Matti Laukkanen

Time-of-Flight Cameras for Depth Imaging

School of Electrical Engineering

Espoo 27.2.2015

Thesis supervisor:

Prof. Ville Kyrki

Aalto University
School of Electrical
Engineering

Contents

1 Introduction

Depth imaging means sensing the three-dimensional geometrical structures of our surrounding world. Compared to traditional 2-D image, the third dimension adds lots of valuable information about the scene. That is probably why stereoscopic vision, capable of sensing dimensions, has developed to most animals, including human. Depth imaging has been researched for decades now, and several different technologies for sensing 3D have been developed since. Many of these technologies, such as stereo camera, which is the most studied and well-known depth imaging method, are inspired by nature. Earlier systems were based on cameras and triangulation, while more modern systems exploit electromagnetic radiation. Current laser-based systems provide superior operating range and depth accuracy compared to older systems.

Traditionally, depth imaging technologies have been used in several applications, including remote sensing of the earth and mapping the environment [1]. But recently, when the semiconductor technology has developed rapidly and prices of devices have dropped, depth imaging has became more popular. In 2010 Microsoft released Kinect[2], a low-cost depth sensor based on structured light. The device was originally intended for gaming and entertainment applications, but researchers as well as amateur programmers around the world quickly became interested in it because of its high accuracy and unbelievably low price tag compared to previous technologies [3]. Ever since it has been researched extensively and used successfully in numerous applications, such as vision sensor for low-cost mobile robot [4]. Low cost sensors enable using depth imaging technology in a whole new areas, including human machine interface and healthcare [5].

The most recent depth imaging technology is Time-of-Flight (ToF) cameras. Since first systems appeared about two decades ago, they have been researched actively [6]. Compared to other systems, ToF cameras have some very promising features: high frame rate, compact size and very low processing power needed. Additionally, performance in different lighting conditions is robust, which makes TOF camera a great challenger for other technologies.

Up to this point, the major drawback has been their low resolution. In 2011, MESA SR4000 [7] had a resolution of only 144x176 pixels, while other systems could provide much higher accuracy. Despite of its low resolution, many researchers have indicated that TOF cameras might be the technology of the future [6, 8, 9, 10]. In a last few years optics and semiconductor technology has developed rapidly. This has led to increase of resolution in ToF sensors, currently one of the largest reported is 512x424 pixels [11]. Especially interesting is the new version of Kinect sensor[12], which is based on ToF technology and offering resolution of 524x488 pixels. Moreover, the price of this new mass-produced sensor is below $200, which is order of magnitude lower than earlier models.

Other disadvantages of Time-of-Flight cameras are their limited operating range and rather low signal to noise ratio (SNR). However, in a few last years many researches have shown advances to overcome these problems. New sensor designs and error reduction methods are constantly developed and the quality of data is increasing.

This paper gives an overview of ToF camera technology and state-of-the-art. The report is structured as follows: Chapter 2 gives an overview of working principles of the Time-of-Flight camera. First, sensor structure and components are described. Then, the theory of depth measurement is presented, and different approaches are considered. Last, representations for 3D data acquired by the camera are discussed briefly. In chapter 3, ToF camera technology is compared to other widely used depth imaging technologies. Working principles of these systems are briefly described and their advantages and weaknesses are considered in comparison to ToF. In chapter 4, error sources affecting to the quality of depth measurement are considered. Most important errors, both random and systematic, are discussed as well as methods to enhance the quality of data. Finally, chapter 5 concludes this report and some future work aspects are considered.

2 Time-of-Flight Camera

In this chapter, the Time-of-Flight camera technology is reviewed. First, the structure of a TOF camera and components in it are considered. Then, the working principles of depth measurement inside a TOF camera pixel are considered, as two different approaches, pulsed light and continuous wave modulation, are presented. Finally, representations of 3D data provided by Time-of-Flight cameras are discussed.

2.1 Components

Time-of-Flight camera physically resembles a traditional digital camera. It consists of a lens collecting light to a image sensor, which is an two-dimensional array of photosensing elements, pixels. Additionally, ToF camera has an light source that is used to illuminate the scene. Camera registers the light reflected from the scene and each pixel in parallel calculates depth information, forming a depth image from the whole scene. All of these components are packed into a compact package with no moving parts. Two examples of state-of-the-art Time-of-Flight cameras are shown in Figure 1.

