Object Surface Reconstruction from One Camera System

Radim Dvorak[1], Martin Drahansky[1], Filip Orsag[1]
[1] *Brno University of Technology, Faculty of Information Technology, Bozetechova 2, 61266 Brno, Czech Republic*
{idvorak, drahansky, orsag}@fit.vutbr.cz

Abstract

In this paper, there is introduced an approach to surface reconstruction of an object captured by one grayscale camera with a small resolution. The proposed solution expects a rectangular grid being projected onto the object and the camera capturing the situation from position different to the grid projector position. The crucial part of the method is exact detection of the grid. The structure of the grid is identified by the centers of the inner space between its lines. The reconstruction process itself is based on a simple math of perspective projection of the captured image. Due to the small resolution of the image some errors arise during object surface reconstruction. We proposed a correction of the calculated coordinates, which is a simple one-dimensional function depending on the distance from the camera. The method performs very well as the results in conjunction with the precision evaluation indicate at the end of the paper.

Keywords: *surface reconstruction, projected grid, structured light, image processing, camera calibration.*

1. Introduction

There are many solutions of a 3D surface reconstruction of an object. They differ from each other and they are typical for specific application. The first solution is based on a method, which constructs the object surface from multiple camera views. Such an example can be an early work from R. Mohr et al. [1], where the points are reconstructed from multiple uncalibrated images. In the work of P. Lavoie et al. [2] the reconstruction is done from a pair of images. On the other hand in the work of F. Pedersiny [3] the multiple camera system was calibrated in order to improve accuracy.

The next solution deals with reconstruction of the surface from one single image. Method presented by A. Saxena et al. [4, 5] is an example of such solution. Distance of the points in the image is calculated using supervised learning. Some other approaches (I.K. Park et al. [6], E. Garcia and J.L. Dugelay [7]) use known apriori information about the shape of the object, such as some significant features of the modeled head.

Another solution of the surface reconstruction is characterized by usage of projected structured light onto the involved object. Published methods deal with the reconstruction from multiple images (P. Lavoie et al. [8], H. Saito and T. Kanade [9]) and from one still image (J. Pages et al. [10, 11], F. Ababsa et al. [12], A. Dipanda and S. Woo [13]). Problems that arise in reconstruction using the structured light are caused by the opposite sides of the object, which become literally invisible. Hence, in this case, several images of the object from different angles must be taken. In order to merge the partial shapes together, some feature points must be set. They can be point out by hand (E. Garcia and J.-L. Dugelay [7]) or automatically (L. Shi et al. [14], K. Lee et al. [15]).

Method proposed in our paper deals with reconstruction of the object surface using the projected grid. It can be characterized as a low cost and minimum dimension solution. The system consists of one grayscale camera with a low resolution (currently 640x480) and a square-structured light source such as a laser module. The known parameters are position of the camera and spacing of the projected grid in a given distance.

The reconstruction itself requires exact grid detection. Because of the low resolution of the camera the centers of grid squares (segments) are detected instead of the lines themselves. Next stage is to define the neighboring segments and the last step is reconstruction of the surface itself.

The paper is organized as follows. All phases of the algorithm are described in section 2. Due to the low resolution of the image, some errors arise during the calculation, especially while the segments are being extracted. Therefore, the calculated positions of the points are corrected by defined correction function that is depicted in section 3. In section 4, some testing examples that indicate the capability of the method are shown. The performed experiments were submitted to the accuracy evaluation and the results are stated in section 5. Some final remarks and future development are proposed in section 6.

2. Algorithm of the surface reconstruction

The scene is defined by a laser grid generator with a structured grid light, camera with given resolution, position and field of view, and an object. It is also necessary to know the fan angle of the laser generator. The laser itself projects a circle in the center of the grid, which can be used as a starting point for the calculation.

The algorithm can be divided into three main parts. The first stage of the algorithm is image preprocessing. The image preprocessing is done to remove the noise and mark the significant parts (lines in this case). Next stage of the algorithm is grid detection. In this stage, the segments are detected and their neighbors are estimated. In the last stage, positions of the points in 3D space are calculated. Details are provided in the following subsections.

