Outdoor Operation of Structured Light in Mobile Phone

Byeonghoon Park, Yongchan Keh, Donghi Lee, Yongkwan Kim, Sungsoon Kim, Kisuk Sung, Jungkee Lee,
Donghoon Jang, and Youngkwon Yoon
Mobile Communications Business, Samsung Electronics Co., LTD.

Abstract

Active Depth Camera is about to be integrated into a mobile phone and now beginning to be introduced into the market. It is expected that this technology will enable brand-new and meaningful user experiences in egocentric ecosystems. In view of practical usage, however, Active Depth Camera does not operate well especially outdoors because signal light is much weaker than ambient sunlight. To overcome this problem, Spectro-Temporal Light Filtering, adopting a light source of 940nm wavelength, has been designed. In order to check and improve outdoor depth quality, mobile phones for proof-of-concept have been implemented with structured light depth camera enclosed. We present its outdoor performance, featuring a 940nm vertical cavity surface emitting laser as a light source and an image sensor with a global shutter to reduce ambient light noise. The result makes us confident that this functionality enables mobile camera technology to step into another stunning stage.

1. Introduction

Egocentric cameras are becoming more popular for eyewear devices. Especially, depth cameras are playing a more and more important role in egocentric devices since depth data can provide absolute scale information for SLAM, surface geometry for AR contents alignment, and real-time 3D modeling. However, the requirement for embedding a depth camera into an eyewear device is very demanding when its size, power consumption and mechanical robustness are taken into account. In practical point of view, these factors are too tough hurdles to overcome in short time so that there seems quite a long way to go for the commercialization of depth cameras with eyewear devices.

On the other hand, a depth camera in a mobile phone can be considered as an alternative for the egocentric device near at hand. Smartphone has been so widespread and indispensable in our lives that people feel very convenient in using it as if it is one of their body parts. Recently a few pioneering groups of companies have successfully demonstrated the concept of depth sensor integration [1, 2] and others are further trying to commercialize mobile phones with depth cameras [3, 4]. These trends make people to think of the use cases such as AR based on a mobile phone or VR with a phone dropped-in, where egocentric cameras are embedded in mobile phones not in eyewear devices. Especially, Active Depth Camera which has a separate infrared light source has an important meaning in that it can produce highly accurate depth information in real-time, so that it can enable wide varieties of egocentric applications.

The configuration of Active Depth Camera inside a mobile phone can be divided into two categories – one is Front Facing which sees a user's face, and the other is Rear Facing which sees a world. The architecture of Active Depth Camera includes Structured Light (S/L), Time-of-Flight (ToF) or Assisted Stereo, which has pros-and-cons originating from its own inherent properties. Accordingly, there exists more appropriate architecture for either Front or Rear Facing. It is generally accepted that S/L is suitable for Front Facing while ToF is for Rear Facing in view of operating range. This is because depth noise of S/L increases quadratically with distance whereas that of ToF goes linearly. Usually, depth noise of S/L is smaller in shorter range (e.g. within 1 meter) when compared to ToF, which has smaller depth noise in longer range (e.g. more than a few meters.) than S/L does.

Recently, the interests on Front Facing Active Depth Camera have been growing up [5], raising the anticipation for the next big innovation of mobile phones. Front Facing configuration can enable or enhance many face-centric use cases: 1. 3D face recognition solves the spoofing problem of iris or face recognition based on 2D camera. By using IR light source, lower recognition rate at low lighting condition of 2D camera is improved. In addition, pose correction for face frontalization is simplified, which gives larger angular freedom, resulting in more convenience in usage. 2. 2D image can be enhanced with the help of depth data. Depth based segmentation of user face from the background makes the calculation load lighter for out-focusing, re-focusing and bokeh effect so that real-time operation is supported. 3. 3D modeling is unique feature of depth camera application. When it is aimed at user face, it generates photorealistic avatar which can be widely used
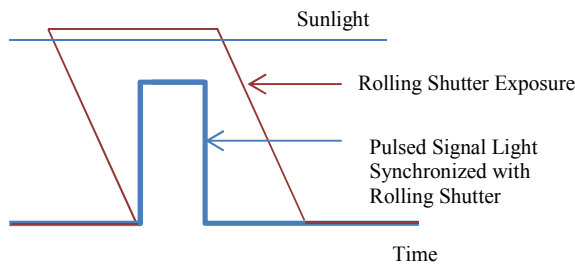
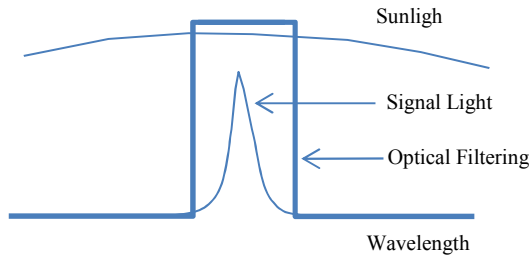Figure 1: Normal rolling shutter timing with global exposure

