Standardization of Intensity-Values Acquired by Time-of-Flight-Cameras

Michael Stürmer, Jochen Penne, Joachim Hornegger
University of Erlangen-Nuremberg
Chair of Pattern Recognition
stuermer@immd5.informatik.uni-erlangen.de

Abstract

The intensity-images captured by Time-of-Flight (ToF)-cameras are biased in several ways. The values differ significantly, depending on the integration time set within the camera and on the distance of the scene. Whereas the integration time leads to an almost linear scaling of the whole image, the attenuation due to the distance is nonlinear, resulting in higher intensities for objects closer to the camera. The background regions that are farther away contain comparably low values, leading to a bad contrast within the image. Another effect is that some kind of specularity may be observed due to uncommon reflecting conditions at some points within the scene. These three effects lead to intensity images which exhibit significantly different values depending on the integration time of the camera and the distance to the scene, thus making parameterization of processing steps like edge-detection, segmentation, registration and threshold computation a tedious task. Additionally, outliers with exceptionally high values lead to insufficient visualization results and problems in processing. In this work we propose scaling techniques which generate images whose intensities are independent of the integration time of the camera and the measured distance. Furthermore, a simple approach for reducing specularity effects is introduced.

1. Introduction

In the past years the modality of Time-of-Flight (ToF) imaging became more and more attractive to a growing community [11, 1, 3, 5, 12, 4, 7]. The distance to the objects in the scene is measured by emitting a modulated near-infrared light signal and computing the phase shift of the reflected light to the phase of the emitted signal. For each image acquired by a ToF camera, information is collected over multiple periods of the reference signal. The time span elapsed during sampling of the received signal for one image is called integration time. The longer the integration time is, the more samples are collected and the signal can be reconstructed more accurately. The drawbacks of long integration times are significantly slower framerates and possible saturation effects. A reasonable compromise has to be made.

In addition to the distance information another dataset is provided by the cameras, containing the amount of the reflected reference signal. These so called amplitude or intensity values clearly depend on the current integration time of the camera in a (ideally) linear way (for the evaluation of this work, valid integration-time values are in the range from $200\mu s$ to $50000\mu s$). Outside of these bounds either noise or saturation made the acquired data unreliable. The observed amplitudes are within completely different value ranges, even if a still scene is captured and only the integration time changes. Additionally, the intensity values are attenuated with increasing distance of the object to the camera. This decay is nonlinear and leads to much brighter values within regions of the image, where the objects are close to the camera. Another effect that may tamper the amplitudes and lead to very bright spots in the images, are uncommon reflecting conditions in the scene. These spots can be compared to specular effects in images taken from traditional cameras.

All these effects lead to strong intra- and inter-image differences of the intensity value ranges, depending on the integration time, the distances to the object and the reflecting conditions and orientations of the surface. Bright specular regions and large value differences for the same color/material not only lead to bad visualizations with low contrast, but also to complex filter parameterizations. Parameters for edge detection, threshold computation, segmentation and classification have to be adjusted according to the integration time.

Still, the amplitude data delivered by ToF-cameras contains valuable information about the scene. Where distances may be too noisy or equal (imagine a black paper on a white desk where the measured distances are equal but the intensities are not), additional structure information can be extracted using this modality. Processing methods can be significantly enhanced when amplitude information is incorporated [1, 2]. So it is advisable to use all available infor-

978-1-4244-2340-8/08/$25.00 ©2008 IEEE

mation from a ToF camera in order to improve the accuracy of the various processing steps. One should however, keep in mind that the original amplitude values are still usefull, when evaluating the quality of the calculated distance values [11, 9].

The API provided by the PMDTec library for their products in ToF-imaging already provides access to distance-weighted intensities [10]. Yet, there is no information available on how the compensation is done and no parameterization or distance denoising in advance is possible. As not every manufacturer provides this kind of preprocessing of the intensity data and possible implementations may vary, a camera independent scaling algorithm is still very usefull and can even be used for evaluating software driver based implementations of amplitude scaling algorithms. Distance weighted scaling of amplitude values was also proposed by Oprisescu *et al*. [9]. The authors weighted the values by the square of the measured distances. In our calibration routine we propose an empirical way of determining the coefficients of the scaling polynomial, so that the resulting images will show the desired value range.

