

Online Improvement of Time-of-Flight Camera Accuracy by Automatic Integration Time Adaption

Thomas Hoegg
Computer Graphics and
Multimedia Systems Group
University of Siegen, Germany
Email: thomas.hoegg@uni-siegen.de

Christian Baiz
Christ-Elektronik GmbH
Memmingen, Germany
Email: christian.baiz@hotmail.de

Andreas Kolb
Computer Graphics and
Multimedia Systems Group
University of Siegen, Germany
Email: andreas.kolb@uni-siegen.de

Abstract—Denoising of Time-of-Flight (ToF) range data is an important task prior to further data processing. Existing techniques commonly work on a post processing level. This paper presents a novel approach for improving data quality on the image acquisition level by automatically determining the best integration time for arbitrary scenes. Our approach works on a per-pixel basis and uses knowledge gained from an extensive analysis of the underlying inherent sensor behavior regarding intensity, amplitude and distance error to reduce the overall error, to prevent oversaturation and to minimize the adaption time. It also works well in presence of various reflectivities and quick changes in the scene. This represents a significant improvement over previous methods.

Keywords—Data analysis, Time-of-Flight, cameras, image sensors, noise reduction, performance evaluation

I. INTRODUCTION

Time-of-Flight (ToF) data denoising has always been an important discipline in image processing. Several techniques have been established during the last years. Methods for outlier removing and outlier correction have been developed to improve acquisition quality of noisy data. Denoising and optimization can be applied at different stages: at image acquisition level and/or during data processing. Previous works have shown that combining both optimization stages gives the best result [1][2]. In this work, we present a new approach to automatically determine the best integration time for arbitrary scenes using the knowledge of underlying inherent sensor behavior and properties. The approach benefits from a detailed sensor data analysis and integrates this knowledge into a novel algorithm that is more flexible and stable than a proportional feedback control system especially in unknown, arbitrary scenes. While prior work [3][4] concentrates on a global optimization of intensities or amplitudes, this work focuses on a per-pixel based improvement. This enables one to use only portions of the image or even apply per-pixel weighting, counteracting sensor properties (e.g. spatial intensity distribution) and allowing to adjust importance of certain image regions (e.g. image center over image border regions). We significantly improve the results compared to previous approaches in regard of adaption performance and thus in reduction of the mean error over time. We have used a PMD CamCube 3.0, however, our findings are applicable for other sensors as well since different ToF sensors depict a similar

behavior [5]. The PMD working principle can be found in [2]. This paper comprises the following contributions:

- A per-pixel online auto integration time estimation algorithm
- An extensive sensor behavior analysis
- A ToF simulation framework based on real physical data for data evaluation

The remainder of this paper is organized as follows. In Sec. II the related work to this paper is discussed. Sec. III shows the newly proposed automatic integration time estimation algorithm. In Sec. IV, a sensor evaluation that explains inherent sensor properties our algorithm takes advantage of is conducted. Sec. V shows our results and Sec. VI concludes this paper.

II. RELATED WORK

This paper presents a novel approach for an online adaption of the integration time for ToF cameras. In this section we will discuss important existing work related to ToF error sources, noise reduction and automatic integration time estimation.

The measurement quality of ToF cameras is influenced by several factors. Foix et al. [6] have shown different kinds of systematic errors, such as depth distortion (wiggling error), pixel-, amplitude-, temperature- and also integration time related errors. In [2], several methods are discussed on how ToF noise can be reduced. They give an overview on how errors and noise occur and state that a longer integration time causes a higher amplitude due to more incident light. This enhances the signal-to-noise ratio and the depth variance. Several other works concentrate on the performance and measurement uncertainty of ToF sensors [5][7].

Noise is an unavoidable source of measurement uncertainty. Its reduction has been studied very well during the last years. Several techniques have been established mainly concentrating on denoising in a post processing step. Methods for detecting and repairing defective areas have been developed and presented. Those approaches work either on raw data or on the final amplitude, intensity or depth image [1][8][9][10].