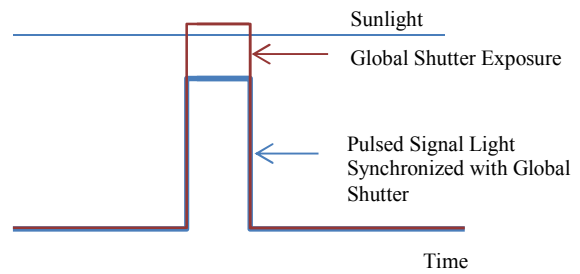Figure 2: Concept of optical filtering to suppress sunlight

Figure 3: Global shutter synchronized with pulsed signal light to suppress sunlight.

Figure 4: Rolling shutter with over-illuminated pulsed signal light to suppress sunlight.

for SNS and video communication. Accurate 3D face model improves the reality of virtual try-on (glasses, hats, and hair-style etc.) and beautification effect not only by texture, but also by geometry.

As can be noticed easily, these use cases are not limited indoors only so that outdoor operation should be supported. However, none of the above mentioned Active Depth Cameras does operate well at outdoor lighting condition because IR light source should compete with IR emission of sunlight. The reasons for the difficulty can be investigated in more detail:

First, the IR intensity of sunlight is much stronger than that of signal light source so that the projected pattern (S/L) or the returned signal light (ToF) is usually buried in the ambient noise light and thus can be hardly detected by the image sensor. Second, the strong sunlight frequently causes pixel saturation which makes it difficult to select a proper exposure condition. That is, if we set exposure time longer, the captured image tends to be saturated. On the contrary, if we change exposure time shorter, the signal light intensity becomes too weak to be detected. Third, when we think of a normal timing of a rolling shutter image sensor with global exposure, the ambient light cannot be suppressed effectively due to the longer camera exposure time than the illumination pulse time as shown in Fig. 1. Because of the DC nature of sunlight, larger amount of sunlight becomes incident to each pixel of the image sensor when compared to the amount of signal light.

2. Related Works

To overcome the strong sunlight and to make Active Depth Cameras operate well outdoors, there have been suggested a few approaches but we find new problems when applying them to the mobile phone system.

2.1. Light Concentration [6]

Sunlight is often 2-5 orders of magnitude brighter than the projected structured light. The key idea is that signal level can be increased by concentrating the available projector light on a small portion of the scene with the help of optical scanner. However, to adopt this technology for mobile phones, there exist a few difficulties. First, optical line scanner needs to be embedded into mobile phone, which is not favorable because of size and reliability concerns. Miniature MEMS scanner may be considered, but it still occupies more footprint than conventional pattern projector module. Second, Light Concentration requires capturing many frames, at least 20 or more to construct one frame of depth map. Normally, the frame rate of depth applications with mobile phone is 5~30fps so that it cannot cope with real-time operation.

2.2. Motion Contrast 3D Scanning [7]

This work tries to improve the long capture time of Light Concentration by using Dynamic Vision Sensor (DVS). The asynchronous event-detection property of DVS helps to reduce the number of frames required for 3D reconstruction. But there is still a way to go for a single shot capturing because of noisy nature of DVS. In addition, DVS is still big-sized and has low-resolution to be deployed at the mobile sector.

2.3. Optical Filtering for suppressing ambient light [8]

The approach shown in Fig. 2 makes use of a narrow spectral bandwidth laser and a narrow-band spectral filter in front of a camera. Since sunlight has broad bandwidth, only a small portion of the whole spectral power can pass through the band-pass filter and be incident to image sensor. On the other hand, the signal light from laser can pass the optical filter with most of optical power intact, resulting in increased Signal-to-Noise Ratio (SNR). Recently, Near-infrared vertical cavity surface emitting laser (VCSEL) array device has become mature, providing very small temperature shift of peak wavelength, ~0.07 nm/°C. In mobile environment, ambient temperature changes in wide range of ~80°C, but 5.6nm shift is expected to be small enough to design a narrow-band spectral filter.

