ISR / ROBOTIK 2010

In Proceedings of Joint 41th International Symposium on Robotics and 6th German Conference on Robotics, Munich, June 2010.

Probabilistic Phase Unwrapping for Time-of-Flight Cameras

David Droeschel, Dirk Holz and Sven Behnke
Autonomous Intelligent Systems Group, Computer Science Institute VI, University of Bonn, Germany

Abstract

Time-of-Flight (ToF) cameras gain depth information by emitting amplitude-modulated near-infrared light and measuring the phase shift between the emitted and the reflected signal. The phase shift is proportional to the object's distance modulo the wavelength of the modulation frequency. This results in a distance ambiguity. Distances larger than the wavelength are *wrapped* into the sensor's non-ambiguity range and cause spurious distance measurements. We apply *Phase Unwrapping* to reconstruct these wrapped measurements. Our approach is based on a probabilistic graphical model using loopy belief propagation to detect and infer the position of wrapped measurements. In experiments, we show that wrapped measurements are identified and corrected allowing to reconstruct the structure of the scene.

1 Introduction

Time-of-Flight (ToF) cameras attracted attention in the field of robotics and automation in the last decade. They are compact, solid-state sensors, which provide depth and reflectance images at high frame rates. The main disadvantage of today's ToF cameras is their complex error model which makes them difficult to handle. They employ an array of light emitting diodes (LEDs) that illuminate the environment with modulated near-infrared light. The reflected light is received by a CCD/CMOS chip for every pixel in parallel. Depth information is gained by measuring the phase shift between the emitted and the reflected light, which is proportional to the object's distance modulo the wavelength of the modulation frequency. This results in an ambiguity in distance measurements. Distances larger than the wavelength λ of the sensor's modulation frequency are *wrapped* into the non-ambiguity range $[0, \lambda)$ and result in artifacts and spurious distance measurements. This effect is one of the fundamental characteristics of ToF cameras. A common way to handle these distance ambiguities is to neglect measurements based on the ratio of measured distance and amplitude, since the amplitude of the reflected signal decreases with the square of the distance to an object (see for example [1]). The limitation of this method is that information is lost due to neglecting measurements. Another limitation is that wrapped measurements are only detected by the ratio of distance and amplitude, not taking into account the gradient surface of neighboring measurements, which results in wrong classifications for distant objects with high infrared reflectivity.

Inferring a correct, *unwrapped* signal from a wrapped signal is known as *Phase Unwrapping* (see **Figure 1**). That is, depth measurements being erroneously projected into the non-ambiguity range of the sensor are identified and projected back into the correct interval. Phase unwrapping is a fundamental problem in image processing [2]. It has been successfully applied in magnetic resonance imaging [3] and interferometric synthetic aperture radar (SAR) [4].

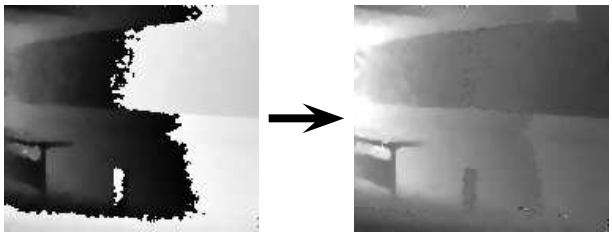


Figure 1: Unwrapping the input depth image (left). The brightness of the pixels encode the measured distance (dark pixels are near, bright pixels far away). The abrupt change from bright to dark in the input image is the phase jump that we want to detect in order to correct the measured distances and to obtain the unwrapped depth image (right).

The goal of phase unwrapping is to infer a number of relative *phase jumps* (or phase shifts) from the wrapped signal. A phase jump is defined between two adjacent pixels in x- and y-direction of the image. Since the phase unwrapping problem in general is ill-posed, most algorithms make *a priori* assumptions to reduce the number of admissible phase jumps. One common assumption is that neighboring measurements are more likely closer together than farther apart. With this assumption, the phase jump that brings the neighboring distance measurements as close together as possible is chosen.

