

Hybrid Sensing Approach For Coded Modulation Time-of-Flight Cameras

Armin Schönlieb*, Plank Hannes*, Christian Steger†, Gerald Holweg* and Norbert Druml*

*Infineon Technologies Austria AG, Graz, Austria

{arminjosef.schoenlieb,hannes.plank,gerald.holweg,norbert.druml}@infineon.com

†Graz University of Technology, Graz, Austria

steger@tugraz.at

Abstract—In recent years, application fields such as industrial automation and indoor robot navigation increased the demand on reliable localization systems. Simultaneous mapping and localization systems often depend on depth imaging in order to reconstruct the scene.

Time-of-Flight sensors prove to be well suited for these applications, however are impaired by different error sources. The measurement principle is based on measuring the phase and consequently the delay of emitted and reflected light. Specular surfaces can cause pixel saturation, while the periodicity of the measured phase leads to ambiguous distances. In this paper, we aim to solve these problems by proposing a new Time-of-Flight depth sensing approach. By combining the emerging coded modulation method with traditional depth sensing, we are able to unify the advantages of both methods. Images captured with coded modulation show a pixel response only within selected distance limits. In contrast traditional continuous wave Time-of-Flight imaging exhibits a superior signal-to-noise ratio. This method enables to mask erroneous distance measurements, allowing Time-of-Flight sensors to produce more reliable depth measurements and gain traction in the industrial environment.

As our evaluation shows, our method is able to remove the influence of specular surfaces, and is capable of masking ambiguous distance measurements. Furthermore, our approach improves the system behavior by enabling more robust exposure time control.

Index Terms—Time-of-Flight, ToF, Coded Modulation,

I. INTRODUCTION

Time-of-Flight (ToF) is an image sensing technology that enables the acquisition of depth information. This technology is a very interesting field of study, for industrial and consumer applications. The depth information can be used for interior mapping and stair recognition. The small form factor and the lack of moving parts are advantages compared to other imaging methods. Furthermore the measurement principle makes it easy to distinguish different surfaces [1].

Problems occur when specular surfaces are present in the scene. In Fig. 1, the glass-wall in the background causes a reflection, which is significantly higher than the surrounding area. ToF imagers are working by measuring the travel time of near infrared light. If such specular surfaces are present, the depth in this area is impaired. One might decrease the exposure time, to reduce the amount of reflected light, but this also decrease the signal strength of unaffected parts of the scenery. Therefore it is hard to determine the optimum exposure time, when specular surfaces are present. A different problem is the

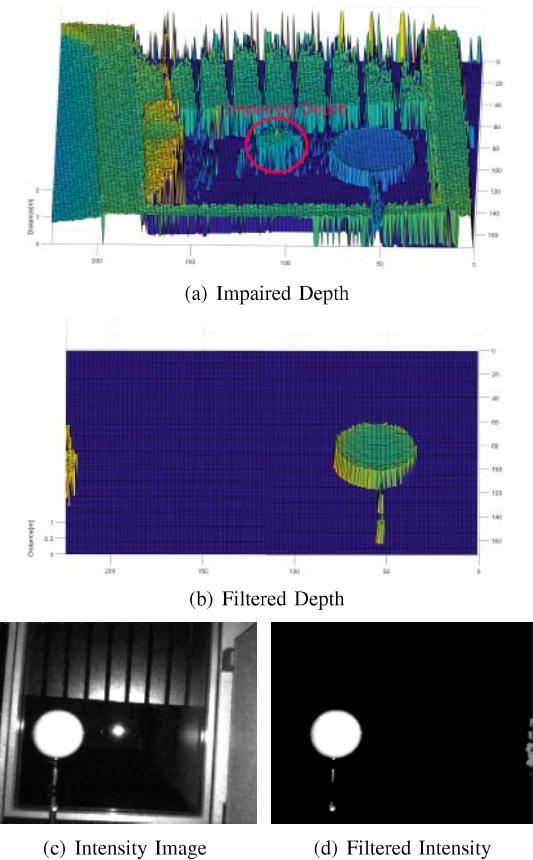


Fig. 1. In (a) the depth image is impaired by a glass-wall. In (b) the depth image is filtered and without erroneous measurement. In (c) the glass wall causes strong reflections in the intensity image, which is removed in (d).

