■ 926

FPGA Based Pattern Generation and Synchonization for High Speed Structured Light 3D Camera

Muhammad Atif*, Sukhan Lee
Intelligent System Research Institute (ISRI), College of Information and Communication Engineering,
Sungkyunkwan University, 2066, Seobu-ro, Jangan-gu, Suwon-si, South Korea.
Corresponding author, e-mail: m.atif@skku.edu*, lsh1@skku.edu

Abstract
Recently, structured light 3D imaging devices have gained a keen attention due to their potential applications to robotics, industrial manufacturing and medical imaging. Most of these applications require high 3D precision yet high speed in image capturing for hard and/or soft real time environments. This paper presents a method of high speed image capturing for structured light 3D imaging sensors with FPGA based structured light pattern generation and projector-camera synchronization. Suggested setup reduces the time for pattern projection and camera triggering to 16msec from 100msec that should be required by conventional methods.

Keywords: FPGA, structred Light 3D camera, HOC

1. Introduction
Structured light 3D acquisition method is one of the renowned methods for high precision 3D measurements. Basic working principle of the structured light system is similar to stereo matching. In stereo matching, we need two passive cameras to grab frame but in structured light system we use only one camera and one projector system to generate 3D depth information. For 3D reconstruction we use projectors that can project patterns on screen, and camera is required to capture those known projected patterns. These known projected patterns are used to find correspondence of common point in projected frame and captured fame, to generate 3D depth information. Basic working mechanism of structured light 3D camera is shown in Figure 1.
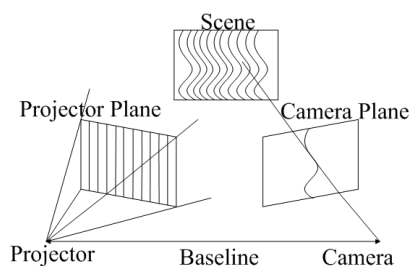

Figure 1. Structured Light 3D camera setup

Distance between the projector and camera is known as baseline shown in Figure 1, triangulation technique is used to compute depth of scene by using intrinsic and extrinsic parameters of the projector and camera which are obtained by calibration. Structured light depth imaging is gaining attention due to its application in diverse areas e.g. robotics, 3D games, micro inspection systems, 3D finger printing and re-engineering. Especially in the field of robotics, structured light depth imaging system replaced traditional passive stereo systems because of its high accuracy, which is used for object recognition and manipulation.

Potential of structured light system lies in it's capability of capturing 3D point cloud of texture-less objects, and providing higher precision and robustness as compared to passive

stereo camera system. Accuracy of depth measurement in structured light system depends upon correct pixel correspondence between the camera and digital mirror device (DMD) of projector. Another feature of structured light system is speed of capturing 3D point cloud. To make system real time it is required to project patterns fast and capture them accurately, this is not possible by using PC due to limited control on display interface of computer. In this paper, we have suggested a technique to project hierarchical orthogonal coded (HOC) [1] patterns, fast by using conventional projector without any modification of hardware of projector, we have achieved scan rate of 3Hz (18 patterns) with more accuracy and precision by generating pixels real time in FPGA. In proposed method we did not use any kind of memory to store patterns, to generate real time pattern pixels.

Y Oike et al [2] presented a real-time 3-D imaging system based on light-section method with VGA pixel resolution. An integrated system controller implemented on FPGA performs sensor control, light projection control, range of data pre-processing and suppression of redundant data transmission for fatal problems of real-time system. S Lee et al [3] suggesteda real-time 3-D camera based on infrared structured light for robots working in home environment. FPGA module can handle high computational cost of HOC based signal separation coding. Moreover, they implemented a compact optic system of the camera to project and receive IR structured light. B. Hong et al [4] proposed a DMD, CMOS sensor and FPGA on a single board with capturing speed of 500 Hz and range information frame rate of 17 Hz. An efficient algorithm for determining phase using the CORDIC function is implemented by A. Peter et al [5] in a time of flight range imaging system. CORDIC arctangent unit for a 3D camera system with four phase shifted patterns implantation of FPGA is used by S. Bellis and W. Marnane [6]. The prototype system described, allows for an external projector to be connected using VGA interface.

