A Characterization of 3D Sensors for Response Robots

Jann Poppinga, Andreas Birk, and Kaustubh Pathak

Jacobs University Bremen*
Campus Ring 1, 28759 Bremen, Germany

Abstract. Sensors that measure range information not only in a single plane are becoming more and more important for mobile robots, especially for applications in unstructured environments like response missions where 3D perception and 3D mapping is of interest. Three such sensors are characterized here, namely a Hokuyo URG-04LX laser scanner actuated with a servo in a pitching motion, a Videre STOC stereo camera and a Swissranger SR-3000. The three devices serve as prototypical examples of the according technologies, i.e., 3D laser scanners, stereo vision and time-of-flight cameras.

1 Introduction

Sensors providing 3D range data are getting more and more important for mobile robotics in general and for Safety, Security and Rescue Robotics (SSRR) in particular. 3D data allows response robots for example to estimates the size of gaps, to construct realistic maps of unstructured disaster environments, or to detect human victims from shape. Concrete examples of research with relevance for SSRR where 3D sensors are used include 3D mapping [1][2][3][4][5], semantic environment classification [6] or the detection of drivable terrain [7].

The main purposes of this paper are twofold. First, updated information concerning the state of the art of according sensors compared to previous discussions in the literature are provided. Second, the focus is on devices that can be directly used on mobile robots, especially in the context of Safety, Security and Rescue Robotics.

The rest of this paper is structured as follows. In section 2, 3, and 4 a concrete 3D laser scanner, stereo camera, respectively time-of-flight camera are introduced and their general properties are discussed. A direct comparison is presented in section 5. Section 6 concludes the paper.

2 3D Laser Range Finder

Laser Range Finders (LRF) in their standard form are the predominant sensor for mapping on mobile robots. But as the interest on 3D mapping increases, the

* Formerly International University Bremen.

J. Baltes et al. (Eds.): RoboCup 2009, LNAI 5949, pp. 264–275, 2010.

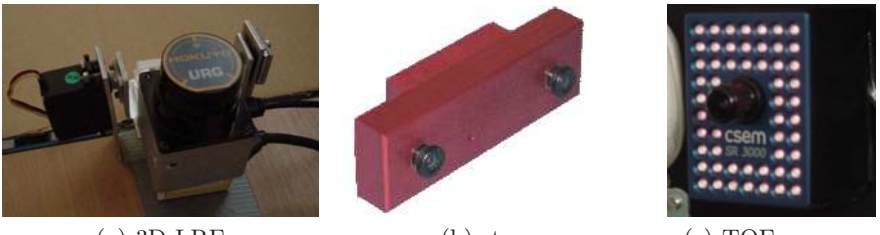
limitations of the standard systems that only scan a horizontal plane become more and more apparent. There are meanwhile quite some 3D laser scanners available off the shelf. A very recent but coarse overview is given in [8]. A detailed discussion of an example sensor can be found in [9]. These systems are designed for geometric applications and hence not really suited for mobile robots, mainly due to their weight and power consumption.

It is hence very popular within the 3D mapping community to take a standard 2D laser scanner and to actuate it to get data from an additional dimension. One option is to mount two scanners perpendicular to each other and to exploit the movement of the robot itself [10]. But most commonly, the sensor is directly driven with some servo-mechanism to also get significant amounts of 3D data on a stationary or slow moving robot. In doing so, different motions of the scanner with respect to the robot's frame are possible. Examples are rolling [11], pitching [12] or yawing movements [13]. The prototypical system presented here uses a pitching motion.

This system is based on a Hokuyo URG-04LX [14]. It has a 240 degrees field of view, which is scanned in 683 steps, i.e., the angular resolution is 0.36 degree. It can cover 0.2 to $4m$ with a resolution of $\pm 10mm$. It is based on a near infrared laser-diode with $\lambda = 785nm$. The URG-04LX is interfaced via USB to the mobile robot. For the servo that moves the sensor, a small board based on a PIC18f2410 micro-controller is used, which is interfaced via RS232. There are of course other 2D laser scanners that could be used as basis for a 3D sensor. Example characterizations of other 2D laser scanners are [15] for the Hokuyo PBS03 and [16] for the popular Sick LMS 200.

