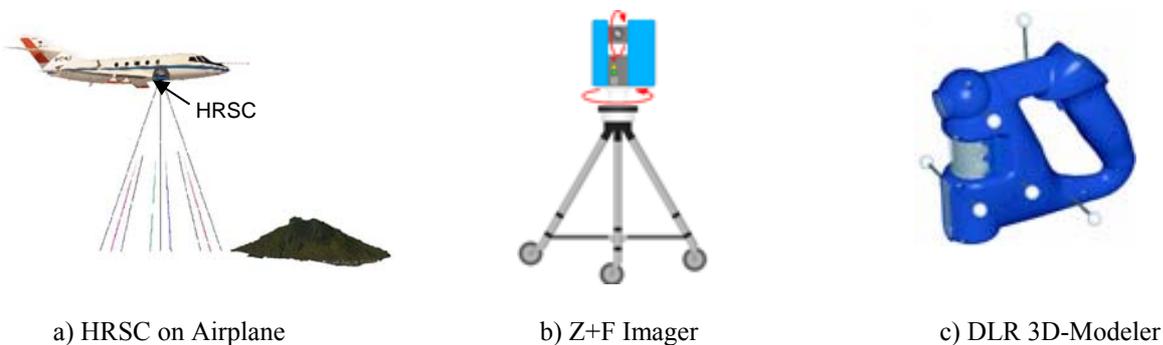


Figure 2: Hardware at three layers

The main contribution of this paper is the exploration in the fast and automatic modeling of the world systematically. This paper focuses on our new development of real-time modelling procedures at the first two layers, for the modelling at the third layer is already published in [6, 7]. Our paper is organized as follows. In section 2, we summarize the existing researches in this area. Then, section 3 describes our main methods at three layers. Their respective experimented models are shown in this section too. After that, the remained problems and possible improvements in the future are addressed in section 4. For the more, the related acknowledgement is announced in section 5. And the end section is referenced materials.

2. Related Work

For the automatic modelling and visualisation of landscape data at the first layer, the principle of LOD (Level of Detail) [8] is widely used. To enhance the 3D roaming velocity, “Buffer-Quadtree” algorithm is introduced by Li [9]. Recently, more and more efforts have been taken to find and reconstruct the man-made and natural structures automatically. For example, buildings, highways, trees etc. [10, 11].

For the modelling of laser range images at middle scale, Allen [12] presented an automated 3D modelling pipeline, which consists of scan registration, surface generation and texture mapping. A set of markers that the scanner can automatically recognize were used to register the scans together. And for the surface generation, they used VripPack [13], which utilises a cumulative weighted signed distance function to integrate groups of aligned range images and employs a modified Marching Cubes algorithm with a lookup table to generate the final

surface. And other notable projects, e.g. modelling of Michelangelo's David [14] and IBM's Pieta project [15], presented also modelling pipelines for large statures, but involved manual pre-aligning of different views to solve the coarse matching problems. Recently, many researchers reported their improvements for fully automatic registration of laser range images. Liu and Hirzinger [16] introduced "Matching Tree" structure to reach a global maximum matching based on local m:n corresponding relationships in a closed form with polynomial runtime. Rabbini and Heuvel [17] used constrained search for finding correspondent objects. Kang and Zlatanova [18] used corner detector to extract feature points from reflectivity images and by construction of a triangle mesh to reduce the search space.

For the modeling of objects at small scale, some smart hand-guided device were developed, e.g. David [19] of TU Braunschweig.

3. Main Method

In this section, we mainly focus on the real time 3D-modellingsprozess at the software aspect, which is performed at three layers: large, medium and small scales. The common character of them is: the input data are 2.5D or 3D point clouds with color information and the outputs are simplified colored mesh. The details of these three methods, the differences between them and the related experiments are showed below.

3.1. Large Scale

The input data of this process are generated with the High Resolution Stereo Camera (HRSC) fixed on an airplane flying over the land and optimised by the photogrammetric pre-processing: correction, refining and stereo-matching. The HRSC camera was originally developed by the Planetary Research Institute of German Aerospace Center for the exploration of the Mars surface. The airborne version HRSC-AX is currently used for capturing earth's landscape and cities from flight altitudes between 1500m to 5000 m.