Figure 1: Two Time-of-Flight cameras, MESA SR4000 [7] and Microsoft Kinect v2 [12]. Kinect has also RGB camera and cooling fan inside, hence the bigger size.

Light emitting diodes (LEDs) are usually used as a light source, because of their fast response time [13]. Most commercial sensors use near infrared (NIR) wavelengths not visible to human eye, usually around 850 nm [6, 11]. NIR region of the spectrum used is because most materials have relatively high reflectance on that band [7] and it does not interfere with human vision.

Pixels of the image sensor can be implemented in several ways, depending on operation principle discussed more in next section. Currently most systems are analogue, using photodetectors and capacitors to collect and store electricity from light pulses and finally analogue signals are converted to digital [9]. Also digital TOF cameras have also been researched, based on single photon avalanche diodes (SPAD) [14] which can register even single photons. Fully digital system helps reducing noise related to analogue signals and digital conversion [9, 11].

2.2 Theory of Operation

The basic working principle of Time-of-Flight cameras is that they illuminate the scene with modulated light source and then measure the returning light that reflects back to the sensor. Because the speed of light is constant, the distance of the object where the light was reflected from can be calculated from the time difference of the emitted and returning light signals. Two different techniques of illumination are used in TOF cameras, either pulsed light or continuous wave (CW) modulated light. [9].

Pulsed modulation

In the pulsed method the depth measurement is straightforward. The illumination unit is switched on and off rapidly, thus generating short light pulses. A timer starts when the light pulse is sent and stops when the reflected light is detected by the sensor. The Distance to the object can then be calculated by [13]

$$d = \Delta t \frac{c}{2}, \tag{1}$$

where Δt is the round-trip time of the light pulse and c is the speed of light. However, the ambient illumination usually contains same wavelengths as the light source of the TOF camera. Therefore the light received by camera is a sum of the emitted light and ambient light. This would lead to distance calculation errors, so the light is recorded also when the illumination unit switched off so that the background can be subtracted from signal. This can be done by using outgoing light signal as a control signal for the detector sensor. Moreover, one short light pulse contains rather low amount of energy [14], and due to unideal realization of the system [9], received signal is noisy. In order to increase SNR, multiple cycles of pulses, typically millions, are recorded for a specific period of time and the depth information is computed from the average. This time interval is called integration time (IT) [15]. The concept of this pulsed modulation method is illustrated in Figure 2. By using the integration time of Δt and two out-of-phase sampling windows C_1 and C_2, the averaged distance is computed by [15]:

$$d = \frac{1}{2} c \; \Delta t \frac{Q_2}{Q_1 + Q_2}, \tag{2}$$

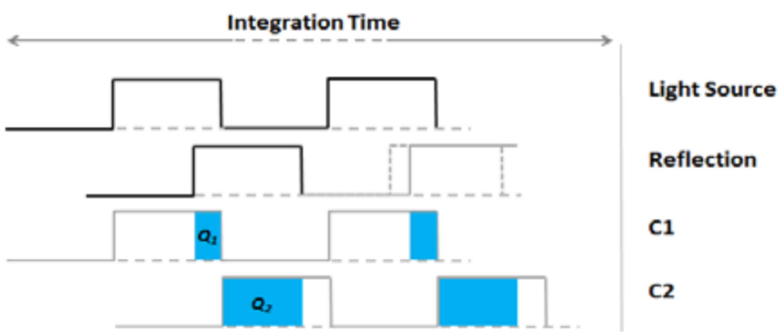

Figure 2: Pulsed Time-of-Flight method. Received signal is sampled in two out-of-phase windows in parallel. [15]

where Q_1 and Q_2 are the accumulated electrical charges received over the integration time period.

Depth resolution of the pulsed method is limited by the speed of the electronics in the camera. From (1) it can be calculated that to achieve depth resolution of 1 mm, one would need a light pulse of approximately 6.6 picoseconds. However, current LEDs/laser diodes limit the rising and falling times and repetition times of the pulses [13]. In addition, achieving such speeds in the receiver circuit is challenging with current silicon-based technology in room temperatures. [15]

Continuous Wave Modulation

Instead of direct measurement of the round-trip time of a light pulse, the CW modulation method is based on measuring phase difference of sent and received signals. The light is modulated by altering the input current to the light source, generating waveform signal [13]. Different shapes of modulation signal can be used, but usually either square or sinusoidal waves are used [8]. CW modulation technique lowers the requirements for the light source, so better depth resolution can be achieved compared to pulsed light.