2π periodicity of the modulation signal. If an object is further away than the unambiguity range of the modulation signal, the distance is folded in the first unambiguity distance. In normal operation mode the folded phase is unwrapped by capturing an additional depth measurement with a different modulation frequency [2].

We present in this work a new ToF sensing method, based on a combination of conventional continuous wave (CW) and coded modulation scheme. With this modulation scheme, it is possible to tune the measurement range of the imager by replacing the modulation signal with code sequences. In previous work, the depth of coded modulation ToF phase images

is calculated with a two phase measurement [3]. We suggest that the two phase depth calculation, is not accurate enough for operation. Therefore we develop a novel sensing method based on coded modulation phase images, and generate with them a filter mask. With this filter mask, erroneous regions in the CW image, can be omitted. This has the advantage, that the accuracy of CW modulation is preserved, but the measurement range can be adjusted with coded modulation.

In our evaluation, we prove the feasibility of our approach, and investigate the influence of different parameters. We validate coded modulation measurement method and consequently the fidelity of the filter mask for different distances, and show the relation to the exposure time. Our contributions to the scientific community are:

- Investigation and method to fuse coded modulation with traditional continuous wave ToF sensing.
- A sensing method to suppress specular surfaces in ToF images.
- A new opportunity to overcome the unambiguity problem of ToF imagers.

II. RELATED WORK

Applying code words to modulation signals is a well-established field in communication technology. With Code Division Multiple Access (CDMA), it is possible to separate different communication partners. In the approach of Fersch et al., this method is applied to light detection and ranging (LIDAR) [4]. Perez et al. introduce this approach for ultrasonic sonar [5]. LIDAR and ultrasonic sonar are measuring directly the time difference of the emitted and reflected signal. With the coded signals they are able to reduce the noise and gain better performance. In comparison, the indirect ToF imager explored in this paper, measures the phase difference of two signals for a certain time.

Buttgen et al., provide the first implementation of coded modulation for ToF [3]. Buttgen et al. present the usage of coded modulation in multi-ToF camera environments. The measurement can be distorted when more than one camera is present. With the coded modulation scheme, it is possible to reduce the distortion, produced by other imagers. m-Sequences with different lengths, are applied for the code sequence. This system was improved by Pardes et al., by using kasami code sequences [6]. Two points of the auto correlation function are used for depth calculation. They are generated by shifting the outgoing light pulses compared to the reference signal. This allows offset cancellation by the measurement. But as pixel variation errors are still present, the measurement is error-prone. If more than two different phase-shifts are used, this error can be compensated [7].

Bhandari and Raskar use different modulation frequencies to overcome multipath effects and to distinguish different layers of glass [8]. These multipath effects are a problem for conventional ToF cameras as the distance is affected by several reflections. To solve this issue different modulation frequencies are used for several depth measurements. The disadvantage of this method is that calibration is required for

various modulation frequencies. Furthermore, the acquisition time increases, in which movement between sensor and scene can cause artefacts. A longer acquisition time requires more energy, which increases the sensor's power consumption. In our work we use a different approach, using a single modulation frequency and a single coded modulation image. Kadambi et al. show that it is possible, to use coded modulation to solve multipath issues [9]. The authors evaluate their theory with a prototyping platform capable to use coded modulation. The work arrives at the conclusion, that it is possible to create a light sweep with coded modulation. With the light sweep certain regions of the scene are measured. This is possible, as only part of the scene is measured with the code. With this technique, it is even possible to measure objects behind translucent materials. m-Sequences with the length of 31 bit are used for these experiments. Guptha et al. provide an analysis for different codes [10]. A simulation and also a measurement with laboratory equipment evaluates the codes.

To unwrap the unambiguity distance is a crucial part of Time-of-Flight sensing, as it is required to measure further distances. Droschel et al. show that it is possible to use the reflectivity, and wrap lower intensities to a further distance. They also demonstrate that this is error-prone due to different reflection coefficients. [2]. Therefore a measurement with two frequencies is used to unwrap these distances.