In this paper, we apply phase unwrapping to the data from a SwissRanger SR4000 ToF camera. A probabilistic approach based on [5] to represent the probabilities of the rel-

ative phase jumps is used. The approach relies on discontinuities in the depth image to infer relative phase jumps. We also applied the idea of the *zero curl constraint* from [5] to assure consistency of the phase jumps in two-dimensional images.
The remainder of this paper is organized as follows: The next section outlines related work based on ToF cameras in the field of robotics. Section 3 describes the probabilistic approach to phase unwrapping. Section 4 presents results showing that the approach enables to correct ambiguous distance measurements such that the structure of the scene can be reconstructed.

2 Related Work

One of the first applications in robotics considering ToF cameras as an alternative to laser scanning has been presented in 2004 by Weingarten *et al.* who evaluated the utility of ToF cameras in terms of basic obstacle avoidance and local path-planning capabilities [6]. In 2005, Sheh *et al.* used a ToF camera for human-assisted 3D mapping in the context of the RoboCup Rescue league [7]. Ohno *et al.* used a ToF camera for estimating a robot's trajectory and reconstructing the surface of the environment in 2006 [8]. Recently, May *et al.* presented and evaluated different approaches for registering multiple range images of ToF cameras in the context of fully autonomous 3D mapping [1].
All the aforementioned approaches have shown that ToF cameras require for specifically taking care of their complex error model. The different systematic and non-systematic errors cause, amongst other effects [9]:

(i) Measurement noise: Data from the ToF camera is subject to noise, especially at larger distances and poorly reflecting objects.

(ii) Jump edges: ToF cameras measure a smooth transition, where the transition between one shape to the other is disconnected due to occlusions [1].

(iii) Distance ambiguity: Measurements larger than the used wavelength are *wrapped* into the sensor's non-ambiguity range and cause artifacts and spurious distance measurements.

The general noise, especially the systematic errors, are usually handled by calibration, as shown by Fuchs and Hirzinger [10], and Lindner *et al.* [11]. Correcting noisy measurements by means of the relation between distance and intensity information, and using so-called *shading constraints*, has been presented by Böhme *et al.* in [12]. For detecting jump edges, sufficient results can be achieved by examining, for every measurement, the opposing angles of the triangle spanned by the camera's focal point, the point itself and its local pixel neighborhood [1]. Distance ambiguities, however, have (to the best of our knowledge) not yet been addressed for ToF cameras.

Especially when mapping larger environments where measured distances exceed the wavelength of the modulation frequency, obtaining an unwrapped depth image becomes crucial [1].

3 Phase Unwrapping

To infer phase jumps between adjacent pixels, we use a graphical model which represents possible locations of relative phase jumps in x- and y-direction (see **Figure 2**). The image pixels are connected to their neighbors by so-called *jump nodes*. These jump nodes represent the probability of a phase jump between two neighboring pixels. To assure consistency of phase jumps in a local neighborhood, we apply the idea of zero-curl constraints [5]. Four jump nodes are connected by a *curl node* that enforces local consistency of the individual jump configurations. Interaction between jump and curl nodes is achieved by passing messages across the graph that represent a node's belief. After convergence of the message passing, the detected phase jumps are integrated into the depth image by carrying out the respective projections, thereby correcting the erroneously wrapped distance measurements. The following subsections describe the above steps in detail.

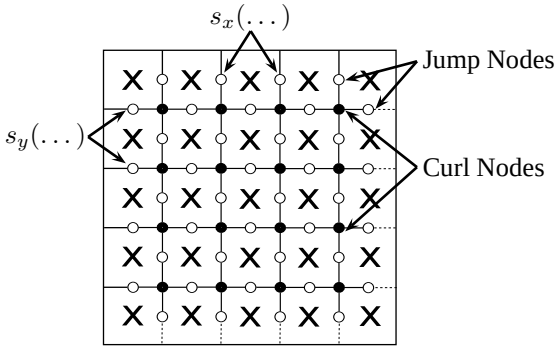


Figure 2: The graphical model representing possible locations of phase jumps. The image pixels (black **x**'s) are connected to their neighbors by jump nodes (white filled circles). Four jump nodes are connected by a curl node (black filled circles) which enforces the zero curl constraint.

