

# Design of 3D Scene Scanner for Flat Surface Detection

A. Lipnickas, K. Rimkus and S. Sinkevičius

**Abstract** This chapter describes the design of a 3D space scene scanning system built from a 2D laser scanner merged with a CCD colour camera; it also presents an algorithm for flat area detection in a 3D point cloud. For that purpose, the RANdom SAMple Consensus (RANSAC) search engine has been adopted for flat area segmentation and planes detection. Due to a fact that human made planes are limited in size, we have proposed data filtering by comparing averaged point triangulation normals to the selected plane normal. The experiments have shown good results for an analysed environment segmentation with the applied angle variation measure up to  $\pm 25^\circ$ . The applied variation threshold allowed to segment flat planes areas considering surface curvedness.

## 1 Introduction

Registration of three-dimensional data is essential for machine vision, object recognition, motion control and robot navigation applications. The main purpose of a visual sensing system is only to detect the presence of any obstacles; a more complex purpose of these machine vision systems is object detection and recognition. There are several techniques to perform sensing and measuring operations, but, depending on the technology used, they can be grouped into passive and active

---

A. Lipnickas · K. Rimkus (✉) · S. Sinkevičius

Department of Control Technology, Kaunas University of Technology, Kaunas, Lithuania  
e-mail: keatas.rimkus@gmail.com

A. Lipnickas  
e-mail: arunas.lipnickas@ktu.lt

S. Sinkevičius  
e-mail: saulsink@gmail.com

sensor systems (Surgailis et al. 2011; Shacklock et al. 2006; Balsys et al. 2009). A passive sensor relies upon ambient radiation; an active sensor, however, illuminates the scene with radiation (often a laser beam or structured light) and determines how this emission is reflected.

Active laser scanning devices are known as LIDAR systems and available for 2D and 3D scene measuring. The operation principle of the LIDAR is based on measuring active signal time-of-flight (TOF) (Shacklock et al. 2006; Wulf and Wapner 2003; Surmann et al. 2001; Scaramuzza et al. 2007) later, the TOF is converted into a distance. To obtain 3D data, the beam is steered through an additional axis (tilt) to capture spherical coordinates  $\{r, \theta, \varphi\}$ : range, pan, tilt}. There are many examples on how to implement such systems (Himmelsbach et al. 2008): rotating prisms, polygonal mirrors, etc. As commercial 3D LIDAR systems are very expensive, many researchers convert commercial 2D laser devices into 3D ones by introducing an extra axis, either by deflecting the beam with an external mirror or by rotating a complete sensor housing (Surmann et al. 2001; Scaramuzza et al. 2007; Klimentjew et al. 2009). For some specific applications an adherent drawback of these systems is the absence of colour information on the points measured. Colour information would allow to detect all kinds of obstacles as well as occupancy of free areas.

When the speed of a 3D scene registration depends on the electromechanical properties of the scanning device, 3D data processing mostly depends on the scanned area and the density of the points measured. Usually, manipulation with 3D data is a very time consuming process, and it requires a lot of computing resources, because three dimensional scenes or objects can consist of thousands to millions of measured points.

Over the last decade, despite the development of new scanning devices, there has been growing interest in the development of new methods which would allow to reduce the number of constructive primitives without a visible loss of the geometrical form or shapes of the objects scanned (Joochim and Roth 2008; Aguiar et al. 2010). Curvature is one of the most important characteristics of the 3D surface and it is an indicator of the sharpness of a 3D object. Non ridge surfaces or sectors of flat surfaces are mostly non informative for the recognition, however, they are scanned with the same measuring point density. These areas are redundant and can be eliminated with the goal of to reducing the amount of data, economizing on the use of computing resources and preserving geometrical accuracy at the same time.

In this chapter, we propose a 3D range scanner composed of a 2D laser scanner with an extra gear supplying the third degree of freedom as well as a charge-coupled device (CCD) camera for colour information perception. The advantages of the technique proposed in this chapter are the following: usage of a calibration technique to couple the CCD camera with the 3D laser scanner as well as application of the RANSAC method for non-informative plane detection and descriptive 3D primitives reduction.

This chapter is organized as follows. In the next section the state of the art is given, followed by a detailed description of the system setup. Section 4 provides the results of the plane detection. Section 5 concludes the chapter and gives an outlook for the future research.