III. TIME-OF-FLIGHT CAMERAS

ToF cameras operate by measuring the time-of-flight of a pulsed near infrared beam of light. The principle can be seen in Fig. 2. The time-of-flight is a function of the distance. The reflected pulses are compared with the emitted pulses by a photonic mixer device (PMD). The PMD is implemented in every pixel. The PMD separates the charges, created by the received light and transfers them into either of two capacitors. This separation operates with the frequency f_{mod} , which is also used for illumination. These charges are then sampled by an analog to digital converter. The difference between the two capacitors is related to the phase shift of the signals. If the phase shift is zero, the phase difference is zero. If the phase shift changes, one of the two charges increases, more than the other. Since the measured charge also dependent on the reflectivity, the phase shift has to be normalized. This can be done, by capturing images of the same scene with four different phase-offset [0°, 90°, 180°, 270°]. The phase-offset is defined as the phase difference between the illumination signal and the reference signal. With these measurements, the four bucket algorithm can be used to calculate φ , the offset-compensated phase-difference of the reflected signal .

$$\varphi = \arctan 2 \frac{Q_{90^\circ} - Q_{270^\circ}}{Q_{0^\circ} - Q_{180^\circ}} \quad (1)$$

Now the distance d can be calculated. The frequency f_{mod} represents the pulse frequency, modulating the infrared light and the pixel reference signal.

$$d = \frac{c}{2} \cdot \frac{\varphi}{2\pi f_{mod}} \quad (2)$$

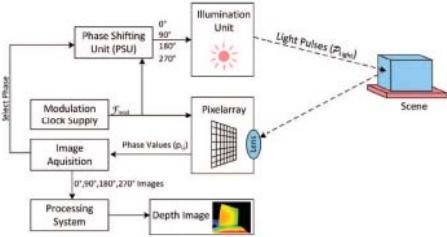


Fig. 2. The working principle of an indirect PMD based ToF depth sensing system. The pixel array measures the phase difference. These phase differences need then to be post-processed for a depth frame. Obtained with changes [11]

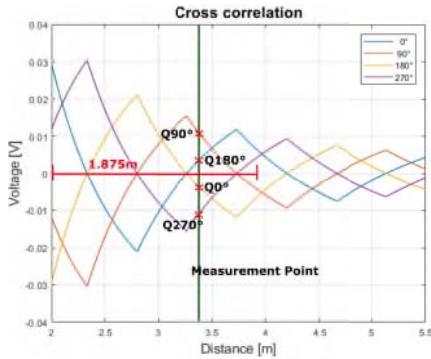


Fig. 3. Correlation plot of a indirect PMD based continuous wave ToF system. Each phase value Q_x is obtained with a separate image, and required to calculate the distance. The unambiguity range, where the depth is unique is 1.875m, for a modulation frequency of 80Mhz.

The intensity A of the light, received by the camera is obtained with the phase-images.

$$A = \frac{\sqrt{(Q_{90^\circ} - Q_{270^\circ})^2 + (Q_0^\circ - Q_{180^\circ})^2}}{2} \quad (3)$$

A continuous wave signal is used for modulation, therefore the signal repeats after a 2π phase shift. This leads to alias measurements. Consequently ToF cameras have an unambiguity range U defined by:

$$U = \frac{c}{2f_{mod}} \quad (4)$$

The correlation plot of a four phase measurement, over a certain distance can be seen in Fig. 3. With these values, the four bucket algorithm can be used to calculate the distance and the intensity. The distance has an unambiguity range of 1.875m, with a modulation frequency of 80.32Mhz. If an object is further away than the unambiguity range, it is folded back in the first unambiguity range. Therefore the distance needs to be unwrapped to match the real distance. The most established way is to measure the scene with different modulation frequencies, yielding in different depth unambiguity ranges. If an object is found in the same distance for both depth frames, the distance can be unwrapped, as shown in Fig. 4. The disadvantage is that at least four additional phase images are required. This yields to excessive power consumption, while simultaneously the frame-rate is limited. Our work is capable to detect and filter this unambiguity problem at the cost of just one additional measurement.

