Efficient and Fast Implementation of Embedded Time-of-Flight Ranging System Based on FPGAs

Zhou, Weiguo; Lyu, Congyi; Jiang, Xin; Zhou, Weihua; Li, Peng ; Chen, Haoyao; Zhang, Tongtong; Liu, Yun Hui

Published in:
IEEE Sensors Journal

DOI (link to publication from Publisher):
10.1109/JSEN.2017.2728724

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Efficient and Fast Implementation of Embedded Time-of-Flight Ranging System Based on FPGAs

Weiguo Zhou, *Student Member, IEEE*, Congyi Lyu, Xin Jiang, *Member, IEEE*, Weihua Zhou, Peng Li, Haoyao Chen, *Member, IEEE*, Tongtong Zhang, and Yun-Hui Liu, *Fellow, IEEE*

***Abstract*— Time-of-flight cameras perceive depth information about the surrounding environment with an amplitude-modulated near-infrared light source. The distance between the sensor and objects is calculated through measuring the time the light needs to travel. To be used in fast and embedded applications, such as 3-D reconstruction, visual SLAM, human-robot interactions, and object detection, the 3-D imaging must be performed at high frame rates and accuracy. Thus, this paper presents a real-time field programmable gate arrays platform that calculates the phase shift and then the distance. Experimental results shown that the platform can acquire ranging images at the maximum frame rate of 131 fps with a fine measurement precision (appropriately 5.1 mm range error at 1.2 m distance with the proper integration time). Low resource utilization and power consumption of the proposed system make it very suitable for embedded applications.**

***Index Terms*— Depth information, near-infrared (NIR), field programmable gate arrays (FPGA), time-of-flight.**

I. Introduction

DEPTH sensors [1]–[3] become increasingly popular because they are getting smaller and cheaper. They are suitable for applications such as augmented reality, gesture and action recognition in human-computer interaction (HCI) applications [4], SLAM [5] for autonomous vehicles and navigation, 3D scanning of human body [6], etc.

Table.I shows the comparisons among existing 3D imaging technologies (structured light [8], stereo vision [9], Time-of-Flight [10]–[14]). The summary is achieved in the aspects of resolution, frame rate, depth range, active illumination and extrinsic calibration. It can be concluded that Time-of-Flight principle cameras show great potential in size, power consumption and frame size. For example, Time-of-Flight technology has the following superior properties:

- No mobile parts are needed;
- Image acquisition can be achieved at high frame rate;
- It enables for small, compact and light weight design;
- Active illumination (near-infrared light) is used;
- Range images and intensity information are captured simultaneously.

Time-of-Flight camera is a powerful sensor that acquires 3D range images at a high frame rate. It can also provide gray images and range images at the same time. However, the resolution of current Time-of-Flight cameras are still low (e.g., 176×144 for Swiss Ranger SR3000 and SR4000 cameras, 204×204 for PMD CamCube camera, and 320×240 for state-of-the-art ESPROS corporation epc660 camera). In addition, with the improvement of the new semiconductor technology, the resolution and noise reduction of the sensor are being improved gradually.

A number of researchers have shown promising results to utilize the Time-of-Flight principle cameras in a variety of applications [15]–[18]. Calibration between color and depth camera pair is achieved to make the depth measurement more accurate [19]. Chip layout of Time-of-Flight camera is simulated for better performance [20]. Errors of Time-of-Flight ranging cameras and corresponding minimum methonds are also discussed [21], [22]. Online improvement of Time-of-Flight camsera accuracy by automatic integration time adoption was also investigated [23].

Due to large bandwidth of data stream to be processed, it is challenging for embedded processors to handle image data of Time-of-Flight cameras in real-time when they are applied to embedded systems. For example, ARM-based platform can only work at around ten frames per second in DME635-Evalkit (ESPROS Photonics Corporation). Thanks to its parallel computational ability [24], FPGA platform has excellent performance compared with other embedded platforms. A real-time FPGA platform for ToF ranging needs to acquire raw images, determine phase shift, and calculate the distance between the objects and sensor [25]. High-quality depth superresolution for Time-of-Flight cameras was proposed [26]. Our method utilizes the relations between different phase offsets observed at multiple time slots in Time-of-Flight sensor. A method has been developed to detect motion blur regions and accurately eliminate them with the minimal computational cost in real time [27].

Manuscript received June 19, 2017; revised July 12, 2017; accepted July 12, 2017. Date of publication July 18, 2017; date of current version August 22, 2017. This work was supported in part by the Shenzhen Peacock Plan Team under Grant KQTD20140630150243062, in part by the Shenzhen Fundamental Research under Grant JCYJ20140417172417120 and Grant JCYJ20140417172417145, in part by the Shenzhen Key Laboratory under Grant ZDSYS20140508161825065, in part by Hong Kong RGC under Project 414813, and in part by the CUHK VC Discretionary Fund under Project 4930725 and Project 4930745. The associate editor coordinating the review of this paper and approving it for publication was Prof. Kazuaki Sawada. (*Corresponding author: Yun-Hui Liu.*)

W. Zhou, C. Lyu, X. Jiang, P. Li, H. Chen, and T. Zhang are with the School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen 518055, China (e-mail: weiguochow@gmail.com).

W. Zhou is with the Department of Energy Technology, Aalborg University, 9220 Aalborg, Denmark.

Y.-H. Liu is with the State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150001, China, and also with the Department of Mechanical Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: yhliu@mae.cuhk.edu.hk).