Main contributions by our developed system are following:
1. Real time pattern pixels are generated without using memory which will help to reduce size and cost of ASIC.
2. Maximum allowable projection speed has been achieved for conventional projector.
3. Accurate synchronization between projected and captured frame for 3D reconstruction.

Rest of the paper includes, Section-2 about Structured light 3D camera system. System flow is discussed in Section-3 and experimental results are presented in Section-4.

2. Structured Light 3D Ccamera

There are two main components of structred light camera: projector and camera. Projector, projects known patterns on the scene and camera captures those frames to find correspondance between DMD pixel to camera pixels. We will discuss about structred light 3D camera in detail in later parts of this section.

2.1 Structured Light Patterns

There are many types of coding techniques used in structured light system: Binary coding, Grey level coding and color coding. Binary coding is well known coding scheme because it is more immune to noise.
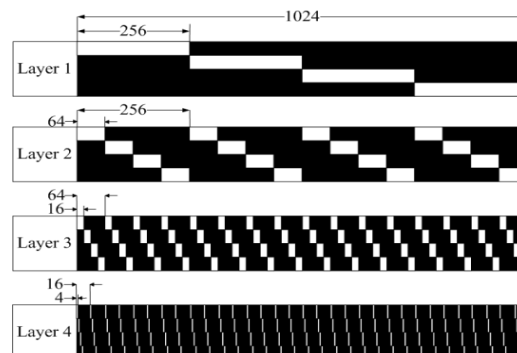

Figure 2. Four Layers of HOC Patterns

In binary coding there are only two states of pixel either zero or one. Every pixel in this setup represents a code which will be further used to determine the correspondence from captured frame to projected frame. One major limitation of this system is that we cannot achieve high resolution, because projected pattern's width always be greater than one pixel. Interpolation techniques are used to interpolate among the detected boundaries of code word. Another way to increase the resolution of 3D depth information is to use more patterns which will increase projection capturing and processing time.

To reduce the length of the codes S. Lee et al [1] developed hierarchical orthogonal coding (HOC), in which orthogonal codes are arranged hierarchically. Patterns which are used to construct 3D point cloud are Hierarchical Orthogonal Code (HOC) patterns. Every layer of the HOC patterns contains four patterns which are orthogonal to each other. Arrangement of codes is performed to decrease length of code. Length of each code is divided into four layers and each layer contains H orthogonal codes which are shown in Figure 2. In this paper we have projected HOC patterns by using FPGA.

2.2 Boundary Inheritance Codec

The decoding of captured patterns is performed by detecting stripe pattern boundaries as well as shadow boundaries, then regions between detected boundaries are formed in each layer and finally, the boundary inheritance and region correspondence inheritance from upper layer to the lower layer are performed to produce accurate and robust boundary-correspondence pairs for 3D triangulation. In this paper we have used Boundary inheritance Codec [7] for 3D reconstruction.

3. Propsed Design Flow

Design flow of complete system is shown in Figure 3. Serial controller, pattern generation, VGA Controller and Camera trigger modules are implemented in FPGA. Frame grabber, 3D reconstruction and triangulation are performed in software. A complete functional hardware prototype was developed by using an FPGA. We had used an off the shelf FPGA development board (XEM3001) from Opal Kelly [8] was selected for the implementation of design. Selected board provides complete system design for prototype development with a Spartan 3 FPGA.
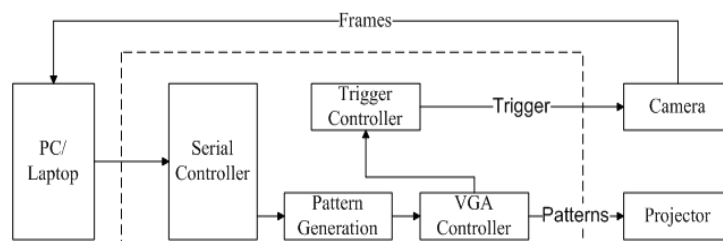

Figure 3. Four Layers of HOC Patterns

3.1 VGA Controller

This section discusses about VGA standard, VGA hardware, and controller implementation in FPGA. Video Graphics Array (VGA) is a video display analog standard that is mostly used with computer monitors and general purpose projectors. There are five signals for normal operation of VGA interface which are Vertical Synchronization Signal (VSynch), Horizontal Synchronization Signal (HSynch), Red, Green and Blue magnitude signal.