2.1. Image preprocessing

We use Gaussian filter (currently with core size of 3×3) to remove the noise and to prepare the image for an adaptive thresholding. The core of the filter is defined as follows:

$$G(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}, \qquad (1)$$

where σ is standard deviation, x and y are variables (coordinates of the pixel in the image, in this case). The standard deviation is computed from the size of the core using this equation:

$$\sigma=\left(\frac{n}{2}-1\right)*0.3+0.8, \qquad (2)$$

where n is the size of the core.

The adaptive thresholding is used for initial grid detection. To certain extent, it eliminates impact of non-uniform illumination in the individual parts of the image.

The average value of pixel neighborhood is computed for each pixel and this value is set as the threshold for the given pixel. If brightness of the pixel is lower than the computed threshold, it is set to 0 (black), otherwise the pixel is set to 255 (white). This process is applied to all pixels in the image, whereas the estimation of the neighborhood size is still open. The best results have been achieved with the size of the neighborhood set to 11 pixels.

However, this fact is given by the image resolution (640×480) and by the average span size of the grid projected on to the object.

2.2. Segment grid detection

Detection of a circle lying in the middle of the projected grid is the first step of the grid detection. The circle is supposed to be the largest cluster of the white points, which comes up from the preprocessing. Therefore, the circle, which may be deformed, is detected as the residual connected segment of white pixels after the sequence of erosions.

In order to determine proper count of the erosions, the average thickness of line is predicted as count of white-black crossings in vertical direction divided by count of crossed lines. Due to this process, all the lines as well as the majority of the grid corners disappear. From the remaining white areas, the largest one is the central circle. You can see an example of the detection in Fig. 1.

Various factors are used for the grid detection in the image consisting only of black and white pixels (i.e. binary image).

The first of them is grid regularity. The grid is composed of parallel lines in both vertical and horizontal directions. These lines are continuous in their whole length. The gaps between lines (segments) are of a square shape in case of a perpendicular grid projection to plane projection surface.

If an object is laid in the space between projection surface and projected grid, the projected lines will be deformed by the object. However, continuity of the grid lines will not be affected (to a certain extent). This fact is used for the computation of neighboring segments.

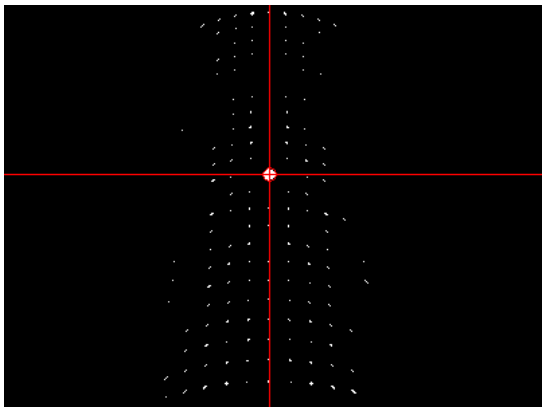

Fig. 1. Detected ellipse and its center in an image processed by a sequence of erosions.

The neighboring segments are close to each other and their joint sides are continuous (see Fig. 2). A simple detection of neighboring segments could be deduced from these facts using the position of corners.

Fig. 2. The grid after application of the adaptive thresholding with marked sides of the segments.

The considered detection requires determination of corners of the segments. The projected grid holds the vertical and horizontal direction of projected lines in the camera image, which can be used to determine the corners. For the computation of the individual corners, the following equations are used:

$$C_{LU} = \min(x_i + y_i) \quad (3)$$

$$C_{LB} = \min(x_i - y_i) \quad (4)$$

$$C_{RU} = \min(-x_i + y_i) \quad (5)$$

$$C_{RB} = \min(-x_i - y_i), \quad (6)$$

where C_{LU} is the upper left corner of the segment, C_{LB} is the bottom left corner of the segment, C_{RU} is the upper right corner of the segment, C_{RB} is the bottom right corner of the segment, x and y are coordinates of the pixel and index i is going through all the pixels in the given segment.

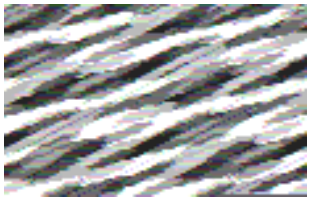

Fig. 3. An example of the detected segment corners and respective centers.