Digital Object Identifier 10.1109/JSEN.2017.2728724

TABLE I
COMPARISON OF DIFFERENT RANGE IMAGING TECHNIQUES [7]

Methods	Correspondence	Active Illumination	Resolution	Frame Rate	Depth Range
Stereo Vision	Yes	No	High resolution. System dependent	Typically 30fps. System dependent	Base-line dependent
Structured Light	Yes	Yes	High resolution. System dependent	Typically 30fps. System dependent	Light-power dependent
Time-of-Flight	No	Yes	Up to 320×240	Up to 135fps at full resolution	0.3 m to 6.25 m Within precision range

The main contributions of this paper are listed as follows:

- We proposed a novel efficient hardware/software co-processing framework for 3D Time-of-Flight range imaging system. In this platform, software (MicroBlaze) is in charge of the initialization and chip configuration, hardware is in charge of the frame data cache, distance calculation, and the final pixel display;
- We carried out detailed analysis of the system performance such as the resource utilization, frame rate and distance precision. To the best of our knowledge, in terms of FPGA platform, there are just some theoretical analysis or simulation results of measurement precision in previous researches. The prototype is the start-of-art and highest fame rate (131 frames per second) Time-of-Flight cameras based on FPGAs. Experimental analysis of distance measurement precision (appropriately 5.1 mm range error at 1.2 m distance) is given;
- An automated integrated time adjusted algorithm according to amplitude was proposed. The value of amplitude of the whole frame reflects the precision of depth information, and our method is to adjust the integrated time to obtain the minimum value of amplitude.

The rest of the paper is organized as follows. Section II explains the basic principle of Time-of-Flight. Section III illustrates the architecture of the proposed hardware implementation including on-chip buffer, hardware shift phase determination, image preprocessing, USB3.0 transmission module, illumination and temperature compensation, calibration and assessment of precision measurement. Section IV analyses the experimental platform and the results obtained by the camera. Finally, conclusions are made.

II. TIME-OF-FLIGHT RANGING PRINCIPLE

Time-of-Flight cameras measure the distance of objects by determining the time modulated light needs to travel between the light source and the objects as illustrated in Fig.1. A near-infrared light modulated in the range of 10-100 MHz is used in most of Time-of-Flight measurement applications.

The transmission time is determined by measuring the phase shift between the emitter and received light with a known frequency. The phase shift, φ, and amplitude, A, illustrated in Fig.1 can be obtained by the Fourier series of the sample

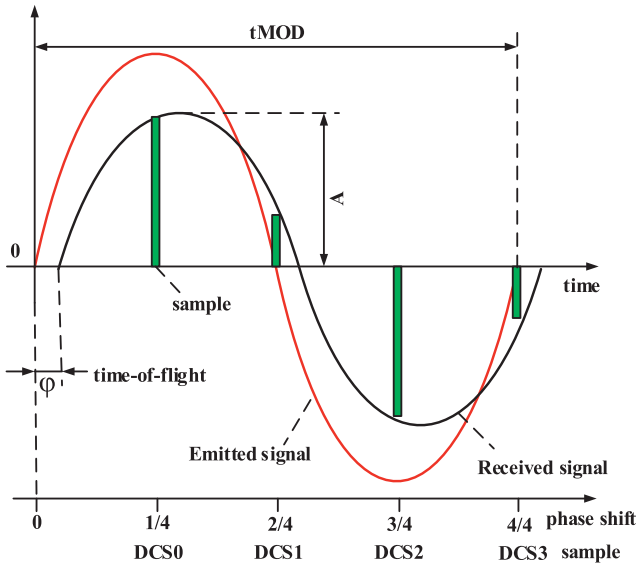


Fig. 1. Time-of-Flight distance measurement using phase offset.

images by

$$\varphi = \tan^{-1}\left(\frac{\sum_{i=0}^{N-1} I_i \sin(2\pi/N)}{\sum_{i=0}^{N-1} I_i \cos(2\pi/N)}\right) \tag{1}$$

$$A = \frac{2}{N}\sqrt{\left[\sum_{i=0}^{N-1}(I_i \cos(2\pi i/N))\right]^2 + \left[\sum_{i=0}^{N-1}(I_i \sin(2\pi i/N))\right]^2} \tag{2}$$

where I_i is the sample images within one period (interval between each sample is $2\pi/N$ rad), and N is the number of sample frames per period. In most cases, N is set to 4, then the time can be obtained by

$$t = \frac{1}{2\pi f_{LED}}[\pi + atan2(DCS2 - DCS0, DCS3 - DCS1)] \tag{3}$$

where f_{LED} is the frequency of modulated light. DCS0, DCS1, DCS2 and DCS3 are the Differential Correction Samples. Function $atan2$ is used to calculate the shift phase in the range of $-\pi \ldots + \pi$. In our case, we use the range from angle 0^0 to 360^0 which corresponds to the distance from 0 m up to the unambiguity distance.

Thus, distance can be calculated according to the phase by

$$d = \frac{c}{2f}\left(\frac{\varphi}{2\pi} + k\right) \tag{4}$$

where f is the modulated frequency of emitted light, c is the speed of the light, and k is an integer that denotes the potential wrapping of phase and is assumed to be 0 typically.

Then, the maximum unambiguous range of the system is

$$d_u = \frac{c}{2f}. \tag{5}$$

From the equation, value of the maximum distance d_u can be increased by decreasing frequency f, but the distance measurement resolution will be weakened at the same time. Two different modulation frequencies are taken in order to extend the maximum unambiguous range. The measurement method acquired a number of possible object locations using different modulation frequency, and offset by integer multiples of unambiguous distance d_u. Then, the ground truth object location is determined when the measurement are mostly in agreement. In the end, the object distance can be calculated as

$$d = \frac{c}{2f_A}\left(n_A + \frac{\varphi_A}{s}\right) = \frac{c}{2f_B}\left(n_B + \frac{\varphi_B}{s}\right) \tag{6}$$

where f_A and f_B are two co-prime integers that represent different frequencies of modulated lights and the ratio between them can be denoted by coprime integers M_A, M_B, namely, $f_A : f_B = M_A : M_B$. φ_A and φ_B denote the fractional remainder after the phase has wrapped around n_A or n_B ($n_A, n_B \in N$) times and s is the maximum range of the integer output (where $s = 2^{Br}$). Then, distance can be acquired using (7) upon minimizing the difference between the left and right side of (6).

$$\begin{aligned} y(n_A, n_B) &= |M_B(sn_A + \varphi_A) - M_A(sn_B + \varphi_B)| \\ &= |(M_B\varphi_A - M_A\varphi_B) + s(M_Bn_A - M_An_B)|. \end{aligned} \tag{7}$$

A simple method is to calculate all possible combinations of the integers n_A or n_B and select the suitable pair that gives the smallest value of y. Finally, the value of d can be determined.

