A Novel Approach to Efficient Error Correction for the SwissRanger Time-of-Flight 3D Camera*

Jann Poppinga and Andreas Birk

Jacobs University Bremen**
Campus Ring 1
28759 Bremen, Germany
{j.poppinga,a.birk}@jacobs-university.de

Abstract. 3D data acquisition gets increasingly important for mobile robotics in general and for Safety, Security and Rescue Robotics (SSRR) in particular. 3D data allows for example to estimate the size of gaps, to construct realistic maps of unstructured disaster environments, or to detect human victims from shape. The SwissRanger SR-3000 time-of-flight 3D camera is a popular device for acquiring 3D range data as it offers the fast update rates of a camera with a resolution of 176 × 144 pixel at a reasonable cost. But the SR-3000 suffers - like any device using phase differences of modulated light - from the fundamental problem of wrap around error, i.e., distances that are a multiple of the wavelength of the modulated ranging signal can not be distinguished. The standard solution to this problem is to use an amplitude threshold, i.e., to discard pixels which are relatively dark and hence assumed to be far away. Here, a significant improvement to the standard method is presented that relates measured brightness and distance and also takes the geometry of the modulated light source into account. It is shown that a significantly higher amount of valid range data can be acquired with this new method.

1 Introduction

The acquisition of 3D range data is of increasing interest for mobile robots. This holds especially for domains like Safety, Security and Rescue Robotics (SSRR) where robots have to perform in complex, unstructured environments, which require 3D data for perception and world modeling. Examples of the usage of according range data include 3D map building [1][2][3][4][5], semantic environment classification [6] or the detection of drivable terrain [7].

A popular choice for according sensors are 3D laser range finders [8][9][10]. There main drawback is that they are typically based on 2D devices, which are supplemented with an actuator for an additional degree of freedom. The overall time to acquisition the data is hence slow. It is typically in the order of several

* This work was supported by the German Research Foundation (DFG).
** Formerly International University Bremen.

L. Iocchi et al. (Eds.): RoboCup 2008, LNAI 5399, pp. 247–258, 2009.

seconds per snapshot. Since a few years, the alternative technology of time-of-flight cameras is emerging [11]. The cameras use a near infrared modulated light source. They measure at the pixel level the phase shift - and hence time of flight - of the light returned by the environment location, which lies on the end of the optical beam corresponding to this pixel. The main advantage of this technology is that it allows high acquisition rates of 20 to 30 Hz and more. Examples of time-of-flight cameras are the SwissRanger SR-3000 [12][13] and its successor, the SR-3100, which are also popular in the RoboCup Rescue League. The SwissRanger SR-3100 is produced by MESA Imaging AG, Zurich, Switzerland. They provide a 176x144 pixel range image at 25 Hz update frequency.

Like any time-of-flight camera, the SwissRangers suffer from the fundamental problem of wrap around error. This means that distances that are a multiple of the wavelength of the modulated ranging signal can not be distinguished. The standard solution to this problem is to use an amplitude threshold, i.e., to discard pixels which are relatively dark and hence assumed to be far away. A significant improvement to this state of the art method is presented. The main idea is to take the geometry of the modulated light source into account. It is shown through extensive experiments that a significantly higher amount of valid range data can be acquired with this new method.

2 The Wrap around Error and Its Standard Solution

The SwissRanger's principle problem is that it cannot detect whether a measurement is erroneous due to being beyond the maximum range. This is because it not only relies on active sensing (illumination), but it also measures distance via phase-shift. Since phase-shift is assumed to be in $[0^o, 360^o[$, a phase-shift of 370^o is measured as 10^o. Obviously, this also corrupts the distance measurement. An object beyond the maximum distance of 7.5 m, which is e.g. at 8 m, is reported as being at 0.5 m. This is called the *wrap-around error*, a fundamental problem of any time-of-flight camera.

The standard solution to this problem is to set an *amplitude threshold*, i.e., to drop pixels which are too dark. As will be shown in the experiment section, this solution is very crude and produces large portions of both false negatives and false positives.

3 A Novel Solution: Adaptive Amplitude Threshold

In the following section we first explain how we identify pixels with a wrapped around measurement, then we use this knowledge and the wrong measurement to extend the SwissRanger's range.