It is necessary to determine the centers of all the segments for the resulting reconstruction. Computation of the center is simply given as an average position of all the pixels in the individual segment:

$$S = \frac{1}{N}\sum_{i=1}^{N} P_i, \quad (7)$$

where P_i is the i^{th} pixel of the segment, N is the number of pixels and the sum of the pixels is defined as sum of the individual components of the coordinates x and y. Detected corners and respective centers of the segments are shown in Fig. 3.

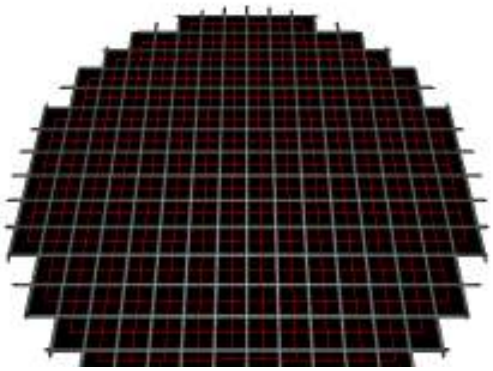

Fig. 4. . Detected segments (black) and their neighbors.

Using the detected corners of the segments, it is possible to determine the mutual neighbors. The neighbor to the right of the segment A is determined as a segment B, which has the corner C_{LU} nearest to the corner C_{RU} of the segment A. The same could be applied to the corner C_{RB} of the segment A and the corner C_{LB} of the segment B. Other neighboring segments could be determined in the same way using the corresponding corners. Result of the segment detection of a plane is shown in Fig. 4.

2.3. Reconstruction of the surface

The camera, which captures the scene with the grid projected on to an object, generates rectangular images. These images are perspective projection of the 3D scene to a 2D rectangle. Because the camera K is inclined towards the projection surface about a defined angle α, there is a trapezoid projected on the projection surface due to the perspective deformation (see Fig. 5).

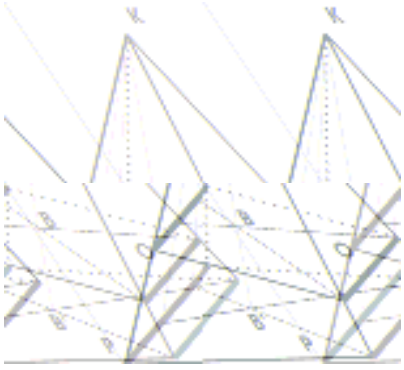

Fig. 5. . A perspective projection of the camera image O to the plain P and one of the points B' projected as a point B.

Due to the fact that lines are projected again as lines, and knowing the lines are parallel, in accordance with Fig. 5, we are able to compute original position of the point B projected as the point B' to the image captured by the camera.

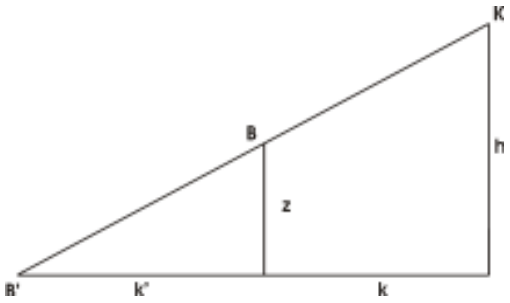

Fig. 6. . A principle of the calculation of the z coordinate of a point.

If some object is placed between the projection surface and camera, each grid point projected on this object will appear above the projection surface. Due to the fact that the camera captures the projection surface sideward the given point will appear on another place in the captured image as well. Computation of the coordinates of all the respective points on the projection surface according to Fig. 5 and the fact that the position of the projected points on the projection surface without an object is known provide us enough information to compute position of the point in the 3D space by determining *z* coordinate of the given point projected on the object.

Fig. 6 shows the principle of computation of the z coordinate of the point in the space. If we know the camera height h and the plane coordinates of the point B then we can easily use the similarity of triangles to compute height z of the point:

$$z = h\frac{k'}{k}. \qquad (8)$$

Meaning of the symbols is clear from the Fig. 6. To begin the reconstruction process it is necessary to setup a source point which the rest of the calculation will be derived from. The segment center lying near the detected middle-grid circle is the best candidate, because the true position of the circle is known (when there is no object placed). Thanks to the known mutual neighbors of all the segments, the other true grid point positions are established recursively from the source point. Consequently, all the necessary information is known and the every surface point can be calculated in a way described earlier.