III. Hardware Design of the Overall System

This section presents the hardware architecture of the overall range imaging system illustrated in Fig.2, which is composed of five components: image acquisition, hardware phase shift computation, image processing, USB3.0 transmission, HDMI (High-Definition Multimedia Interface) display. The illumination part emits the modulated near-infrared light. The epc660 sensor chip receives the reflected light from the surface of objects. The main control part, a Xilinx Kintex-7 series FPGA chip, in our design is in charge of the configuration of the chip and capturing the pixel data of the sensor, processing the image data and then transferring the processed data to the host system (PC) via the high-speed USB3.0 32bits parallel interface.

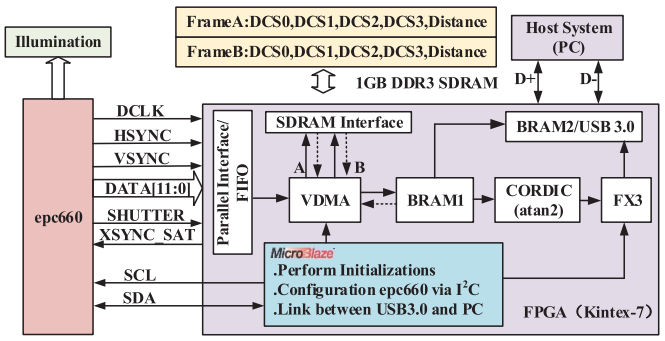


Fig. 2. General FPGA architecture proposed in this work.

A. Sensor Chip Configuration

The ecp660 (ESPROS Photonics Corporation) chip is an integrated Time-of-Flight sensor which consists of a Charge-coupled Device (CCD) pixel field and complete control logic. The chip can achieve a distance resolution in millimetre with the measurement up to 100 meters, 131 frames per second at its full frame size (320×240). Furthermore, it can boost up to more than 1000 frames per second in advanced mode (Binning and ROI mode).

There are two kinds of interfaces used. One is I^2C serial which is used for mode selection, configuration and temperature reading, and the other is high-speed TCMI interface which is adapted to transfer the frame pixel data to processing unit (FPGA in our design). Due to the powerful processing ability of the FPGA, pixel value of the whole frame can be calculated simultaneously with the best exposure that requires the most suitable parameters for better distance calculation. The main advantage of such a solution is that it can acquire the DCS (Differential Correction Samples) frames at equidistant time in a multi-integration-time (multi-exposure) mode to obtain a higher precision range image.

B. High Speed Frame Data Buffer

The time-shared application frame buffer FPGA architecture depicted in Fig.2 is employed in our design. The data stream is blocked into the FPGA controller as long as epc660 starts the transfer burst. After the burst transmission was completed, the internal buffer is released by DMA controller and the system prepares for another burst transmission. Therefore, the data bus ratio between epc660 and FPGA must be carefully designed within such hardware architecture.

The bus ratio can be set by modifying different TCMI DCLK frequency (24, 48 or 96 MHz) operating on the epc660 chip. The burst can transfer faster, so the TCMI clock works at a higher frequency. In order to increase the transmission speed, a wider or faster SDRAM can be used. The FPGA has a 128-bits wide SDRAM bus running at 160 MHz clock, and the epc660 TCMI DCLK is set to 40 MHz. Then, 4× speed improvement from inner buffer of FPGA to external SDRAM can be performed. High speed external SDRAM can not only improve the processing speed, but also save the FPGA resources such as LUTs, FFs, RAMs, etc.

Algorithm 1: Process of Time-of-Flight Calculation

1: $d_u = \frac{c}{2f_{Mod}}$
2: $Im = DCS2 - DCS0$
3: $Re = DCS3 - DCS1$
4: $\varphi = \arctan(Im, Re)\ (-\pi, \pi]$
5: **if** $\varphi < 0$ **then**
6: $\quad \varphi = \varphi + 2\pi$
7: **end if**
8: $d = \varphi \frac{d_u}{2\pi}$
9: $d = d + d_{offset} + d_{temperature}$
10: $d = d - \left|\frac{d}{d_u}\right| \cdot d_u$
11: $A = \frac{\sqrt{Re^2+Im^2}}{2}$, Gray-scale amplitude
12: $d_x = d.vector_x$, 3D Point Cloud
13: $d_y = d.vector_y$, 3D Point Cloud
14: $d_z = d.vector_z$, 3D Point Cloud
15: Image Processing Algorithm Implementation
16: User Interface Application Algorithm

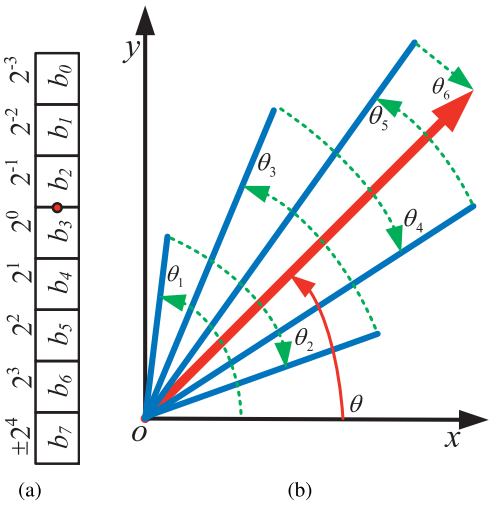


Fig. 3. (a) Fixed-point representation of number with 2^{-3} resolution. The number can be denoted at U8.3 or S8.3 formate. It depends on whether the most significant bit of the number is positive or negative. (b) CORDIC algorithm represented by angle rotation, and the resultant $\theta = \sum_{i=1}^{n} \theta_i$, iterate until θ_i is smaller than the predefined threshold.