The main advantage of the Hokuyo URG-04LX is its compactness (l:50mm, w:50mm, h:70mm), small weight (160 g) and low power consumption (2.5 W). These advantages are traded in with the relative short range of 4 m. Our system uses a standard servo for a pitching motion of the scanner (figure 1(a)). The main difference of this system compared to other 3D scanners, for which larger scanners like the Sick LMS 200 are very popular as sensor basis, is its compactness in size, weight and power combined atthe cost of a shorter range. But the characteristic aspects that serve as basis for the evaluation and comparison to other 3D laser range finders are very generic.

Like laser scanners in general, the URG-04LX has a relatively slow update rate of 100 $msec/scan$. The overall time for a 3D scan depends on the range and



(a) 3D LRF (b) stereo (c) TOF camera

Fig. 1. The three different 3D sensors that are analyzed here

the resolution of the actuation of the sensor. The low cost servo in our system allows precise steps with a minimum angular resolution of 3 deg. It hence takes about 3 seconds for the 30 scans that cover a pitching motion of 90 deg, i.e., 683×90 data points. This relative slow rate for a total 3D scan is typical. The system described in [12] for example takes 3.4 seconds to produce a scan with 256×181 data points.

3 Stereo Vision

Stereo vision is a well known technique for 3D measurements in general [17] as well as for mobile robots in particular [18,19]. A common criticism for stereo vision, especially when compared to laser range finders, are its high computational requirements. This true if the generation of the disparity image is done in software. But alternatives exist like the stereo-on-chip (STOC) camera (figure 1(b)) from Videre Design [20], which has an embedded processor.

The device connects to a PC a using IEEE 1394 (Firewire) interface and consumes 2.4 Watts. It produces a 3D point cloud at a resolution of 640×480 at 30 frames per second. The cameras are CMOS imagers, rigidly mounted on an anodized aluminum chassis; the base-line is 9 cm. They have a global shutter, i.e., all pixels are exposed simultaneously. The left and right pixels are interleaved in the video stream. The device needs to be calibrated and this information is then stored on it. Both monochrome and color images can be obtained. The device board runs a version of the SRI Small Vision System (SVS) [21] stereo algorithm, which is based on area correlation.

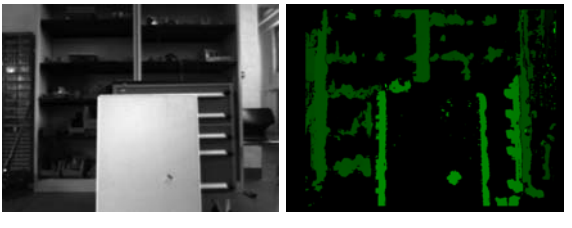


Fig. 2. The left camera image, and the corresponding disparity image of a flat white surface with almost no visual features

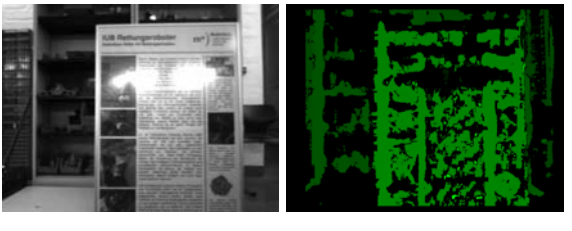


Fig. 3. The left camera image, and the corresponding disparity image of a flat surface with visual features