More specifically, we propose a variety of scaling approaches of the amplitude values depending on the set integration time and the measured distance. The aim is to achieve a stable standardized value range for amplitude images acquired by ToF-cameras independent of the integration time and the distance of the camera to the scene. In section 2 an explanation of the different effects on the amplitudes is given and the methods for reducing the bias are described. In section 3 results are presented. All experiments where done using a Swissranger SR-3100. The integration time of the camera may be set via the API to integer values from 0 to 255 where 0 corresponds to an integration time of $200\mu s$ and each increment adds another $200\mu s$.

2. Theory

2.1. Integration time based deviation

The distance values acquired using ToF-imaging with a modulated reference signal are calculated by computing the phase shift of the reflected signal to the current signal within the camera [6, 8]. This is done by sampling the received signal at constant positions over many intervals of the modulation frequency and accumulating the corresponding values. The samples A_1, A_2, A_3 and A_4 are collected in so-called bins on the camera chip. They are proportional to the number of photons incident on the corresponding chip positions. From these samples the signal is reconstructed and the distance to the scene is calculated. The amplitude a of the reconstructed signal is computed by:

$$a = \frac{\sqrt{(A_1 - A_3)^2 + (A_2 - A_4)^2}}{2}. \tag{1}$$

The longer these single signal values are accumulated, the more the noise is reduced and the accumulated signal can be reconstructed more accurately. In other words the amplitude values can be used as a quality measure of the computed distances. Theoretically, the assumption applies, that double integration time values lead to double intensity values. A linear scaling is thus straight forward and could be trivially implemented. The accumulated amplitude value can then be expressed as:

$$A_{final} = A_{spec}c, \tag{2}$$

where A_{final} is the acquired amplitude over an integration time t_i, A_{spec} is the amplitude of a single sample of the signal and c is the number of samples collected during the integration time t_i. Yet this assumption only holds for low integration times where no or neglectable saturation effects appear. For higher integration times (or if the incident light on the chip is too intense) the increase in the amplitudes becomes smaller (see fig. 1). Thus saturation effects have to be included in the computation of the final amplitude value:

$$A_{final} = A_{spec}c - S_{saturation}(A_{spec}, c, o). \tag{3}$$

The nonlinear saturation term $S_{saturation}$depends both on the amplitude value of the pixel and on the constant offset o of the received signal. Depending on the reflection properties of the scene, at a certain integration time value all the bins of a ToF-pixel become saturated. From the resulting four equal values A_1, A_2, A_3, A_4 no reliable signal can be reconstructed, the amplitude decreases to zero and distance values become invalid. This effect is often called as over saturation.

The longer the integration time, the higher the amplitude values become. So the value range of the whole image is shifted. This leads to incomparable intensity-images for different integration times, even for identical scenes. Consequently, parameters for processing these images have to be adjusted. Ideally, one should compensate every value depending on the saturation level $S_{saturation}$ of the corresponding pixel. As this information is not provided by any ToF-camera (as far as the authors know), a global scaling of the image is proposed to reduce the saturation effects as far as possible. A lookup table (LUT) S_{itime} is computed, providing a scalar multiplier for each integration time value. All the points within one image are scaled by the corresponding multiplier. So, given a integration time t_i, the integration time independent amplitude image A_{itime}can be computed by:

$$A_{itime} = A_{final}S_{itime}(t_i). \tag{4}$$

Using the acquired reference sequence (fig. 1), it can be seen that integration time values above $27.2ms$ (corresponds to a value of 135 in the API) lead to over-saturation

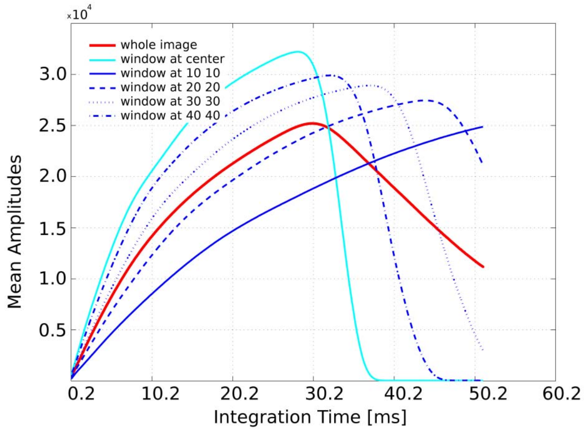

Figure 1. Mean intensity values of one scene over different integration times and at different regions of the image.