C. Hardware Shift Phase Computation

Algorithm 1 depicts the processing flow to implement the designed system. The algorithm needs to be executed for every pixel of the image. From this, some math operations such as add, sub, milt, div, atan, square, variance, average, which are not easy for FPGA to implement. As seen from equations in section II, real numbers are calculated during the computation process. In hardware, the processing of real numbers is solved by scaling the integer by a fraction, because the hardware more suitable for fixed-point number operations. In that case, the number is represented by negative power of two, for example, 2^{-k}. In this method, the number to be processed can be represented by this equation $B_{FP} = B \times 2^{-k} = \pm b_{N-1}2^{N-1-k} + \sum_{i=0}^{N-2} b_i 2^{i-k}$, where the integer can be either unsigned or signed (usually two's complement) which is shown in Fig.3(a). Real numbers can be represented approximately with a fixed-point representation in Verilog format.

For the location of binary point is fixed, the method can be more simply by the representation of fixed-point format. Thus, all arithmetic operations are the same as for integers, apart from aligning the binary point for addition and subtraction. This corresponds to an arithmetic shift by a fixed number of bits because the location of the binary point is fixed, which is effectively free in hardware.

During the computation of the hardware phase computation, the computation cost of trigonometric sines and cosines is high and complicated due to their nonlinear functions and multipliers. There are a lot of methods for calculating the formula in hardware efficiently. The simplest method is to adopt a look-up table indexed n and store the results of trigonometric sines and cosines. Although many efficient methods are proposed, FPGA resource in terms of LUTs, RAMs and other logic resources are still very costly to carry out these operations.

Coordinate rotation digital computer (CORDIC), an efficient hardware implemented algorithm, is employed to calculate the hardware phase. Only addition, subtraction, and bit-shift operations are required to perform this algorithm. As illustrated in Fig. 3(b), the computation operations can be replaced by angle rotation which is quite suitable for hardware implementation. It is an iterative technique for calculation, and the iteration is operated by rotating a vector (x, y) by angle θ_k:

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} \cos\theta_k & -\sin\theta_k \\ \sin\theta_k & \cos\theta_k \end{bmatrix} \begin{bmatrix} x_k \\ y_k \end{bmatrix}. \tag{8}$$

and the right side can be rewritten as:

$$\begin{aligned} \begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} &= \frac{1}{\cos\theta_k} \begin{bmatrix} 1 & -d_k \tan\theta_k \\ d_k \tan\theta_k & 1 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \end{bmatrix} \\ &= \sqrt{1+2^{-2k}} \begin{bmatrix} 1 & -d_k 2^{-k} \\ d_k 2^{-k} & 1 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \end{bmatrix} \end{aligned} \tag{9}$$

where d_k is the direction of rotation, representing the anticlockwise and clockwise directions by 1 and -1, respectively. Angle can be calculated by $\tan\theta_k = 2^{-k}$. And then the function $atan2$ is worked out during the calculation of hardware shift phase which maps the distance between the sensor and object to angle [-π,π]. In our case, the result phase was represented in 24 bits fixed format to achieve an appreciate compromise between the accuracy and resource utilization.

D. Image Preprocessing (Parallel Processing)

In order to remove the random noises of the images, some low-level image processing algorithms such as morphology dilation, morphology erosion, median filter are implemented.

Without caching, nine pixels must be read for each window position (each clock cycle for stream processing), and each pixel must be read nine times as the window is scanned through the image. The mostly used form of caching is row buffering. Window filter can be considered as an operation of the nine pixel within the window. In that method, previous two rows pixel data need to be cached, and then the same pixel does not need to be read for the second time.

1) Morphological Filter: The basic operations of morphological filter are erosion and dilation. For erosion, an object pixel is remained only if the structuring element fits completely within the object. Considering the structuring element as a window, the output is considered an object pixel only if all of the inputs are one. Erosion is therefore a logical AND of the pixels within the window: $A \odot B = \{x|B(x) \subset A\}$

Because the object size becomes smaller as a result of the processing, it is called erosion operation. The operation can be represented by $p1 = p11\&p12\&p13, p2 = p21\&p22\&p23, p3 = p31\&p32\&p33, p = p1\&p2\&p3$ which takes advantage of hardware parallel processing ability.

As for dilation, each input pixel is replaced by the shape of the structuring element within the output image. This is equivalent to outputting an object pixel if the flipped structuring element hits an object pixel in the input. In other words, the output is considered as an object pixel if any of the inputs within the flipped window is a one, that is dilation is a logical OR of the flipped window pixels: $A \oplus B = \{x|B(x) \cap A \neq \emptyset\}$. Similarly for the erosion operation, dilation can also be represented by $p1 = p11|p12|p13, p2 = p21|p22|p23, p3 = p31|p32|p33, p = p1|p2|p3$ to improve the processing speed.

2) Median Filter: Median filter is used to remove the random noises on the image which satisfies the following equation. $f(x, y) = Median\{g(x-i, y-j)\}(i, j) \in S$, where $g(x, y)$ is the gray value of the current pixel, $f(x, y)$ is the post-processed gray image value, S is the slice template and i, j are the horizontal and vertical size of the template.

E. High Speed USB3.0 Transmission

Cypress EZ-USB FX3 (USB3.0) was adopted to transfer the obtained depth images to a host system (PC) for further applications. FPGA is in charge of sending the clock, control, data signals to SlaveFifo of FX3. Then, FX3 products socket buffer transfer to the host system. The transmission speed of the system can be reached at 359 MB/s which is adequate for the bandwidth of the range imaging system.

F. Compensation (Temperature & Light Illumination) and Calibration

Time-of-Flight sensor chip can separate self-emitted and reflected modulated light for the background light. Illumination influence and temperature compensation could affect the accuracy of the imaging system. There are four temperature sensors located in the four corners of the chip. Temperature of sensors located on the chip can be read via I^2C interface. Value of the sensor temperature can be used to compensate the temperature effect. The calibration is done in front of a white large flat wall, make sure the image of wall can cover the field of the view include the vertical and horizontal direction. and set a appropriate distance, e.g., 1.2 meters. Warm-up the sensor chip at least 15 minutes. Then, the offset value parameters D_{offset} can be determined. Then, adjust the measurement distance to the setted value, and then new calibration measurement are obtained.