3.1 Identification of Erroneous Pixels

The proposed approach is to sanity-check the reported distance by relating it to the brightness. The basic idea is that a wrapped around pixel p will be reported

with a low distance value d_p (measured from the image plane), but its brightness b_p is significantly darker than it were in case of a correct small distance. Since the illumination is not uniform, the pixel's position in the image array x_p, y_p is also taken into account. These two factors are used to calculate a minimum expected brightness for each pixel (equation 1).

$$b_p > \frac{\mathrm{bw}_p \mathrm{AAT}}{d_p^2}, \tag{1}$$

where AAT is the *advanced amplitude threshold* and bw_p is the brightness of pixel p when viewing a white wall at roughly one meter, approximated by

$$\mathrm{bw}_p := B - ((x_p - offset_x)^2 + (y_p - offset_y)^2) \tag{2}$$

(B being the *brightness constant*). The approximation is reasonably close, except for the very edges of the image (see figure 1).

The reasoning for this is that the measurement is based solely on light emitted by the camera's illumination unit, hence the perceived brightness of an object decreases quadratically with the distance. The second observation is that the brightness produced by the imager decreases towards the edges of the images (see figure 1(a)). This is an effect, which is common to time-of-flight cameras due to the usage of multiple near infrared LEDs as light sources. They are grouped in a rectangular pattern around the cameras' lens. Hence, the center of the image is better illuminated than the border regions. To compensate for this effect, the brightness is also normalized. We approximate the brightness distribution by the above function 2. For the accuracy see figure 1(b).

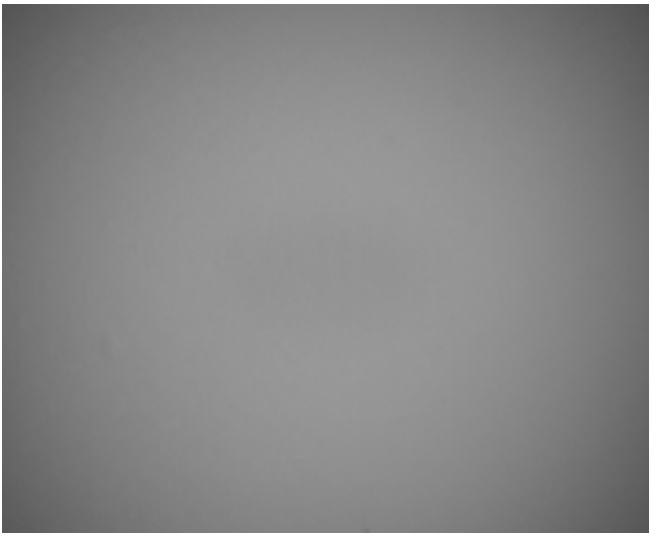
In the SR-3000, according to our measurements, the center of the brightness pattern is not in the center of the image. Instead, it is shifted upward. Consequently, the offset in y direction is shifted, namely $offset_y = 61$. In x direction it is at half the image height, i.e., $offset_x = 88$, as one would expect. In the SR-3100, the image is brightest in the center. In experiments calibrating the parameters B (brightness constant) and the AAT in various scenes, we determined the values $12,000$ and 0.2 respectively as working very well.