Vertical and horizontal synchronization signals are digital signals, voltage range of these signals are 0-5V dc. These signals are used to synchronize the timing with the projector. Each projected or displayed frame constitute of a fixed number of lines and each line is made from fixed number of pixels for a particular resolution. Number of rows and number pixels depends upon resolution, if resolution of the projected frame changed then number of pixels and number of lines will also be changed. Transmission of each frame is carried out in such a way

lines are top to bottom and pixels are transmitted left to right sequence. Frame rate of a particular resolution is determined by the VSynch signal and resolution of the frame is determined by the HSynch signal. Each frame contains some blanking interval while the transmission of frame. Blanking intervals are named as front porch and back porch. These blanking periods are exists horizontally and vertically, which means there will be some more lines in each frame and some extra pixels in each line.
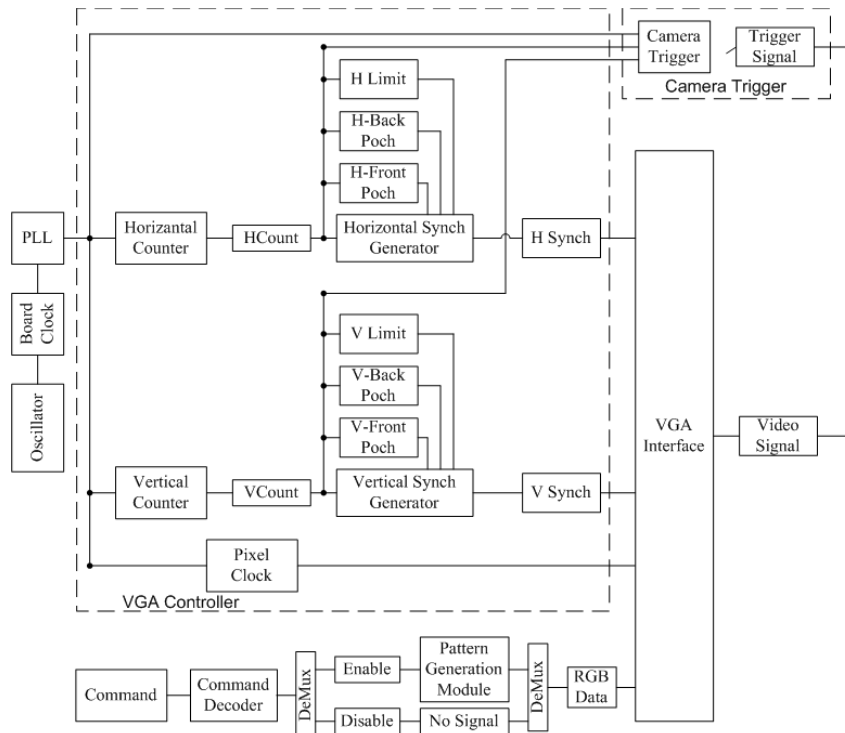
Figure 4. VGA Controller Implementtaion in FPGA

Purpose of VGA module is to provide five signals to DAC chip. DAC chip will transmit that signal to projector. Horizontal synch and vertical synch are generated by synchronous counters. Synch pulse is generated when counter value reaches to the end of front porch and synch pulse ends at the end of back porch. VGA implementation is shown in Figure 4. Polarity of the synchronized pulse depends on resolution of projected frame. Horizontal counter is incremented on each clock cycle which is generated from onboard PLL.