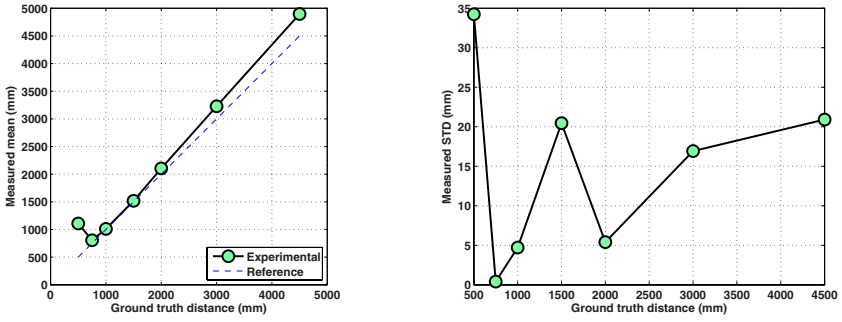


Fig. 4. The mean and the standard deviation of distance measured with a Videre STOC using the points lying within the central 150 × 150 sq. mm. The results are based on averaging 20 measurements at each distance.

Figures 2 and 3, illustrate the main disadvantage of stereo systems, namely their dependency on sufficiently distinctive regions for matching. As the disparity image in figure 2 shows, a flat featureless board is not detected by the device. Only the boundary of the board is detected. This should be compared to the disparity image in figure 3 where the board has some text on it. This time the interior of the board is captured quite well.

The dependence on environment conditions is also illustrated by an experiment to estimate the discrepancy from the ground truth for the points lying within the central 150×150 sq. mm. This experiment was performed on a freshly calibrated device. The results are shown in figure 4. The ranges of objects too close to the cameras, namely at a distance less than the focal length, are incorrect. As the distance of the object increases, the range error steadily increases as one can expect. Nevertheless, note the irregular behavior in the standard deviation, which can be related to environment conditions.

4 Time of Flight Cameras

The Swissranger SR-3000 (figure 1(c)) is a time-of-flight camera, i.e., a technology that is much less established than laser scanners or stereo cameras. The device is produced and marketed by the "Centre Suisse d'Electronique et de Microtechnique" (CSEM). An earlier version of this sensor, namely the SwissRanger-2, is characterized in some detail in [22]. The technological principles on which this sensor is based are described in [23]. Roughly speaking, this type of sensor uses an array of cells similar to an imager of a camera to measure the phase-shift of emitted modulated infrared light. By this, a time-of-flight based distance measurement can be done simultaneously in each cell of the array.

The sensor connects to a Computer via USB. It is delivered with an API and example applications for Windows, Linux and Mac OS X. The sensor generates color encoded distance images as well as intensity images. The first correspond

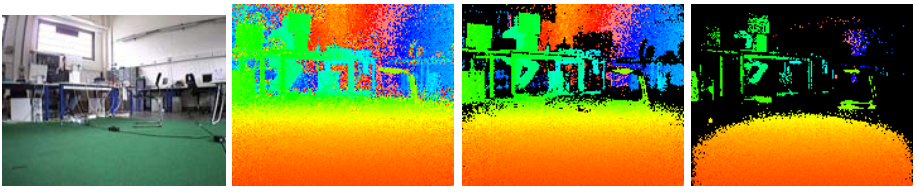


Fig. 5. At the left, the underlying scene shot with a normal camera; next to it the effects of the Amplitude Threshold (AT) on range images with a SR-3000

to the measured phase-shift, the second to the amplitude of the signal. The measured distances are encoded by hue, range from red (near) to violet (far). Black pixel indicate that no useful information was measured, mostly because too little modulated light was received. The underlying technology of the SR-3000 is relatively young and far less established than laser range finders or stereo cameras. Though promising, there are still many drawbacks.

The first fundamental problem is the *wrap-around* error. As phase-shift is measured, $c = c + k \cdot 2\pi$ for arbitrary k. This means if the phase of 2π corresponds to for example 8 m then two points in a distance of 0.2 m and 8.2 m lead to the same measurement.

In theory, there is the option to use the so-called *Amplitude Threshold* parameter to remedy this problem. When used with this parameter, the camera only returns distance values for pixels with a certain minimal brightness. Note that the range for the unused pixels is set to zero. This parameter is potentially also useful for eliminating other kinds of errors as discussed later on. Roughly speaking, a higher amplitude threshold leads to less data points but the data points are of higher quality.