3.2 Additional Error Sources

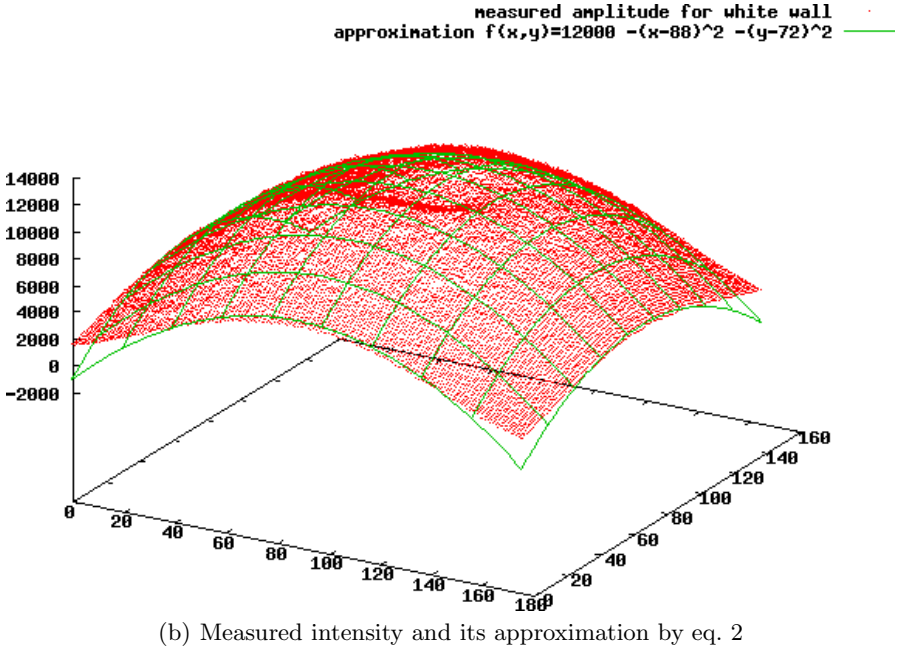
Material that reflects near-IR light poorly naturally appears very dark to the SwissRanger. Objects of such material can be filtered out alongside with the wrapped-around pixels. As the measurement of these objects tend to be noisy and sometimes also offset, filtering the resulting pixels out is desirable.

Furthermore, bright sun light can cause areas of random noise in the SwissRanger. This also applies to indirect perception through glass bricks or when seen on other objects. Most of these pixels are filtered out by the proposed method, except for those which by coincident have a roughly correct distance.

As a third case, pixels that appear to close due to light scattering are excluded. As mentioned in [12], bright light from close obstacles can fail to get completely absorbed by the imager. The remaining light can get reflected back to the imager



(a) Grayscale image of a white wall as captured by the Swiss-Ranger



(b) Measured intensity and its approximation by eq. 2

Fig. 1. Irregular distribution of near-IR light from the camera's illumination unit

e.g. by the lens. This way, rather many pixels can get tainted, showing too close values. A large portion of these will be excluded by the presented criterion.

3.3 Correction of Wrong Values

Once the incorrect pixels due to wrap around errors are identified, they do not have to be dropped. Since the principles of the error source are regular and well-defined, they can be undone. We call this process *unwrapping*.

For this, first the perceived distance of a given wrapped-around point is calculated as the Euclidean distance d from its coordinates to the origin. Then, the maximum range (7.5 m) is added to d. Finally, the new coordinates are calculated from the distance and the known angles of the beam corresponding to the pixel in question.

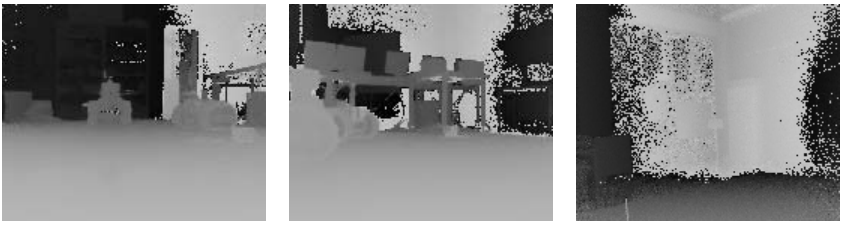
This method assumes that all wrong pixels have wrapped around once. Under certain conditions, pixels can also wrap around more than once. However, in typical indoor scenes obstacles further away than 15 m are only rarely encountered. Other filtered out pixels (e.g. too dark, see previous section) are misinterpreted as being far away. In practice, rather few pixels are affected. Still, it has to be considered for each application if a dense and partially incorrect image with a doubled range or a sparse, short range but almost perfectly correct one is preferable.