3.2 Camera Trigger Module
Camera trigger module accepts three signals to generate camera trigger signal, clock signal which is generated by onboard PLL module, remaining two signals are vertical count and horizontal count which are generated by VGA controller module. After frame completion a trigger pulse for camera is generted to open shutter of camera. Trigger module implementation is shown in Figure 4.

3.3 HOC Pattern Generation
HOC patterns are sixteen patterns and those patterns are divided in four layers which are represented as L1, L2, L3 and L4 in Figure 2. Implementation of HOC pattern in FPGA is shown in Figure 5. A buffer is assigned to first row of each pattern. On each pixel clock pixel information is read out from buffer and transferd to RGB pin. This process is performed on first pattern of each layer for 2nd, 3rd and 4th pattern; we just need to shift register pixels. HOC patterns have a systematic relationship to each other. We can get the one pattern by shifting the register pixels to right. For first layer we need to shift 256 pixels, layer 2, 3 and 4 requires 64, 16 and 4 pixels shifting respectively.
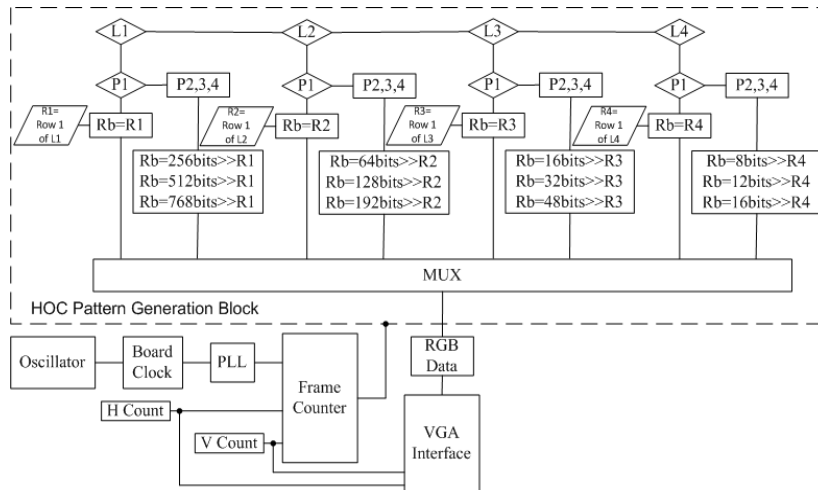
Figure 5. HOC Patterns generation Module

3.5 Simulation Results

Simulation for VGA interface and camera trigger is carried out in ModelSim to generate HSynch, VSynch and camera trigger signal. HSynch signal generated after an interval of 20.67usec which is shown Figure-6a, Vsynch and camera trigger signal is generated after 16.667msec which are shown in Figure 6b.
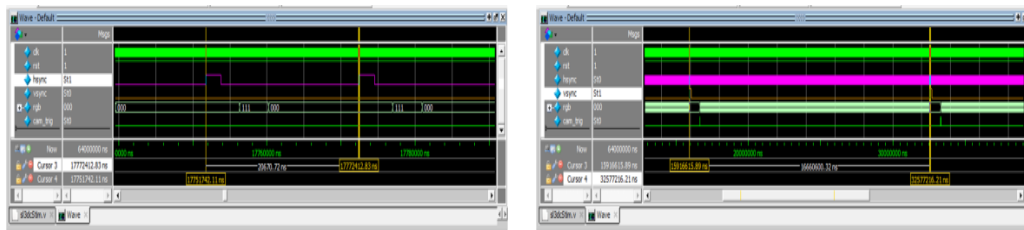
a. Horizontal Synch Pulse Signal   b. Vertical Synch Pulse Signal
Figure 6. Simulation Results of VGA and Camera Trigger Signals

4. Experimental Result

We have performed extensive experiments to compare the results of FPGA and PC based projection. There are two kinds of experiments which are performed to test implemented method, in first type of experiments we have compared improvment in projection and capturing time and second kind of experiments 3D output is compared by projecting the patterns through PC and FPGA. Experimental results are represented in later part of this section.