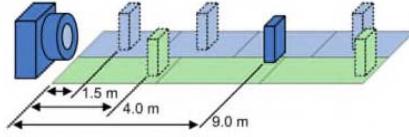


Fig. 4. Example for depth unwrapping of objects, obtained with changes [12]. The green and blue measurements use different frequencies, resulting in different unambiguity ranges. A combination of these measurements enables to calculate the actual distance.

IV. CODED MODULATION FOR TOF CAMERAS

A. Principle of Coded Modulation

The principle of indirect continuous wave ToF, is to measure the phase difference between a reference signal and a reflected signal. State-of-the-art ToF imagers use a square wave signal as reference signal. Research is also conducted for sinusoidal and triangular wave signals [10], but they are not practicable for operation, as the signal form is not feasible to generate. Coded modulation in Time-of-Flight imaging is about altering the sequence of the emitted light-pulses and the corresponding pixel modulation signal. This enables to customize the phase-distance function. This works by replacing the square wave signal is replaced with a code sequence, which is repeated for the phase image exposure time. Fig .5 compares the coded modulation and continuous wave reference signal. For the code design, different kind of codes like maximum length code, gold code or barker code can be used. It depends on the application which code fits best. For pseudo random pulses, E.q. 6 describes the behavior of the PMD. The description is done by an autocorrelation function of the reference signal and the reflected signal, time shifted by τ . T_c describes the time of one chip period, the smallest period in the code sequence. The length of the code L is described by the number of bits n :

$$L = 2^n - 1 \quad (5)$$

$$\phi_{cc}(\tau) = \frac{1}{LT_c} \cdot \int_{\frac{LT_c}{2}}^{-\frac{LT_c}{2}} c(t)c(t+\tau)dt \quad (6)$$

The code sequence repeats periodically, and is summed up in a discrete way. This leads to the following equation:

$$\phi_{cc}(kT_c) = \frac{1}{L} \sum_{i=1}^L C_i \cdot C_{i+k} \quad (7)$$

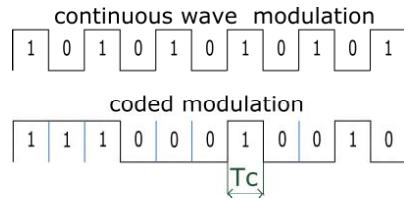


Fig. 5. The illumination and pixel reference signal are exchanged with a repeating code, to create coded modulation images. Continuous wave (top) and a coded modulation illumination signal (below). The time T_C is the smallest signal part in the code sequence.

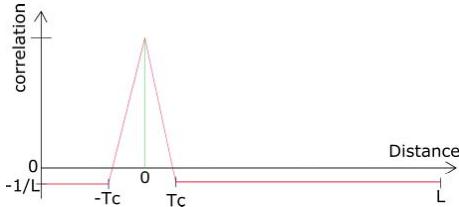


Fig. 6. Auto correlation function of an maximum length code. The correlation is at $\pm T_c$ around the zero point higher than the minimum. The area between $\pm T_c$ is our measurement range.

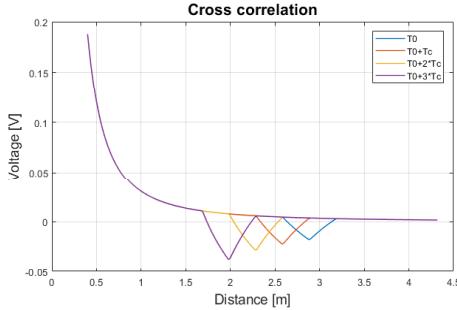


Fig. 7. Correlation plot of a four phase-offset coded modulation measurement. A Barker 11 code is used for the correlation function. The phase-offset of the codes is one T_c . For each depth image, only two phase images are contributing with a correlation value (violet and yellow, yellow and red, red and blue).