There are various styles to demodulate the received signal and extract amplitude and phase information from it. A traditional way is to calculate cross correlation function of the original modulation signal and returned signal [8]. Cross correlation can be calculated by measuring the returned signal at selected phases, which can be implemented using mixers and low-pass filters in the detector. This, however requires very complicated circuitry [10].

Another, more efficient approach is to sample the modulated returned light synchronously using special pixel structure. Received modulated light is simultaneously mixed with reference signal and sampled at four different phases (0°, 90°, 180°, 270°), as illustrated in Figure 3 [8]. Advantage of this synchronous sampling technique is simpler circuit design and smaller pixel sizes, which allows putting more pixels in to a sensor, resulting in higher resolution. In the literature, this kind of pixel architecture is referred as photonic mixer device (PMD)[3, 13] or lock-in-pixel[9, 10].

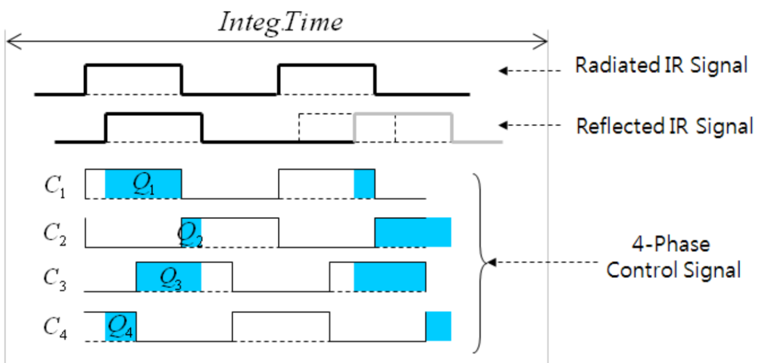

Figure 3: Demodulation principle of CW method. $C1$ to $C4$ are control signals with 90° phase delay from each other. [16, p. 3]

As with the pulsed method, multiple samples are measured and averaged to increase the SNR. Using four equally spaced sampling windows Q_1 to Q_4 timed by the reference signal (see Figure 3), received signal is sampled at different phases for a integration time period. Assuming that the modulation signal is sinusoidal wave without harmonic frequencies, discrete Fourier transform (DFT) equations can be used to calculate phase ϕ, amplitude A and offset B as follows [8, 15]:

$$\phi = arctan\left(\frac{Q_3 - Q_4}{Q_1 - Q_2}\right) \tag{3}$$

$$A = \frac{\sqrt{(Q_1 - Q_2)^2 + (Q_3 - Q_4)^2}}{2} \tag{4}$$

$$B = \frac{Q_1 + Q_2 + Q_3 + Q_4}{4}. \tag{5}$$

From the phase ϕ, the distance can be finally calculated by [15]:

$$d = \frac{c}{4\pi f}\phi. \tag{6}$$

The intensity, i.e. amplitude A of the light decreases proportionally to the travelled distance in a known way. Hence, the received amplitude value from (4) can be used as a confidence measure for the distance measurements [7]. Additionally, the reflected signal is often superimposed to background illumination, which causes error to the measurement. Thus, the offset (5) is used to distinguish modulated light component from the background light [8].

When calculating distances from phase difference as in (6), one important thing has to be considered. Since the modulation signal is periodical, its phase wraps around every 2π. This means that also distances can be measured unambiguously only in a certain range:

$$L = \frac{c}{2f}, \tag{7}$$

which depends only on the modulation frequency f. For example, using typical 20 MHz modulation frequency, unambiguous range is 7.5 meters. Therefore, reflected signal coming from distance of 8.6 meters would be mapped to 1.1 meters, since they are in same phase. Operating range can be extended by using lower modulation frequency, but it results to decreasing of resolution[17]. Another way is to take multiple measurements with different modulation frequencies. First, coarse distance is calculated using lower frequency, and after that higher frequency is used to increase accuracy of the measurement [15]. However, using multiple frequencies shortens the integration time that can be used, thus decreasing the SNR. [16, p. 37].

To summarize, pulsed method gives better operating range (up to 1500 m [18]), but depth resolution is worse than in CW modulated method. On the other hand, CW method has an ambiguity problem, which limits the operating range. Most commercial TOF cameras are based on continuous wave modulation [9, 18].