4.1 Experimental Setup

In the following experiment, we used a projector camera system, as shown in Figure 7 which consists of an Optoma DLP projector (Optoma ML750); a PGR Flea3 1394 digital camera mounted a TV LENS 12 mm 1:1.3, an off the shelf FPGA development board (XEM3001) from Opal Kelly [8] and a computer. The position of the camera was about 18 cm to the left of the projector.
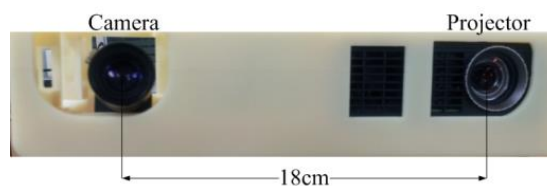

Figure 7. Structred light 3D Camera

4.2 Projection and Capture Time

We have carried out extensive experiments to estimate projection/capturing time of patterns for PC and FPGA based projection. Results are shown in Figure 8.

Figure 8. Projection and capturing time for PC and FPGA

Experiments were carried out for a range of shutter intervals from 2 to 14msec. Ensemble average of 500 continuous captures for each shutter interval is carried out to observe mean capture time of each exposure time. Results shows that projection and capturing time ranges from 1800 to 2350msec and 350msec for PC and FPGA based projections respectively. Results shows that projection of patterns is 6.7 times faster than projection through PC and it projection/captuing time reamins same for different exposure time of camera. Projection and capture time by FPGA based projection lead us to acheive scan speed of 3Hz for 18 patterns (16 HOC, White and Ambient).

4.3 Shutter Interval and Distance

To perform performance testing from captured smooth surface, select region of interest and then measure mean, standard deviation, maximum and minimum errors to that fitted plane [9]. Following steps are carried out for performance evaluation: 3D point cloud is captured, points in the region of 15x15cm are observed. Plane co-efficients are estimated by using RANSAC [10]. Euclid distance is calculated from fitted plane to the 3D points. Error parameters are determined by using points of the selected plane to the fitted plane. We have carried out experiment to observe effect of shutter interval in range of 3msec to 14msec for both cases PC and FPGA which are shown in Figure 9a. Error in case of FPGA is less as compared to projection through PC due to synchonized opening of shutter of camera because of hardware trigger generated by FPGA while in PC based projection a software trigger is generated which is not synchronized with projection and leads to more errors.

a. Fixed Distance   b. Fixed Shutter time
Figure 9. Standaraddevivation of Error

Effect of distance on performance of 3D sensor is shown in Figure-9b. As distance increases between object and camera, error increases. Standard deviation of error for PC based projection is quite large as compared to projection through FPGA. We have measured errors in 3D point cloud from 75cm to 200cm for each distance we have carried out 25 samples to make an average which is shown in Figure-9b. Shutter interval for said experiment is 3msec

keeping the iris fix and gain of camera to zero. A composite result of shutter interval along with distance is shown in Figure 10. Maximum error in case of PC is 0.6mm for 200cm distance and 3msec shutter interval. While in case of FPGA error is 0.37mm which is 61% better than projection through PC. FPGA based projection out performs as compared to projection through PC while keeping all the conditions same for both cases.
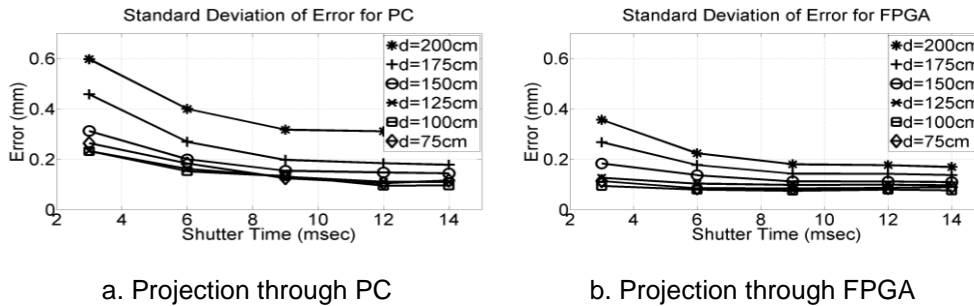

a. Projection through PC   b. Projection through FPGA
Figure 10. Standarad devivation of Error fo different distances and different shutter time