The input of this 3D-modellingsprozess is the 2.5 D point cloud resulted from the photogrammetric pre-processing: correction, refining and stereo-matching. The input data format is double-Tiff-format: one Tiff-file contains the height-information of every pixel, and the other holds the RGB colour-information. One of these input files is normally several Gigabytes, which overrides the capabilities of fast all of the 3D modelling and visualisation tools. To reduce the data amount and prepare for the online visualisation of the coloured landscape model, we use a real time and fully automatic 3D-modellingsprozess. It is consist of six steps:

1. **Tiling with overlap:** The whole area will be divided in tiles with a proper size. To avoid great distort out of merging of meshes, we set a overlap between neighboured tiles.
2. **Mesh generation:** A triangle mesh will be generated in each tile by connection the shorter diagonal of every rectangle grid.
3. **Mesh simplifying:** An efficient mesh-simplifying will be applied in this step. Geometric characters become more outstanding, for example, edge becomes sharper. And the data amount can be reduced over 90% without distinguish distortion. A modified quadric error metric algorithm is used to contract vertex-pairs with minimal contraction error iteratively [20].

4. **Mesh cut:** Exactly along the middle axis of the overlapping area, every mesh will be cut neatly in this step. That means, the projection of all resulted boundary edges from one cut action should be on a straight line.
5. **Mesh merger:** Due to the 3D property of the meshes, there are tiny gaps between the neighbored meshes. To make the adjacent meshes seamless, we apply an efficient merger algorithm by taking the following “stitch-to-point” and “stitch to edge” actions.
6. **Texture mapping:** The RGB information of each tile will be saved as compressed image format. And the mapping from texture to 3D mesh will be achieved by projection the 3D meshes to the texture coordinate system.

The result is a group of simplified meshes with corresponding textures for an efficient online visualisation.

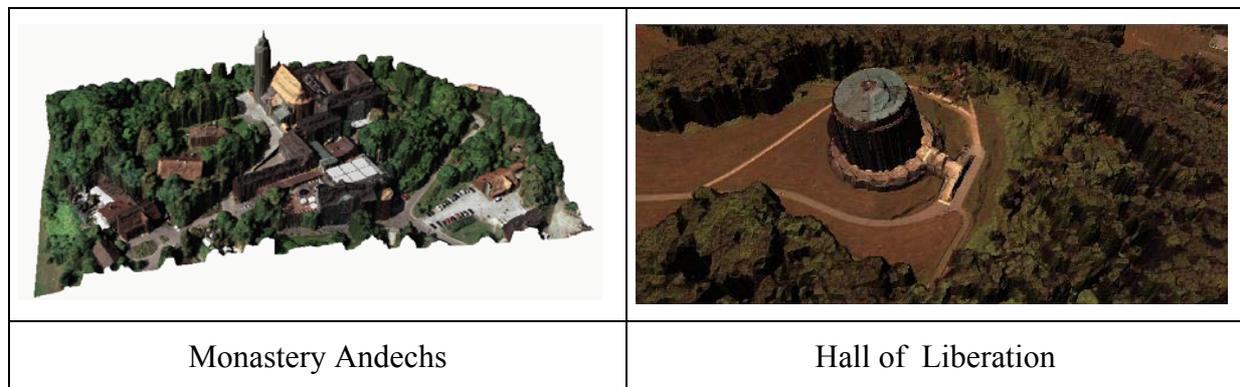


Figure 3: Reconstructed landscapes at the first Layer

On a computer with Intel Core2 Duo Processor E6600 (2.4GHz) and 2048 MB RAM, experiments have been done for the reconstruction of diverse regions in Bavaria, Germany, e.g. city of Kelheim with Befreiungshalle ("Hall of Liberation"), Monastery Andechs, Neuschwanstein Castle etc. The processing time is showed in Table 1.