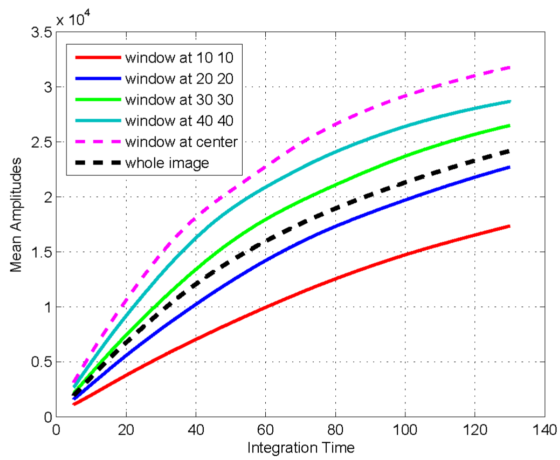

Figure 2. The mean amplitudes that were used for setting up the LUT for scaling (extract from fig. 1)

and the mean values are not usable any more for setting up a reliable vector for scaling of the amplitudes. So the proposed mapping is only set up for integration times up to $27.2ms$ (see fig. 2). For higher values no valid scaling factor could be provided. This stated integration time range heavily depends on the reflectivity of the material used for calibrating the method and on the surroundings of the scene like additional light sources. Ritt [11] also proposed a formula for intensity correction based on integration times. However, he assumes a linear amplitude gain.

Furthermore, the LED-array of the ToF camera acts as the illuminant of the scene and shows a strong focus on the center of the image, thus leading to a inhomogeneous illumination of the scene. Hence the intensities in the middle of the image are significantly higher than at the borders. As a consequence, the saturation levels of the pixels differ within one image, making it impossible to obtain an amplitude image which is uniformly independent of integration time. So,

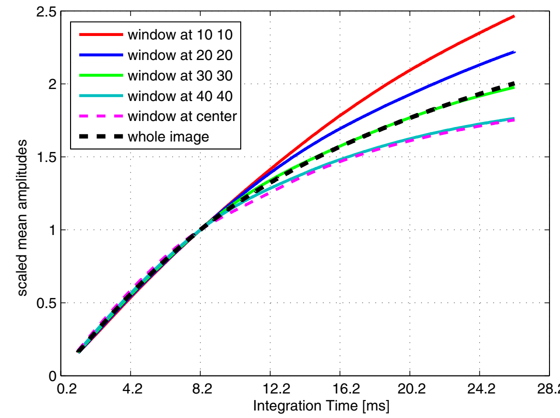

Figure 3. Normalized mean amplitudes, acquired using different integration times and at different image regions.

when looking for a single scalar multiplier a compromise has to be made. The deviation of the amplitude gain in different image regions can be seen in (fig. 2). The size of the window over which the mean value was computed, was set to 20×20 pixels. A single scaling function $S_{itime}(t_i)$ for the entire image was derived using the mean amplitude values of the whole image. The amount of scaling was chosen so that these amplitudes would remain unscaled at $8.2ms$. These normalized amplitudes were then inverted to set up the mapping.

2.2. Distance based deviation

The amount of light emitted per unit time by the reference light source is constant. If the area that is illuminated becomes larger, i.e. by capturing a scene with a greater range of depths, fewer photons are reflected from a specific point in the scene. As the intensity values are proportional to the incident light on the chip, the amplitudes decrease if the scene is farther away. Simple geometric considerations and taking into account that the emitted light has to travel both to the object and back to the sensor, lead to an analytic estimate of the decay:

$$A_{dist} = I_{src} \frac{1}{(2d)^2 + 1}, \tag{5}$$

where A_{dist} is the acquired amplitude value, I_{src} is the light density of the source and d is the distance to the scene. From this simplified formula, which neglects any further parameters like camera sensitivity, one can derive a scaling polynomial of order two for scaling the amplitude values depending on the distances. While Ritt [11] gives a completely analytic approach for scaling without reference measurements, we propose a way for compensating for the attenuation due to the distance d, which is based on acquired data and which can easily be recalibrated doing only a few simple processing steps.