G. Assessment of Measurement Precision and Auto Integration Time Adjusted Method

The better is the reflected light signal, the more precise the distance measurement. Then the quality indicator for measurement distance is the peak-to-peak amplitude value of the modulated light. Amplitude value can be achieved by (2) in which N is set to 4, that is depicted in **Algorithm 1**.

Then according to the amplitude computation, integration time can be adjusted for the next measurement recycle. The value of amplitude of the whole image obtained by the proposed system is utilized to assess the quality of the measurement system. Based on experimental results, amplitude range between 100 LSB and 1200 LSB (Least Significant Bit) is regarded as good signal strength. Less than the range will be regarded as weak illumination, otherwise will be regarded as overexposure. The integration time is directly proportional to amplitude, so it is can be automatically adjusted to obtain a more precise depth image.

IV. Experiment

A. System Setup

The hardware of the overall Time-of-Flight camera system which consists of four parts is illustrated in Fig.4. The first part is a near-infrared led illumination board which emits near-infrared light. The second part is the Time-of-Flight sensor chip which detects the reflected light from the object. The third part is the main control and processing board, FPGA board, which is in charge of the configuration, control, processing and display tasks. The bottom part is the high speed USB3.0 transmission board which transfers the post-processed frame data to PC. In the system, the Time-of-Flight sensor module and USB3.0 module are both connected with FPGA via FPGA Mezzanine Card (FMC).

In order to measure the precision of the designed camera, measurement platform shown in Fig. 4(b) is set up. Expect the setup in Fig.4, some supernumerary instruments are used. Newport Corporation's Vision IsoStation is employed to guarantee the system to be more steady. GCM-830304M optical digital translation stage (size is 120×120, range is 150 mm, minimum resolution is 0.01 mm) is used to move the object and display the offset on the LCD screen. Besides, the absolute distance between the sensor and the object can be obtained by the laser range finder. Before the measurement operation, the calibration is done in front of a large white wall at a certain distance (1.2 m in this case).

B. Results and Discussion

1) Resources and Power Consumption Analysis: Fig.6(a) illustrates the utilization of the hardware resource. Even for the largest consumption resource, i.e. the block random-memory (BRAM), its consumption only accounts for 31.51% of the whole chip. The whole power consumption which contains the dynamic and static parts is only 0.625 W. The compact and low-power consumption design makes the system suitable in a variety of applications such as micro unmanned aerial vehicle, light-weight Automatic Guided Vehicle (AGV), portable wearable device, etc.

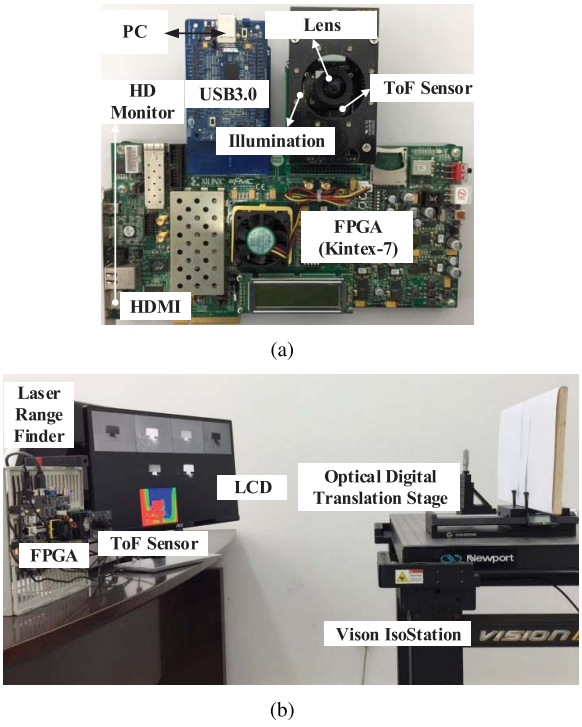


Fig. 4. (a) Real-time Time-of-Flight range imaging system based on FPGA and USB3.0 designed in our work, include led illumination module, epc660 sensor chip module, FPGA module and USB3.0 module. (b) Experimental setup of the proposed system which constitutes of four part. Time-of-Flight sensor, FPGA board, LCD screen and optical digital translation stage.

2) Depth Image Analysis: Fig.5 shows the remarkable images obtained by the designed system. In order to observe the range data more clearly, we display the value as the pixel data, namely, $PixelData = d/d_u \cdot 2^{24}$, where d_u is the unambiguous range (6.25 m in this figure), and d is the measurement distance. In the right side of the range image is the legend whose pixel data represent the corresponding depth. The first four frames represent the Differential Correction Samples (DCS) during one period corresponding to the four sample images illustrated in Fig. 5(a)-5(d). The resultant range image processed and calculated from the system is illustrated in Fig. 5(e). That is a false-color image that displays from red to green and blue according to the range.

Real-time performance compared with different processing platform, ARM, PC and FPGA are shown in Tab. II. Kinect V2.0 is the most famous commercial RGB-Depth ranging system which can only acquire 30 depth images per second. For embedded system case, to the best of our knowledge, the start-of-art ARM-based platform (ESPROS Photonic Corporation) can only obtain 15 frames per second. Ultrafast performance of the sensor (131 fps) can be implemented using the FPGA platform designed in this paper which showed improvement in real time performance.

3) Bandwidth Analysis: For the sake of testing the maximum performance of the proposed system, utmost performance of epc660 is set via I^2C serial bus, that is, 131 fps at full (320×240) resolution. In that case, the bandwidth is 320×240×131×24 bit/s = 241,459,200 bit/s ≈ 241 Mbps ≈ 30 MB/s. High speed transmission system (upper limited speed is 359 MB/s) is still enough to handle this task.