2.4. Spectro-Temporal Light Filtering

Our work is based on Optical Filtering method and further selecting a proper peak wavelength of light source at a water absorption peak of sunlight, ~940nm. The advantage of this approach is that it is very practical for implementation, only by changing wavelength design of the existing mobile S/L system. Since the quantum efficiency of an image sensor at 940nm drops ~1/3 of that of 850nm, bigger pixel size or pixel binning of image sensor is recommended to compensate the number of photo-generated electrons at each pixel.

Preferably, an image sensor with global shutter which is synchronized with laser pulse can further suppress the ambient light by reducing integration time and increasing laser power, i.e. Light Concentration in time domain as shown in Fig. 3. Because of the short integration time, most of the continuous-wave sunlight can be filtered out while short laser pulse can be filtered in during the exposure time. In order to reach a threshold energy level to be detected at the image sensor's pixel, laser power needs to be increased during the short pulse time. Alternatively, we can use a
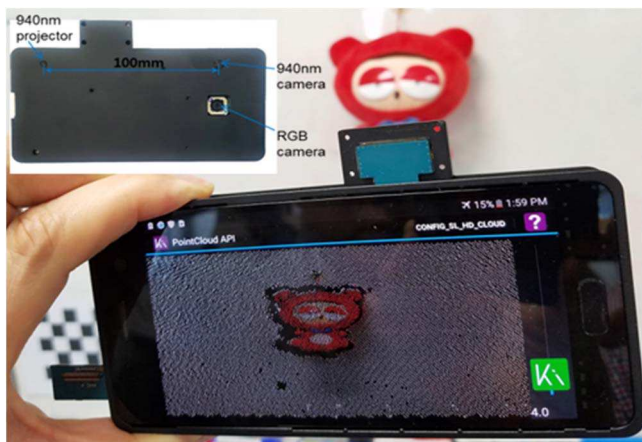

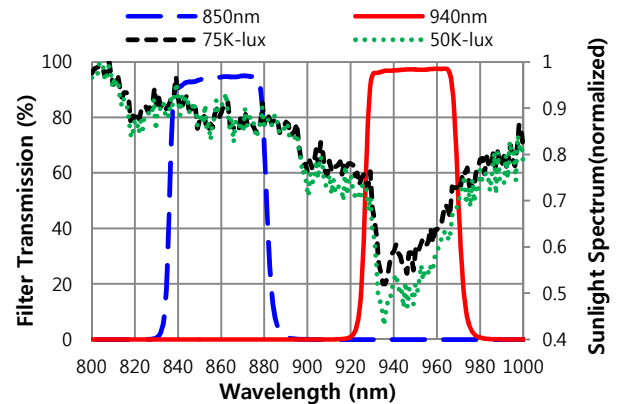Figure 5: Proof-of-concept phone. Rear cover is shown at inset.

Figure 6: Bandpass filter spectra for center wavelength of 850nm and 940nm, respectively. Sunlight spectra are also measured at different ambient light conditions.

rolling shutter image sensor by introducing over-illumination concept as shown in Fig. 4. In this approach, the rolling shutter time has been minimized to suppress sunlight whereas the pulse width of signal light has been increased to cover the entire exposure lines. The power consumption of the light source will increase inevitably, but it can be minimized if we make the readout time short by using fast readout technology.

Although the proposed approach may not completely cope with very severe outdoor lighting conditions, it has been proved to operate successfully within short range (<1m) and at moderate sunlight condition (<50,000 lux), which is enough to enable most Front Facing applications of mobile phone at outdoor usage.

3. Implementation Details

S/L type depth camera consists of IR image sensor and pattern projector as well as decoding processor. To evaluate ambient light effects on depth accuracy for different wavelengths, we have built two kinds of depth sensor systems, 850nm and 940nm. Prototype device is shown in Fig. 5.

For sensor sides, in order to limit ambient light noise outside the transmission windows, passband width should be minimized as narrow as possible. By considering the angle of incidence and the wavelength shift of light source due to thermal heating or cooling, passband is designed as an optimal value in accordance with mobile operating conditions.