After solving the equation, we can distinguish two areas of the auto correlation function:

$$I(x, y) = \begin{cases} 1 - \left(\frac{L+1}{L} \cdot \frac{|\tau|}{T_c}\right) & \dots |\tau| \leq T_c, \\ -\frac{1}{L} & \dots |\tau| > T_c \end{cases} \quad (8)$$

Fig. 6 describes the correlation function. The zero point the auto correlation shows a maximum. If one signal is shifted by the reflection the correlation maximum decays, and minimizes after one chip time.

B. Comparison of CW and Coded Modulation ToF Cameras

In order to omit the influence of the reflectivity on the depth image, ToF depth algorithms need more than one phase image. In conventional CW operation mode, this is achieved by shifting the phase-offset by 90° to the reference signal. For coded modulation, a similar approach is proposed by the literature. Here one of the signals can be delayed or rotated to create a phase-offset. Fig. 7 presents this method with four phase offsets. The correlation functions in this image are inverted due to the sensor design. The position of the correlation peak can be adjusted by rotating the code sequence. In this example, each measurement is shifted by a multiple of T_c . The correlation function, within the boundaries of the correlation peak, has the same characteristics as the CW correlation function. Therefore the distance can also be calculated using Eq. 1. On the one hand the advantage of this method is that the measurement range can be tuned precisely, on the other hand the disadvantage of this method is that two of the four phases do not correlate. This means they have no information in it (violet and yellow, yellow and red, red and blue). Fig. 8 describes this problem with the phasor

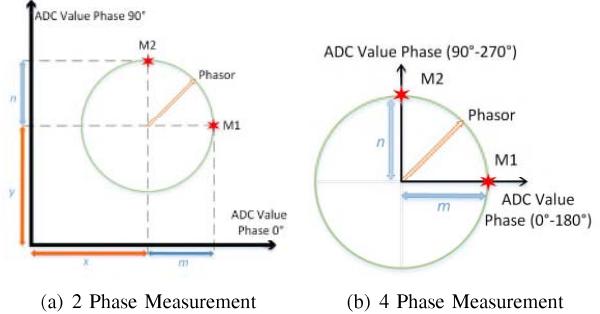


Fig. 8. In (a) the two phase measurement is presented. The values m and n denote the real and imaginary part of the phasor. Due to the pixel variation errors x and y , it is not possible to reliably determine the phasor from just two phase measurements. The values m and n denote the phase value which is different for each pixel. In (b) it is possible to compensate x and y where additionally capturing phase-images with 180° and 270° phase-shift. They have the same x and y value and enable to compensate this error.

TABLE I
COMPARISON OF THE STANDARD DEVIATION OF CONTINUOUS WAVE MODULATION AND CODED MODULATION.

	Continuous Wave	Coded Modulation
Standard Deviation Depth	0.0627	2.0432
Standard Deviation Amplitude	3.0050	6.1719

principle [13]. In (a) the two phase measurement is presented n and m represent the real distance, which is measured by the ADC for two phase-offsets. One is 0° , the other is shifted by 90° . The overall phase value measured is a combination of m , n , x , y where x and y are pixel variation errors. The pixel variation error is caused by the ADC and manufacturing process variation of each individual pixel. Therefore normal CW depth calculation is based on the four phase algorithm, visualized in image (b). The pixel variation error x and y is canceled out. To demonstrate the lack of information, caused by the pixel variation error, we conduct a measurement to evaluate the standard deviation of both measurements defined in Table I. We measure a 32 times higher standard deviation in the depth frame for 100 measurements. This error is significant, and is the reason why the two phase algorithm is not adapted for ToF cameras.

C. Hybrid Sensing Approach

In this Section, we present a new approach to coded modulation ToF sensing. Our goal is to combine the advantages of traditional CW modulation with coded modulation in a hybrid sensing approach. With the coded modulation technique, it is possible to measure a well-defined distance range in the depth image. In comparison, we are using CW to gain the best accuracy. We first take a conventional CW four phase measurement of the scene. This depth or intensity image is then multiplied with a filter mask, generated with a single coded modulation frame. Because information in a coded modulation frame is only present where the correlation is present we are able to tune the measurement limits. This tuning is done with our code design. The correlation graph only depends on the code itself, its length and the smallest pulse T_c . To calculate this filter mask, we implemented a threshold to the coded modulation phase image according to Eq 9.