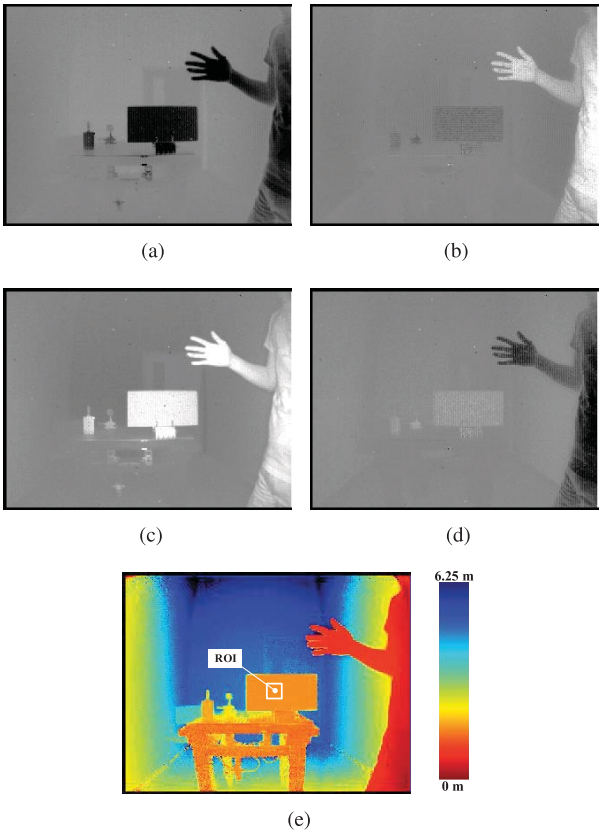


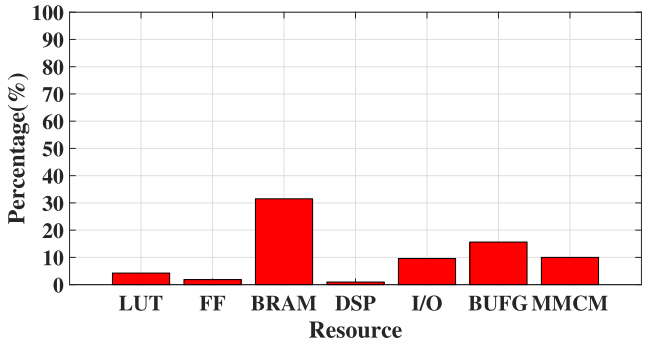
Fig. 5. Amplitude DCS and ranging images obtained by the real-time FPGA-based Time-of-Flight camera, where DCS represents Differential Correction Samples. IR LED modulation frequency (f_{LED}) adopted in this case is 24 MHz, integration time is 2400μs, then the unambiguous distance $d_u = c/(2f_{LED})$ equals to 6.25 m, where $c = 3\times10^8 m/s$ is the speed of light. ROI is set to 16×16 in the center of depth frame. (a) First DCS Amplitude Image. (b) Second DCS Amplitude Image. (c) Third DCS Amplitude Image. (d) Fourth DCS Amplitude Image. (e) Depth-colored Image.

TABLE II
REAL-TIME PERFORMANCE COMPARISON BETWEEN OTHER RELATED PLATFORMS

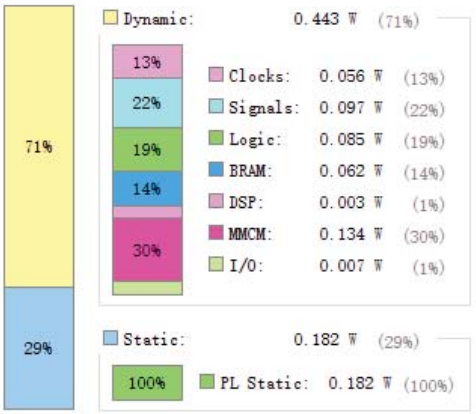
Platform	PC (Kinect V2.0)	Embedded System	
		ARM[1]	**FPGA** (this work)
Frame Rate	Depth Channel,30 fps	15 fps	**131 fps**

4) Measurement Precision Analysis: In order to reduce the effect of distortion and the boundary effect, pixels located in the center of the pixel field by the region of interest (ROI) are sampled, and the size is set to 16×16. The color and surface reflectivity of the object can affect the intensity of reflected light, and hence the measurement distance. Then, make sure that the pixel within in the ROI belongs to the same object in experiment. Moreover, the range error versus the range measurement with different integration time is illustrated in Fig.7. The measurement errors are sampled every 32.5 cm distance, and 20 samples are gathered in the 6.25 meters unambiguous range. The measurement errors versus the different integration time (800μs, 1600μs, 2400μs, 3200μs, 4000μs in this paper) are also investigated. In this case, 2400μs integration time shows a better performance than others, and the minimum

[1]https://www.espros.com/time-of-flight-technology/



(a)



(b)

Fig. 6. (a) Hardware resource utilization of the overall system. (b) On-chip power consumption of the overall system.

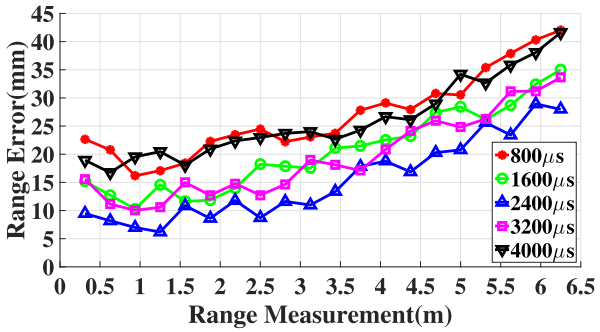


Fig. 7. Absolute range error with respect to integration time is obtained from ROI (size 16×16 in the center of the pixel field) area using range imaging measurement proposed in our work. IR LED modulation frequency (f_{LED}) adopted in this case is 24 MHz, integration time is 1600μs, then the unambiguous distance $d_u = c/(2f_{LED})$ equals to 6.25 m, where $c = 3 \times 10^8 m/s$ is the speed of light.

error is 5.1 mm at 1.2 m distance. Further, the measurement error will increase with the measurement distance increasing. Then an automatic adjusted algorithm of integration time is proposed depending on the amplitude.