## 2 State of the Art

The idea of reducing the number of the constructive primitives of a 3D object has gained an increasing interest in today's computer vision, pattern recognition and image processing research field. In Hinker and Hansen (1993) the authors have proposed a method that combines coplanar and nearly coplanar polygons into a larger polygon which can be re-triangulated into fewer simpler polygons. The geometric optimization described in Hinker and Hansen (1993) performs best on geometries made up of larger objects; for example, isosurfaces generated from three dimensional hydrodynamics simulations and/or 3D range scanners. Kalvin and Taylor (1996) have presented a simplification method called "superface". The superface, like the method mentioned above, simplifies polygons by merging coplanar faces and then triangulating the perimeter of these merged faces. The merging procedure is controlled by considering an infinite number of possible approximating plane solutions. The merging of a new face into current superfaces is stopped when a solution set of approximating planes disappears. Garland et al. (2001) have presented a method which can be applied not only to surface simplification tasks, but also to the detection of collision and intersection. This method is based on iterative clustering, i.e., pairwise cluster merging. Such a merging technique produces a hierarchical structure of clusters.

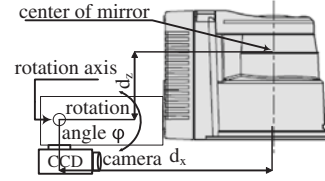
Another technique for the representation of highly detailed objects is presented by Lee et al. (2000). They have defined the domain surface using subdivision surfaces which are applied as descriptors for the representation of smooth surfaces. The displaced subdivision surface consists of a control mesh and a scalar field that displaces the associated subdivision surface locally along its normal vectors. The authors of Lee et al. (2000) have demonstrated that this method offers significant storage savings.

3D object detection algorithms are often divided into two groups: *model-based* and *model-free*. Model-based approaches try to solve detection and classification simultaneously by fitting models to the data. The novelty of the technique presented in this work is based on adoption of the RANSAC technique for detecting big flat areas based on a *plane-model* with redundant points by incorporating information of constructive primitives (colour and/or tangential normal). The described technique has been successfully applied to a scene consisting of various 3D objects to eliminate flat areas such as walls, flooring, boards etc.

## 3 Experimental Setup

In our work, we use a 2D scanning laser range finder UBG-04LX-F01 for area scanning with laser class 1 safety of wavelength 785 nm (Fig. 1); a servomotor was used for supplying rotational movement and a video camera for supplying colour information of a visible scene.

**Fig. 1** Schematic of the 3D sensor



The scan area of the 2D laser scanner is a  $240^\circ$  semicircle with the maximum radius of 4 m. The pitch angle is  $0.36^\circ$  and the sensor outputs the distance measured at every point (682 steps). Accuracy ranges are: in range of 0.06–1 m:  $\pm 10$  mm, and in range of 1–4 m: 1 % of distance (Hinker and Hansen 1993). The servomotor Hitec-HS-422HD with step angle  $0.02^\circ$  and torque 3.0 kg cm is used to supply an additional coordinate  $\varphi$  (see Fig. 1).

The designed system integrates a digital camera (type FFMV-03MTC-60), mounted just below the central axis of a laser range finder. Because of two independent systems used for colour 3D scene scanning, they had to be calibrated to match the same measured and sensed points. The calibration procedure is presented in the next section.

The accuracy of 3D data colour mapping mostly depends on the accuracy of camera calibration. To estimate the camera's parameters we have applied Jean-Yves Bouguet's well-known and widely-used Camera Calibration Toolbox (Bouguet 2006) in our work.

The most widely used model of the camera is a pinhole model (Joochim and Roth 2008; Serafinavičius 2005). The equation of the camera model is (1):