3.1 Jump Nodes

Jump nodes represent the probability that a phase jump occurs between the two neighboring pixels. A phase jump in x-direction, i.e., between pixels (x, y) and $(x + 1, y)$, is denoted as $s_x(x, y)$. Jumps in y-direction, i.e., between pixels (x, y) and $(x, y + 1)$, are denoted as $s_y(x, y)$.
A phase jump can occur either in positive direction (-1), in negative direction (+1) or not at all (0). Considering the x-shifts, positive direction at pixel (x, y) thereby means that there is a phase jump between pixels (x, y) and $(x + 1, y)$.

Negative means that the phase jump occurs between pixels $(x+1, y)$ and (x, y). The direction of a phase jump is important for the correction of distance measurements as it decides which and how measurements need to be corrected. The possible jump directions (-1,0, and 1) are called *jump configurations*.

Jump nodes are represented by a 3-element vector storing the probabilities for every jump configuration. The initial probabilities for a jump $s_x(x, y)$ at pixel (x, y) for configuration $i \in \{-1, 0, 1\}$ and given the wrapped depth image Φ are calculated by

$$P\left(s_{\{x,y\}}(x, y) = i \mid \Phi\right) \propto f_d\left(x, y, i\right), \qquad (1)$$

where f_d is a term incorporating the discontinuity of depth measurements.

The basic assumption behind the discontinuity term f_d is that neighboring measurements are more likely closer to each other than farther apart. This term increases the probability $P\left(s_{\{x,y\}}(x, y) = i \mid \Phi\right)$, when the phase jump for configuration i brings the respective depth measurements closer together than the other configurations. Here, we follow the formulation of [5]:

$$f_d(x, y, i) = \begin{cases} e^{-(\phi(x+1,y)-\phi(x,y)-i)^2/2\sigma^2}, & \text{for } s_x \\ e^{-(\phi(x,y+1)-\phi(x,y)-i)^2/2\sigma^2}, & \text{for } s_y \end{cases} \qquad (2)$$

where σ^2 is the variance in the depth values between neighboring pixels in the wrapped image, and $\phi(x, y)$ is the measured, wrapped phase for pixel (x, y) (scaled into the interval $[0, 1)$ for simplicity but without loss of generality).

3.2 Curl Nodes

Four jump nodes are connected by a curl node which enforces the zero-curl constraint [5]. A curl node assures local consistency of the phase jumps, by summing up the shift configurations of the jump nodes around it. For example, the sum of the 4-pixel loop around (x, y) is $s_x(x, y) + s_y(x+1, y) - s_x(x, y+1) - s_y(x, y)$ (see **Figure 3**). A zero-curl constraint is violated when a jump is not matched by another jump in the opposite direction, i.e., when the sum around a pixel loop is $\neq 0$. Therefore, the set of phase jumps for an image must satisfy the constraint

$$s_x(x, y) + s_y(x+1, y) - s_x(x, y+1) - s_y(x, y) = 0. \qquad (3)$$

If all zero-curl constraints are satisfied, consistency of the inferred phase jumps in the complete image can be assumed.

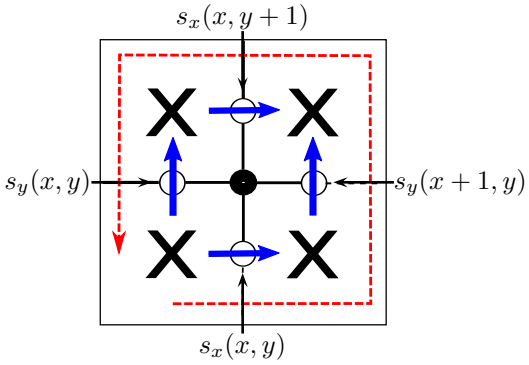


Figure 3: A curl node assures local consistency of the phase shifts by summing up the shift configurations of the jump nodes around it (red dashed arrow), taking into account the direction of the jumps (blue arrows). A zero-curl constraint is violated when a jump is not matched by another jump in the opposite direction.