4.4 Qualitative Evaluation

We have performed qualitative evaluation of proposed method by capturing 3D RGB data of stationary object for both the cases of PC and FPGA projection. In qualitative analysis we carefully measured outliers in case of PC and FPGA based projection for scenes shown in Figure-11. Results of multiple captured frames are shown in Table-1 which shows that less outliers and more 3D points are generted by FPGA based projection.

Figure 11. 3D Capture by FPGA and PC based projection

Table 1. 3D Points and Outliers

Object	3DPoints		Outliers	
Projection	PC	FPGA	PC	FPGA
Industrial Object	259368	262565	157	96
Step Module	272696	276455	143	131
Venus with Toy	267498	268904	135	106

5. Conclusion

In this paper we have suggested a technique to increase scan speed of structred light 3D camera by projecting HOC patterns fast by using FPGA, we have achieved scan speed of 3Hz for 18 patterns which was less than 0.5Hz by using PC. Precision of 3D point cloud is icreased by 61% along with more 3D points and less outliers due to synchronized projection and capturing. The same procudre can be used to project gray and hybrid and color patterns other than binary patterns to achieve higher speed and also accurate point cloud of moving object can be obtained by using suggested method.

4. Conclusion

Provide a statement that what is expected, as stated in the Introduction chapter can ultimately result in Results and Discussion chapter, so there is compatibility. Moreover, it can also be added the prospect of the development of research results and application prospects of further studies into the next (based on result and discussion).

References

[1] Sukhan Lee, Jongmoo Choi, Daesik Kim, Jaekeun Na, Seungsub Oh. *Signal Separation Coding for Robust Depth Imaging Based on Structured Light*. Proceedings of the 2005 IEEE International Conference on Robotics and Automation. 2005; 4430-4436.
[2] Y Oike, H Shintaku, S Takayama, M Ikeda. K Asada. *Real-time and high-resolution 3D imaging system using light-section method and smart CMOS sensor, Sensors*. Proceedings of IEEE, 2003; 1: 502-507.
[3] Sukhan Lee, Jongmoo Choi, Seungsub Oh, Jaehyuk Ryu, Jungrae Park. *A real-time 3D IR camera based on hierarchical orthogonal coding*. Proceedings 2006 IEEE International Conference on Robotics and Automation. ICRA 2006. Orlando, FL, 2006: 2035-2040.
[4] H Byung-Joo, P Chan-Oh, S Nam-Seok, C Jun-Dong. A real-time compact structured-lightbased range sensing system. *Journal of SemiconductorTechnology and Science*. 2012; 193-202.
[5] APP Jongenelen, DG Bailey, AD Payne, DA Carnegie, AA Dorrington. Effcient fpga implementation of homodyne-based time-of-flightrange imaging. *Journal of Real-Time Image Processing*. 2012; 21-29.
[6] SJ Bellis, WP Marnane. *A cordic arctangent fpga implementation for a high-speed 3d-camera system*. In Proceedings of the Roadmap to Reconfgurable Computing, 10th International Workshop on Field-Programmable Logic and Applications, FPL '00. 2000; 488-494.
[7] S Lee, LQ Bui. Accurate estimation of theboundaries of a structured light pattern. *J. Opt. Soc.Am. A.* 2011; 954-961.
[8] OK. Inc. Front panel user manual. 2013.
[9] K Khoshelham, EO Elberink. Accuracy and resolution of kinect depth data for indoor mapping applications In Sensors. 8238.
[10] MA Fischler, RC Bolles. *Random sampleconsensus: A paradigm for model fitting with applications to image analysis and automated cartography*. Commun. ACM. 1981: 381-395.