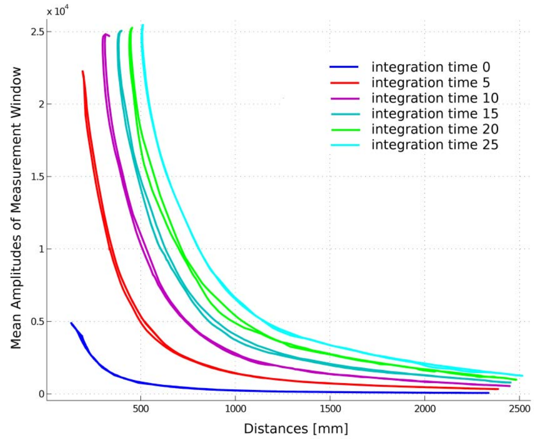

Figure 4. Attenuation of the amplitudes over the measured distance at different integration times.

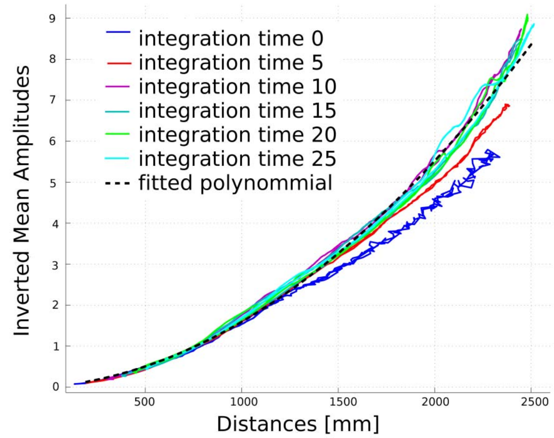

Figure 5. Scaled inverse decay of amplitudes over distances for different integration times.

Figure 6. Image with specular reflections. On the left the original image, on the right the intensity-corrected one.

A reference decay of the amplitudes depending on the distance values is acquired (see fig. 4). This is done for several integration times to ensure the decay of the amplitudes due to the distance is independent from the integration time. The measurements with smaller integration times start at shorter distances, because the starting distance is chosen to be the lower bound before over-saturation occurs and this bound is influenced by the integration time. The reference measurements for each integration time are inverted and normalized so that amplitudes at distance value of $750mm$ become equal (see fig. 5).

In order to improve our accuracy, we first fitted a cubic polynomial to each individual reference decay that we obtained for a specific integration time. A cubic polynomial was chosen so that we can have a slight oversampling compared to the original (see. eq. 5). Because the decay curves are not tightly clustered (see. fig. 5) we fitted a second cubic polynomial through the whole set of available valid points, which then is used as the final scaling functional. For this final fitting the decay with a integration time of $0.2ms$ is not used, as the values deviate quite strongly from the rest and for very short integration times the acquired distances are not as reliable as for longer ones. The advantage of the proposed method is, that it is based on real data and thus takes into account any other type of attenuation. Furthermore, the scaling can be recalculated at any time if conditions change.

2.3. Special reflection properties

Due to uncommon reflection conditions in some cases, small image parts may show very high intensity values. These can lead to bad visualizations of amplitude images. An example image showing such a behaviour is given in fig. 6. A very straight forward approach to cope with this problem is, to set all values that are above a certain percentage of the maximum intensity (e.g. 90%) to the same value. Yet in images with general high values this leads to a significant loss of information. Regions with large intensities are set to one equal value and all structure information within is lost. Better approaches to cope with this problem are based on threshold computation depending dynamically on the standard deviation and the mean value of the intensity images. An even more promising approach is, to combine information on the intensity values with the number of pixels with that intensity value. This can be done by histogram operations. Given the historgram H_a of an amplitude image, a threshold t_{thresh} has to be defined and the cutoff value in the image is chosen to be the highest intensity with at least t_{thresh} occurences. The intuition behind it, is that we only have a small number of specular pixels.

3. Experiments and Implementation

To estimate a usable scaling factor of the amplitudes with respect to the integration time, an image sequence with increasing integration time values was acquired (see fig. 1). The SR-3100 was aimed at a white wall with reasonable reflecting properties. After each increment of the integration time by $200\mu s$, a number of 50 Frames was skipped, because the mean intensities showed some set-

Figure 7. Intensity image of two different scenes (left: scene 1, right: scene 2)