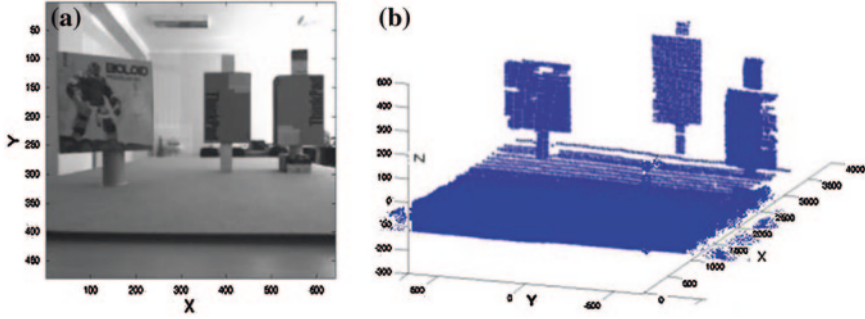
$$\begin{bmatrix} u \\ v \\ 0 \end{bmatrix} = \frac{f}{z} \cdot \begin{bmatrix} k_u & 0 & \frac{u_0}{f} \\ 0 & k_v & \frac{v_0}{f} \\ 0 & 0 & \frac{1}{f} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad (1)$$

where  $u, v$  are the image plane coordinates,  $x, y, z$  are the world coordinates,  $k_u, k_v$  are scale factors along the axes of pixel coordinates,  $u_0, v_0$  are the pixel coordinates of the principal point (orthogonal projection of the optical centre on the image plane),  $f$  is the focal length.

3D laser range finder data are expressed in spherical coordinates. The sensor model can be written as (Wulf and Wapner 2003):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} c_i c_j & -c_i c_j s_j & s_i & c_i d_x + s_i d_z \\ s_j c_i & c_i & 0 & 0 \\ s_i & s_i c_i s_j & c_i & s_i d_x + c_i d_z \end{bmatrix} \begin{bmatrix} \rho_{ij} \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad (2)$$

where,  $c_i = \cos(\varphi_i)$ ,  $c_j = \cos(\theta_j)$ ,  $s_i = \sin(\varphi_i)$ ,  $s_j = \sin(\theta_j)$ ,  $\rho_{ij}$  is the  $j$ -th measured distance with corresponding orientation  $\Theta_j$  in the  $i$ -th scan plane, which



**Fig. 2** Undistorted image (a) and 3D scanned scene (b)

makes the angle  $\varphi_i$  with the horizontal plane (Fig. 1). The offset of the external rotation axis from the centre of the mirror in the laser frame has components  $d_x = 90$  [mm] and  $d_z = 20$  [mm]. The  $[x, y, z]^T$  are the coordinates of each measured point relative to the global frame (with its origin at the centre of the rotation axis, the  $x$ -axis pointing forward and the  $z$ -axis toward the top).

In the case of mapping  $[x, y, z]^T$  point with its colour, it is possible to add depth value ( $z$ ) to the visual image ( $u, v$ ) and, conversely, it is possible to add point colour information to the measured data points ( $x, y, z$ ). The only difficulty here is to find the corresponding formula which would map respective points in both measuring systems.

In order to map camera points with a laser scanner system, the tangential and radial distortion of the camera had to be corrected. The calibration and parameters of the internal camera model were determined according to the instruction given in Bouguet (2006). Camera calibration results (with uncertainties) are as follows:

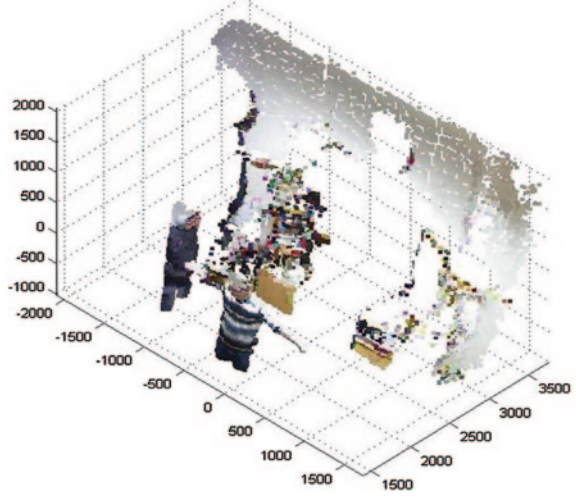
- Focal Length:  $[f_{cx}, f_{cy}] = [490.23 \ 488.58] \pm [0.72 \ 0.71]$ ;
- Principal point:  $[a_{u0}, a_{v0}] = [261.76 \ 233.23] \pm [0.49 \ 0.51]$ ;
- Distortion:
 
$$d_c = [-0.36235 \ 0.144 \ 0.001 \ -0.001 \ 0.000] \pm [0.001 \ 0.002 \ 0.001 \ 0.001 \ 0.000];$$
- Pixel error:  $\text{err} = [0.164 \ 0.169]$ .