Region	Area Size (km ²)	Resolution (m)	Input Size n x m (pixels ²)	Output (Triangles)	Runtime (min)
Kehlheim	11.33	0.20	24480 x 11570	3,623,330	26,5
Andechs	0.81	0.15	6000 x 6000	463,333	3,6
Wiekirche	1.44	0.20	6000 x 6000	465,557	3,7

Table 1: Processing time of diverse regions at the first layer

3.2. Medium Scale

On this layer, we use Z+F Imager to get laser range scans inside the building. The sensing distance is from 2m to 50 m with resolution from 5mm to 10 cm. To protect the cultural heritage, no marker should be pasted on the wall. So we applied a marker-free automatic matching process to align different views together. For the resulted point-cloud is normally very large, we applied a similar method as the above modelling process at large scale. The important modules are listed below:

- Automatic Matching
- Mesh Generation
- Mesh Optimisation

3.2.1. Automatic Matching

This is a coarse to fine matching process, which consists of three steps:

- **Coarse matching** solves the pre-alignment problem automatically
- **Fine matching** aligns two views accurately
- **N-view bundle adjustment** is our last step to align multiple views in one coordinate system.

At the coarse matching stage, we applied dynamic programming technique to “matching tree” structure [16] to reduce the runtime complexity from $O(n^4)$ to $\Theta(n^{1.5})$. We called it “Dynamic Matching Tree” algorithm. The basic form to solve the sub-problem of dynamic programming is:

$$M_{i,j} = \max\{M_{i-1,j-1} \oplus w_{i,j}, M_{i-1,j}, M_{i,j-1}\} \quad (1)$$

$M_{i,j}$ is the maximal matching formed by the first i object-nodes (representative points for objects) from the one view $\{X_1, X_2, \dots, X_i\}$ and the first j object-nodes from the other view $\{Y_1, Y_2, \dots, Y_j\}$. And $w_{i,j}$ is the weight of the correspondent pair (X_i, Y_j) . $w_{i,j} = 0$ means, there is no corresponding relationship between X_i and Y_j . The goal of us is to find $M_{I,J}$, if there are I objects in one view and J objects in the other. Then we make a table and file the table from left to right, top to down.

X \ Y	0	1	...	j	...	J
0	0	0	...	0	...	0
1	0					
⋮	⋮					
i	0					
⋮	⋮					
I	0					$M_{I,J}$

Table 2: Table of the Dynamic Matching Tree algorithm

The value of $M_{i,j}$ is chosen as the maximum from the three items: $M_{i-1,j-1} \oplus w_{i,j}$, $M_{i-1,j}$ and $M_{i,j-1}$. The first item $M_{i-1,j-1} \oplus w_{i,j}$ implies, that we should not add the matching weight $w_{i,j}$ of the node paar (X_i, Y_j) to $M_{i-1,j-1}$ directly. Validation to the root-nodes should be done according to the distance and direction conditions described in [16]. Thus

$$M_{i-1,j-1} \oplus w_{i,j} = \begin{cases} M_{i-1,j-1} + w_{i,j} & \text{accepted} \\ \max\{M_{i-1,j-1}, M'_{i-1,j-1} + w_{i,j}\} & \text{rejected} \end{cases} \quad (2)$$

A “split-tree” action is involved, if the correspondent pair (X_i, Y_j) is rejected by the root of $M_{i-1,j-1}$. $M'_{i-1,j-1}$ is the split part from $M_{i-1,j-1}$, which forms a new matching tree with the node (X_i, Y_j) . Details of split-tree action are not the focus of this paper, for it is described in [21].

After coarse matching, we use a modified ICP (*Iterative Closest Point*) to achieve an efficient fine matching by projecting sampling points from one segment to its correspondent segment. Finally, we use bundle adjustment to align multiple views accurately.

3.2.2. Mesh Generation

In this stage, we apply an special designed meshing algorithm to convert these large point-clouds to homogeneous triangle mesh.

The input points and the vertices of the mesh are held in separate point sets, each implemented in a hierarchical data structure that allows fast insertion and search of local point neighborhoods. The structure generates only a small memory overhead, so it is perfectly suited for holding large data sets. Details were described in our previous published papers [6, 7].

3.2.3. Mesh Optimisation

In this step, holes will be filled and the mesh will be simplified.

The reasons of the visible holes in triangle mesh are diverse, which can be categorized into two fields: missing triangles or false triangulation. If there are vertices with false normal or triangles with false triangulating direction, they will be displayed as holes. Therefore, we made our task in two steps: first, filter the illegal triangulation out of the mesh; und then, we fill the holes with a recursive algorithm regarding of the 3D relationship between vertices and edges. Very large holes with thousands of vertices lying on the different planes are filled successfully with our recursive algorithms [22].