3.3 Message Passing

Inference of phase jumps is done by applying belief propagation (sum-product algorithm) in the graphical model. Messages, representing a node's belief of a configuration, are passed bi-directionally through the graph on the vertices between jump and curl nodes in a forward-backward-up-down type schedule.
Messages are represented by 3-element vectors, where the elements are the probabilities for a specific configuration. Referring to **Figure 4**, messages from the jump nodes to the curl nodes are calculated by incoming messages from the curl nodes.
Messages from curl nodes to jump nodes (**Figure 4.a**) are calculated by incoming messages from jump nodes.

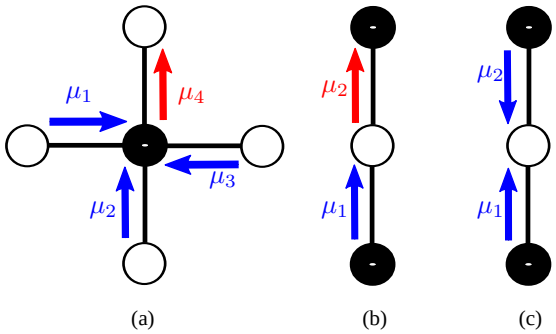


Figure 4: (a) Messages from curl nodes to jump nodes (red arrow) are computed by incoming messages originating from jump nodes in the last iteration (blue arrows). (b) Messages from jump nodes to curl nodes (red arrow) are computed by incoming messages from curl nodes (blue arrow). (c) Messages from curl nodes to jump nodes are combined to approximate the marginal probability of a phase jump.

For example, the outgoing message μ_4 (for every configuration i) depends on the incoming messages μ_1, μ_2 and μ_3 (for the configurations j, k and l):

$$\mu_{4i} = \sum_{j=-1}^{1} \sum_{k=-1}^{1} \sum_{l=-1}^{1} \delta(k+l-i-j)\mu_{1j}\mu_{2k}\mu_{3l} \quad (4)$$

with

$$\delta(x) = \begin{cases} 1, & x = 0 \\ 0, & \text{otherwise.} \end{cases}$$

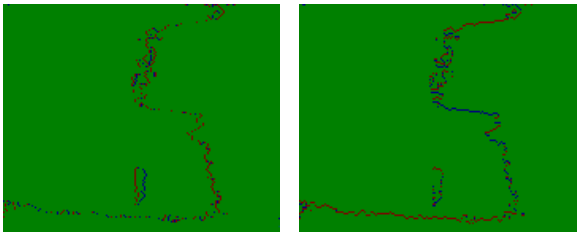
The outgoing message μ_2 for configration i (**Figure 4.b**) is calculated from the incoming message μ_{1i} from the curl node by

$$\mu_{2i} = \mu_{1i} f_d\,(x, y, i)\,. \quad (5)$$

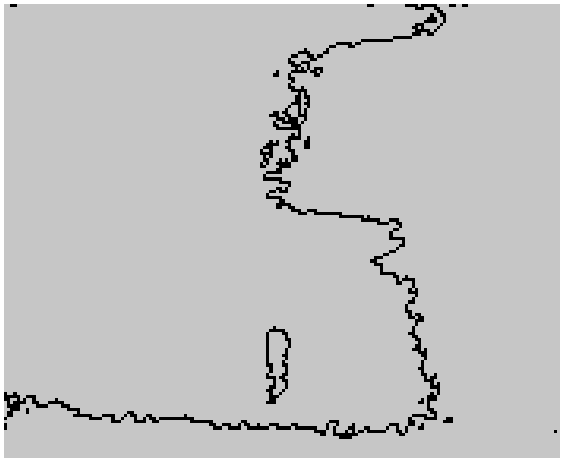
The marginal probabilities for the phase jumps (**Figure 4.c**) are approximated by

$$\hat{P}(s_x(x,y) = i \mid \Phi) = (\mu_{1i}\mu_{2i}) / \sum_{j=-1}^{1} (\mu_{1j}\mu_{2j}). \quad (6)$$

The message vectors are normalized in every iteration. The belief propagation converges when all zero-curl constraints are satisfied. Propagating the beliefs results in consistent and correct phase jump locations along x and y-directions.