The estimated distortion parameters allow correcting the distortion in the original images. Figure 2 displays a distortion-free image and a 3D scanned scene.

The same undistorted image (Fig. 2a) is used for mapping 2D visual information to a 3D scanned area shown in Fig. 2b. For the mapping, 24 points on box corners in the scene were selected in 2D images and 3D point clouds.

The data mapping model is the classical pinhole camera model described in (1). The parameters of this model were obtained by applying the least squares (lsq) fitting technique. Following a lengthy application of the trial and error method to determine the model and its parameters, the following optimal solution was obtained: the mean-average-error  $\text{MAE} = 1.9$  and mean-square-error

**Fig. 3** Colour information mapping to 3D scene

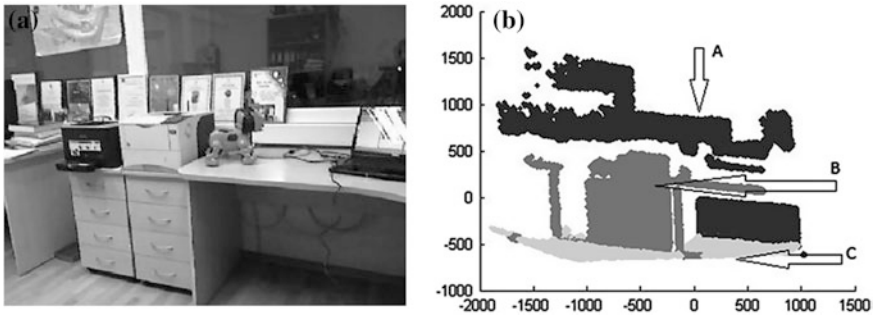


MSE = 6.62. The mapping functions for a 2D image plane  $(u, v)$  are with the corresponding coefficient vectors  $c_u$  and  $c_v$ :

$$\begin{aligned} c_u &= [111.2 \ 479.1 \ 254.7 \ 39.0], \\ c_v &= [175.8 \ 550.4 \ 235.3 \ 19.8], \\ u_i &= c_u(3) + \frac{c_u(2) \cdot (x_i + c_u(1))}{z_i - c_u(4)}, \end{aligned} \quad (3)$$

$$v_i = c_v(3) + \frac{c_v(2) \cdot (y_i + c_v(1))}{z_i - c_v(4)} \quad (4)$$

where, the first coefficient of  $c_u$  and  $c_v$  describes a physical mounting mismatch in millimetres between the centres of 3D and 2D systems, the second parameter describes a new focal length similar to  $f_{cx}, f_{cy}$ , the third parameter describes a new principal point  $a_{u0}, a_{v0}$ , and finally the fourth parameter describes the correction of the depth measurement. The mismatch between the camera calibration results and mapping model parameters are due to the fact that the undistorted images are not the same size as used for calibration. With the models obtained (in Eqs. 3, 4) it becomes possible to map each measured 3D point  $(x_i, y_i, z_i)$  to a corresponding point in the image plane  $(u_i, v_i)$  and vice versa (see Fig. 3). As it is seen from Figs. 2 and 3, man-made scenes usually consist of big and flat areas. In order to reduce the amount of computation for data analysis, flat areas can be removed as non-informative parts (Aguilar et al. 2010). For that purpose, the RANSAC search technique was applied for plane detecting in a 3D data point cloud.



**Fig. 4** Gray scene view analyzed by the RANSAC method (a) and the detected planes {A, B, C} (b)

## 4 Planes Estimation Using the RANSAC

The observed plane in a 3D point cloud is described by plane equation:

$$C(1) * X + C(2) * Y + C(3) * Z + C(4) = 0 \quad (5)$$

where,  $C$  is  $4 \times 1$  array of plane coefficients.