2.3 Data formats

Three-dimensional data can be stored and presented in various ways. The data types differ in memory usage, computational effectiveness and amount of information they can preserve. *Depth image* and *point cloud*, both commonly used with ToF cameras, are presented and their common uses and limitations are considered. Examples of these data types are shown in Figure 4. Other representations exist as well, but they are out of scope of this work. For further information, please see [1, Ch. 4]

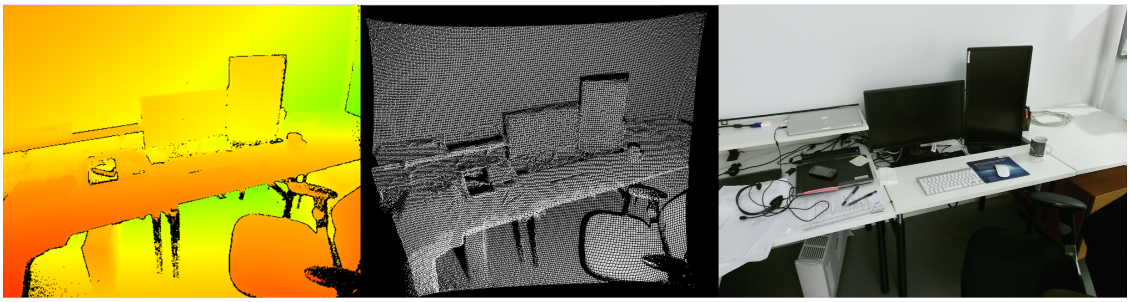

Figure 4: Depth image, point cloud and RGB image taken from same scene.

Depth image, *range image* or sometimes called *depth map* is a data type that ToF cameras usually output directly. Similarly to traditional intensity image, depth image is a two-dimensional array with the width of x and the height of y pixels, but each pixel containing a depth value instead of intensity. Depending on the device, pixels of the output image can also contain intensity and amplitude information, which can be used as a confidence measure [7].

Because depth image is acquired from a single viewpoint, it can only have one depth value for each x, y position. If there are occluding, i.e. overlapping objects in the scene, only the nearest point from the camera can be stored in a depth image. Therefore, depth image is considered being actually 2.5D instead of 3D [1, p. 72].

Point cloud is an unstructured data format, where the 3D data is stored as a set of points, each having x, y, z coordinates. Points can also include additional data, such as RGB values. The cloud can hold more complex geometric shapes than range image, allowing to combine data from multiple viewpoints to a single structure. Since each point needs to hold at least three floating point values, the memory consumption of point cloud is higher than in a depth image.

Point cloud can be converted to a depth image by projecting the data points from single perspective to a 2D plane, but all occluding points are lost. On the other hand, depth image can also be converted to point cloud if the intrinsic parameters, such as focal length and lens distortion of the camera are known. Most commercial Time-of-flight cameras provide both depth images and point clouds as output.

3 Comparison to Other Technologies

In addition to Time-of-Flight cameras, there are several other technologies traditionally used for acquiring depth information from the scene. These systems can be divided into two categories, passive and active. Passive systems, such as stereo cameras, do not use any source of illumination and are only relying on ambient lighting. On the other hand, active systems (where ToF camera also belongs to) illuminate the scene with a light source, and the depth information is computed in various ways. [1, p. 11]

Each system has its limitations, and best choice always depends on the application. Properties that need to be considered include operating range, resolution, physical size and performance in different environmental conditions. In this chapter, some most popular depth imaging techniques are discussed. Different approaches for both active and passive systems are briefly presented, and their advantages and disadvantages compared to ToF cameras are considered.

3.1 Passive 3D Imaging systems

Passive 3D imaging systems, generally based on traditional 2D gray-level or RGB cameras, are the oldest and the most studied depth imaging technology. Several different approaches fall into the passive category, some employing only a single camera while others use multiple cameras. They all have in common the fact that they use only ambient lighting, so the depth information is extracted from the details in 2D images. Single-camera methods, such as shape-from-focus or structure-from-motion use sequences of images, and geometry of the scene is computed from variations between the frames. The most well-known multi-camera method is stereo vision, where two fixed cameras image the same scene from slightly different perspectives. Stereo camera system is illustrated in Figure 5. Corresponding features (e.g. edges, corners) are searched from both images and the depth information can then be determined using *triangulation*. For more details on passive methods, see [1, ch. 2].