(a) Rack and door (b) Corner window (c) Gates

Fig. 2. Three test scenes comparing the proposed method to the standard method; here normal images of the scenes from a web-cam next to the SwissRanger are shown

Table 1. Results on all six scenes

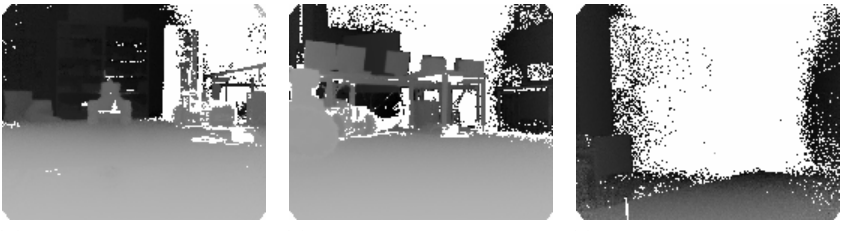
Scene	Discarded Pixels [%]		
	Proposed method	AT 160	AT 240
Rack and door	16.7	46.3	60.3
Corner window	21.7	43.3	54.7
Gates	56.3	92.8	100
Close boxes	38.1	35.0	–
Dark material	13.1	28.0	43.1
Sun	44.9	73.1	84.6
mean	31.8	53.1	68.5
median	44.8	29.9	60.3



(a) Unfiltered distance image: wrap-around in the corner and behind the door

(b) Unfiltered distance image: wrap-around in the corner, noise on the window

(c) Unfiltered distance image: mostly wrapped around, noise on windows



(d) Proposed method: few wrong pixels remain – 16.7 % d. p.

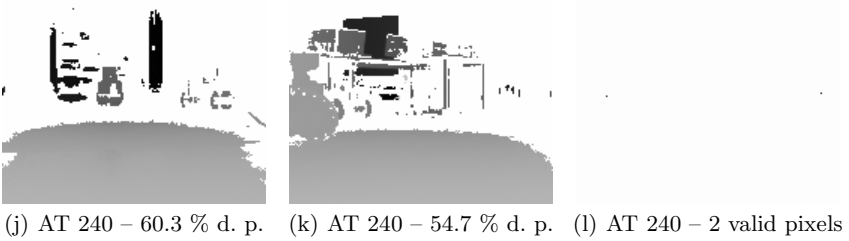
(e) Proposed method: no wrapped-around pixels remain – 21.7 % d. p.

(f) Proposed method: no wrap-around, very little noise – 56.3 % d. p.

(g) AT 160: some wrong pixels remain, correct ones are discarded – 46.3 % d. p.

(h) AT 160: some wrong pixels remain, correct ones are discarded – 43.3 % d. p.

(i) AT 160: almost all valid pixels removed while some wrap-around is kept – 92.8 % d. p.



(j) AT 240 – 60.3 % d. p.

(k) AT 240 – 54.7 % d. p.

(l) AT 240 – 2 valid pixels

Fig. 3. Three test scenes comparing the proposed method to the standard method of the SwissRanger: setting an amplitude threshold (AT). Here, two AT values presented: 240, high enough to exclude every wrapped around pixel and 160, which only almost high enough, but not as over-restrictive. For each filtered image, the fraction of discarded pixels (d. p.) is given.

4 Experiments and Results

4.1 Only Exclusion

We applied the approach presented in the previous section to snapshots taken around our lab with the SwissRanger SR-3100. To compare it to the standard abilities of the device, we also used its default error exclusion technique, the *amplitude threshold* (AT). This term designates merely the exclusion of pixels which are too dark, regardless of their distance. Depending on the situations, different AT work best. These, however, have to be manually chosen every time. Unfortunately, when using a widely applicable AT value, this standard technique excludes far too many pixels (60.3 % on average). As an alternative, we use a lower AT value which preserves more pixels (it only discards an average of 44.8 %). As can be expected, quite some of these are wrong. In figure 3, three example images are presented together with a complete distance image of the scene and a color image that was taken by a web-cam which is mounted on the robot right next to the SwissRanger.

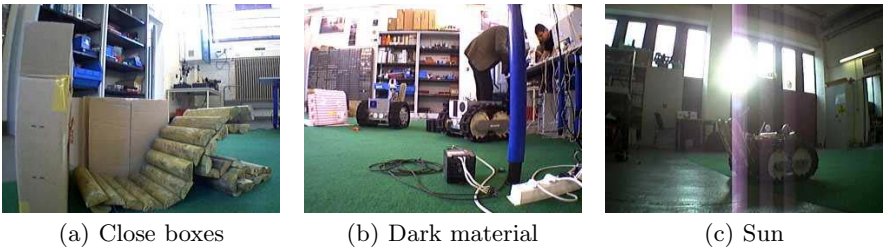
As mentioned earlier, not only wrapped around pixels can be filtered out. This is illustrated in figure 5. Here, the improvement in performance towards the conventional method is not as big as with the wrap-around, but still apparent. For a summary of all results, refer to table 1.