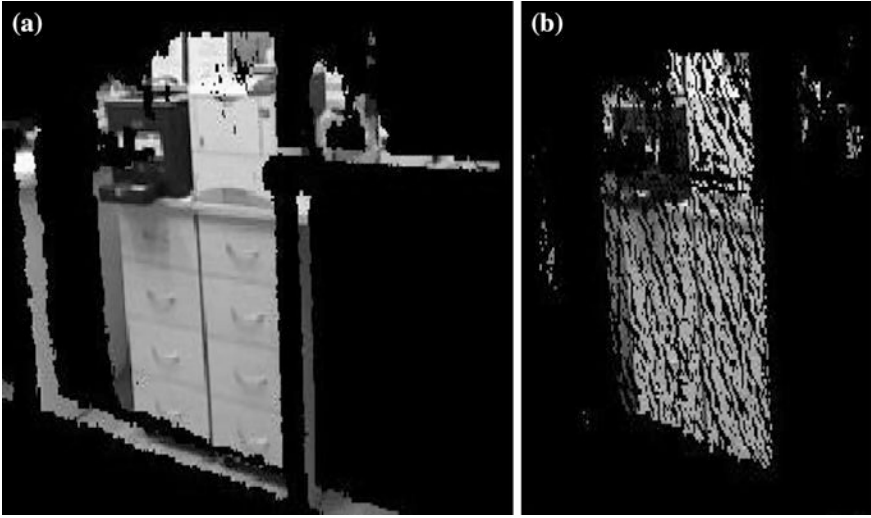
For the purpose of detecting flat planes we have used a scanned 3D scene shown in Fig. 4a. The result was three fitted planes shown in Fig. 4b. As it is seen from Fig. 4b, plane A corresponds to the wall, C—to the floor and plane B corresponds to the front of the cabinet. Physically, plane B is not a single plane since its points lie on two desks and the cabinets. The challenge here is how to avoid the assignment of desk points to the same plane as the front of the cabinets. The answer is to measure the curvedness of the point neighbourhood or compare the face's normals of triangulated elements to the normal of the defined plane.

A strong causal relationship between the measured curvedness of the scanned 3D scene and the computed plane was not determined due to considerable noise which accompanied the measuring procedure. But the comparison of point triangulation normals to the normal of the plane allows the separation of desk points from the cabinet points (Fig. 5). The angle is measured as a dot product of two vectors, i.e. plane normal and point triangulation normal. Point triangulation normal is calculated as the average of all triangulated neighbourhood elements connected to that point.

Depending on the selection of the filtering threshold, the number of selected points will vary as well e.g., by making a small threshold gate for possible variation of triangulated normals to the measured plane, only a small portion of data points will be assigned; and vice versa, for too higher threshold, no points will be filtered out. Therefore, the investigation of data filtering based on the angle variation has been carried out. Figure 5 shows the plane B of the scene in Fig. 4a before (a) and after (b) data filtering with acceptable variation up to  $\pm 25^\circ$ .

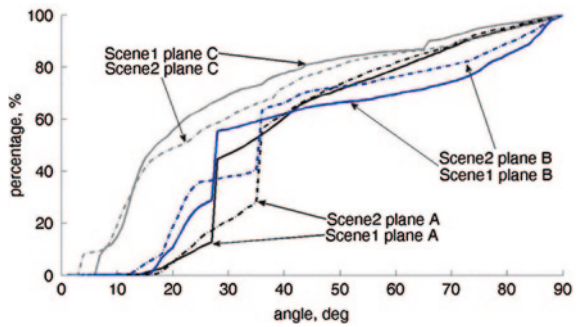
Figure 6 shows the dependency between angle variation (abscissa) in degrees and the remaining points after filtering out the data in percentage (ordinate). The





**Fig. 5** Detected plane B by RANSAC method (a) and in (b) the plane B with filtered out normal's bigger than  $\pm 25^\circ$

**Fig. 6** The dependency between angle variation (*abscissa*) in degrees and remaining points after filtering out the data in percentage (*ordinate*)



curves in the graphs correspond to the planes from two scenes. As it is seen from Fig. 6, curve behaviour largely depends on the nature of the scene and objects presented in the scene.

From our experience we can say that the most appropriate angle variation is acceptable only up to  $\pm 30^\circ$ .

## 5 Discussion and Conclusions

In this chapter, we have presented a 3D space scene scanning system built of the 2D laser scanner and the CCD camera. A mapping function for colour information mapping is derived. The designed system allows straightforward building of colour depth maps as well as assigning colour information to 3D points measured.