As shown in Fig. 6, we used IR band pass filter on the top of image sensor with 45nm of full-width at half-maximum for 850nm system and 43nm for 940nm one, respectively. Peak transmission depends on dielectric film configuration, but it is over 93% typically. Outside the passband, transmission meets the condition of OD-4 at least.

The advantage of 940nm system comes from reduced solar spectral power at that wavelength due to water vapor

absorption [9, 10]. It is known that 940nm sunlight intensity is about one fifth ratio compared to that of 850nm [11], however in our measurements, it was reduced to 64~70% of 850nm spectra at 940nm depending on the sunlight illuminance of 75K~50K-lux. This originates from different atmosphere condition of dry spring weather. By lowering solar spectral power as noise, we get higher SNR, that is, much better depth quality can be obtained. We expect that the SNR improves ~1.6 times at 940nm wavelength compared with that of 850nm. The experiment uses global-shutter type CMOS image sensors with 640x480 resolutions for 850nm system and 640x400 for 940nm system, respectively. Since the quantum efficiency decreases about one third at 940nm, we binned four pixels together as a single pixel from the original higher sensor resolution in order to compensate low light absorption. We have also designed two kinds of lens system according to different image field sizes.

Relative sensitivity (RS) values of both cameras are calculated by multiplying aperture size ($1/F/\#^2$), quantum efficiency, and pixel area (square of pixel size). While captured image of 850nm system shows partially saturated image near nose and forehead as shown in Fig. 7(a), 940nm system does not saturate at the same conditions of 0.5ms exposure time and x1 gain because it has a smaller relative sensitivity.

In addition, reduced ambient sunlight helps not to saturate up to 1.4ms of camera exposure time at 60K-lux sunlight. With less than 0.2ms exposure time, 850nm camera does not saturate, but active structured light level becomes too weak to distinguish its pattern from noise.

This turns out more severe at even higher sunlight illuminance, therefore 850nm camera is definitely not appropriate for outdoor usage. When we capture images of a flat wall in Fig. 7(b) and (d), structured pattern seems too weak or partially saturated with 850nm camera, but looks good with 940nm camera. To calculate standard deviation of depth data, we obtained point clouds by measuring distances of this wall from the POC device.

	850nm camera	940nm camera
Resolution	640x480	640x400
Field Size [mm]	2.4	4.53
Pixel [um]	3	6
Lens F/#	1.7	2.4
Quantum Eff.	0.32	0.12
Relative Sensitivity	11.1	8.3
Normalized RS	**1.0**	**0.75**

Table 1: Parameters for calculating normalized relative sensitivities.

Further increase of RS in 940nm camera can be obtained by a large aperture lens design. For example, F/# of 2.0 lens system provides normalized RS of ~1.1, or 32% of QE at
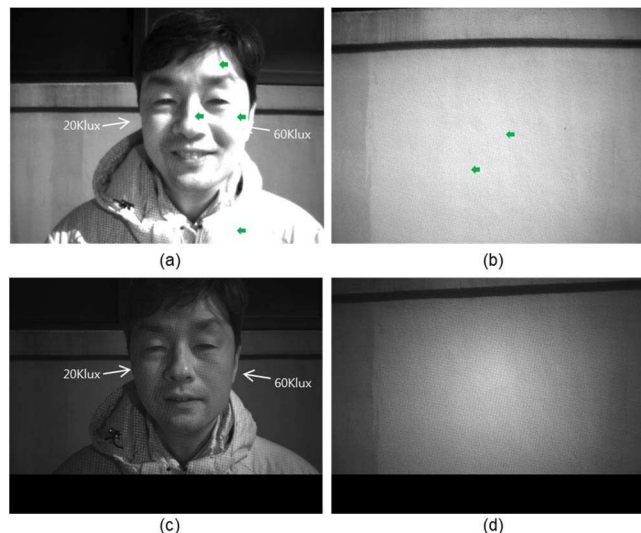

Figure 7: Outdoor IR images with the same 0.5ms of exposure time under the 60K-lux ambient sunlight for different wavelengths of bandpass filter (a) 850nm face, (b) 850nm wall, (c) 940nm face, and (d) 940nm wall. Green marks indicate saturated part.
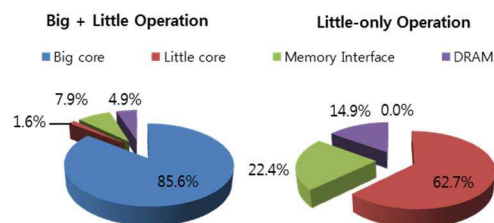

Figure 8: Power consumption budget for depth decoding with respect to big+little core operation and little core only decoding.