While most of the works for data denoising concentrate on post processing, data improvement can also be achieved by optimizing the integration time as proper saturation yields a

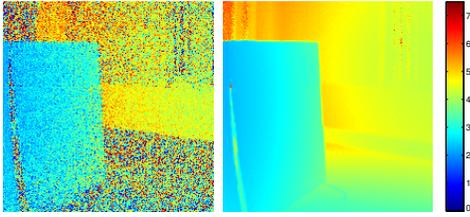


Fig. 1: The figure shows two different measurements of the same office scene in jet color encoding. The left image uses a low integration time of $50 \mu\text{s}$ while the right image uses a more adequate integration time of $2000 \mu\text{s}$.

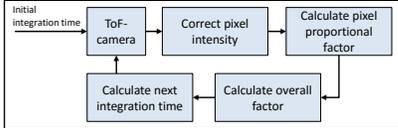


Fig. 2: The principle schema of the integration time estimation algorithm.

high signal-to-noise ratio and thus reduces noise. May et al. [3] present an approach for dynamic integration time estimation by approximating an overall mean intensity using a proportional controller. Gil et al. [4] propose an automatic integration time adaptation approach for visual servoing of mobile robots by approximating a mean amplitude. This is related to [3], but optimized for robots.

III. THE PROPOSED METHOD

Optimizing a ToF sensor's integration time for the current scene is an important aspect in reducing sensor errors stemming from both under- and oversaturation. In this section we present an approach applicable to scenes without any scene pre-knowledge, but by applying sensor specific characteristics. Thus, scenes can contain randomly placed objects with different reflectivities.

Fig. 1 shows how the appropriate choice of integration time significantly reduces noise-related errors in a scene. This section will explain how our proposed algorithm works.

A. Algorithm

The basic idea of our algorithm is to calculate the optimal integration time for the next frame on a per-pixel basis instead of averaging the whole image as prior work does. To accomplish this, we use our understanding of the sensor's behavior regarding intensity, amplitude and distance error. The algorithm schema can be found in Fig. 2.

The intensity shows to be much less susceptible to ambiguous behavior compared to the amplitude. Also, its linear behavior is much more consistent (see Sec. IV-B). Additionally, its deviation from linear behavior, which can also be described as degree of oversaturation, can be partly compensated (see Sec. IV-E), which makes it ideal for approximating a desired intensity by estimating a proportional factor. However, oversaturation causes the intensity to deviate from linear behavior which makes approximations by a proportional factor inaccurate when oversaturation is present. But these effects

can be partly compensated using a correction function (see Sec. IV-E). The amplitude has a direct correlation with the distance error (see Sec. IV-C), but we can only reliably approximate intensities not amplitudes. However, there is a linear correlation between the amplitude and intensity (see Sec. IV-D). So, by defining an ideal amplitude, correlating with the smallest error, we can derive an ideal intensity that is to be approximated.

With this knowledge, we are able to calculate the optimal integration time (the integration time at which the error is lowest) t_{opt} per-pixel in the following manner:

- 1) Determine the optimal amplitude A_{opt} , i.e. the amplitude where the error is the lowest. This is explained in detail in Sec. IV-C.
- 2) Determine the optimal intensity I_{opt} that corresponds to the optimal amplitude A_{opt} using the amplitude-intensity mapping function

$$m(x) = \sum_{i=0}^1 a_i x^i, \quad (1)$$

a linear function fit to amplitude and intensity data (see Sec. IV-D).

- 3) Determine the intensity correction function

$$h(x) = \sum_{i=0}^3 a_i x^i, \quad (2)$$

a polynomial of 3rd degree, fit to the intensity's measured and correct values. The extraction of correct intensity values and its connection to measured ones is explained in detail in Sec. IV-E.