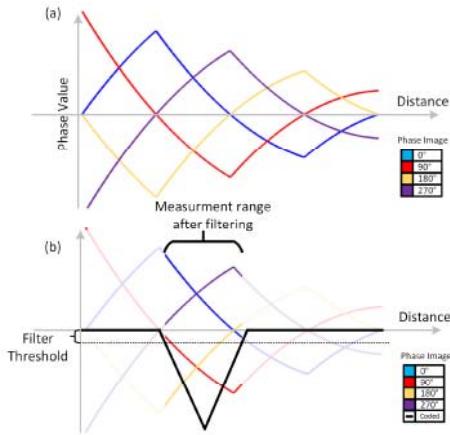


Fig. 9. Principle of filtering with an additional coded modulation frame. First an traditional CW measurement is conducted. An additional coded modulation frame is taken. With the filter threshold, the measurement range of the coded modulation frame is guarded.

For different exposure times or signal-to-noise requirements, the filter-threshold value can be adapted. To reduce the noise of the sensor and the analog to digital converter we subtract a reference image obtained without illumination.

$$F(i, y) = \begin{cases} 0 & \text{if } Q_{code} > \text{threshold}, \\ 1 & \text{if } Q_{code} < \text{threshold} \end{cases} \quad (9)$$

Fig. 9 shows the phase-value function over the distance for the CW principle (a). The coded modulation image is used to set a meaningful measurement range, thus eliminating unwanted and erroneous distance measurements from the final image (b). With this approach, for every environment a special code can be designed, to mask out reflections. Also for fast depth image acquisition, a code can be used which correlates in a certain unambiguity range, shown in Fig. 10. This is done by changing the code length, and the modulation frequency. These parameters need to be changed, in order to fit the unambiguity range of the CW modulation frequency. With only one additional frame, the power consumption is reduced by 37.5%. This allows to measure energy-efficient depth data of one ambiguity range. This fact makes it very useful for robots working on battery. Also for fast moving objects, the reduced number of frames is advantageous, as motion artifacts are reduced. We believe that this approach very useful for industrial and robotic operation.

Saturated pixel, are pixel in which one capacitance is full and the measurement is therefore wrong. Therefore it is not possible to solve this problem with a single phase-image, as also there the saturated pixel is visible. For saturated pixel we use a two phase coded modulation measurement. The second phase image is subtracted from the first phase image. With this, the invalid pixel are canceled out. We shift the second phase image by two times T_c .

V. RESULTS

In this section, we present the results of our evaluation, and application examples. In Fig. 11 the processing steps, done

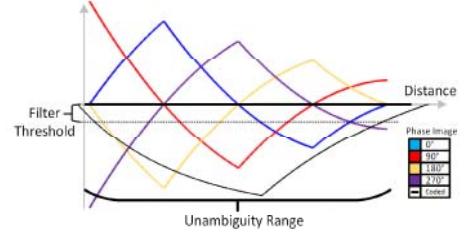


Fig. 10. Application of our method: A coded modulation image is used to remove ambiguous depth values from the image. For this we use a correlation function, which correlates in the first ambiguity range.

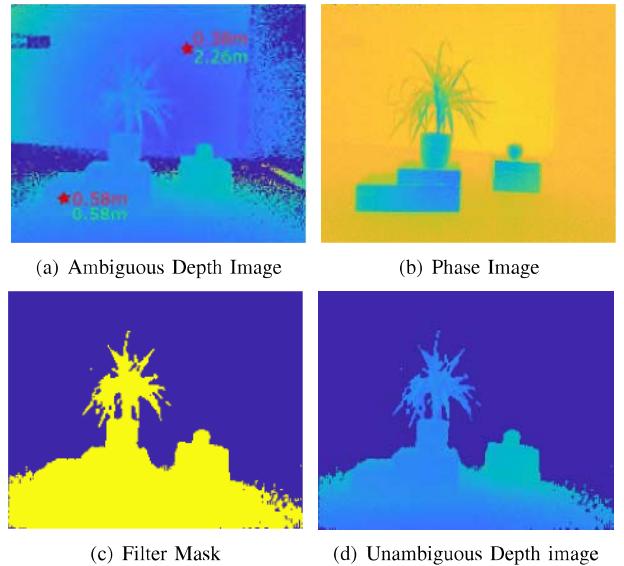
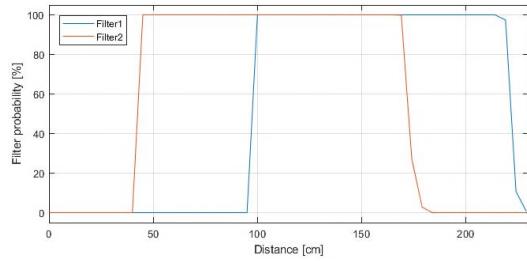