Stereo vision systems has been researched for several decades, so they are rather well-known. With modern technology, stereo vision system can be implemented by using even cheap consumer cameras and personal computers, so the costs can be

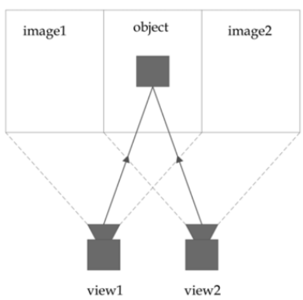

Figure 5: A stereo camera system. [17]

very low [15]. The depth resolution and range of a SV setup can also be rather easily changed by selecting suitable camera equipment and adjusting baseline of the cameras [9].

Compared to TOF cameras, stereo vision as well as other passive systems have a few major drawbacks. First, they are highly sensitive to changes in lighting and can not operate at all in darkness, since they do not employ a light source like TOF cameras. Secondly, the computation of depth values is based on details and features in the image. This means that passive methods perform poorly in non-textured and uniformly coloured surfaces that don't have much details. Third, the correspondence problem, i.e. finding corresponding features in different images is very complex and computationally intensive process. Last, setting up and calibrating a stereo camera system requires great amount of work and knowledge. Several commercial implementations exist, but they usually are fixed and not easily configurable [1, ch. 2.9.1].

In contrast, Time-of-flight cameras perform well in low-detailed scenes and in different lighting conditions. Depth computation is done directly on the chip, so they can provide depth data at high frame rate without need for heavy computing. Operating range can be changed easily by adjusting integration time, and no calibration is needed [17].

3.2 Structured Light

Structured light (SL) approach is also based on triangulation principle, but it uses active illumination unit that projects specific pattern of light on to a scene. Camera sees the distorted pattern from a different viewpoint, and the geometry of the surface can then be calculated from deformations in observed pattern. Many alternative light patterns can be used, such as lines or speckles. [9]

The most successful structured light has been Kinect [2] by Microsoft. Since its launch in 2010, numerous applications have been developed and its performance has been extensively studied. Several comparisons to Time-of-Flight cameras have also been done, with varying conclusions [3] [19] [9]. A few years ago TOF cameras still had such a low resolution that Kinect outperformed state-of-the-art TOF camera models in accuracy, when operated in short ranges in a controlled environment. However, the operating range of Kinect is very limited (3.5 meters)[19], so it suitable only for limited applications.

Structured light systems generally, not limited only to Kinect, have also some other disadvantages compared to TOF cameras. First, they suffer from partial occlusion because the light projector and camera are in different viewpoints. This leads to missing depth information in parts of the image [9]. Secondly, SL systems are very sensitive to lighting conditions. Decrease of illumination power and triangulation accuracy limit the operating range, so generally SL systems do not work well in outdoors. Increasing the illumination power affects to compactness, and eye-safety aspects also has to be considered. Last, because of the triangulation method, the computation of depth values is significantly more demanding than in TOF cameras.

3.3 Laser range finders

Laser range finder (LRF), also called *Lidar* or *laser scanner*, is based on the same principle as TOF camera, measuring the time-of-flight of a reflected light. Instead of illuminating the whole scene, it uses a laser beam to illuminate and measure only a single point at a time. Typically a laser scanner device also has a mirror rotating at high speed, so it can measure distances in a 2D plane. To acquire complete 3D information, the laser scanner needs to be tilted up and down to sequentially scan the scene row by row to form a point cloud. Thus, this kind of actuated system clearly differs from cameras, as it *scans* the scene instead of simultaneously capturing the whole view.

Compared to other 3D measurement technologies, laser scanners generally have very high resolution and long operating range up to hundreds of meters. Laser scanners are widely used in many high accuracy demanding applications, such as remote sensing [1, ch. 9.4] or in mobile robotics for localization and mapping purposes [6] [9]. Stoyanov et al. experimented by using actuated laser scanner to define ground truth data for comparison of other depth sensors [19]. They indicated that although being most accurate of tested sensors, actuated laser scan data still had rather large variance, mainly because of inconsistencies in timing of the samples and actuating the sensor.