Having determined I_{opt} , m and h , we are able to calculate the optimal integration time t_{opt} per-pixel:

- 1) Calculate the corrected intensity from the current intensity, using the intensity-correction function:

$$I_{\text{corr},x,y} = h(I_{\text{curr},x,y}) \quad (3)$$

- 2) Calculate the proportional factor:

$$f_{x,y} = \frac{I_{\text{opt}}}{I_{\text{corr},x,y}} \quad (4)$$

- 3) Calculate the optimal integration time:

$$t_{\text{opt},x,y} = f_{x,y} \cdot t_{\text{curr}} \quad (5)$$

This yields an individual optimal integration time for each pixel. Now we have to calculate an overall optimal integration time for the whole image. For this task we use a weight map, incorporating a compensation for sensor-specific behavior and regarding pixels near the image center as more important than near image borders (see Sec. IV-A). The weight map based calculation of the optimal integration time for the whole image is done in the following manner:

- 1) Calculate the pixel weight from the weight map and the gain factor:

$$w'_{x,y} = w_{x,y} \cdot g \quad (6)$$

- 2) Calculate the optimal integration time for the whole image as a weighted average of the pixel specific optimal integration times:

$$t_{\text{opt}} = \frac{\sum w'_{x,y} \cdot t_{\text{opt},x,y}}{\sum w'_{x,y}} \quad (7)$$

Compared to prior work on integration time estimation, our proposed algorithm has several advantages. It estimates the integration time on a per-pixel basis. This enables one to use only portions of the image or even apply per-pixel weighting, counteracting sensor properties and allowing to adjust importance of certain image regions (see Sec. IV-A). Additionally, it uses knowledge gained from an extensive analysis of the underlying inherent sensor behavior regarding intensity amplitude and distance error. This knowledge is used to minimize the overall error and to prevent oversaturation or at least escape from it quickly. This works well in presence of highly various reflectivities and quick changes in the scene.

IV. SENSOR ANALYSIS

The main goal of this paper is the reduction of the overall error due to inappropriate integration times in arbitrary scenes. As previous work has shown (see Sec. II), this can be achieved by preventing or at least by minimizing under- and oversaturation to increase the number of usable data points. To get a better understanding on how amplitudes, intensities and the distance error correlate with each other, an initial PMD sensor analysis is necessary. For this purpose we use eight metal plates with different colors and reflectivities: black, metallic, blue, green, red, silver, yellow and white. These plates are fixed on a planar wall at a distance of approximately 1 meter and are recorded with a distance and intensity calibrated PMD camera with an integration time range between 50 and 8000 μs with a step size of 1 μs . These measurements are used for the sensor data evaluation in the next sections. The polar ground truth for the error analysis has been calculated per pixel, using the median of various hundred measurements with different integration times while omitting under- and oversaturated values, and a final 2D median filtering step. Using a reference plane fitted to temporally averaged Cartesian distance data yielded very similar results.

A. Spatial Intensity Distribution

Pixels near the image border yield a much lower intensity and amplitude, even when the measured distance is approximately the same (see Fig. 3). Such areas are much harder to properly saturate. Often very high integration times are needed. But choosing such high integration times causes an oversaturation in the image center while properly saturating the image borders. If we weight all pixels equally when estimating the optimal integration time, the integration time will be optimized to what most pixels need to be properly saturated. Since most pixels reside outside the image center, the calculation will yield an integration time that optimizes the image borders and oversaturates the image center. If oversaturation causes a

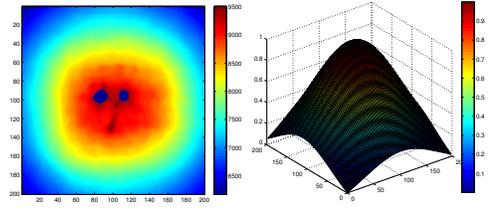


Fig. 3: The left image shows an intensity image from our test scenario (see Sec. IV). Two dark blue circular spots in the center have been invalidated due to very high oversaturation, caused by a total reflection. The right image shows a weight map with values of a Gaussian distribution that is used to weight pixels in the algorithm (see Sec. III-A).

growth in error to the same degree as non-optimal saturation, then optimizing areas where most pixels reside, namely image borders, will actually reduce the overall error. However, if a system detects and excludes oversaturated values, such an approach will reduce the number of valid points. Also, the image center often captures more important objects than the image borders and should thus be regarded as more important. As a result, we have incorporated a weight map into our approach (see Sec. III-A). This allows us to weight each pixel according to its importance. As Fig. 3 shows, we have chosen to use a Gaussian value distribution for our weight map, normalized between 0 and 1.