For mesh reduction, we use a modified quadric error decimation algorithm to reduce the points and triangles in the mesh. The topology of the mesh remains well and the reduction can arrive over 95% (See Figure 4).

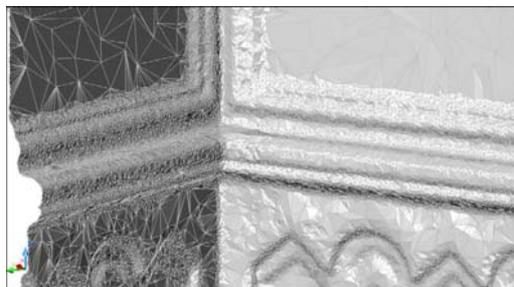
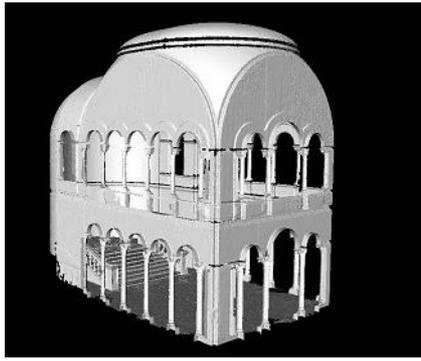
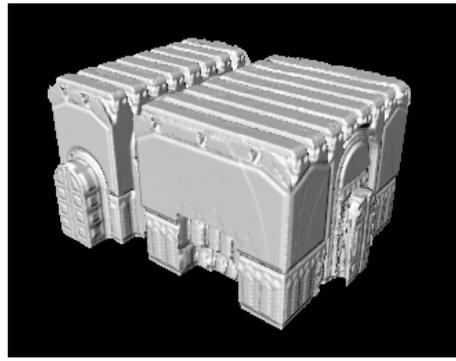


Figure 4: Result from mesh reduction : details of a room corner

Various experiments has been done on the reconstruction of the Neuschwanstein Castle, Church of Seefeld in Bavaria, etc. (See Figure 5).



a) Throne Hall



b) King's Working Room

Figure 5: Partial models of the Neuschwanstein Castle

3.3. Small Scale

We use the hand-guided multi-sensory 3D-modeler of our institute to capture the models of small object with resolution of 1~5 mm. During the scanning, a 3D mesh of the sensed object is generated online and the texture will be mapped to the model after one button-click. Scan position is tracked by a real time IR-optical tracking system. Details of the online meshing and texture mapping is presented in papers [6,7] and filling of holes is described in [22].

4. Conclusion

For the further development, firstly we should verify the model fusion algorithms on objects coming from different scales. And secondly, an improvement of texture mapping technique at middle scale is required. Finally, a self oriented hand-guided device is the goal in the future.

5. Acknowledgement

Thanks to Bernhard Strackenbrock, Johann Heindl and Dr. Heiko Hirschmueller for the useful discussions, delivery of related input data and testing of our implementations.

References:

- [1] H. Hirschmüller, F. Scholten, G. Hirzinger. "Stereo Vision Based Re-construction of Huge Urban Areas from an Airborne Pushbroom Camera (HRSC)", in Lecture Notes in Computer Science: Pattern Recognition, Proceedings of the 27th DAGM Symposium, 30 August - 2 September 2005, Vienna, Austria, Volume 3663, pp. 58-66
- [2] F. Wewel, F. Scholten, and K. Gwinner. "High resolution stereo camera (HRSC)-multispectral 3D-data acquisition and photogrammetric data processing," Canadian J. Remote Sensing, 26:466-474, 2000.
- [3] C. Froehlich: Aktive Erzeugung korrespondierender Tiefen- und Reflektivitätsbilder und ihre Nutzung zur Umgebungserfassung. PhD Thesis, Fakultät fuer Elektrotechnik und Informationstechnik, TU Munich, 1996.
- [4] K. Strobl, E. Wahl, W. Sepp, T. Bodenmueller, M. Suppa and G. Hirzinger: The DLR Hand-guided Device: The Laser-Stripe Profiler. In Proc. of the IEEE International Conference on Robotics and Automation (ICRA), 1927-1932, New Orleans, 2004.