940nm guarantees normalized RS of 2.0.

Bright camera system, thus high SNR, has a significant advantage of low power consumption for mobile devices. Therefore, recent research has been focused on new infrared image sensor to boost QE at 940nm [12].

4. Experiments

In order to verify the effect of ambient sunlight on the accuracy of depth sensor in mobile devices, we have built a proof-of-concept device, which has camera and projector with baseline of 100mm, and decodes depth by using mobile application processor at a frame rate of 30fps. Big-little ARM-core architecture suggests us to use little cores to extract point clouds at the lowest power consumption. When we decode depth from Exynos 7420 application processor, little-only operation consumes 27mW/frame, while that of big+little case is 57mW/frame. Practical power consumption for eyewear or head-mount
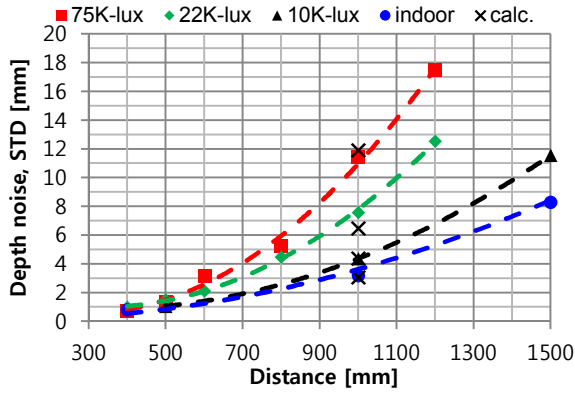
Figure 9: Depth noise characteristics of 940nm system as a function of actual distance to the target wall. Ambient illuminance of indoor condition is about 500lux.

device will be converged to ~10mW/frame, where small memory usage is required as analyzed in Fig. 8.

As a light source of IR projector, VCSEL array was used. Since it has a small thermal wavelength shift with relatively high wall-plug-efficiency, it can minimize required bandwidth of IR filters for mobile application. Its optical pulse power is about 3.0W at the operating current of 4.0A.

We separated the term "depth accuracy" into "depth noise" and "depth error" for clarity. Depth noise can be represented as a standard deviation (STD) of depth values obtained from point clouds. STD is extracted from the center 50% depth data out of total field-of-view. Fig. 9 shows measured STD curves according to the absolute distance. We controlled exposure time of the sensor from 0.3ms to 1.0ms in order to avoid image saturation. Peak power of optical pulse was set as 3.0W. As ambient light grows brighter, STD increases accordingly. The STD curves resemble to a typical STD curve as in a triangulation measurement, showing square function of distance:

$$\sigma_Z \propto \frac{1}{f \cdot b} \cdot \sigma_\Delta \cdot Z^2 \quad (1)$$

where f is the focal length of camera, b is the baseline of POC device, Z is the distance, and σ_Δ is the STD of the measured disparity [13].

Depth noise of structured light sensor can be expressed analytically by considering the intensity of active structured signal light over ambient sunlight intensity [6]. Since σ_Δ is proportional to the inverse of SNR, depth noise can be expressed as:

$$\sigma_Z \propto \sigma_\Delta \propto \frac{\eta}{I_L} = \frac{\sqrt{\sigma_R^2 + (I_L + I_A)}}{I_L} \quad (2)$$

where I_L and I_A are the measured pixel grey values, given in digital numbers, of active structured signal light and

ambient noise light, respectively. σ_R is the sensor readout noise. Additional noise sources can be included such as laser speckle noise and decoding noise for more accurate modeling, but for simplicity, we are going to consider intensity noises only at limited conditions.

Pixel grey values are also proportional to photo-generated electron numbers on the sensor by the relation [14]:

$$I_L \propto N_e = P_{LS} \cdot \rho \cdot \left(\frac{D}{2Z}\right)^2 \cdot k_{lens} \cdot QE(\lambda) \cdot \frac{A_{pixel}}{A_{image}} \cdot \frac{\lambda}{hc} \cdot T_{int} \quad (3)$$

where P_{LS} is the power of light source, ρ is the reflectivity of objects, D is the lens aperture, and k is the loss factor.