3. Error correction

When the *z* coordinate of the point is calculated, the error arises due to the low angle between the line depicted by camera and the point and the normal vector of the grid plane. Therefore, the error is directly influenced by the distance of the point from the camera - the closer the point is the bigger the error is.

The error is in form of the polygonal function as shown in Fig. 7, where the approximated curve is drawn as well. Correct value of the z coordinate is 10 mm. Therefore, if we know the error, we are able to correct the calculated values to the exact ones.

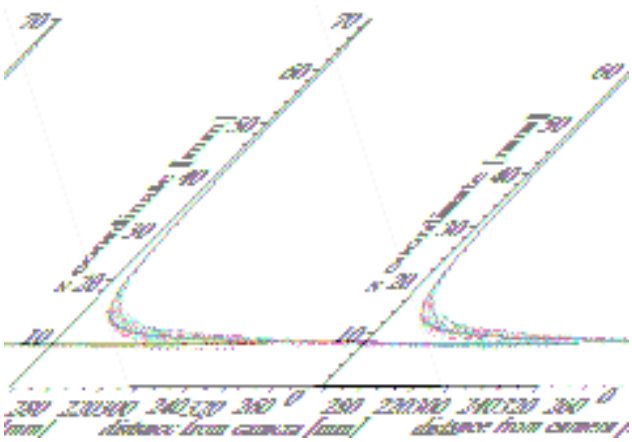

Fig. 7. . An example of the calculated positions and its approximated function.

The proper form of the corrected function is given as:

$$\mathrm{corr} = \mathrm{A}\frac{\mathrm{x} - \mathrm{Cam}_z + \mathrm{B}}{\mathrm{C} * \mathrm{x} - \mathrm{D} * \mathrm{Cam}_z + \mathrm{E}}, \qquad (9)$$

where x is distance of the point from the camera, Cam_z is height of the camera and A, B, C, D are coefficients that depend mainly on the grid spacing.

4. Experimental results

We did experiments with artificial scenes to find out performance of the method. We chose such objects, from which correctness of the calculated coordinates can be easily determined.

The following settings were used in the experiments. The camera was placed 200 mm above the projection plane. Spacing of the segments of the grid is 5.5 mm which is relatively dense according to the camera distance and its low resolution.

4.1. Plane

The first experiment was done with one plane placed perpendicularly to the laser view vector. The plane was positioned 10 mm above the projection plane. The detected grid with detected center circle is shown in Fig. 8. The outer segments are ignored because they are incomplete.

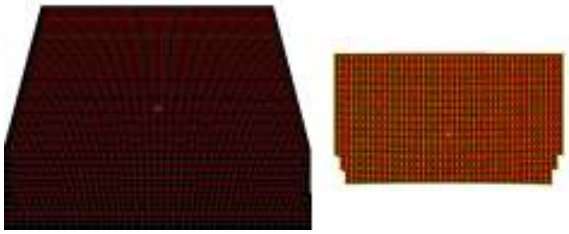

Fig. 8. Detected grid with detected center circle of a plane on the left and reconstructed plane in 3-D space on the right.

The reconstructed surface of the plain is shown in Fig. 8. You can clearly see that there are some minor deformations. The plane is slightly deflected, especially at the bottom of the image.

4.2. Cylinder

The second experiment was done with a cylindrical object placed in the scene with the same settings as in the case of the plane. The projected grid with the circle was successfully detected and the result is shown in Fig. 9 on the left.

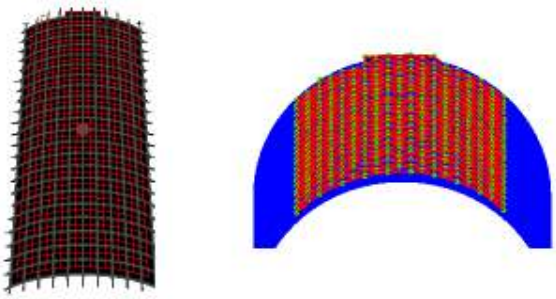

Fig. 9. Grid as detected on a cylindrical surface on the left and reconstructed cylinder surface with densified point grid on the right.