The main problem is to find a suited amplitude threshold (AT) for the current conditions in a particular environment. This is illustrated in figure 5. First of all, there is some random noise on some objects. This most prominently occurs on dark objects and around the edges of objects. Second, the amplitude depends on distance and it hence can remedy wrap-around errors. But the amplitude also depends on the reflectance properties of the objects. If the AT is chosen too low, correct data points are discarded.

An other important parameter is the integration time, which is also called exposure time. It is the time used for acquiring each frame. The perceivable brightness of the illumination unit increases with the integration time. The integration time of course determines the frame rate and also the power consumption as it influences the brightness of the illumination LEDs. The quality of the images can increase with the exposure time, but also the minimal distance increases. An autoillumination feature for the device can be used to automatically determine integration time and illumination intensity. The user can supply a certain range for the integration time and the best value within the boundaries is chosen. This setting yields good results, so it should be used unless there are reasons not to do so, e.g., the urgent need to save energy or to achieve a certain frame rate.

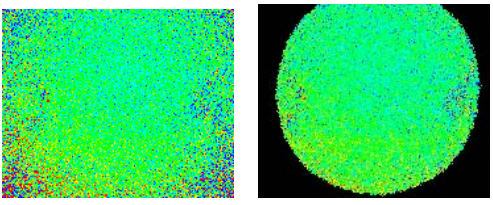


Fig. 6. Though the SR-3000 already has an extremely limited field of view, artifacts from the focused modulated light sources are apparent. When using a short exposure time and AT=0, there are is a significant amount of noise in the corners of the range image (left). When increasing the AT to 500, any information from these regions is completely discarded and a almost perfect circular shape caused by the narrow cone of the modulated illumination can be seen (right).

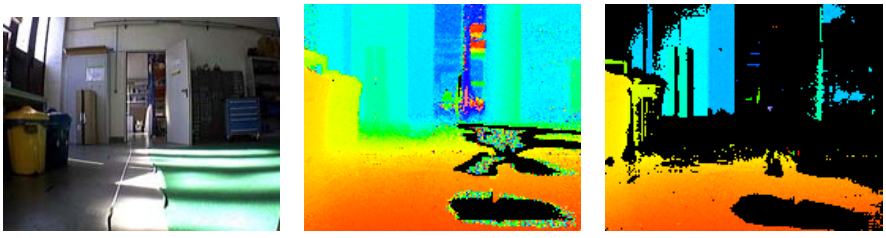
There are two possible ways around the errors caused by wrap-around of the phase-shift. The simpler one as mentioned before is to apply an amplitude threshold. This works quite often, as the amount of light reflected by an obstacle is proportional to its distance to the source of light. But when dealing with obstacles with different reflective properties, this fails. The second possibility is to use two modulation frequencies in an alternating way. As each modulation frequency entails a specific non-ambiguity range, the pixels of alternating frames wrap around at different distances. Thus, the errors in one frame can be evened out by comparing it to another frame taken at a different modulation frequency.

While frames retrieved by this method still contain some errors and are also not totally dense, this method is in general advisable. The down side is that the frame rate is reduced, especially as two frames have to be dropped immediately after the change of the modulation frequency.

An other significant drawback of the sensor is its extremely small field of view. Even with its already very limited standard view of $47^o \times 39^o$, artifacts from the LEDs as quite focused illumination sources are apparent. The main portion of the modulated light emitted by the LEDs is centered in the middle of the sensor. So when using a relatively low exposure time, the corners of the picture do not get enough light for decent measurements. An example is given in figure 6(a). Please note that the color scale for this image is extremely short; the distance between red and violet is 40 cm. This is done for illustration purposes. When the only the brightest pixels are used, i.e., AT is accordingly adjusted, the circular shape of the lighting becomes visible (figure 6(b)).

Bright light. The SR-3000 is very sensitive to ambient light conditions. Especially sunlight requires some manual tuning of AT (see figure 7).