For indoor case, since I_L is larger than $\sigma_R^2 + I_A$, depth noise can be reduced to by using Eq. (2) and (3):

$$\sigma_Z \propto \frac{\sqrt{\sigma_R^2 + (I_L + I_A)}}{I_L} \propto \frac{1}{\sqrt{I_L}} \propto \frac{1}{\sqrt{P_{LS} \cdot \rho \cdot T_{int}}} \quad (4)$$

while in outdoor case, since I_A is larger than $\sigma_R^2 + I_L$, depth noise can be also reduced to:

$$\sigma_Z \propto \frac{\sqrt{\sigma_R^2 + (I_L + I_A)}}{I_L} \propto \frac{\sqrt{I_A}}{I_L} \quad (5)$$

which means that depth noise is converged to the intensity ratio of ambient light over structured light.

According to Eq. (5), for example, depth noise could be estimated by assuming values like 0.1 for I_L and 10, 22, 75 for I_A, respectively. Results are summarized in Table 2 and indicated by bold x in Fig. 9.

at 1000mm	measured	calculated	factor
10K-lux	4.4	4.4	1
22K-lux	7.6	6.5	1.48
75K-lux	11.4	11.9	2.73

Table 2: Weighting factors are multiplied to measured values.

When we measure depth noise as a function of exposure time at the given optical power and distance, it shows the effect of ambient sunlight more clearly as shown in Fig. 10. For indoor measurements, both 850nm and 940nm systems have enough optical power in the range of 0.2~1.6ms exposure time, so that depth noise is already entered into stable state with 1.5W and 3.0W power. Eq. (4) explains this phenomenon where it decreases rapidly in short exposure time and does not decrease anymore after threshold. In case of 940nm outdoor measurements, depth noise increases due to strong ambient sunlight. It increases further if we reduce structured light power from 3.0W to 1.5W. This means that required optical power for outdoor operation has a specific criteria. For example, if we limit the depth noise less than 5mm, we should increase the operating power up to approximately 2.5W.
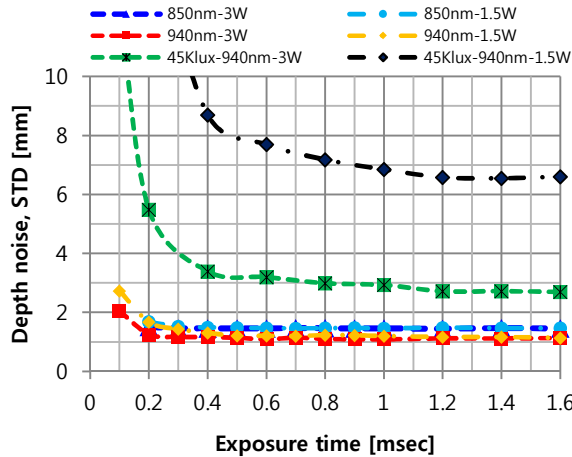
Figure 10: Saturation of standard deviation for different ambient light conditions at 600mm distance.

Depth error is estimated by averaging 50% depth data out of total field-of-view with reference to ground-truth distance of a flat wall. In outdoor case, it was difficult to place POC device at a precise distance, there happen some fluctuations of depth error plot in Fig. 11. Depth error does not seem to increase as the ambient light grows brighter. Typically, depth error can be kept within 1% of distance after good calibration.

We captured point clouds of human faces outdoors as shown in Fig. 12. As ambient light gets strong, the number of point clouds were reduced, however, it still remains in good face geometry. When we make a 3D model by using the point clouds of Fig. 12, photo-realistic result can be obtained after proper post-processing, such as noise removal, hole-filling and color registration. 3D contents generation including 3D gallery, 3D printing, and 3D display would be popular application fields, enabled by this technology. Some indoor scanning examples of 3D reconstruction taken by our POC device are shown at Fig.
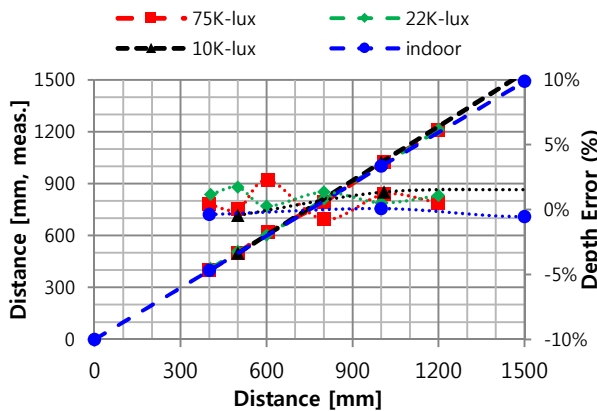