(a) (b)



(c)

Figure 5: The resulting phase jump configurations after belief propagation has converged in x-direction (a) and y-direction (b). The color of the pixels indicate the inferred jump configurations: -1 (blue), 0 (green) and 1 (red). (c) The combined phase jumps.

The resulting jump configurations are shown in **Figures 5.a** and **5.b**. After convergence or when a fixed number of iterations is exceeded, the phase jumps are integrated. The resulting image provides the exact x and y-coordinates of the locations where phase jumps occur in the wrapped input depth image. The phase jump locations for the example depth image image in **Figure 1** are shown in **Figure 5.c**.

4 Experiments

The following experiments demonstrate the results of the approach. The experiments have been carried out in an indoor environment (c.f. **Figure 6.a**). The employed ToF camera is a SR4000 from Mesa Imaging [13], operated with $30\,\mathrm{MHz}$ modulation frequency, which results in a non-ambiguity range of $5\,\mathrm{m}$. The wrapped depth image of the first experiment is shown in **Figure 6.b**. The wrapping effect can also be observed in the resulting 3D point clouds in **Figures 7.a** and **7.c**. Objects beyond $\lambda \approx 5\,\mathrm{m}$ are wrapped into the non-ambiguity range and result in spurious artifacts.

The application of the described phase unwrapping method results in the unwrapped distance image in **Figure 6.c**. Compared to the wrapped distance image, objects beyond $5\,\mathrm{m}$ do not result in close distance measurements. However, a remaining distortion in the unwrapped depth image can be seen at the upper bound of the non-ambiguity range, which is mainly due to the camera's internal filter, which smoothens the wrapped depth image.

Figure 7 shows the 3D point clouds that have been generated based on the wrapped and unwrapped depth images. The results show that the structure of the room with distances larger than 5 meters was reconstructed successfully.

A second experiment is shown in **Figure 8**. Also here, the depth image could be unwrapped successful with the described approach, resulting in a reconstructed structure of of the room.

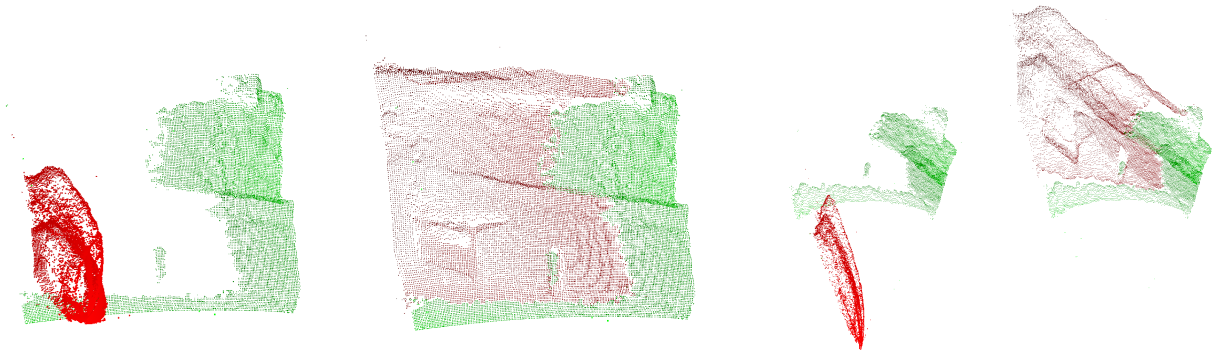
5 Conclusion and Future Work

We have presented a probabilistic approach for phase unwrapping specifically designed for handling ambiguities in ToF camera data. By means of a graphical model with loopy belief propagation, the approach takes into account discontinuities in the measured distances to obtain accurate estimates of the locations where phase jumps have taken place.

The results show that the approach enables to correct ambiguous distance measurements such that the structure of the scene can be reconstructed correctly from wrapped depth images. This is an important result for the use of ToF cameras in the field of robotics, since current camera models, such as the SR4000 from Mesa, have a non-ambiguity range that is shorter than the maximum mea-

(a) (b) (c)

Figure 6: (a) An image of the scene. (b) Wrapped depth image from the ToF camera. A pixel's grey-scale value corresponds to the measured depth, from dark (close) to bright (far). The dark parts of the image indicate that distance measurements larger than 5 meters are wrapped. (c) The unwrapped depth image.



(a) (b) (c) (d)

Figure 7: (a + c) The wrapped point clouds from two different perspectives. (b + d) The unwrapped point clouds. The color of the points indicate the result of the algorithm. Wrapped measurements are colored red.

surable distance of commonly used laser range scanners. Simply sorting out the wrapped measurements based on the ratio of distance and amplitude does not work reliably in natural scenes, e.g., when highly reflective objects are sensed.

Especially in the context of 3D mapping, the ambiguity of the phase-shift based distance measurements hindered from using ToF cameras for modeling larger environments where measured distances exceed the wavelength of the sensor's modulation frequency [1]. Correctly unwrapping acquired depth images enables to model larger environments. However, it remains a matter of future work to actually apply probabilistic phase unwrapping for mapping a larger environment.

A limitation of the presented method is that even in the absence of noise, in situations where the actual phase jump between two adjacent pixels is larger than the modulation wavelength, a decision based on the discontinuities in the measured distances cannot be made. Another limitation is that phase jumps can only be detected based on the gradient surface of the neighboring measurements, which poses a problem, for example, when all measurements in the field-of-view are wrapped.

A possible extension to overcome this limitation could be, to acquire measurements of the same scene with multiple modulation frequencies, since measurements with different modulation frequencies result in different wrappings. However, simply filtering based on different wrappings would, *on its own*, only work in the absence of noise and is expected to be not appropriate for reliably identifying phase jumps. It remains a matter of future work to integrate multiple modulation frequencies with the currently used distance continuities into the probabilistic graphical model and the propagation of beliefs.

Acknowledgment

This work has been supported partially by grant BE 2556/2-2 of German Research Foundation (DFG).