-
- [5] M. Suppa, S. Kielhoefer, J. Langwald, F. Hacker, K. H. Strobl, and G. Hirzinger: The 3D-Modeller: A Multi-Purpose Vision Platform, in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2007), Rome, Italy, 2007.
- [6] T. Bodenmueller and G. Hirzinger: Online Surface Reconstruction From Unorganized 3D-Points for the DLR Hand-guided Scanner System. In Proc. of the 2nd International Symposium on 3D Data Processing, Visualization and Transmission, Thessaloniki, Greece, 2004.
- [7] G. Hirzinger et al. "Photo-Realistic 3D-Modelling - From Robotics Perception Towards Cultural Heritage," International Workshop on Re-cording, Modelling and Visualization of Cultural Heritage, Ascona, Switzerland, 2005.
- [8] D. Luebke, M. Reddy, J. Cohen, A. Varshney, B. Watson, and R. Huebner: Level of Detail for 3D Graphics, Morgan Kaufmann Pub-lisher, 2003.
- [9] Q. Li, X. Zuo, B. Yang: 3D Highway Landscape Modeling and Visualization," The International Archives of Photogrammetry and Remote Sensing, Vol. 34, Part (5), Kunming, 2002.
- [10] A. Gruen, O Kübler, P. Agouris (Eds.): Automatic Extraction of Man-Made Objects from Aerial and Space Images (I/II), Birkhäuser Verlag, Basel, 1995/1997.
- [11] M. Gerke, B.M. Straub and A. Koch: Automatic Detection of Buildings and Trees from Aerial Imagery Using Different Levels of Abstraction, Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung, Band 10, Eckhardt Seyfert (Hrsg.), pp.273-280, 2001.
- [12] P. Allen, A. Troccoli and P. Blaer: Automating the 3D Modeling Pipeline, International Symposium on the CREST Digital Archiving Project, Tokyo, 2005.
- [13] B. Curless and M. Levoy: A Volumetric Method for Building Complex Models from Range Images, in Proc. SIGGRAPH 96, pages 303–312, ACM Press, 1996.
- [14] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk: The digital Michelangelo project: 3D scanning of large statues, in Proc. Siggraph 2000, pages 131–144, 2000.
- [15] F. Bernardini, H. Rushmeier, I. M. Martin, J. Mittleman, and G. Taubin: Building a digital model of Michelangelo's Florentine Piet'a, IEEE Comp. Graph. and Applic., 22(1):59–67, Jan/Feb 2002.
- [16] R. Liu and G. Hirzinger: Marker-free Automatic Matching of Range Data, in Proc. Of 2nd Panoramic Photogrammetry Workshop, Berlin, 2005.
- [17] T. Rabbani, Frank van den Heuvel: Automatic Point Cloud Registration Using Constrained Search for Corresponding Objects, in Proc. of 7th Conference on Optical 3-D Measurement Techniques, Part 1, pp. 177-186, Vienna, 2005
- [18] Z. Kang and S. Zlatanova: Automatic Registration of Terrestrial Scan Data Based on Matching Corresponding points from Reflectivity Images, in Proc. of Urban Remote Sensing Joint Event (6th URS/4th Urban), Paris, 2007
- [19] S. Winkelbach, S. Molkenstruck and F. M. Wahl: Low-Cost Laser Range Scanner and Fast Surface Registration Approach, in Proc. of DAGM 2006
- [20] R. Liu, D. Burschka, G. Hirzinger and B. Strackenburg: Real Time Fully Automatic 3D-Modelling of HRSC Landscape Data, in Proc. of Urban Remote Sensing Joint Event (6th URS/4th Urban), Paris, 2007
- [21] R. Liu, D. Burschka and G. Hirzinger: A Novel Approach to Automatic Registration of Point Clouds, In Proc of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Barcelona, 2007.
- [22] R. Liu, D. Burschka and G. Hirzinger: On the Way to Water-tight Mesh, in Proc. Of ISPRS International Workshop "3D-ARCH 2007", Zurich, Switzerland, 2007.