B. Intensity and Amplitude Behavior

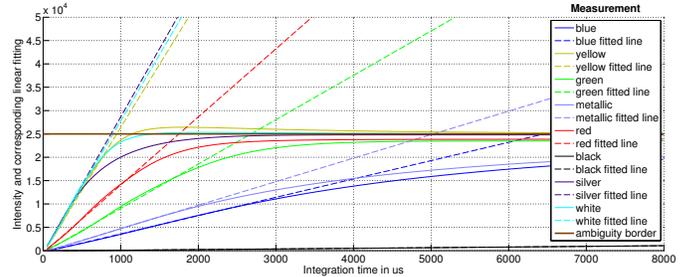


Fig. 4: Intensity to integration time and corresponding lines fitted to the intensity's linear behavior. Values above the ambiguity border (brown line) are considered ambiguous.

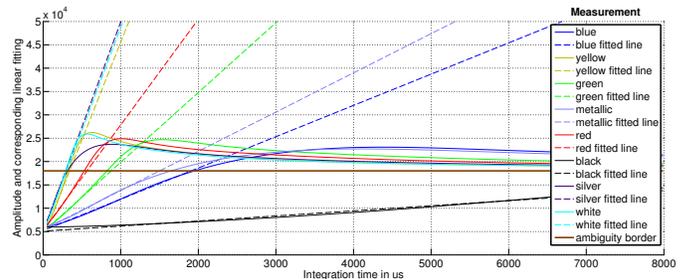


Fig. 5: Amplitude to integration time and corresponding lines fitted to the amplitude's approximately linear behavior. Values above the ambiguity border (brown line) are considered ambiguous.

Fig. 4 shows the intensity as function of the integration time and linear fits to the intensity-curves based on their initial

linear behavior. When oversaturation occurs (starting with values around 10000), the intensity starts deviating from its linear behavior (dashed lines) and reduces its gradient, which in some cases (e.g. white or silver) becomes even negative. Areas where the intensity function is not strictly increasing are considered ambiguous, as they cannot be mapped back to a distinct integration time, and reside above the ambiguity border (brown line). Also note the black measurement's extremely low reflectivity.

Fig. 5 shows the amplitude as function of the integration time and, as with the intensity functions, linear fits to the curves based on their initial linear behavior. It can be seen that with an increasing integration time, the amplitude values reach a peak and then decrease in value. Compared to the intensity, the amplitude has a much larger area of ambiguity. Also, the amplitude behaves less linearly than the intensity as its gradient becomes a bit larger before the values reach the peak.

In both figures, the effects of oversaturation can be seen as deviation from linear behavior. This has also been observed by May et al. for the Swiss Ranger SR-2 ToF camera and denoted as "oversaturation gap" [3, p.3].

As already explained, values above the ambiguity border cannot be mapped back to a distinct integration time. This essentially means that, from their value alone, they cannot be distinguished between being slightly or strongly oversaturated. Overall the intensity shows to be much less susceptible to ambiguous behavior compared to the amplitude. Also, its linear behavior is much more consistent.

C. Amplitude-Error Correlation

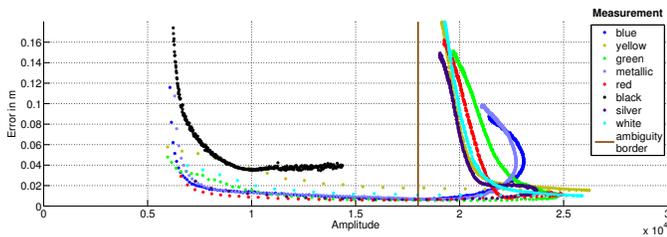


Fig. 6: Error to amplitude. Amplitude values on the right of the ambiguity border (brown line) can have both small and very large errors, but cannot be differentiated due to the ambiguous behavior explained in Sec. IV-B and Fig. 5.