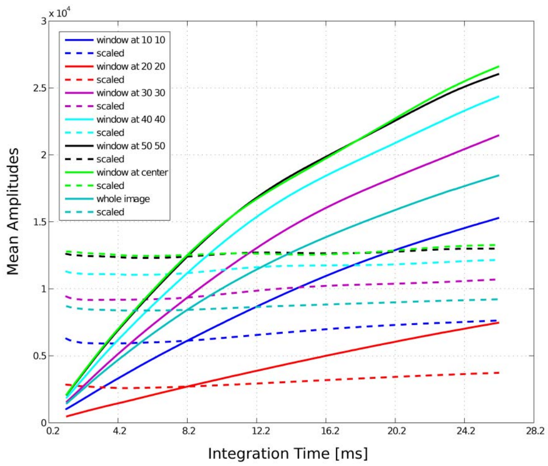

Figure 8. Mean intensities for scene 1, before and after integration time correction.

tling effects which are suspected to be due to temperature changes within the camera or other physical conditions [11]. To validate the method proposed in section 2.1, image sequences of two still scenes with changing integration times were evaluated. Sample intensity images of the scenes are shown in fig. 7. The original and the corrected mean intensity values of the whole scene and of windows of size 20×20 pixels are plotted over different integration times in figs. 8, 9. Though there is still some deviation in the mean values of the images (presumably mainly due to saturation effects in the ToF-chip), this effect is not nearly as strong as before and the value range is stable enough to allow a reliable parametrization of the earlier mentioned processing methods.

To evaluate the scaling of the amplitude values depending on the distance values, the reference measurements were scaled using the determined scaling polynomial (fig. 10). The reference sequences where set up by changing the distance between the camera and a white board and computing the mean- distance and amplitude of a measurement window of 20×20 pixels. For a perceptive evaluation two example images are given with unscaled and scaled amplitudes with respect to the acquired distance values (see figs. 11, 12).

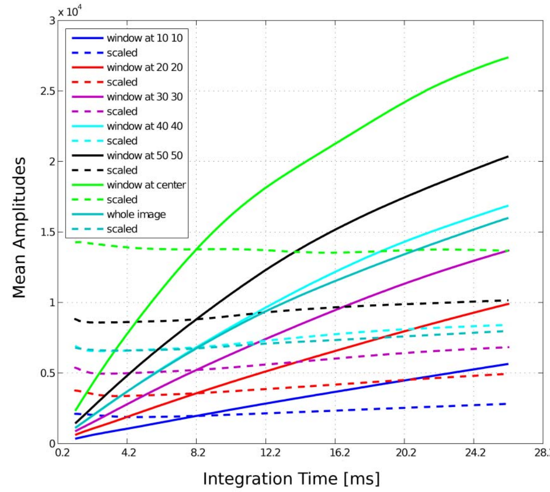

Figure 9. Mean intensities for scene 2, before and after integration time correction.

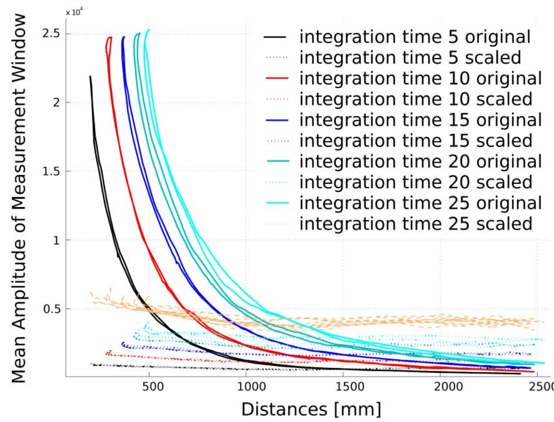

Figure 10. The reference decay due to the distance values and the scaled results. The dashed lines around the amplitude value of 0.5 additionally incorporate the integration time scaling.

Figure 11. Two coffee cups at different distances to the camera. On the left the original image, on the right the rescaled one.

A problem that has to be addressed in future work, is the denoising of the distance values by which the scaling polynomial is evaluated. For the scaling only the measured distance is available. Any noise within the distance image will be ported to the amplitudes. The results are amplitude images with a better contrast of background scenes but worse noise, if the integration time is set to low values and the

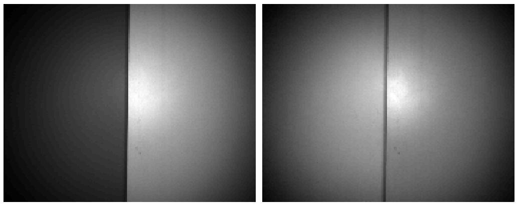

Figure 12. Two identical boards at different distances to the camera. On the left the original image, on the right the scaled one.