A further problem is that objects close to the camera can make objects near them on the picture appear closer to the camera than they are; especially if the far objects are dark and the material of the near object is bright, i.e., highly reflective. An other form of irregular distortion appears when an object is close to the camera. Then, a faint ghost image of it will often appear on the opposite

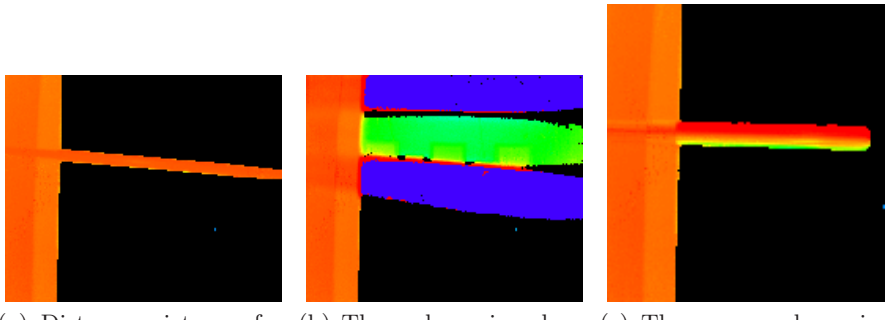


(a) Photograph of a scene with some sunlight.

(b) Random noise due to bright light, AT=0

(c) Corrected image with AT=300

Fig. 7. The SR-3000 is very sensitive to ambient light conditions. When confronted with a scene including a bit of sunlight (left), a significant amount of noise occurs with AT=0 (center). Only when tuning AT, here to 300, the problem is solved (right).



(a) Distance picture of a static rod.

(b) The rod moving down fast (about 10 m/s). The large blue and green bars are movement errors.

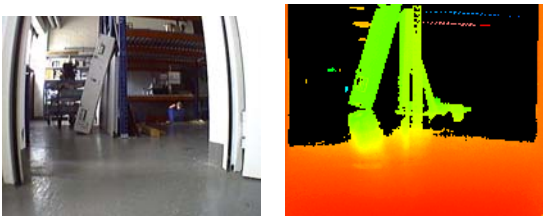
(c) The same rod moving up slowly (about 0.25 m/s). The errors become smaller but are still present.

Fig. 8. Errors caused by a moving object (AT=500)

of the frame. Yet an other source for distortions are moving objects (figure 8). When sampling them, there are always distortions at their edges, even with low exposure times. Many of these problems are also mentioned in the manual of the sensor [24].

Last but not least, the SR-3000 behaves much like a regular camera and it is hence can be mislead by reflections. When a reflected surface appears in its view, the SR-3000 will not measure the distance to where the reflection occurs, but the perceived distance to the objects visible on the reflective surface. The measured distance is then the distance from the camera to the reflective surface plus the distance from the surface to the object. An example is shown in figure 9.

As a rough quantitative comparison between the SwissRanger to the stereo camera, the same measurements concerning accuracy were performed: the discrepancy from the ground truth for the points lying within the central 150×150



(a) Photograph of a scene with reflections due to a glossy floor.

(b) Distance information with AT 500

Fig. 9. Reflections, here on a floor, cause the measurement of false distances

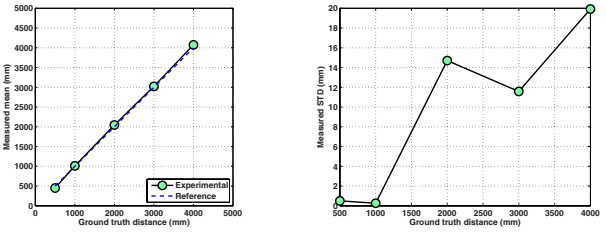


Fig. 10. The mean and the standard deviation of distance measured with a SwissRanger SR-3000 using the points lying within the central 150 × 150 sq. mm. The results are based on averaging 20 measurements at each distance.

sq. mm was determined. The results are shown in figure 10. As can be seen, the measurements are quite accurate, especially for longer distances.