However, actuated laser scanner systems have some significant drawbacks. First of all, moving parts cause many problems. Mobile parts are vulnerable to mechanical stress and they add weight and complexity to the system as mentioned above. Row by row scanning also increases the time needed to capture the whole 3D scene, making actuated laser scanners unsuitable for high-speed dynamic applications [9]. Recently, manufacturers such as Velodyne [20] has brought to markets more advanced multi-channel LRF sensors, which can scan up to 64 rows simultaneously and are more compact in size. These multi-channel sensors still suffer from poor vertical field of view, and although prices have been dropped lately, they are still rather expensive compared to recent TOF camera models.

4 Error Sources

The accuracy of depth measurement is limited by the errors added to the signal. Many different errors originating from various sources and ways to compensate them have been identified by various researchers [6, 8, 9]. Errors can be classified in to two categories: systematic errors, which can generally be calibrated out, and random errors, whose presence can not be predicted but which can filtered out. Some most important error sources affecting ToF depth measurement are presented in this chapter. For more detailed overview of errors, please see [9].

4.1 Systematic Errors

Circular error, also referred as *wiggling error*[9] originates from unideal modulation process. In calculations the modulation signal is assumed sinusoidal but in reality it has irregularities, resulting in offset that depends on the distance measured. Effect of wiggling error is shown in Figure **??**. Offset follows a sinusoidal shape when plotted against distance, thus the name circular error.

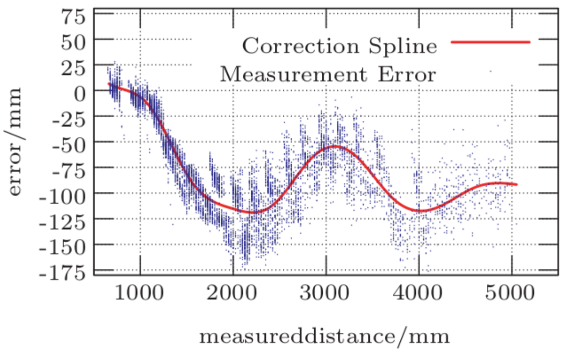

Figure 6: Circular error, measured with SR3000 camera. Blue dots are measurements, and red line is a polynomial model fitted to dataset for correction. [6]

Circular error can be compensated either by comparing measurement to a reference ground truth, or by modelling the error using multiple samples [9]. The first mentioned requires additional sensor, and the one is more time-consuming as it requires large amount of measurements.

Amplitude related errors result either from low amplitudes or overexposed pixels. Amplitude of the received light is closely related to the depth, so generally higher amplitude means better depth accuracy. There are a few main reasons to amplitude related errors. First, inhomogenous light source causes different level of illumination in different parts of scene. Usually the light power decreases when moving from center towards the edges, as seen in Figure 7. Lower amplitude values result to more distant measurements in the edges, although the object is flat wall. On the other hand, the center of the image seems to be closer because the pixels are overexposed. Objects being near to the camera or too long integration times can cause overexposure, which in worst case can lead pixels to be saturated [9].

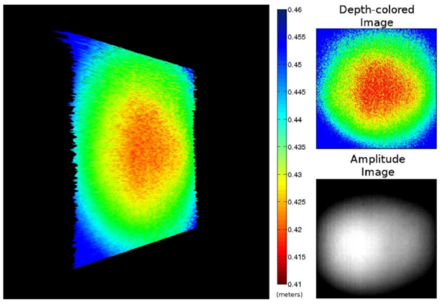

Figure 7: Amplitude related error due to inhomogenous light. Object is a flat wall at 0.43 meters. [9]

Another cause for amplitude errors is different reflectivity of materials []. For example, while white paper reflects approximately 80% of the light at 850nm wavelength, black rubber tire reflects only 2% [7]. If the scene includes multiple materials with very different reflection factors, depth measurements are invalid [16].

Temperature related errors exist because semiconductors technology, especially photodetectors used in TOF cameras, are highly sensitive to temperature changes [21]. When turned on, temperature starts gradually rising, causing drift to the whole depth image. Reports suggest that cameras should be let warm up to measurements to stabilize [9, 18]. In SR4000 camera, error stabilized after waiting 40 minutes to a few millimeters depending on a measured distance and integration time [18].

Also external temperature affects to the measurement. Kahlmann et al. measured object at different temperatures between -10°C and 40°C taking only single frames, so that camera did not warm up [21]. Their results showed that the temperature of the object clearly affects to measured distance, drift being around 8mm/°C. Drift was nearly constant, so it can be eliminated by calibration.