V. Conclusion

This paper presents an real-time FPGA-based system for high-speed range imaging using the phase shift for obtaining the distance. The obtained depth images are transmitted to PC via high speed parallel USB3.0 interface for subsequent applications. The camera sensor adapted in our design is epc660 which represents the state-of-the-art performance Time-of-Flight sensor. The phase determination algorithm is implemented on a Xilinx high performance Kintex-7 series FPGA chip. Compensation and image preprocessing algorithms are also implemented on the system so as to obtain a more accurate and robust range image. Cypress's EZ-USB FX3, the latest-generation USB3.0 peripheral controller, is used to connect FPGA and PC. Then, resource utilization and power consumption are discussed. The ranging errors under different integration time are also compared. Experimental results illustrated that the system produces good accuracy and high speed in ranging measurement. produces good accuracy and high speed in ranging measurement. 5.1 mm measurement errors at about 1.2 m distance, and operated at high frame rate as 131 fps with little on-chip resource and power consumption.

Compared with other Time-of-Flight principle camera platforms, e.g. CPU-based, DSP-based and ARM-based, FPGA-based platform has huge prospect in academic, industrial and entertainment applications because of its low power consumption, small size and high frame rate. Time-of-Flight cameras are especially suitable for real-time on-line computation applications in which the sensors require moving. In the future works, we plan to utilize this high-speed embedded platform to implement the vision tasks such as 3D reconstruction, gesture control, object detection, visual SLAM and so on.

References

[1] Z. Xiong, Y. Zhang, F. Wu, and W. Zeng, "Computational depth sensing: Toward high-performance commodity depth cameras," *IEEE Signal Process. Mag.*, vol. 34, no. 3, pp. 55–68, May 2017.

[2] P. G. Gulden, D. Becker, and M. Vossiek, "Novel optical distance sensor based on MSM technology," *IEEE Sensors J.*, vol. 4, no. 5, pp. 612–618, Oct. 2004.

[3] E. Stoykova *et al.*, "3-D time-varying scene capture technologies—A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 11, pp. 1568–1586, Nov. 2007.

[4] V. P. Minotto, C. R. Jung, and B. Lee, "Multimodal multi-channel on-line speaker diarization using sensor fusion through SVM," *IEEE Trans. Multimedia*, vol. 17, no. 10, pp. 1694–1705, Oct. 2015.

[5] C. Forster, Z. Zhang, M. Gassner, M. Werlberger, and D. Scaramuzza, "Svo: Semidirect visual odometry for monocular and multicamera systems," *IEEE Trans. Robot.*, vol. 33, no. 2, pp. 249–265, Apr. 2017.

[6] Z. Liu, J. Huang, S. Bu, J. Han, X. Tang, and X. Li, "Template deformation-based 3-D reconstruction of full human body scans from low-cost depth cameras," *IEEE Trans. Cybern.*, vol. 47, no. 3, pp. 695–708, Mar. 2017.

[7] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (ToF) cameras: A survey," *IEEE Sensors J.*, vol. 11, no. 9, pp. 1917–1926, Sep. 2011.

[8] X. I. Wong and M. Majji, "A structured light system for relative navigation applications," *IEEE Sensors J.*, vol. 16, no. 17, pp. 6662–6679, Sep. 2016.

[9] K. McGuire, G. de Croon, C. D. Wagter, K. Tuyls, and H. Kappen, "Efficient optical flow and stereo vision for velocity estimation and obstacle avoidance on an autonomous pocket drone," *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 1070–1076, Apr. 2017.

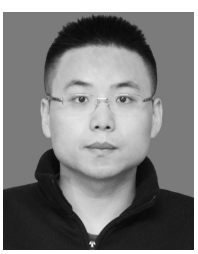
[10] B. Buttgen and P. Seitz, "Robust optical time-of-flight range imaging based on smart pixel structures," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 6, pp. 1512–1525, Jun. 2008.

[11] S. Kawahito, I. A. Halin, T. Ushinaga, T. Sawada, M. Homma, and Y. Maeda, "A CMOS time-of-flight range image sensor with gates-on-field-oxide structure," *IEEE Sensors J.*, vol. 7, no. 12, pp. 1578–1586, Dec. 2007.

[12] S. Foix, G. Alenyà, J. Andrade-Cetto, and C. Torras, "Object modeling using a ToF camera under an uncertainty reduction approach," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2010, pp. 1306–1312.

[13] H. Shim and S. Lee, "Recovering translucent objects using a single time-of-flight depth camera," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 5, pp. 841–854, May 2016.
[14] A. Bhandari and R. Raskar, "Signal processing for time-of-flight imaging sensors: An introduction to inverse problems in computational 3-D imaging," *IEEE Signal Process. Mag.*, vol. 33, no. 5, pp. 45–58, Sep. 2016.
[15] A. Corti, S. Giancola, G. Mainetti, and R. Sala, "A metrological characterization of the Kinect V2 time-of-flight camera," *Robot. Auto. Syst.*, vol. 75, pp. 584–594, Jan. 2016.
[16] K. Kim and H. Shim, "Robust approach to reconstructing transparent objects using a time-of-flight depth camera," *Opt. Exp.*, vol. 25, no. 3, pp. 2666–2676, 2017.
[17] A. Bhandari *et al.*, "Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization," *Opt. Lett.*, vol. 39, no. 6, pp. 1705–1708, Mar. 2014.
[18] D. Piatti and F. Rinaudo, "SR-4000 and camcube3.0 time of flight (ToF) cameras: Tests and comparison," *Remote Sens.*, vol. 4, no. 4, pp. 1069–1089, 2012.
[19] J. Jung, J. Y. Lee, Y. Jeong, and I. S. Kweon, "Time-of-flight sensor calibration for a color and depth camera pair," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 7, pp. 1501–1513, Jul. 2015.
[20] M. Lambers, S. Hoberg, and A. Kolb, "Simulation of time-of-flight sensors for evaluation of chip layout variants," *IEEE Sensors J.*, vol. 15, no. 7, pp. 4019–4026, Jul. 2015.
[21] P. Fürsattel *et al.*, "A comparative error analysis of current time-of-flight sensors," *IEEE Trans. Comput. Imag.*, vol. 2, no. 1, pp. 27–41, Jan. 2016.
[22] A. P. P. Jongenelen, D. G. Bailey, A. D. Payne, A. A. Dorrington, and D. A. Carnegie, "Analysis of errors in ToF range imaging with dual-frequency modulation," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 5, pp. 1861–1868, May 2011.
[23] T. Hoegg, C. Baiz, and A. Kolb, "Online improvement of time-of-flight camera accuracy by automatic integration time adaption," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2015, pp. 613–618.
[24] W. Zhou *et al.*, "Real-time implementation of panoramic mosaic camera based on FPGA," in *Proc. IEEE Int. Conf. Real-Time Comput. Robot. (RCAR)*, Jun. 2016, pp. 204–209.
[25] A. P. 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, 2012.
[26] Y. Kim, B. Ham, C. Oh, and K. Sohn, "Structure selective depth superresolution for RGB-D cameras," *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5227–5238, Nov. 2016.
[27] S. Lee, "Time-of-flight depth camera motion blur detection and deblurring," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 663–666, Jun. 2014.