5 Comparison of the Results

The three sensors are quite similar in several aspects. First of all, all three can be relatively easily incorporated on a mobile robot as they have similar low demands in terms of space, payload and power (table 1). They also deliver roughly the same amount of data points (table 2). Note that the stereo system uses an embedded hardware to do the matching of regions to produce the disparity image. Hence all three sensors directly deliver 3D data points via their interface. No additional computation is needed for the raw data, which is of interest here. Obviously, the representations differ, but any robotics application will require some coordinate transformation of this raw data anyway. The systems are also quite similar in terms of cost, which is a few thousand Euros each.

The main differences are the update frequency and the quality of the data. The stereo and the time-of-flight camera beat the 3D laser scanner by far in terms of sampling frequency. The 3D laser scanner in contrast delivers much higher quality data. This does unfortunately not only hold with respect to Gaussian noise, which could be easily compensated by averaging, which would be supported by

Table 1. General physical properties

	size	weight	power
3D URG40-LX	l:50mm, w:50mm, h:70mm	425 g	2.5 W
STOC	l:132mm, w:39mm, h:44mm	261 g	2.4 W
SR-3000	l:42.3, w:50, h:67mm	162 g	12 W[1]

Table 2. Data acquisition

	number of data points	sampling rate	field of view	range
3D URG40-LX	683×90	0.3 Hz	$240^o \times 90^o$	0.2 - 4 m
STOC	640×480	30 Hz	$70^o \times 52^o$	0.75 - 3 m
SR-3000	176×144	$\leq$ 50 Hz[2]	$47^o \times 39^o$	0.6 - 8 m[3]

the higher sampling rates of the two camera sensors. But in addition to high Gaussian noise, the stereo and the time-of-flight camera suffer from fundamental drawbacks that can not just be remedied by higher update frequencies. The differences in terms of quality of the data of the three sensors can not easily be presented in detailed quantitative terms as they strongly depend on environmental conditions. An according exhaustive discussion would by far exceed the limits of this paper. Hence a qualitative analysis, which should be at least as useful, is given here.

Stereo requires the matching of identifiable regions. Featureless objects are simply not detected, no matter how often or from which position they are viewed. Hence, a significant amount of regions is not coped with. This holds especially for plain walls, which can be found in many corridors, offices or other "structured" environments. At least, stereo does not provide any completely false depth information. If no match can be performed, this is indicated accordingly. The error of depth estimates significantly increases with range. But near objects, if detected, are well measured. Stereo hence can serve as a fast 3D sensor on a mobile robot.

The basic idea of a time-of-flight camera as a solid state, 3D range finder is at first glance very promising. But the technology is still in its infancy. To get useful data out of the SR-3000, some parameters have to be tuned by hand. Some problems are of a fundamental nature due to the underlying technology, e.g., the wrap-around error or wrong measurements due to reflections. Others, like the strong sensitivity to ambient light conditions or the extremely narrow field of view, may be remedied by software or future hardware generations. This sensor can nicely supplement a stereo vision system as their strengths are complementary. A time-of-flight camera performs best on homogeneous flat surfaces,

[1] The exact power consumption depends on the integration time.
[2] The exact sampling frequency depends on the integration time.
[3] The range depends on the frequency setting for the modulation. Here the values for 20 MHz are given for which the best results were achieved.

where stereo tends to work badly. The accuracy of data points, which are not corrupted by structural errors is very high as indicated in figure 10. Nevertheless, a usage of just this sensor for 3D data acquisition on a mobile robot in an arbitrary environment is at least non-trivial.

Laser scanners are a very mature, well established technology. The currently available off the shelf 3D scanners are mainly targeted for geometric applications. They are hence not really suited for mobile robots, mainly due to their weight and power consumption. Also, the cost is typically much higher than for the sensors presented here. But a 3D sensor can be easily constructed from any of the popular 2D sensors supplemented with an actuator. The main disadvantage in general, no matter whether off the shelf or based on an own design, is that the acquisition of a single scan takes several seconds. This is a time period where the motion of the robot usually can not be ignored anymore, i.e., either the robot stops for taking a full scan or the motion is compensated for in the acquisition process. The quality of the data points is much higher than compared to stereo and time of flight cameras. This holds with respect to two aspects. First, the mean and standard deviation of measurements compared to ground truth is much smaller. Second, the amount of completely false or non-classified points is very small. This is mainly due to the fact that laser scanners sample single points, which in turn causes their main disadvantage, namely the relatively long time it takes to generate a single scan.