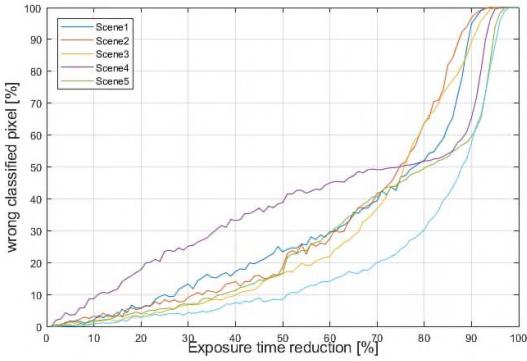
Fig. 11. Application example of the coded modulation filter with one coded modulation measurement, filtering the second unambiguity range. In (a) a CW depth image with one frequency can be seen. Although the wall in the background is further away, the measurement indicates that it has the same distance as the plant, this is due to the unambiguity range. The coded modulation phase image, which measures the first unambiguity range in (b). The created filter mask in (c). The filtered depth image and the first unambiguity range are in (d). In (d) only the first unambiguity range is visible and the wrong distances are omitted.

for the filtering of one unambiguity range on a depth image are presented. In the first image (a), a CW four phase depth image can be seen. The plant in the foreground and the wall in the background show nearly the same color coding, indicating the same distance. In the second frame, a coded modulation phase image is visible, measuring the first unambiguity range. A barker 11 code is used for generating this image. We choose Barker codes, as they show good auto correlation behavior. After subtracting the reference frame, we use a filter threshold to filter the sensor noise. Afterwards we generate a filter mask, visible in (c). Fig. 11(d) shows the filtered result.

We evaluate our algorithm for two scenarios. At first we evaluate the filter generation for different distances. Furthermore we want to evaluate the pixel classification, for reduced exposure time, mostly affecting the energy efficiency. An analysis of our algorithm is presented in Fig. 12. In (a) we use a linear translation stage, to move the camera away from a canvas with well-defined reflectivity. The filter mask



(a) Filter Generation For Different Distances



(b) Exposure Reduction Sweep

Fig. 12. Evaluation measurements of our algorithm. The evaluation of the distance depended filter generation (a). A linear translation stage is used to move the camera away from a wall. For different distances our filter mask is generated. As it can be seen with different code configurations we are able to measure different regions of the scene. Reducing the exposure of the coded modulation image reduces the power consumption at the expense of the filter mask quality (b).

generation is evaluated for a defined number of distance steps. It can be seen that for different distances, different filter masks are present. This confirms our theory that certain distances generate a correlation, which can be measured. Furthermore is the slope of the filter very steep, leading to a good separation. In (b), we conducted a series of measurements, where we reduced the exposure time of the correlation image, captured with coded modulation. With this measurement, it is possible to evaluate the energy consumption and update-rate, compared to the number of wrong classified filter mask pixels. For this we compare the number of valid pixels with a 100% exposure, while reducing the exposure time. It is possible to observe in the range of 0% to 50% reduced exposure time, a proportional relation. The lower the exposure gets, the higher the number of wrong classified pixel get. This is due to the lack of reflected light. The factor of this linear reflectivity can be adjusted by the filter threshold, used for mask generation. For long exposure and high-reflective scenes, a higher threshold can be used and vice versa. The advantage of this method is that the same calibration and post processing algorithm as for CW depth frames can be used. Therefore no additional calibration, like wiggling, temperature or system calibration is required.