The RANSAC method applied for the planes detection in 3D point clouds always gives some planes found in the scene. In reality, these scenes do not continue into infinity; therefore, points from different objects are automatically assigned to the plane. In this chapter we are proposing data filtering method based on the comparison of averaged point triangulation normals to the selected one. The results have shown that due to the scanning noise of the 3D scanning system, the acceptance threshold should be set to  $\pm 30^\circ$ . The presented approach can be applied for solving extraction problems of the constructive primitive, i.e. the problem of the topological simplification of the scanned scenes. It can be used as part of complex solutions for surface re-meshing tasks as well as for reducing constructive primitives. A simplified 3D object consisting of the reduced number of constructive primitives requires less computation and the storage.

Future work will consist of augmenting the discrimination properties of our method with a classification framework for 3D object identification.

**Acknowledgment** This research is funded by the European Social Fund under the project “Microsensors, microactuators and controllers for mechatronic systems (Go-Smart)” (Agreement No VP1-3.1-ŠMM-08-K-01-015).

## References

- Aguiar CSR, Druon S, Crosnier A (2010) 3D datasets segmentation based on local attribute variation. In: IEEE/RSJ international conference on intelligent robots and systems, pp 3205–3210
- Balsys K, Valinevičius A, Eidukas D (2009) Urban traffic control using IR video detection technology. *Electron Electr Eng—Kaunas: Technologija* 8(96):43–46
- Bouguet J (2006) Camera calibration toolbox for matlab. [http://www.vision.caltech.edu/bouguetj/calib\\_doc](http://www.vision.caltech.edu/bouguetj/calib_doc). Accessed 5 Dec 2013
- Garland M, Willmott A, Heckbert PS (2001) Hierarchical face clustering on polygonal surfaces. In: Proceedings of ACM symposium on interactive 3D graphics, pp 49–58
- Himmelsbach M, Muller A, Luttel T, Wunsche HJ (2008) LIDAR-based 3D object perception. In: Proceedings of 1st international workshop on cognition for technical systems
- Hinker P, Hansen C (1993) Geometric optimization. *Visualization* 93:189–195
- Joochim C, Roth H (2008) Development of a 3d mapping using 2d/3d sensors for mobile robot locomotion. In: Proceedings of IEEE international conference of technologies for practical robot applications, pp 100–105
- Kalvin A, Taylor R (1996) Superfaces: polygonal mesh simplification with bounded error. *IEEE Comput Graphics Appl* 16:64–77
- Klimentjew D, Arli M, Zhang J (2009) 3D scene reconstruction based on a moving 2D laser range finder for service-robots. In: Proceedings of IEEE international conference on robotics and biomimetics, pp 1129–1134
- Lee A, Moreton H, Hoppe H (2000) Displaced subdivision surfaces. In: Proceedings of 27th annual conference on computer graphics and interactive techniques, pp 85–94
- Scaramuzza D, Harati A, Siegwart R (2007) Extrinsic self calibration of a camera and a 3d laser range finder from natural scenes. In: Proceedings of IEEE international conference on intelligent robots and systems, pp 4164–4169
- Serafinavičius P (2005) Investigation of technical equipment in computer stereo vision: camera calibration techniques. *Electron Electr Eng—Kaunas: Technologija* 3(59):24–27

- Shacklock A, Xu J, Wang H (2006) Visual guidance for autonomous vehicles: capability and challenges. In: Shuzhi SG, Frank LL. *Autonomous mobile robots: sensing, control, decision making and applications*, CRC Press, pp 8–12
- Surgailis T, Valinevičius A, Eidukas D (2011) Stereo vision based traffic analysis system. *Electron Electr Eng—Kaunas: Technologija* 1(107):15–18
- Surmann H, Lingemann K, Nuchter A, Hertzberg J (2001) A 3D laser range finder for autonomous mobile robots. In: *Proceedings of 32nd international symposium on robotics*, pp 153–158
- Wulf O, Wapner B (2003) Fast 3d scanning methods for laser measurement systems. In: *International conference on control systems and computer science*, vol 1, pp 312–317

Issues and Challenges in Artificial Intelligence

Hippe, Z.S.; Kulikowski, J.L.; Mroczek, T.; Wtorek, J. (Eds.)

2014, VIII, 182 p. 64 illus., Hardcover

ISBN: 978-3-319-06882-4