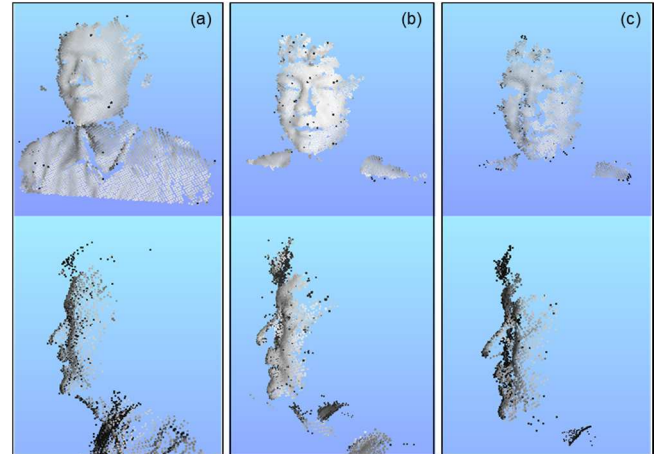Figure 11: Depth error plot for different distances. Dots mean depth error in percentage.

Figure 12: Point clouds at different ambient light conditions. (a) Indoors, (b) 10K-lux, (c) 60K-lux. All images are captured at 400mm distance.

Figure 13: Examples of 3D reconstruction of indoor objects such as office interior, desk and wall curtains.

13. It shows a nice color-to-depth matching result. However, for outdoor scanning of real world other than human faces, relative sensitivity should be increased further, as well as system optimization is needed.

5. Conclusion

In this paper, we implemented S/L type Active Depth Camera in the mobile phone and evaluated its performance at actual outdoor environment. In suppressing strong ambient sunlight, Spectro-Temporal Optical Filtering revealed a good approach to obtain higher SNR and better depth quality.

For further improvements, efficient IR image sensor and highly efficient VCSEL for 940nm need to be developed at the hardware side, as well as the reduction of depth decoding power utilizing DSP core at the software side.

Perception of environment by egocentric camera would be enhanced with the help of Active Depth Camera. Therefore, our efforts to implement depth sensor into the mobile phone and optimize outdoor performance would have meaningful impacts that will bring synergy among all

kinds of environment capturing devices including drones, assistive robots, wearable devices and home appliances.

References

[1] Google Project Tango, 2014.

[2] Intel RealSense Reference Phone, Intel's Developer Forum, 2015.

[3] Lenovo Phab 2 Pro, 2016, http://shop.lenovo.com/us/en/tango/

[4] Asus Zenfone AR, 2017, https://www.asus.com/Phone/ZenFone- AR-ZS571KL/

[5] Apple iPhone8 rumors, 2017, http://www.businessinsider.com/apple-iphone-8-features-facial-recognition-laser-sensor-2017-1

[6] M. Gupta, Q. Yin, and S. K. Nayar. Structured light in sunlight, IEEE ICCV, 2013.

[7] N. Matsuda, O. Cossairt, and M. Gupta. MC3D: Motion Contrast 3D Scanning, IEEE ICCP, 2015.

[8] D. Padilla, P. Davidson, J. Carlson, and D. Noyick, Advancements in sensing and perception using structured lighting techniques: An ldrd final report, Sandia National Lab Report, 2005.

[9] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan, Structured Light 3D Scanning in the Presence of Global Illumination, In CVPR, 2011.

[10] C. Mertz, S. J. Koppal, S. Sia, and S. Narasimhan, A low-power structured light sensor for outdoor scene reconstruction and dominant material identification, Proc. IEEE PRO-CAMS, 2012.

[11] C. Hill and R. L. Jones, Absorption of solar radiation by water vapor in clear and cloudy skies: Implications for anomalous absorption, J. Geophys. Research. 105:9421~9428, 2000.

[12] M. Ihama, et. al., Organic CMOS Image Sensor with Thin Panchromatic Organic Photoelectric Conversion Layer : Durability and Performance, ISSCC, 2017.

[13] K. Khoshelham and S. O. Elberink, Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications, Sensors, 12:1437~1454, 2012.

[14] R. Lange, 3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD -Technology, In Ph.D. thesis, 2000.