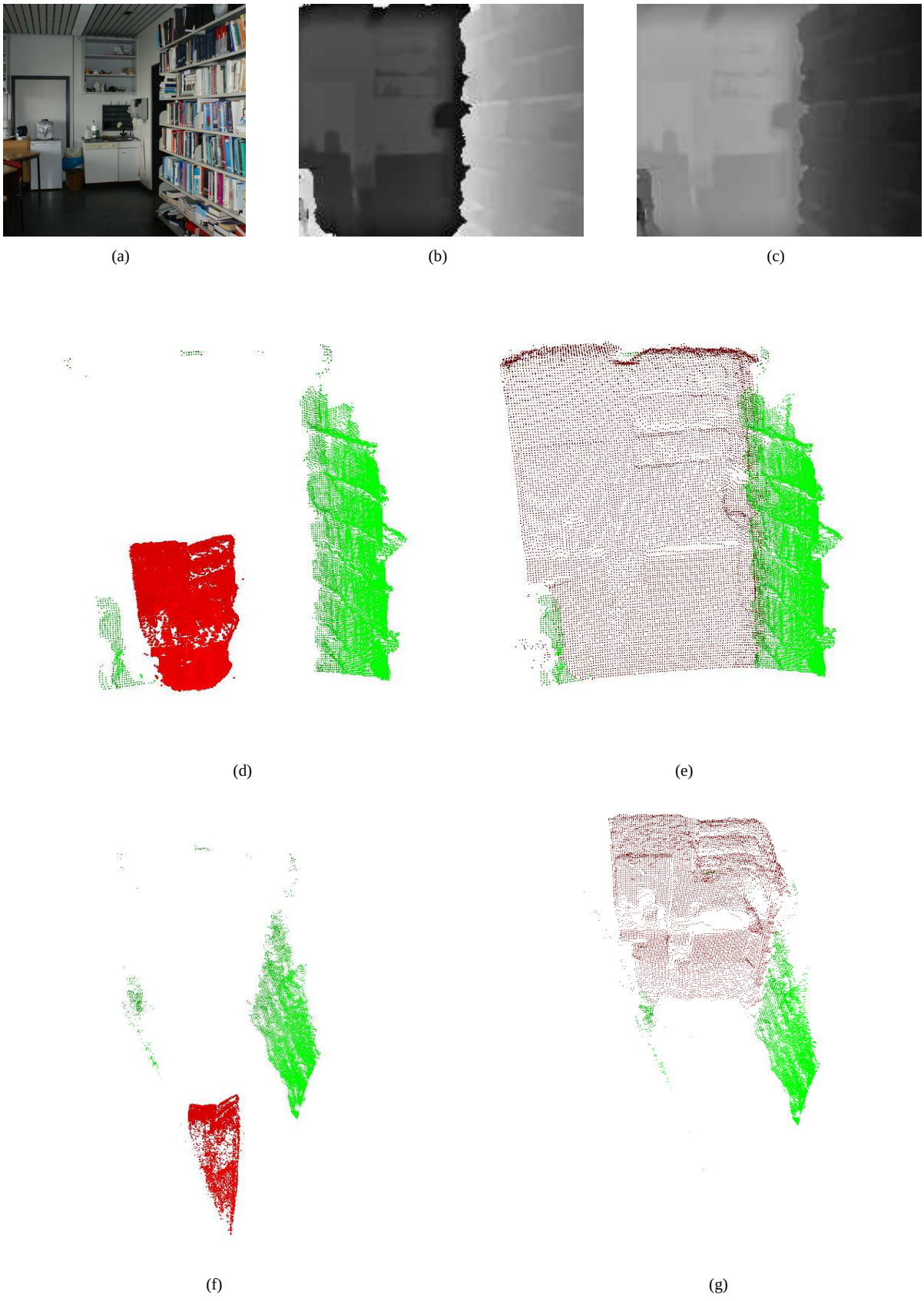


Figure 8: (a) An image of the scene. (b) Wrapped depth image from the ToF camera. (c) The unwrapped depth image. (d + f) The wrapped point clouds. (e + g) The unwrapped point clouds.

References

[1] S. May, D. Droeschel, D. Holz, S. Fuchs, E. Malis, A. Nüchter, and J. Hertzberg. Three-dimensional mapping with time-of-flight cameras. *Journal of Field Robotics, Special Issue on Three-Dimensional Mapping, Part 2*, 26(11-12):934–965, December 2009.

[2] D. C. Ghiglia and M. D. Pritt. *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. Wiley-Interscience, 1998.

[3] Z. Liang and P. Lauterbur. *Principles of Magnetic Resonance Imaging (A Signal Processing Perspective)*. IEEE Press, 1999.

[4] P. Thompson, D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and C. V. Jakowatz. *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*. Kluwer Academic Publishers, Norwell, MA, USA, 1996.

[5] B. J. Frey, R. Koetter, and N. Petrovic. Very loopy belief propagation for unwrapping phase images. In *Advances in Neural Information Processing Systems 14 [Neural Information Processing Systems: Natural and Synthetic (NIPS)*, pages 737–743, Vancouver, British Columbia, Canada, 2001.

[6] J. W. Weingarten, G. Grüner, and R. Siegwart. A State-of-the-Art 3D Sensor for Robot Navigation. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2155–2160, Sendai, Japan, 2004.

[7] R. Sheh, M. W. Kadous, and C. Sammut. On building 3d maps using a range camera: Applications to rescue robotics. Technical report, University of New South Wales, Sydney, Australia, 2006.

[8] K. Ohno, T. Nomura, and S. Tadokoro. Real-time robot trajectory estimation and 3d map construction using 3d camera. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5279–5285, Bejing, China, 2006.

[9] R. Lange. *3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology*. PhD thesis, University of Siegen, 2000.

[10] S. Fuchs and G. Hirzinger. Extrinsic and Depth Calibration of ToF-Cameras. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, AK, USA, 2008.

[11] M. Lindner, A. Kolb, and T. Ringbeck. New Insights into the Calibration of TOF Sensors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Workshop on ToF Camera based Computer Vision (TOF-CV)*, Anchorage, AK, USA, 2008.

[12] M. Böhme, M. Haker, T. Martinetz, and E. Barth. Shading Constraint Improves Accuracy of Time-of-Flight Measurements. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Workshop on ToF Camera based Computer Vision (TOF-CV)*, Anchorage, AK, USA, 2008.

[13] Mesa Imaging AG. Swissranger tof cameras. http://www.mesa-imaging.ch.