Figure 13. Removing of "specular" reflections. Top left: original image. Top right: setting the upper 10% to their lowes value.

distances show a strong noise. When denoising the distance values, this has to be done very carefully to prevent adding new systematic errors. The spatial filtering proposed by Oprsiescu *et al*. [9], for example, look very promising.

As an example of the proposed simple reduction of high reflective regions, result images of the different mentioned approaches are presented (see fig. 13). On the upper left, the original image without any correction is given. On the upper right the top 10% of the values are set to 90% of the maximum value of the intensities. As mentioned in section 2.3, all information within the foreground of the image is lost and set to one fix value. On the lower left a threshhold was computed from the mean intensity plus four times the standard deviation of the image. On the lower right the proposed approach using histograms is used. The structure within the bright intensities is preserved.

4. Conclusion

The intensity images delivered by ToF cameras are biased in various ways. The most significant ones are almost linear gains depending on the integration time and the distance of the scene to the camera. Furthermore, specular reflections may also be observed. We proposed methods for compensating each of these effects based solely on the acquired data. Stable parameterization of threshold computations was possible, a task that was not solvable in advance using the unscaled original images. A recomputation of the scaling parameters can be achieved.

Future work will include: a) a more accurate compensation of the integration time bias by also taking into account the nonlinear saturation behaviour b) denoising of the distance data to reduce the noise that is ported to the amplitudes when scaling with respect to the distances and c) a more thorough evaluation of the methods using several camera models of various manufacturers. We are also interested in evaluating the intrinsic distance scaling methods provided by some manufacturers.

References

[1] M. Böhme, M. Haker, T. Martinez, and E. Barth. A facial feature tracker for human-computer interaction based on 3D TOF cameras. In *Dynamic 3D Imaging, International Workshop in Conjunction with DAGM'07*, pages 26–34, Ratzeburger Allee 160, 2007. University of Lübeck, Germany.

[2] J. Feulner. Reconstruction of static scenes using TOF cameras. Master's thesis, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany, August 2007.

[3] S. Fuchs and S. May. Calibration and Registration for Precise Surface Reconstruction with TOF cameras. In *Dynamic 3D Imaging, International Workshop in Conjunction with DAGM'07*, pages 35–43, Ratzeburger Allee 160, 2007. University of Lübeck, Germany.

[4] M. Holte, T. Moeslund, and P. Fihl. View Invariant Gesture Recognition using the CSEM SwissRanger SR-2 Camera. In *Dynamic 3D Imaging, International Workshop in Conjunction with DAGM'07*, pages 53–60, Ratzeburger Allee 160, 2007. University of Lübeck, Germany.

[5] E. Kollorz and J. Hornegger. Gesture recognition with a time-of-flight camera. In *Dynamic 3D Imaging, International Workshop in Conjunction with DAGM'07*, pages 86–93, Ratzeburger Allee 160, 2007. University of Lübeck, Germany.

[6] R. Lange and P. Seitz. Solid-state time-of-flight range camera. *IEEE Journal of Quantum Electronics*, 37, Mar. 2001.

[7] T. Oggier, B. Büttgen, and F. Lustenberger. SwissRanger SR3000 and First Experiences based on Miniaturized 3D-TOF Cameras. Zürich, 2006.

[8] T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, and N. Blanc. An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger). In *Optical Design and Engineering. Proceedings of the SPIE*, volume 5249, pages 534–545, Feb. 2004.

[9] S. Oprisescu, D. Falie, M. Ciuc, and V. Buzuloiu. Measurements with ToE Cameras and Their Necessary Corrections. In *Signals, Circuits and Systems, 2007. ISSCS 2007. International Symposium on*, pages 1–4, 2007.

[10] PMDTec GmbH, http://www.pmdtec.com. *MiniSDK, Programming Manual*, May 2005.

[11] P. Ritt. Vermessung und Optimierung einer PMD-Kamera für den Einsatz in der Minimal Invasiven Chirurgie. Master's thesis, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany, 2007.

[12] T. Schamm, S. Vacek, J. Schröder, J. M. Zöllnerl, and R. Dillmann. Obstacle detection with a PMD-camera in autonomous vehicles. In *Dynamic 3D Imaging, International Workshop in Conjunction with DAGM'07*, pages 70–77, Ratzeburger Allee 160, 2007. University of Lübeck, Germany.