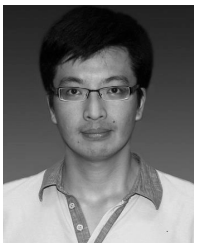
Weiguo Zhou (S'17) received the B.E.E. and M.S.E.E. degrees in mechanical engineering from the Hefei University of Technology and the Harbin Institute of Technology, China, in 2012 and 2014, respectively. He is currently pursuing the Ph.D. degree in mechanical engineering with the Harbin Institute of Technology, China. His research interests include FPGA implementation, 3-D time-of-flight range imaging, human interaction, machine vision, and neural networks.



Congyi Lyu received the B.E. degree in automation from Henan Polytechnic University, Jiaozuo, Henan, China, in 2009 and the master's degree in health technology engineering from Hongkong polytechnic University, Hong Kong, in 2011. He is currently pursuing the Ph.D. degree with the School of Mechatronical Engineering, Beijing Institute of Technology, Beijing, China. His research interests include object tracking, FPGA-based image processing.

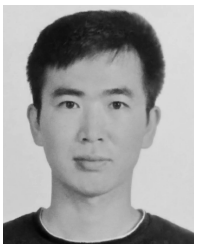
Xin Jiang (S'06–M'07) received the B.Eng. degree in control engineering from the Dalian University of Technology, Dalian, China, in 2000, the M.Eng. degree in mechatronics and precision engineering, and the Ph.D. degree in aerospace engineering from Tohoku University, Sendai, Japan, in 2004 and 2007, respectively. He was with the Department of Mechanical Engineering, Tohoku University, Sendai, Japan, from 2007 to 2015, as an Assistant Professor. Since 2015, he has been with the Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China. He is currently an Associate Professor with the School of Mechanical Engineering and Automation. His research interests are flexible manipulator, deformable object manipulation, service robotics.

Weihua Zhou received the B.Eng. and M.Eng. degrees in electrical engineering and its automation, electric machines and electric apparatus from Northwestern Polytecnical University, Xi'an, China, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with the Department of Energy Technology, Aalborg University, Aalborg, Denmark. He is working on the advanced coordinative control of the wind power conversion system.



Peng Li received the B.Eng. degree in mechanical engineering from North Eastern University, Shenyang, China, in 2004, and the Ph.D. degree in mechatronics from the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, in 2010. From 2009, he was with the Department of Mechanical and Automation Engineering, Chinese University of Hong Kong, Hong Kong, as a Research Assistant, where, since 2010, he has been a Postdoctoral Fellow and Research Associate. He is currently with the Department of the Mechanical Engineering and Automation, Harbin Institute of Technology, as an Assistant Professor. His research interests include developing assistive surgical robots, medical devices, and biomimetic robots.

Haoyao Chen (M'09) received the bachelor's degrees in mechatronics and automation from the University of Science and Technology of China in 2004, and the Ph.D. degree in the robotics and automation from the University of Science and Technology of China and the City University of Hong Kong in 2009. He is currently an Associate Professor with the Harbin Institute of Technology, Shenzhen Graduate School, and the State Key Laboratory of Robotics and System. His research interests lie in machine vision, multi-robot systems, motion controls and biological processing automation.



Tongtong Zhang received the B.Eng. degree in mechanical engineering and automation from Beihua University, Jilin, China, in 2015. He is currently pursuing the M.Eng. degree in mechatronic engineering with the Harbin Institute of Technology, China. His research interests include machine vision, pattern analysis and FPGA based image processing.

Yun-Hui Liu (S'90–M'92–SM'98–F'09) received the B.Eng. degree in applied dynamics from the Beijing Institute of Technology, Beijing, China, in 1985, the M.Eng. degree in mechanical engineering from Osaka University, Osaka, Japan, in 1989, and the Ph.D. degree in mathematical engineering and information physics from the University of Tokyo, Tokyo, Japan, in 1992.

He was with the Electrotechnical Laboratory, Ministry of International Trade and Industry, Ibaraki, Japan, from 1992 to 1995. Since 1995, he has been with The Chinese University of Hong Kong (CUHK), Shatin, Hong Kong, where he is currently a Professor with the Department of Mechanical and Automation Engineering and the Director of the CUHK T Stone Robotics Institute. He is also visiting state Key Laboratory of Robotics Technology and System, Harbin Institute of Technology, Harbin, China, and the Director of Joint Center for Intelligent Sensing and Systems, National University of Defense Technology, Hunan, China, and CUHK. He has published more than 200 papers in refereed journals and refereed conference proceedings and was listed in the Highly Cited Authors (Engineering) by Thomson Reuters in 2013. His research interests include visual servoing, medical robotics, multifingered robot hands, mobile robots, sensor networks, and machine intelligence.

Dr. Liu received numerous research awards from international journals and international conferences in robotics and automation and government agencies. He is the Editor-in-Chief of *Robotics and Biomimetics* and an Editor of *Advanced Robotics*. He served as an Associate Editor of IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION and the General Chair of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems.