



Fully automatic 3D digitization of unknown objects using progressive data bounding box

Souhaïel Khalfaoui, Antoine Aigueperse, Ralph Seulin, Yohan Fougerolle,
David Fofi

► To cite this version:

Souhaïel Khalfaoui, Antoine Aigueperse, Ralph Seulin, Yohan Fougerolle, David Fofi. Fully automatic 3D digitization of unknown objects using progressive data bounding box. Three-Dimensional Image Processing (3DIP) and Applications II, Jan 2012, San Francisco, California, United States. pp.829011, 10.1117/12.909164 . hal-00683925v2

HAL Id: hal-00683925

<https://hal.science/hal-00683925v2>

Submitted on 12 Jul 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fully automatic 3D digitization of unknown objects using progressive data bounding box

Souhaïel Khalfaoui, Antoine Aigueperse, Ralph Seulin, Yohan Fougerolle, and David Fofi

Le2i Laboratory, UMR-CNRS 5158, University of Burgundy,
71200 Le Creusot, France

ABSTRACT

The goal of this work is to develop a complete and automatic scanning system with minimum prior information. We aim to establish a methodology for the automation of the 3D digitization process. The paper presents a method based on the evolution of the Bounding Box of the object during the acquisition. The registration of the data is improved through the modeling of the positioning system. The obtained models are analyzed and inspected in order to evaluate the robustness of our method. Tests with real objects have been performed and results of digitization are provided.

Keywords: 3D Digitization, Automation, Automatic scanning

1. INTRODUCTION

The three dimensional (3D) model of an object is widely used in many applications such as industrial inspection, computer games, and augmented reality. When the model is artificially generated by the user, the technique is called computer assisted designed (CAD). The 3D model can also be obtained by digitization. From the digitized model, the real object can be inspected and/or analyzed. With one acquisition, it is not possible to cover the whole geometry of the object. Thus, the scanner has to be moved around it to acquire all the missing data. In practice, the viewpoints (scanner positions) are selected by a specialized human operator. The quality of the final result strongly depends on the selected viewpoints and thus on the human expertise. Therefore, it is necessary to define a new digitization strategy objects while minimizing the impact of the human factor. This independence can be obtained by automating the digitization process. The automation of the 3D acquisition process requires the knowledge of the different components of the measuring system (scanner, positioning system, etc.). Automatic scanning can be decomposed into two phases: the determination of the different poses (position and orientation) for the set *sensor-object* and the sensor trajectory generation from previous poses to achieve the next optimal location. Currently, some commercialized systems allow object scanning using a set of pre-programmed scanner positions (teach-in process). In such case, the views are not optimized in function of the object geometry. It is more appropriate to say: pre-programmed 3D measurement process.

*Scott et al.*¹ provided a complete survey paper on the view planning problem. Depending on the nature of reasoning, the developed methods to compute the best views can be classified into two main approaches: volumetric methods and surface methods.

1.1 Volumetric methods

Volumetric approaches are based on the analysis of the workspace state and use voxelization to represent occupied and empty areas. *Abidi*² adapted the concept of entropy in information theory to the problem of volumetric modeling of an unknown object. A utility function is proposed to qualify the amount of acquired data at each stage of the positioning and the acquisition process. The local maximum of this function determines the *Next Best View* (NBV). One advantage is that there is no restriction on the shape, its size, and the location of the object.

*Reed*³ proposed an automatic method in two phases: a modeling phase by constructing a coarse surface mesh from fixed positions, and a planning phase that analyses the model to calculate the next sensor pose. *Reed* determined the visibility volume which is the volume of the space in which a sensor has a direct view on a particular target.

Determining the NBV requires the consideration of three constraints (sensor visibility constraint, occlusion constraint, and sensor positioning constraint) which are associated with different volumes.

*Massios and Fisher*⁴ introduced a quality criterion in addition to the visibility criterion. A voxel representation is used to label voxels as: empty, seen, unseen, or Occlusion Plane. The application of the sensor positioning constraints to the coarse model of the object determines the areas for possible views. This representation is then projected onto the visibility sphere and the quality criterion is applied to evaluate each view. The direction with the best result is selected.

An extension of this algorithm was proposed by *Munkelt et al.*⁵ and is composed of two phases. The first step is the initialization, during which the algorithm maximizes the number of visible voxels respecting a given quality. The second phase maximizes the number of occluded and visible voxels. The stopping criterion is based on a low amount of new information obtained from scanning in the direction of the NBV.

*Connolly*⁶ used an *octree* structure to subdivide the space. The determination of the NBV requires an update of this structure that contains four types of node: parent, empty, occupied, and unseen. Using this representation, two algorithms were proposed: the Planetarium and the Normal algorithm. The Planetarium searches for the area with maximum unseen voxels and is able to detect occlusions caused by the *occupied* voxels. The Normal algorithm performs a counting of the normals of the *unseen* voxels that have *empty* neighbors, resulting on a viewing vector.

1.2 Surface methods

The Surface methods use the surface representation of the model to obtain cues about the regions of missing data. *Maver and Bajcsy*⁷ used the edges to compute the occluded area of each view which was described as a polygon and assumed to be planar. Each edge of the occluded area was classified as active or inactive. A histogram of the occluding region and the camera angle was then used to select the next scanning direction. *Maver et al.*⁸ used the *max-min principle* as heuristic. The NBV must ensure a minimum of acquired data and acquire a maximum of new information.

The method proposed by *Kok-lim Low*⁹ takes into account several constraints for the view planning (Acquisition and Quality constraints). From a first acquisition, a partial model of the scene is created. The acquired surfaces are labelled *real*. The determination of NBV involves the assessment of the cost of a view regarding various constraints. If all these views have a score below a predefined threshold, the view planning is completed. The view with the highest score is chosen as the NBV.

The idea proposed by *Wenhardt et al.*¹⁰ relies on a probabilistic approach based on Kalman filter. The NBV is determined through the minimization of some optimality criteria such as: D-Optimality, E-Optimality, and T-Optimality.

*Yuan*¹¹ introduced the concept of Mass Vector Chains (MVC) to characterize closed models. A MVC is defined as the sum of the normal vectors of all surface patches weighted by its projected area in the normal's direction. It was demonstrated that for closed model, the sum of the MVC, named MVS for Mass Vector Sum, is null. The existence of holes in the model causes the MVS to be not null. The opposite of the vector \overrightarrow{MVS} is used as the NBV to complete the model. The algorithm stops only if the object is closed, hence complete. Thus, this method can not be used in the case of complex objects (holes, narrow cavities, etc.). *He and Li*¹² have added a termination condition using a threshold variation of the surface. The sensor characteristics such as desired resolution, scanning resolution, field of view, and working distance were also included to optimize the scanning process.

Borangui et al.^{13,14} use pre-programmed path to obtain a first model. The second phase detects the concave regions in the model and defines a new path to scan the hidden areas. The view planning defines the robot configuration and the turntable angle by computing the inverse kinematics of a system with 7 DoF.

*Loriot*¹⁵ established rapid methods for view planning to ensure maximum coverage of the object's surface. The planning is done online, since no information about the object and its shape is available. Digitization process is composed of two steps :

- First phase: *MVC method*. The NBV is the view in the direction opposite to the normal average of previous views (figure 1.b). A new acquisition which does not provide additional information indicates the end of this phase.
- Second phase: Support for non-digitized areas. *Loriot* identifies a "Normal to every hole" by estimating a mean vector orthogonal to the curve describing the hole's contour and indicating the direction in which the scanner must be positioned.

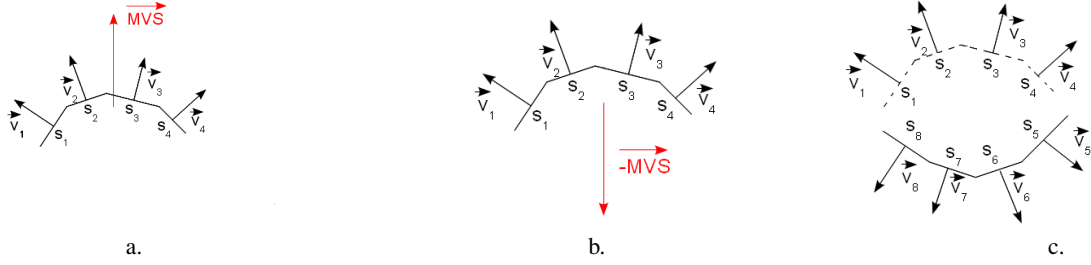


Figure 1. Nest Best View estimation:¹⁵ a. Mass Vector Sum determination b. NBV determination c. New acquired data

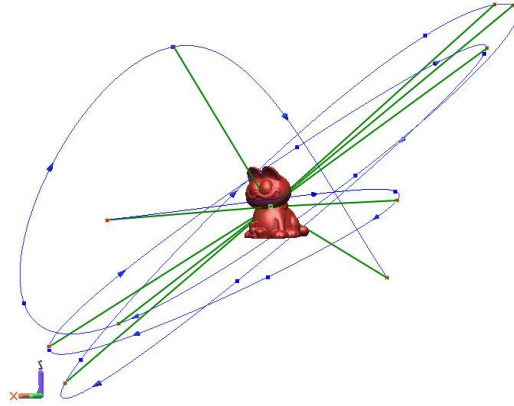


Figure 2. Scanner trajectory during the first phase of the algorithm of Loriot.¹⁵

This method is very expensive in robot's movements, especially in the first phase of the algorithm as the method of the MVC supports many round trips around the object (figure 2). Furthermore, this approach does not take into account the distribution of views in the work space since some view points are particularly close to each other.

2. PROPOSED METHOD

We presented two main approaches: volumetric and surface based methods. From this overview we can conclude that no system is actually able to digitize autonomously a whole unknown object. Some of the most advanced solutions are at most partially automated. Hence, the aim is to define a new and simple strategy that leads to automatic selection of view points and decreases the acquisition time and the numbers of readjustments.

How can we define *the scanner pose (position, orientation)*? Under the assumption that the object is "close" enough to its bounding box, one can arbitrarily define scanning poses from the bounding box of the object, as illustrated in figure 3. The target points are the centers of the faces, the midpoints of the edges and the vertice. The scanner is positioned at a distance d which respects the intrinsic parameters of the scanner. Each phase of the acquisition process is then composed of three steps:

- Bounding box and pose identification.
- Positioning of the scanner and acquisition.
- Updating the dimensions of the bounding box.

The first pose is determined arbitrarily and is the top view. The bottom view is physically unreachable because of the positioning system. The distribution of the view points around the object is uniform. Therefore, using a mid point subdivision, the number of acquisitions is equal to 25.

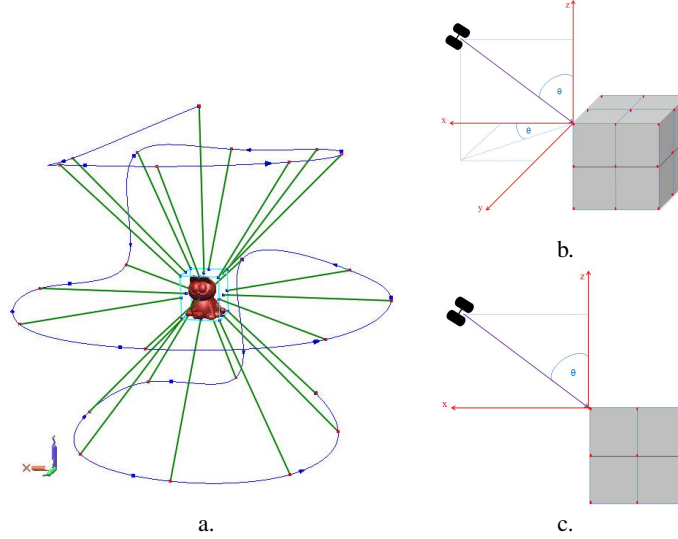


Figure 3. Determining the position of the scanner using a subdivision of faces in 4: a. Scanner trajectory during the acquisition process b. The target point is a vertex c. The target point is the midpoint of an edge. The red dots are the different possible orientations.