In our effort to improve quality by automatic integration time estimation, we have to analyze the error behavior in regard to the integration time and resultant intensity and amplitude. Fig. 6 shows the correlation between the average error and the amplitude. Amplitude values above 18000 can have both small and very large errors, but cannot be differentiated due to the ambiguous behavior explained in Sec. IV-B and Fig. 5. However, it can be seen that for unambiguous amplitude values, the error is smallest between 10000 and 18000. This is true for all measurements, even black.

May et al. show that the "most precise mean accuracy could be acquired with an integration time located near the amplitudes maxima" [3, p.3]. This is true for both their Swiss Ranger SR-2 and our PMD CamCube 3.0, however, as already shown in Fig. 5 and Fig. 6, values near the maxima are ambiguous

and thus cannot be used. May et al. also omit these values but attribute this to the fact "that the image has a non-neglective saturation at the amplitudes maximum" [3, p.4].

Choosing an amplitude a bit too low only causes the error to rise slightly while choosing an amplitude too large causes values to exceed 18000. Such values cannot be differentiated between being only slightly or strongly oversaturated (see Fig. 6) and may carry large errors.

It can be argued that every value in-between is a viable candidate for the optimal amplitude. E.g., if omitting a substantial portion of the image due to oversaturation in order to optimize the remaining values is conceivable, choosing a value close to 18000 is adequate. However, if preventing oversaturation is the main goal, a value close to 10000 is better suited for this task.

D. Amplitude-Intensity Mapping

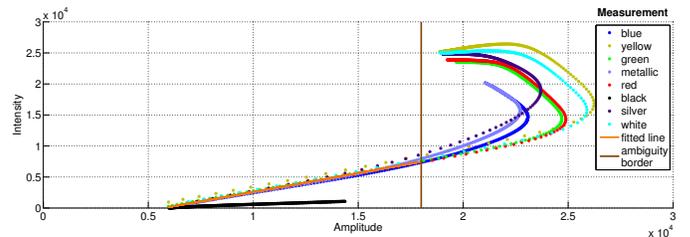


Fig. 7: Intensity to amplitude. In non-oversaturated areas, the intensity and amplitude are directly proportional to one another. The orange line represents the mapping function between amplitude and intensity values. Values to the right of the ambiguity border (brown line) are considered ambiguous (see Fig. 5).

Fig. 7 shows the correlation between the intensity and the amplitude. It can be seen that there is a linear relationship between the amplitude and the intensity, as long as we stay within value ranges outside of oversaturation (amplitude ≤ 18000). By fitting a line (orange) for amplitude values below 18000 (values above cannot be mapped by a function as they are ambiguous; see Fig. 5), we establish a mapping between the amplitude and intensity. The black measurement exhibits a unique behavior that distinguishes it from the other colors. This stems from its unique intensity behavior that was already observed in Fig. 4. Since this behavior is an exception and cannot be easily compensated, the black measurement is not applicable for amplitude-intensity matching and thus omitted from the fitting process.

E. Intensity Correction

We define the correct intensity as the value an intensity would have reached if the oversaturation did not have any effects regarding the linear behavior. Fig. 8 shows the correlation between the measured and the correct intensity. The correct intensity is derived from the intensity and its corresponding line that has been fitted to non-oversaturated values (see Fig. 4). The difference in values shows the enormous discrepancy between oversaturated measurements and what their values would have been if they had not been oversaturated.

To compensate for this discrepancy, we have chosen to fit a 3rd-degree polynomial to the data left of the ambiguity border

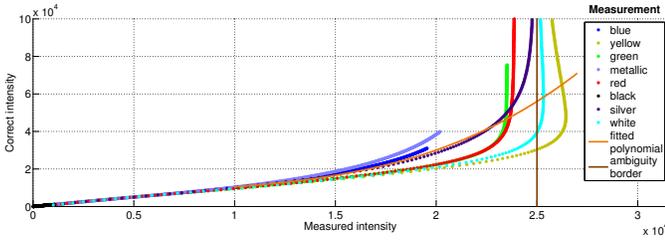


Fig. 8: Measured intensity to correct (linearized) intensity. The orange line represents the correction polynomial for intensity values.