VI. CONCLUSION

In our paper, we present a novel Time-of-Flight depth sensing approach, based on coded modulation. We show that the two phase depth calculation is not accurate enough for conventional ToF operation. We combine the measurement of

CW and coded modulation, to gain the advantages of both methods. We tune the correlation of the coded modulation to limits, where the measurement needs to be done. From this phase image we generate a filter mask, based on a threshold value. Afterwards, we filter the depth image with this filter mask. For specular surfaces, we use two coded modulation frames. With this approach, we can use the well-known CW measurement and post process it with the coded modulation approach. With this processing technique, it is possible to omit reflections in the scene and adapt the light intensity for better signal to noise ratios. These new possibilities enable the use of ToF cameras in various new environments. It is also possible to resolve the unambiguity problem with just one measurement, leading to a 37.5% reduction of power consumption, compared to state-of-the-art ToF sensing methods. Future work will deal with real time adjustment of the code, to optimize for different scenes.

ACKNOWLEDGMENTS

This project has received funding from the European Union Horizon 2020 research and innovation programme under grant agreement n°769944.

REFERENCES

- [1] C. Ye, "Navigating a Portable Robotic Device by a 3D imaging sensor," in *2010 IEEE Sensors*, pp. 1005–1010.
- [2] D. Droschel, D. Holz, and S. Behnke, "Multi-frequency Phase Unwrapping for Time-of-Flight cameras," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 2010, pp. 1463–1469.
- [3] B. Buttgen, M. A. E. Mechat, F. Lustenberger, and P. Seitz, "Pseudonoise Optical Modulation for Real-Time 3-D Imaging With Minimum Interference," vol. 54, no. 10, pp. 2109–2119.
- [4] T. Fersch, R. Weigel, and A. Koelpin, "A CDMA Modulation Technique for Automotive Time-of-Flight LiDAR Systems," vol. 17, no. 11, pp. 3507–3516.
- [5] M. C. Perez, J. Urena, A. Hernandez, A. Jimenez, D. Ruiz, F. J. Alvarez, and C. D. Marziani, "Performance comparison of different codes in an ultrasonic positioning system using ds-cdma," in *2009 IEEE International Symposium on Intelligent Signal Processing*, Aug 2009, pp. 125–130.
- [6] J. A. Paredes, T. Aguilera, F. J. Ivarez, J. A. Moreno, and D. Gualda, "Spreading sequences performance on time-of-flight smart pixels," in *2015 IEEE 9th International Symposium on Intelligent Signal Processing (WISP) Proceedings*, pp. 1–4.
- [7] J. Gruenwald, "Investigation of systematic errors in time-of-flight imaging," Jan 2013.
- [8] A. Bhandari and R. Raskar, "Signal Processing for Time-of-Flight Imaging Sensors: An introduction to inverse problems in computational 3-D imaging," vol. 33, no. 5, pp. 45–58.
- [9] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar, "Coded time of flight cameras: Sparse deconvolution to address multipath interference and recover time profiles," vol. 32, no. 6, pp. 1–10.
- [10] M. Gupta, A. Velten, S. K. Nayar, and E. Breitbach, "What Are Optimal Coding Functions for Time-of-Flight Imaging?" vol. 37, no. 2, pp. 1–18.
- [11] H. Plank, A. Schoenlieb, C. Ehrenhofer, C. Steger, G. Holweg, and N. Druml, "Synchronization of time-of-flight 3d sensors for optical communication," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6.
- [12] A. P. Jongenelen, D. G. Bailey, A. D. Payne, D. A. Carnegie, and A. A. Dorrington, "Efficient fpga implementation of homodyne-based time-of-flight range imaging," *J. Real-Time Image Process.*, vol. 7, no. 1, pp. 21–29, Mar. 2012.
- [13] M. Gupta, S. K. Nayar, M. B. Hullin, and J. Martin, "Phasor Imaging: A Generalization of Correlation-Based Time-of-Flight Imaging," *ACM Transactions on Graphics*, vol. 34, no. 5, pp. 1–18, Nov. 2015.