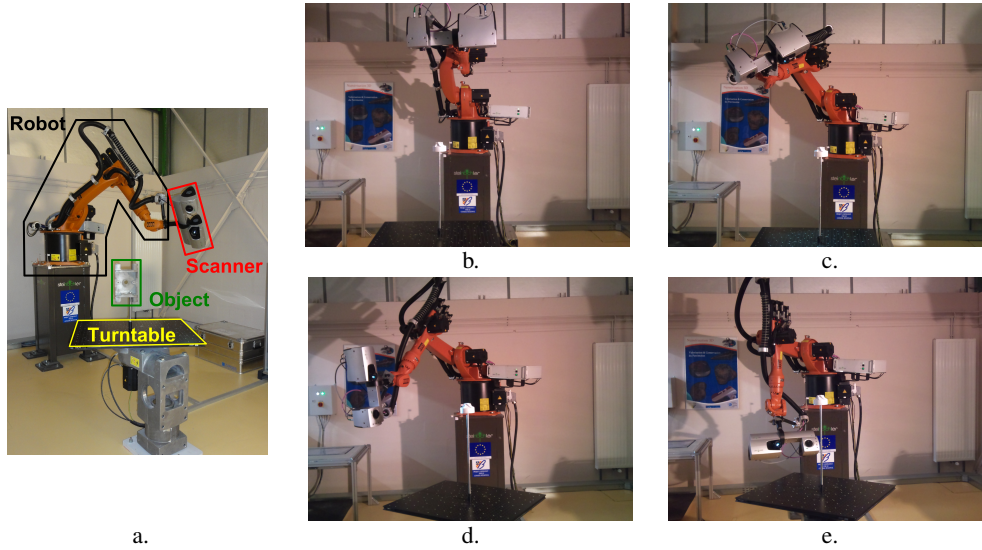


Figure 4. Robot position during the scanning process: a. Automation work cell b. Top view (first scan) c. First level view (high) d. Second level view (middle) e. Third level view (down).

This set of view points can be increased considering a fine distribution. The acquisition process is based on an arbitrary number of positions and for each step the new data are independent of what was previously acquired. Note that the only information which links two successive acquisitions is the size of the bounding box and not the geometry of the scanned part, which avoids expensive computations on aquired data and leads to a simple and intuitive algorithm.

To obtain a complete 3D model, the previously acquired data and the new ones must be registered. In the literature, many iterative algorithms have been proposed.¹⁶ In order to decrease the computational time, the views can be pre-aligned using the coordinates of the positioning system. A fine registration step is then performed on the pre-aligned meshes to minimize errors in the overlapping regions.

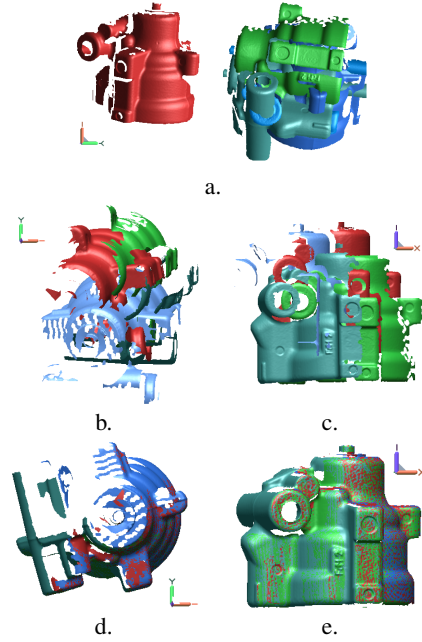


Figure 5. Results of the registration: a. Acquired data b. Coarse registration (XY plane) c. Coarse registration (XZ plane) d. Fine registration (XY plane) e. Fine registration (XZ plane).

3. EXPERIMENTAL RESULTS

The proposed method has been implemented on a robotic cell (see figure 4.a) which is composed of a robotic arm is a KUKA KR16 manipulator, a 6 DoF arm with 6 rotational joints and one turntable. The sensor is a CometV 3D scanner manufactured by Steinbichler GmbH with a 1.4 Mega Pixels camera and a set of lens that yields a working volume of $480 \times 360 \times 250$ mm with a working distance of 850 mm. During the acquisition process, the robot moves around three main positions (see figure 4). At each step, the new acquired data are registered with previous scans. In our application, a positioning error is generated because of the accumulation of errors along the estimation of some parameters. To evaluate the influence of the coarse registration, we use data obtained from four views (exemplified in figure 5.a). As shown in figures 5.b and 5.c, after the coarse registration all the point clouds have the same orientation, but there is a gap between them. This spacing is due to the estimation error of the transformation between the robot flange and the scanner origin. A fine registration is performed by ICP algorithm on the pre-aligned meshes. As shown in figures 5.d and 5.e, the fine registration is ensured. The above tasks are repeated until the 25th acquisition. A hole-filling step is performed to obtain the final model. Then, several operations are performed onto the point cloud model (meshing, cleaning, smoothing, etc.). Inspection task is performed and a deviation map is obtained. As shown in figure 6, the obtained model (figure 6.b) is then compared to a CAD model of the object (figure 6.c) if it exists or to the manually obtained model in order to study the robustness of our approach. Figure 6.d illustrates an example of a simulated 3D errormap. The values of the over-thickness are displayed in a red to blue color scale that allows a rapid identification of zones where there is no over-thickness or negative over-thickness. The obtained mean error is around $0.03mm$ which is a very good result and means that our method allows a precise reconstruction of 3D models.

To evaluate the surface coverage, since a reference model is not available, the areas of the two meshes (with and without holes) are compared in order to compute a value of surface coverage which is around 97 %. The final models obtained can be seen in figure 7. Practically full surface coverage was achieved for all the test objects. These results also show the improvements when compared to manual scanning that were reported to take, for the simplest of the objects, at least 100 minutes. The improvements with respect to a manual scanning were estimated to be between 2x to 5x faster (around 17 minutes).

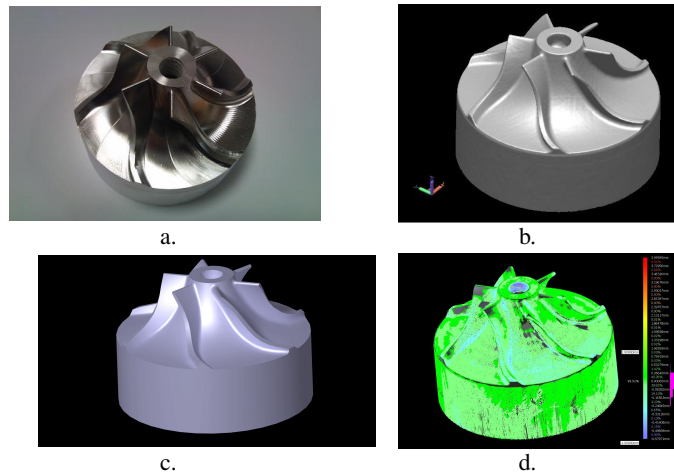


Figure 6. Object scanning and analysis: a. Real object b. Acquired data c. CAD model d. Deviation map.

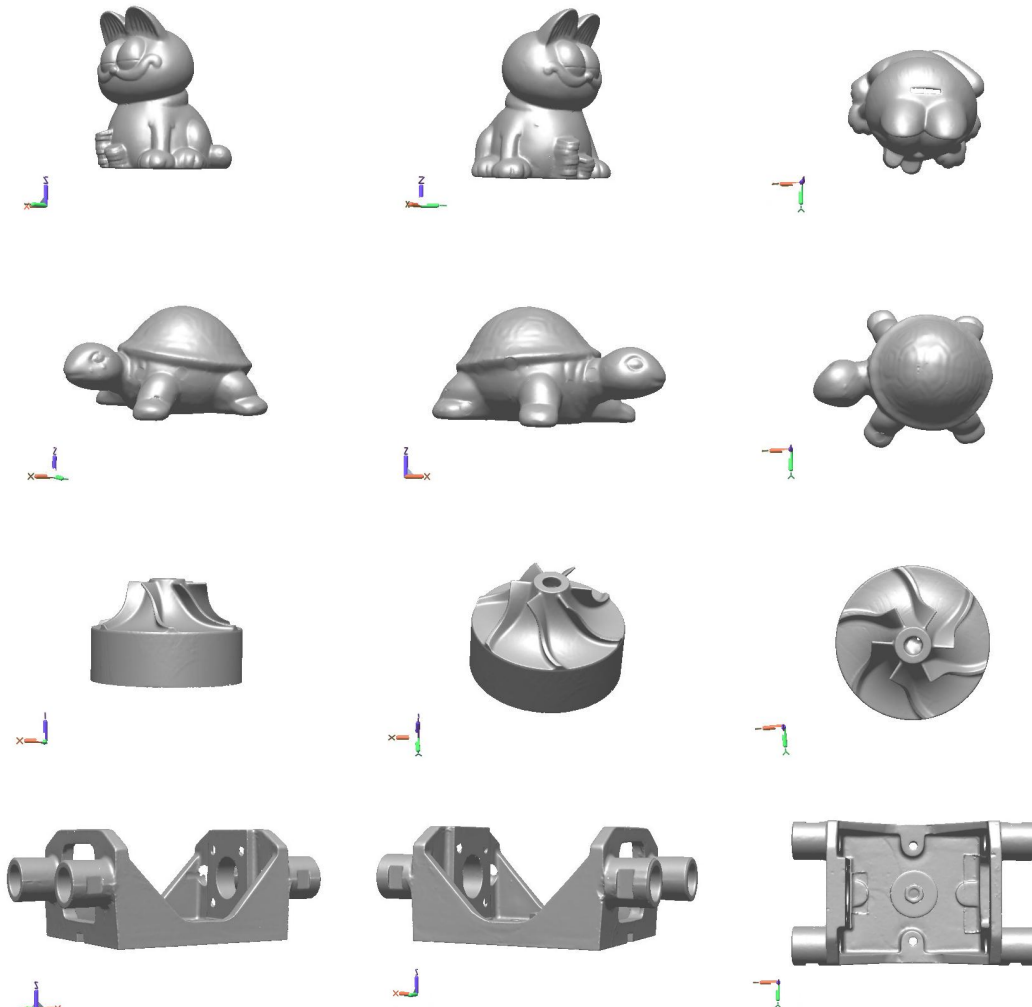


Figure 7. Some digitized objects.

4. CONCLUSIONS AND FUTURE WORKS

In this paper we presented a new digitization strategy based on an arbitrary number of acquisitions. At each step, the acquired data are merged in order to adjust the dimensions of the bounding box. This method is characterized by its simplicity and efficiency and leads to very good results (in term of surface coverage). The acquisition of a complete 3D model of an object is fast and takes around 20 minutes. The human operator decides only of the number of acquisitions and their distribution around the object. This method requires several heuristics (number of acquisition, arbitrary trajectory definition, etc). Furthermore, it does not take into account the physical limitations of the acquisition system such as geometric limitations and singularities of a robotic system since it computes only a unique NBV at each iteration. The works in progress are related to the management of physical constraints of our system. The aim is to define a new strategy that adapts the number of acquisitions to the geometry of the object. Hence one can define new criteria to optimize the selected views. The scanner trajectory will be defined automatically using the view points previously calculated to maximize the surface coverage and their distribution will be adapted to the shape of the object. Then, the only role of the human operator will be to start the acquisition process.

ACKNOWLEDGMENTS

This work was done within the framework of the LE2I laboratory and financially supported by the *Regional Council of Burgundy*. The authors would like to thank *Mr. Mickaël Provost*, Head of *Vecteo company* (www.vecteo.com), for his technical support and effective collaboration.

REFERENCES

- [1] Scott, W. R., Roth, G., and Rivest, J.-F., "View planning for automated three-dimensional object reconstruction and inspection," *ACM Computing Surveys* **35**(1), 64–96 (2003).
- [2] Abidi, B., "Automatic sensor placement," in [*SPIE Conf. on Intelligent Robots and Computer Vision XIV*], **2588**, 387–398 (October 1995).
- [3] Reed, M. K., *Solid Model Acquisition from Range Imagery*, PhD thesis, Columbia University (1998).
- [4] Massios, N. A. and Fisher, R. B., "A best next view selection algorithm incorporating a quality criterion," in [*BMVC*], (1998).
- [5] Munkelt, C., Denzler, J., and Kuhmstedt, P., "Incorporation of a-priori information in planning the next best view," in [*Proceedings of the International Workshop Vision, Modeling, and Visualization (VMV)*], 261–268 (November 22-24 2006).
- [6] Connolly, C. I., "The determination of next best views," in [*Proceedings of the International Conference on Robotics and Automation*], **2**, 432–435, IEEE Computer Society, St. Louis, Mo., USA (March 25-28 1985).
- [7] Maver, J. and Bajcsy, R., "Occlusions as a guide for planning the next view," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**(5), 417–433 (1993).
- [8] Maver, J., Leonardis, A., and Solina, F., "Planning the next view using the max-min principle," in [*CAIP '93: Proceedings of the 5th International Conference on Computer Analysis of Images and Patterns*], 543–547, Springer-Verlag, London, UK (1993).
- [9] Low, K. L., "An adaptive hierarchical next-best-view algorithm for 3d reconstruction of indoor scenes," tech. rep., Proceedings of 14th Pacific Conference on Computer Graphics and Applications (Pacific Graphics) (2006).
- [10] Wenhardt, S., Deutsch, B., Angelopoulou, E., and Niemann, H., "Active visual object reconstruction using d-, e-, and t-optimal next best views," in [*CVPR*], 1–7 (2007).
- [11] Yuan, X., "A mechanism of automatic 3d object modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**(3), 307–311 (1995).
- [12] He, B. W. and Li, Y. F., "A next-best-view method with self-termination in active modeling of 3d objects," in [*IROS'06*], 5345–5350 (2006).
- [13] Borangiu, T., Dogar, A., and Dumitrache, A., "Constraints based motion planning for an automatic flexible laser scanning robotized platform," in [*Automation, Quality and Testing, Robotics, 2008. AQTR 2008.*], **2**, 65–69, IEEE International Conference. (may 2008).

- [14] Borangiu, T., Dogar, A., and Dumitrache, A., "A heuristic approach for constrained real time motion planning of a redundant 7 dof mechanism for 3d laser scanning," in [*Information Control Problems in Manufacturing. 13th IFAC Symp. on Information Control Problems in Manufacturing - INCOM'09*], (2009).
- [15] Lorient, B., Seulin, R., and Gorria, P., "Non-model based method for an automation of 3d acquisition and post-processing," *Electronics Letters on Computer Vision and Analysis* **7**(3), 67–82 (2008).
- [16] Salvi, J., Matabosch, C., Fofi, D., and Forest, J., "A review of recent range image registration methods with accuracy evaluation," *Image Vision Comput.* **25**, 578–596 (May 2007).