TABLE I: The table shows the estimated sensor and correction parameters.

Parameter	Value(s)
Optimal amplitude (A_{opt})	18000
Optimal intensity (I_{opt})	7353
Amplitude-intensity-mapping function ($m(x)$)	0.6067281 -3475.016 $6000 \leq x \leq 18000$
Intensity-correction function ($h(x)$)	2.2026e-10 4.417871e-05 0.4276858 1100.069 $10000 \leq x \leq 25000$
Gain factor	1.11

and to only correct values above 10000 (see Eq. 2). This polynomial serves as correction function that allows us to approximate an intensity's actual value in case of oversaturation.

V. ALGORITHM EVALUATION

The previous sections have presented our algorithm (see Sec. III-A) and depicted an extensive analysis of underlying inherent sensor behavior (see Sec. IV). This section compares our approach to the work proposed by May et al. [3] and details differences in quality, quantity and adaption speed. For further evaluation purposes, we use the parameters estimated in Sec. IV. These parameters can be found in Table I.

A. Evaluation Framework

We perform two kinds of evaluation. In the first measurement, we use a static scene with metal plates of different colors/reflectivities (black, metallic, blue, green, red, silver, yellow and white) in a distance of approximately 1 m (see Sec. V-B). However, since the results have been very similar for all plates, we exemplarily present the plots for red plate. In the second measurement, we have recorded an arbitrary office scene with the camera orientation changing in-between every frame by about 15° around the x-, y- and/or z-axis, representing a highly dynamic scene scenario (see Sec. V-C). For both scenes, each frame has been recorded with the full spectrum of integration times, ranging from 50 to 8000 μ s. This allows us to use a scene for reproducible tests with different and differently parametrized auto integration time estimation approaches. Our simulation framework chooses the integration time for the next frame according to the optimal integration time calculated in the previous frame. To achieve a fair and meaningful comparison, we have performed the simulation with various initial integrations times.

B. Static Metal Plate Scene Evaluation

Fig. 9 compares the approach of May et al. (top row) to ours (bottom row). The static scene has been recorded for 20 frames. We have used 5 different initial integration times, covering strong under- and oversaturation as well as average saturations. We compare the development of integration times over the course of the algorithm, the mean error to the ground truth (see Sec. IV) and the number of well saturated values (amplitude values between 250 and 18000) in the image. It can be seen that May's algorithm slowly converges to the optimum after about 8 frames while our approach is already close to the optimum after 3 frames. Additionally it can be seen, that our algorithm has a smaller mean error (0.013 m) compared to May et al. (0.027 m). Also the number of well saturated pixels is higher for our algorithm (19000) compared to May et al. (17000).

Our approach's fast adaption to the proper integration time, especially in areas of oversaturation (the first 3 frames) and the resultant lower error can be attributed to the fact that we correct the intensity values before calculating the proportional factors.

C. Dynamic Scene Evaluation

Fig. 10 compares the approach of May et al. (top row) to ours (bottom row) in the highly dynamic office scene over the course of 35 frames and with 5 different initial integration times. Our approach converges within 8 frames to the global optimal integration time, while the algorithm of May et al. needs up to 20 frames. However, the mean error is approximately the same for both approaches (0.047 m). The number of well saturated pixels is higher for our algorithm (35700) compared to May et al. (33500).

This shows us several things. As explained in Sec. IV-A, if oversaturation causes a growth in error to the same degree as non-optimal saturation, then optimizing areas where most pixels reside, namely image borders, will actually reduce the overall error. Also, just because amplitudes reach beyond the ambiguity border does not necessarily mean that they are oversaturated.

Overall our approach shows a faster adaptability, especially in oversaturated scenes, which can, like with the static scene, be attributed to the intensity correction.

VI. CONCLUSION

In this paper we presented a novel online integration time adaption algorithm that works on a per-pixel basis and uses knowledge gained from an extensive analysis of the underlying inherent sensor behavior regarding intensity, amplitude and distance error to reduce the overall error, to prevent oversaturation and to minimize the adaption time. It also works well in presence of various reflectivities and quick changes in the scene. The per-pixel character enables us to use only portions of the image or even apply pixel-specific weighting, counteracting sensor properties (e.g. spatial intensity distribution) and allowing to adjust importance of certain image regions. Overall, this represents a significant improvement over previous methods. We also introduced a simulation framework

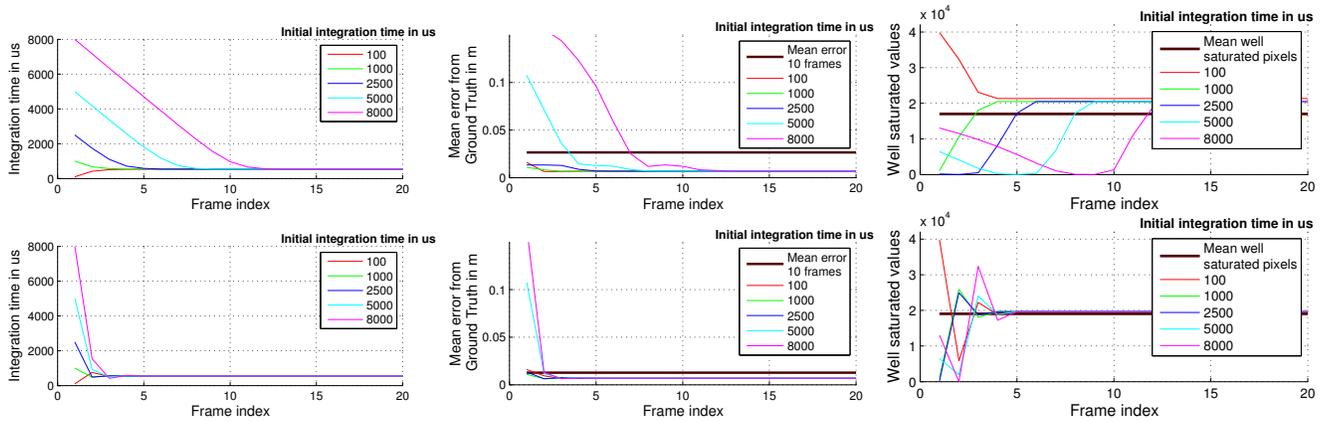


Fig. 9: Comparison between the approach of May et al. (top row) and ours (bottom row). A static scene capturing a red metal plate from a distance of 1 m is recorded for 20 frames. We compare the integration time (left), the mean error (center) and the number of well saturated pixels (right).

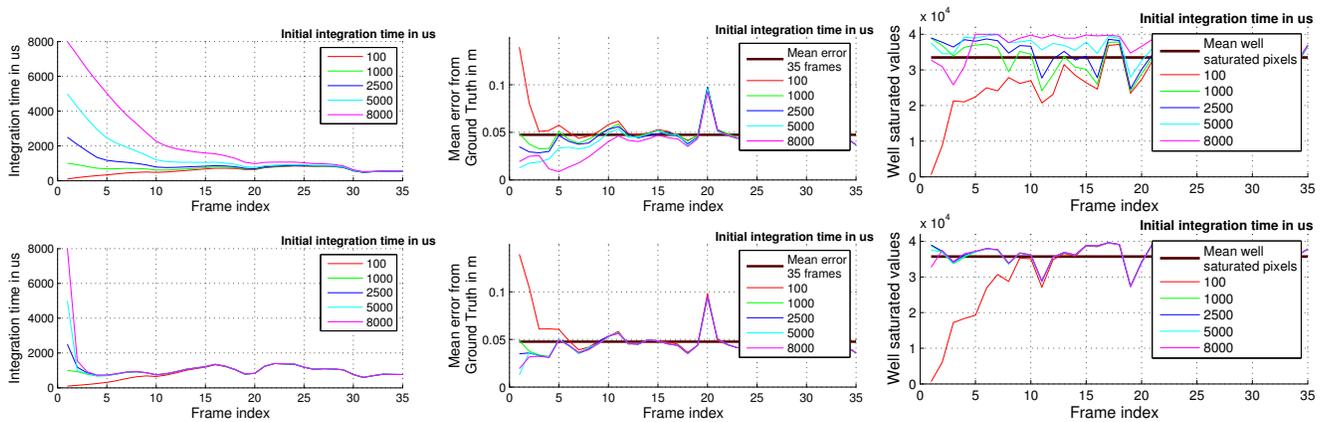


Fig. 10: Comparison between the approach of May et al. (top row) and ours (bottom row). A highly dynamic office scene is recorded for 35 frames. We compare the integration time (left), the mean error (center) and the number of well saturated pixels (right).

that allows us to perform reproducible and comparable tests with different and differently parametrized auto integration time estimation approaches.

Future work will concentrate on applying and optimizing this method to other ToF sensors.

ACKNOWLEDGMENT

This work was supported by the German Research Foundation (DFG) as part of the research training group GRK 1564 'Imaging New Modalities'.

REFERENCES

- [1] F. Lenzen, K. I. Kim, H. Schäfer, R. Nair, S. Meister, F. Becker, C. S. Garbe, and C. Theobalt, "Denosing strategies for time-of-flight data," in *Time-of-Flight Imaging: Algorithms, Sensors and Applications*, ser. LNCS, M. Grzegorzec, C. Theobalt, R. Koch, and A. Kolb, Eds., vol. 8200. Springer, 2013, pp. 24–25, 1.
- [2] D. Lefloch, R. Nair, F. Lenzen, H. Schaefer, L. Streeter, M. J. Cree, R. Koch, and A. Kolb, *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*. Springer Berlin Heidelberg, September 2013, vol. 8200, ch. Technical Foundation and Calibration Methods for Time-of-Flight Cameras, pp. 3–24.
- [3] S. May, B. Werner, H. Surmann, and K. Pervolz, "3d time-of-flight cameras for mobile robotics," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, Oct 2006, pp. 790–795.
- [4] P. Gil, J. Pomares, and F. Torres, "Analysis and adaptation of integration time in pmd camera for visual servoing," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, Aug 2010, pp. 311–315.
- [5] B. Langmann, K. Hartmann, and O. Loffeld, "Depth camera technology comparison and performance evaluation," in *ICPRAM 2012 - Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods, Volume 2, Vilamoura, Algarve, Portugal, 6-8 February, 2012*, 2012, pp. 438–444.
- [6] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (tof) cameras: A survey," *Sensors Journal, IEEE*, vol. 11, no. 9, pp. 1917–1926, Sept 2011.
- [7] T. Edeler, S. Hussmann, and F. Knoll, "Uncertainty analysis for optical time-of-flight sensors based on four-phase-shift range calculation," in *Sensors Applications Symposium (SAS), 2014 IEEE*, Feb 2014, pp. 382–387.
- [8] L. Jovanov, A. Pizurica, and W. Philips, "Wavelet based joint denoising of depth and luminance images," in *3DTV Conference, 2007*, May 2007, pp. 1–5.
- [9] T. Edeler, K. Ohliger, S. Hussmann, and A. Mertins, "Time-of-flight depth image denoising using prior noise information," in *Signal Processing (ICSP), 2010 IEEE 10th International Conference on*, Oct 2010, pp. 119–122.
- [10] K. Joong-Sik, J. Hoon, and K. Whoi-Yul, "Adaptive noise reduction method using noise modeling for tof sensor," in *Network Infrastructure and Digital Content (IC-NIDC), 2014 4th IEEE International Conference on*, Sept 2014, pp. 99–102.