The reconstruction of the cylinder was done with a grid twice denser than in the previous example. The corners of the segments were used in order to do so. The points were calculated simply from the neighboring segment corners. Result together with the original cylinder is shown also in Fig. 9.

4.3. Head

The model of artificial head was chosen for the third experiment. The detected grid and the reconstructed surface in front view and in profile are shown in Fig. 10. The original model of the head is also shown in figure in order to point out the correctness of the calculations.

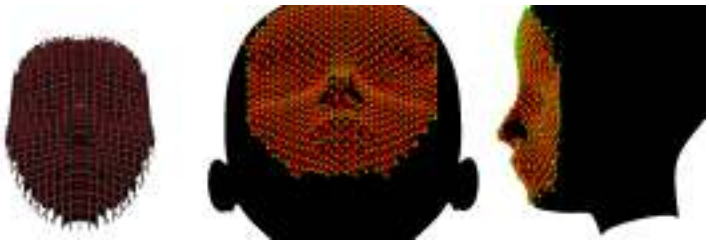

Fig. 10. The detected grid on the head (left) and the surface reconstruction (center, right).

It is crucial for the calculation to properly detected the grid structure and set the proper mutual segments. For more complex objects, such as the head, it is harder to achieve it. Therefore we have implemented the self-repairing routines for the grid structure. These are correction of the mutual segments based on information of the distance and the character of the border lines. If some errors are detected in the structure during the repair, the false connections are deleted. It can be done within some tolerances because it is not necessary to have all the segments connected with all its neighbors. The result is almost proper connected grid pattern.

The minor errors in this experiment are evident around the nose and mouth areas. It is caused mainly by the small denseness of the projected grid.

5. Evaluation of the results

Although it is important to know accuracy of the reconstruction process, this information is missing in the most of the papers dealing with this topic. In case there are some results presented some other important facts are absent. These are mainly the camera (image) color space and resolution, the distance of the camera from the projection plane and, in case of the structured light, density of the grid. We present all the above information once more in Table 1.

Table 1. Settings of the system.

Image	grayscale
Image resolution	640 × 480
Approx. object distance	190 mm
Grid resolution	50 × 50
Grid fan angle	31.5°

The calculations were done for the plain experiment and the experiment with a cylindrical object. We used objects with known analytical description to be able to precisely calculate the error of our method.

At the beginning we introduce the metrics we used. The first one is an average error of all the calculated points (ERRAVG). The second one is a standard deviation of the error (ERRDEV). The maximum error is marked as ERRMAX. All of them are shown in Table 2.

It is clear from the error evaluation that in spite of the low resolution of the camera and relatively dense grid, the errors are relatively small. The differences between the presented two cases are caused mainly by nature of the scene. The plane has a narrow surface and the grid projected on it is almost not deformed, therefore the overall error is smaller and vice versa.

Table 2. The error calculated for the given experiments.

Exp.	ERRAVG (mm)	ERRDEV (mm)	ERRMAX (mm)
Plain	0,579	0,420	1,787
Cylinder	1,175	0,521	2,162

Another reason causing the error is quality of the correction function, which is one-dimensional and depends only on the distance of the camera. We suppose that better correction of the errors would be achieved if a two-dimensional correction function were used.

6. Conclusion

We proposed and implemented a method of surface reconstruction of an object using one camera and a projected laser grid. The important restriction was use of a low resolution grayscale camera. Because the projected structured grid is quite dense for the given resolution, the inner parts of the grid are used for the purpose of the structure detection too. The crucial part of the method is proper connection of the grid. To do so some minor correction algorithms were applied.

The error evaluation shows the weak point of the method, which is accuracy of detection of the segment centers. Small error in this phase (in order of pixel fragments) essentially influences precision of the calculation of the z coordinate. Following the simplicity, we used simple correction function with one variable – distance of a point from the camera. Despite the simplification, the resulting errors are within acceptable limits.

In the future we are going to propose more sophisticated correction function in three-dimensional space. We are preparing testing of the method with real scenes as soon as a hardware prototype of the system will be available.

7. Acknowledgements

This research is supported by the following two grants: "*Security-Oriented Research in Information Technology*", MSM0021630528 (CZ) and "*Information Technology in Biomedical Engineering*", GA102/09/H083.

8. References

[1] Mohr, R., Quan, L., Veillon, F.: Relative 3D Reconstruction Using Multiple Uncalibrated Images, The International Journal of Robotics Research 14, 1995, pp. 619 – 632.

[2] Luvoie, P., Ionescu, D., Petriu, E.: 3D reconstruction using an uncalibrated stereo pair of encoded images, In: Proceedings of International Conference on Image Processing, 1996, pp. 859 - 862 .

[3] Pedersini, F., Sarti, A., Tubaro, S.: Multi-Camera Acquisitions for High-Accuracy 3D Reconstruction, In: Proceedings of the European Workshop on 3D Structure from Multiple Images of Large-Scale Environments, 1998, pp. 124 - 138.

[4] Saxena, A., Chung, S.H., Ng, A.Y.: 3-D Depth Reconstruction from a Single Still Image, International Journal of Computer Vision 76, 2008, 53 – 69.

[5] Saxena, A., Sun, M., Ng, A.Y.: Make3D: Learning 3D Scene Structure from a Single Still Image, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 , 2009.

[6] Park, I.K., Zhang, H., Vezhnevets, V., Choh, H.-K.: Image-based Photorealistic 3-D Face Modeling, In: Proceedings of Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004, pp. 49 – 54.

[7] Garcia, E., Dugelay, J.-L.: Low cost 3D face acquisition and modeling, In: Proceedings of International Conference on Information Technology: Coding and Computing, 2001, pp. 657 – 661.

[8] Lavoie, P., Ionescu, D., Petriu, E.: A high precision 3D object reconstruction method using a colorcoded grid and NURBS, In: Proceedings of International Conference on Image Analysis and Processing, 1999, pp. 370 – 375.

[9] Saito, H., Kanade, T.: Shape Reconstruction in Projective Grid Space from Large Number of Images, In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999.

[10] Pages, J., Salvi, J., Matabosch, C.: Implementation of a robust coded structured light technique for dynamic 3D measurements, In: Proceedings of International Conference on Image Processing, 2003.

[11] Pagès, J., Salvi, J.: Coded light projection techniques for 3D reconstruction, J3eA 4, 2005.

[12] Ababsa, F., Roussel, D., Mallem, M.: Structured light 3D free form recovering with sub-pixel precision, Machine Graphics & Vision International Journal 12, 2003, pp. 453 – 476.

[13] Dipanda, A., Woo, S.: Towards a real-time 3D shape reconstruction using a structured light system, Pattern Recognition 38, 2005.

[14] Shi, L., Yang, X., Pan, H.: 3-D Face Modeling from Two Views and Grid Light, Image Analysis and Processing – ICIAP 2005, 2005, pp. 994 – 1001.

[15] Lee, K., Wong, K., Or, S., Fung, Y.: 3D Face Modeling from Perspective-Views and Contour-Based Generic-Model, Real-Time Imaging 7, 2001, pp. 173 - 182.

Authors

Radim Dvorak is a Ph.D. student in the Department of Intelligent Systems at the Faculty of Information Technology, Brno University of Technology, Czech Republic. He received a master degree in information technology program from the Brno University of Technology on subject of deformable bodies' dynamics. He is interested in modeling, simulation, biometric systems and computer graphics. For more information – see please http://www.fit.vutbr.cz/~idvorak/.

Martin Drahanský graduated in 2001 at the Brno University of Technology, Faculty of Electrotechnics and Computer Science in the Czech Republic and simultaneously at the FernUniversität in Hagen, Faculty of Electrotechnics, Germany. He achieved his Ph.D. grade in 2005 at the Brno University of Technology, Faculty of Information Technology in the Czech Republic. Now he works as associate professor at the BUT FIT, DITS. His research topics include biometrics, security and cryptography, artificial intelligence and sensoric systems. For more information – see please http://www.fit.vutbr.cz/~drahan/.

Filip Orság graduated in 2001 at the Brno University of Technology, Faculty of Electrotechnics and Computer Science in the Czech Republic. He achieved his Ph.D. grade in 2004 at the BUT FIT in the Czech Republic. Now he works as assistant professor at the BUT FIT, Department of Intelligent Systems. His research topics include biometrics, security and cryptography, artificial intelligence and robotics. For more information – see please http://www.fit.vutbr.cz/~orsag/.