6 Conclusion

The acquisition of 3D data is increasingly important for mobile robots, especially for systems operating in unstructured environments like response robots . Three different technologies are characterized here based on three prototypical devices, namely a 3D laser scanner, a stereo camera and a time-of-flight camera. The results and their discussion provides general guidelines for system developers as well as potential end users.

Laser scanners are the most mature and reliable technology. Off the shelf 3D scanners are mainly targeted for geometric applications and not really suited for mobile robots. Turning a common 2D device into a 3D scanner is relatively easy. Here, an example based on a Hokuyo URG04-LX and a servo for a pitching motion was presented. Such systems can deliver very high quality data, but at the cost of relatively slow update rates. Stereo cameras in contrast have very high sampling frequencies, especially when using embedded hardware like the Videre STOC presented here. In addition to the typically higher error at larger distances, stereo systems suffer from the drawback that they require regions that can be matched. Featureless objects, especially plain walls, are not detected. Last but not least, the CSEM Swissranger SR-3000 as a time-of-flight camera combines in theory all the advantages of a laser scanner and a stereo camera. But in reality, the technology is still in its infancy. The update rate is high, but the quality of the data is poor. The device requires a high amount of parameter tuning. Ideally, all three sensors are simply used together for 3D data acquisition

if space, payload, power and budget constraints permit. To quite some extent, they supplement each other.

Acknowledgments

We gratefully acknowledge financial support by the *Deutsche Forschungsgemeinschaft* (DFG).

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.D., Haehnel, D.F., 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. Nuechter, 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. Point of Beginning 2006 3D Laser Scanner Hardware Survey (2006), `http://www.pobonline.com/POB/Protected/Files/PDF/POB0506/LaserScanningSurvey.pdf`
9. Langer, D., Mettenleiter, M., Frohlich, C.: Imaging laser scanners for 3-d modeling and surveying applications. In: IEEE International Conference on Robotics and Automation (ICRA), vol. 1, pp. 116–121 (2000)
10. Thrun, S., Burgard, W., Fox, D.: A real-time algorithm for mobile robot mapping with applications to multi-robot and 3d mapping. In: ICRA, pp. 321–328 (2000)
11. Wulf, O., Wagner, B.: Fast 3d-scanning methods for laser measurement systems. In: International Conference on Control Systems and Computer Science, CSCS14 (2003)
12. 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)
13. 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)

14. Automatic, H.: URG-04LX Scanning Laser Range Finder (2006), http://www.hokuyo-aut.jp/products/urg/urg.htm
15. Alwan, M., Wagner, M., Wasson, G., Sheth, P.: Characterization of infrared range-finder pbs-03jn for 2-d mapping. In: IEEE International Conference on Robotics and Automation, ICRA, pp. 3936–3941 (2005)
16. Ye, C., Borenstein, J.: Characterization of a 2d laser scanner for mobile robot obstacle negotiation. In: IEEE International Conference on Robotics and Automation (ICRA), vol. 3, pp. 2512–2518 (2002)
17. Barnard, S.T., Fischler, M.A.: Computational stereo. ACM Comput. Surv. 14(4), 553–572 (1982)
18. Murray, D., Little, J.J.: Using real-time stereo vision for mobile robot navigation. Autonomous Robots 8(2), 161–171 (2000)
19. Elfes, A.: Using occupancy grids for mobile robot perception and navigation. Computer 22(6), 46–57 (1989)
20. Videre-Design (2006), http://www.videredesign.com/
21. Konolige, K.: The Small Vision System (2006), http://www.ai.sri.com/~konolige
22. 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)
23. Lange, R., Seitz, P.: Solid-state time-of-flight range camera. IEEE Journal of Quantum Electronics 37(3), 390–397 (2001)
24. AG, M.I.: Swissranger SR-3000, Manual V1.02. (July 2006)