Fixed-pattern noise (FPN) is related to imperfect manufacturing process of pixels. Each pixel has different characteristics of sensing light, causing constant offset [6]. Also the placement of the pixel in the chip has effect. Rows or columns of pixels are connected in series, which results to small signal delays in consecutive pixels. FPN is always constant, thus it can be rather easily calibrated out.

4.2 Random Errors

Noise consists of several different sources, the most substantial being *photon shot noise* [8]. Shot noise limits theoretically reachable SNR and also the depth resolution. Knowing the statistical properties of shot noise, the absolute limit for *depth resolution* can be derived as[8]:

$$\sigma_d = \frac{L}{\sqrt{8}} \frac{\sqrt{B}}{2A}, \tag{8}$$

where A, B and L are from (4), (5) and (7). From the equation it can be seen that a large offset, i.e. bright background illumination increases the noise. Only way to increase this accuracy is averaging [8]. However, using more frames decreases the frame rate, and if the object is moving between frames, averaging does not work. Also median filtering can be used to filter out noise, but it does not work in every case, e.g. in very detailed scene [10].

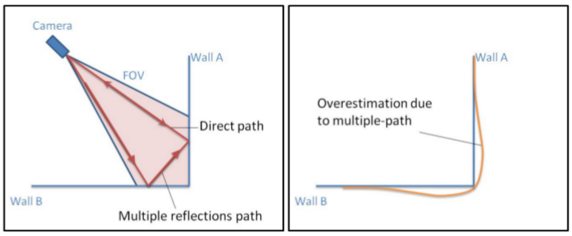

Figure 8: Multipath reflections on a corner. Due to interfering light waves, corner appear rounded.[7]

Multipath reflection errors appear because light reflects from multiple surfaces before returning to receiver. In corners, multiple light rays interfere and cause wrong depth measurement in the pixel, resulting in rounded-off corners as seen in Figure 8. Concave objects may reflect light in such way that it never reaches back to sensor, causing discontinuities called *jump edges* in the depth image. Several approaches have been proposed to filter out jump edges, such as comparing to neighbouring pixels [6]. Dealing with multipath reflections originated from corners seems to be still an open question.

Light scattering errors come from multiple light reflections between lens and image sensor inside the camera. As illustrated in Figure 9, some part of light scatters to neighbouring pixels, resulting to depth errors. Scattering error occur only when very near objects are present in the scene [9].

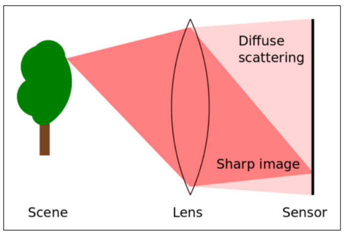

Figure 9: Scattering effect. Light is scattered inside lens and spreads all over imaging sensor. [22]

Scattering error can be reduced by filtering with amplitude and phase data. Some results show that up to 90% of scattering error can be eliminated [22]. However, this approach requires access to raw data, which some camera models does not provide.

5 Conclusion

In this report, Time-of-Flight camera technology and its operating principles were reviewed. Sensor design and two different approaches for depth measurement, pulsed light and continuous wave modulation, were presented and their advantages and limitations were discussed. ToF camera was also compared to other widely used depth imaging technologies, emphasizing its advantages. Finally, most prominent error sources affecting to depth measurements were discussed and some methods to overcome these problems were briefly presented.

ToF cameras provide depth images at high frame rate without need for heavy data processing. Compact size and robust performance in dynamic conditions are also its advantages. Up to this point, the most significant drawback has been its low resolution. However, as the technology has been developing rapidly, ToF cameras have dramatically improved during last few years. Lately, when Microsoft Kinect v2, the first mass-produced ToF device aimed to consumer markets was released, also the prices dropped. High resolution sensor being available to even wider audiences, new applications are expected to emerge in the near future. At the same time when the whole world is changing as robots are substituting humans in more and more jobs, there definitely is a need for new, better 3D vision sensors.

Time-of-Flight technology still has some fundamental problems, such as limited range and a couple of error sources decreasing the data quality. Nonetheless, better understanding of ToF technology helps solving these problems, as constant advances in the research field have shown. It is certainly not mature technology yet, but it clearly has potential.

References

[1] Nick Pears, Yonghuai Liu, and Peter Bunting. *3D Imaging, Analysis and Applications*. Springer, 2012. ISBN: 9978-1-4471-4062-7.