4.2 Analysis

The exclusion criterion (eq. 1) consists of two factors: Wrap around is detected by comparing the reported brightness to inverted squared reported distance to detect wrap-around. In order to be more accurate, brightness is normalized with a function modeling the illumination unit's geometry. Figure 6 shows the contributions of the two components.

4.3 Wrap Around Correction

The information, which pixels are invalid due to wrap-around can be used to correct their value. This is demonstrated in figure 7. One problem here is that



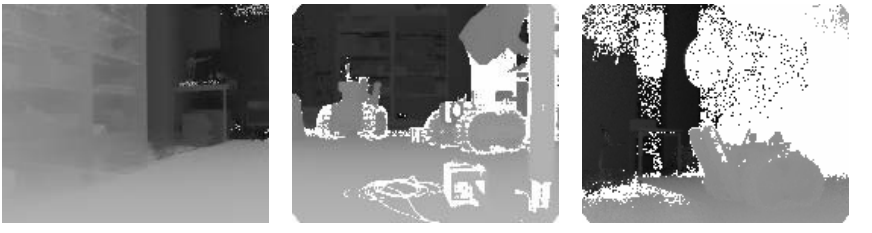
(a) Close boxes (b) Dark material (c) Sun

Fig. 4. Three more test scenes showing other sources of error that can be filtered out; here normal images of the scenes from a webcam next to the SwissRanger are shown

(a) Unfiltered distance image: Dark obstacles in the back appear too close

(b) No wrap-around, but noisy and too close trouser due to dark material

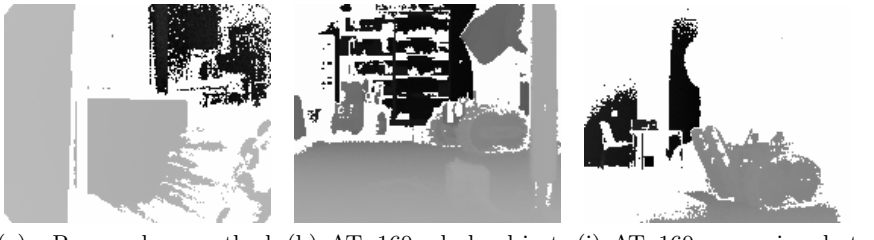
(c) Sun causes error and random noise



(d) Unfiltered distance image with objects removed, for comparison

(e) Proposed method: dark pixels removed – 13.1 % d. p.

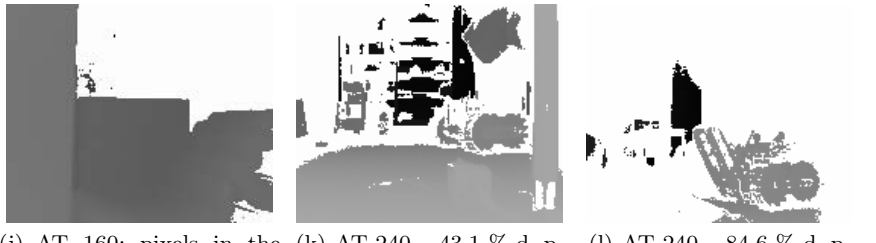
(f) Proposed method: very little noise remains – 44.9 % d. p.



(g) Proposed method: correct pixels of both back- and foreground kept – 38.1 % d. p.

(h) AT 160: dark objects discarded, on edges more than in center – 28.0 % d. p.

(i) AT 160: no noise, but parts of floor and locker removed – 73.1 % d. p.

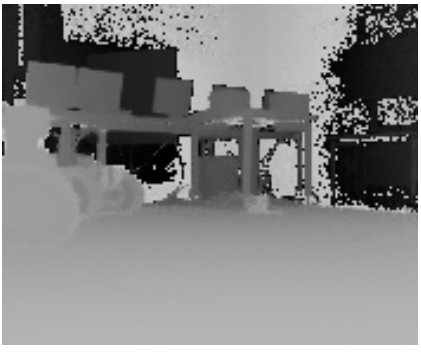


(j) AT 160: pixels in the front preserved, those in the back gone – 35.0 % d. p.

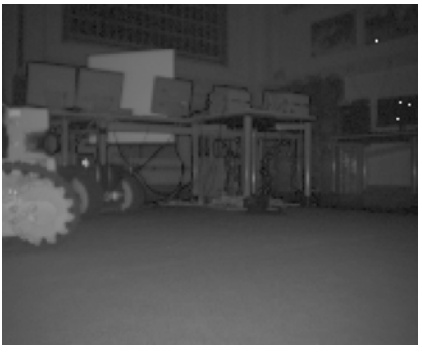
(k) AT 240 – 43.1 % d. p.

(l) AT 240 – 84.6 % d. p.

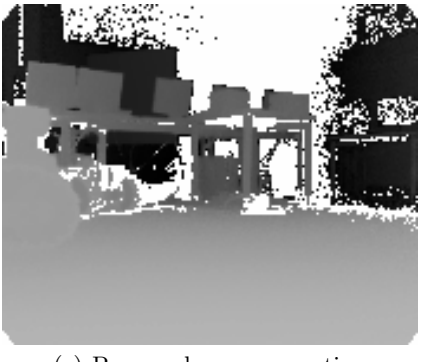
Fig. 5. Three more test scenes showing other sources of error that can be filtered out. From left column to right: light scattering, too low reflectivity, and sun light. Here, an AT of 160 is sufficient, most AT 240 images are still given for completeness.



(a) Full image



(b) Intensity image



(c) Proposed error correction

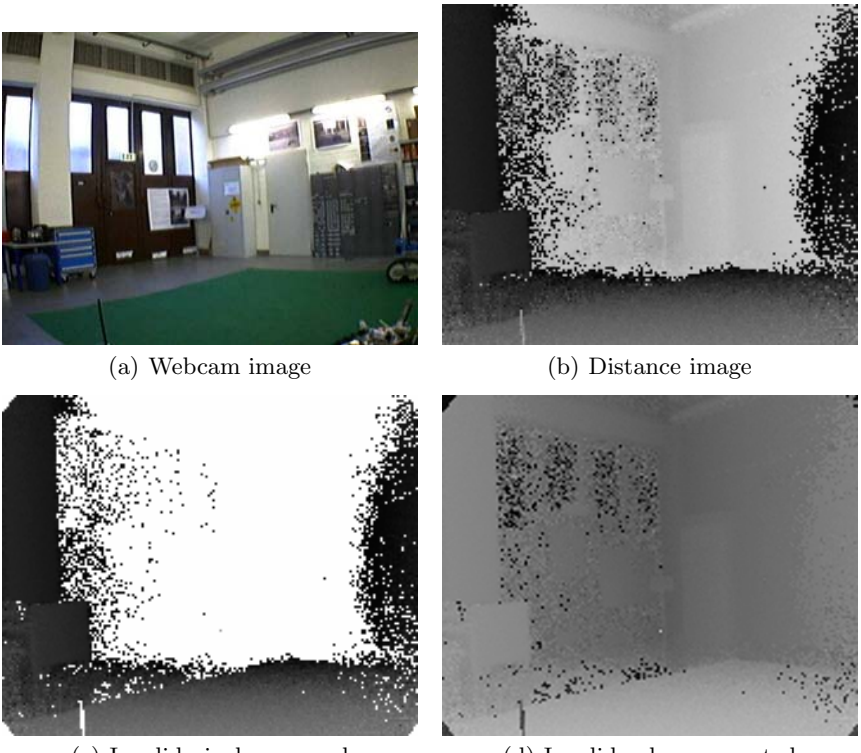
(d) Conventional error correction: AT 160

(e) Only relating distance and brightness: Wrap-around removed, but too many excluded pixels, esp. on the edges

(f) Only using position dependent amplitude threshold: Not increasing strictness towards the edges (as in (d) and (e)), but keeping wrap-around

Fig. 6. Using different error exclusion criteria on the scene in figure 2(a)

some of the pixels classified as invalid are not wrapped around but excluded due to other reasons, e.g. because they are too dark. These very few pixels hence get a wrong value in the corrected image.



(a) Webcam image (b) Distance image

(c) Invalid pixels removed (d) Invalid values corrected

Fig. 7. An example for correcting wrapped-around pixels

5 Conclusion

Time-of-flight cameras are a promising technology for 3D range data acquisition. But they suffer from a principle drawback of wrap around error, i.e., distances at multiples of the wavelength of the modulated light emitted by the camera can not be distinguished. The state of the art solution is to use a fixed threshold on the amplitude of the returned modulated light to discard far away and hence wrapped around pixels. Choosing a suited value for the amplitude threshold (AT) is very difficult as it is strongly environment dependent. Furthermore, the approach tends to discard a significant number of valuable information. Here, a novel approach is presented, which takes the geometry of the light source into account. This allows a physically more plausible estimate of the to be expected amplitudes per distance than a fixed threshold. This leads to much higher amounts of valid data in time-of-flight camera images as illustrated with a SwissRanger SR-3000. Furthermore, the quality of the correction even allows to detect and compensate for wrap around and hence to increase the nominal range of this time-of-flight camera.

Acknowledgments

Please note the name-change of our institution. The Swiss Jacobs Foundation invests 200 Million Euro in **International University Bremen (IUB)** over a five-year period starting from 2007. To date this is the largest donation ever given in Europe by a private foundation to a science institution. In appreciation of the benefactors and to further promote the university's unique profile in higher education and research, the boards of IUB have decided to change the university's name to **Jacobs University Bremen (Jacobs)**. Hence the two different names and abbreviations for the same institution may be found in this paper, especially in the references to previously published material.

Furthermore, the authors gratefully acknowledge the financial support of the *Deutsche Forschungsgemeinschaft* (DFG) for their research.

References

1. Howard, A., Wolf, D.F., Sukhatme, G.S.: Towards 3D mapping in large urban environments. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Sendai, Japan (2004)
2. Thrun, S., Haehnel, D., Montemerlo, M., Triebel, R., Burgard, W., Baker, C., Omohundro, Z., Thayer, S., Whittaker, W.: A system for volumetric robotic mapping of abandoned mines. In: Proc. IEEE International Conference on Robotics and Automation (ICRA), Taipei, Taiwan (2003)
3. Hähnel, D., Burgard, W., Thrun, S.: Learning compact 3D models of indoor and outdoor environments with a mobile robot. Robotics and Autonomous Systems 44(1), 15–27 (2003)
4. Davison, J., Kita, N.: 3D simultaneous localisation and map-building using active vision for a robot moving on undulating terrain. In: IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, December 8-14 (2001)
5. Liu, Y., Emery, R., Chakrabarti, D., Burgard, W., Thrun, S.: Using em to learn 3d models of indoor environments with mobile robots. In: 18th Conf. on Machine Learning, Williams College (2001)
6. Nüchter, A., Wulf, O., Lingemann, K., Hertzberg, J., Wagner, B., Surmann, H.: 3D mapping with semantic knowledge. In: Bredenfeld, A., Jacoff, A., Noda, I., Takahashi, Y. (eds.) RoboCup 2005. LNCS (LNAI), vol. 4020, pp. 335–346. Springer, Heidelberg (2006)
7. Poppinga, J., Birk, A., Pathak, K.: Hough based terrain classification for realtime detection of drivable ground. Journal of Field Robotics 25(1-2), 67–88 (2008)
8. Wulf, O., Wagner, B.: Fast 3D-scanning methods for laser measurement systems. In: International Conference on Control Systems and Computer Science (CSCS 14) (2003)
9. Surmann, H., Nuechter, A., Hertzberg, J.: An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor environments. Robotics and Autonomous Systems 45(3-4), 181–198 (2003)
10. Wulf, O., Brenneke, C., Wagner, B.: Colored 2D maps for robot navigation with 3D sensor data. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), vol. 3, pp. 2991–2996. IEEE Press, Los Alamitos (2004)

11. Weingarten, J., Gruener, G., Siegwart, R.: A state-of-the-art 3D sensor for robot navigation. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), vol. 3, pp. 2155–2160. IEEE Press, Los Alamitos (2004)
12. CSEM: The SwissRanger, Manual V1.02. CSEM SA, 8048 Zurich, Switzerland (2006)
13. Lange, R., Seitz, P.: Solid-state time-of-flight range camera. IEEE Journal of Quantum Electronics 37(3), 390–397 (2001)