[2] Microsoft. *Kinect*, 2010. http://www.xbox.com/en-US/xbox-360/accessories/kinect. Accessed on 15.1.2015.

[3] Benjamin Langmann, Klaus Hartmann and Otmar Loffeld. Depth Camera Technology Comparison And Performance Evaluation. In *Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods*. pages 438-444. 2012.

[4] Willow Garage. *TurtleBot*, 2011. http://www.willowgarage.com/turtlebot. Accessed 15.1.2015.

[5] Erik E. Stone and Marjorie Skubic. Evaluation of an inexpensive depth camera for passive in-home fall risk assessment. In *5th International Conference on Pervasive Computing Technologies for Healthcare*, pages 71-77. 2011.

[6] Stefan May, David Droeschel, Dirk Holz, Stefan Fuchs, Ezlo Malis, Andreas Nüchter and Joachim Hertzberg. Three-dimensional mapping with time-of-flight cameras. In *Journal of Field Robotics*, Vol 26: pages 934–965. 2009.

[7] Mesa Imaging AG. *SR4000 User Manual*. 2011.

[8] Robert Lange and Peter Seitz. Solid-State Time-of-Flight Range Camera. In *IEEE Journal of Quantum Electronics* 37(3). pages 390-397. 2001.

[9] Sergi Foix, Guillem Alenyà and Carme Torras. Lock-in Time-of-Flight (ToF) Cameras: A Survey. In *IEEE Sensors Journal* 11(9). pages 1917-1926. 2011.

[10] Burak S. Gokturk, Hakan Yalcin and Cyrus Bamji. A Time-Of-Flight Depth Sensor - System Description, Issues and Solutions. In *IEEE Conference on Computer Vision and Pattern Recognition Workshop, 2004 (CVPRW'04)*. 2004.

[11] Cyrus S. Bamji, et al. A 0.13 μm CMOS System-on-Chip for a 512x424 Time-of-Flight Image Sensor With Multi-Frequency Photo-Demodulation up to 130 MHz and 2 GS/s ADC. In *IEEE Journal of Solid-State Circuits 5(50(1)*. 2015.

[12] Microsoft. *Kinect for Windows v2*, 2014. http://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx. Accessed on 15.1.2015.

[13] Xuming Luan. Experimental investigation of photonic mixer device and development of TOF 3D ranging systems based on PMD technology. *Diss., Department of Electrical Engineering and Computer Science, University of Siegen* 2001.

[14] Cristiano Niclas, Claudio Favi, Theo Kluter, Frédéric Monnier and Edoardo Charbon. Single-Photon Synchronous Detection. In *IEEE Journal of Solid-State Circuits* 44(7). pages 1977-1989. 2009.

[15] Larry Li. Time-of-Flight Camera - An Introduction. *Texas Instruments - Technical White Paper*, 2014.

[16] Miles Hansard, Seungkyu Lee, Ouk Choi and Radu Horaud. *Time of Flight Cameras: Principles, Methods, and Applications*. Springer, 2012. ISBN: 978-1-4471-4658-2.

[17] Stephan Hussmann, Thorsten Ringbeck, Bianca Hagebeuker. A Performance Review of 3D TOF Vision Systems in Comparison to Stereo Vision Systems. In *Stereo Vision,* Asim Bhatti(Ed.). InTech Open Access Publisher. 2008.

[18] Dario Piatti and Fulvio Rinaudo. SR-4000 and CamCube3. 0 time of flight (ToF) cameras: Tests and comparison. In *Remote Sensing 4(4)*. pages 1069-1089. 2012.

[19] Todor Stoyanov, Athanasia Louloudi, Henrik Andreasson and Achim J. Lilienthal. Comparative Evaluation of Range Sensor Accuracy in Indoor Environments. In: *ECMR*. pages 19-24. 2011.

[20] Velodyne. *Velodyne VLP-16*, 2014. http://velodynelidar.com/lidar/hdlproducts/vlp16.aspx. Accessed 19.1.2015.

[21] Timo Kahlmann, Fabio Remondino and H. Ingensand. Calibration for Increased Accuracy of the Range Imaging Camera Swissranger. In *Image Engineering and Vision Metrology (IEVM) (36:3)* pages 136-141. 2006.

[22] Henrik Schäfer, Frank Lenzen and Christoph S. Garbe. Model based scattering correction in time-of-flight cameras. Optics express 22(24) pages 29835-29846. 2014.