The Twenty-Sixth International Symposium on Artificial Life and Robotics 2021 (AROB 26th 2021),
The Sixth International Symposium on BioComplexity 2021 (ISBC 6th 2021),
ONLINE, January 21-23, 2021

Fitness Calculation with a FPGA Implementation

Shiqian Luo[1],Yuichiro Toda[1], and Mamoru Minami[1]

[1]Okayama University, Japan
(Tel: 81-86-251-8233, Fax: 81-86-251-8233)
[1]pw3n9sdf@s.okayama-u.ac.jp

Abstract: Due to the increasing stress of labor deficiency, robots that could handle heavy load of work automatically are in need. Visual servoing system is widely used in target's pose recognition for it's adaption in all kinds of complex situations. However, conventional visual servoing system requires prerequisite information of targets like size color or apex that can specify an object, which limits their usability in recognition when target changes. Consequently, we want to propose a new method named Projection-based Method, It can recognize arbitrary target in real time with no prior knowledge, besides no preparation is needed when targets change. But the recognition system has been implemented on personal computer in the previous research, which limits their on-the-move application where portability is desired. The portability of the system will ensure the ease of use in movable robots. For instance, the AUV (Autonomous Underwater Vehicle) with FPGA (Field Programmable Gate Array) could be employed in some extreme environment such as deep sea where human operator has difficulties in conducting tasks while AUV with FPGA could achieve a 3D perception just like human. In this paper, we describe our approach for arbitrary objects pose recognition system with FPGA technology, the proposed system is implemented on Zynq UltraScale+ MPSoC ZCU104. Two static photos were used as models in experiment to verify the effectiveness of this system.

Keywords: FPGA Implementation, Projection-Based Method, Arbitrary Targets Recognition

1 INTRODUCTION

As sharp decrease in birth rate happens not only in japan which has already been known for aging society, virtually the world is experiencing growth in the number and proportion of older persons in their population. According to data from World Population Prospects: the 2019 Revision, by 2050, one in six people in the world will be over age 65 (16%), up from one in 11 in 2019 (9%). By 2050, one in four persons living in Europe and Northern America could be aged 65 or over. It can be seen in the near future that there would be a further huger deficiency happening in labor force that has already been a hot potato to some country now. Consequently, robots that replace humankind to do tough and monotonous task are urgently in need and intellectual robot is poised to become one of the most significant issue from the early 2oth century. Making robots self-driving is the first step. Stereovision is the extraction of 3D information from digital images, such as those obtained by a CCD camera. By comparing information about a scene from two vantage points, 3D information can be extracted by examining the relative positions of objects in the two panels. This is similar to the biological process Stereopsis. Comparing to conventional recognizing technology of robots, vision servo help robots adapt better in complex environment such as place after earthquake where huge amount of debris remains. However, unlike the sophisticate human brain that could distinguish target automatically, robots need prerequisite information of targets, such as color, apex and size, which cause time wasting at some extent. And that kind of time wasting could fulfill target recognition in real time. So a method named projection-based method has been proposed. It can recognize arbitrary target in real time with no prior knowledge, besides no preparation is needed when targets change. In addition, FPGA (field-programmable gate array) technology has been implemented on this system. A FPGA is an integrated circuit designed to be configured by a customer or a designer after manufacturing. It contain an array of programmable logic blocks that can be inter-wired in different configurations. Logic blocks can be configured to perform complex combinational functions, or merely simple logic gates like AND and XOR. With FPGA technology, the recognition system can achieve high performance at a very early phase of the development.

2 BASIC PRINCIPLE OF PROJECTION-BASED METHOD

2.1 Basic Principle of Projection-based Method

Firstly, a real target is projected onto left and right camera images-that is natural projection into left and right camera's images. Secondly, the natural projection is thought to be completed based on true pose and then the target object in left image is selected as model. Thirdly, the selected model is inversely projected into 3D space with an assumed pose. Fourthly, the 2Dmodel in 3D space is re-projected back to the right camera image with the same assumed pose. Finally, if the re-projected model coincides with the real target image and the assumed pose represents the real target ’ s pose in 3D

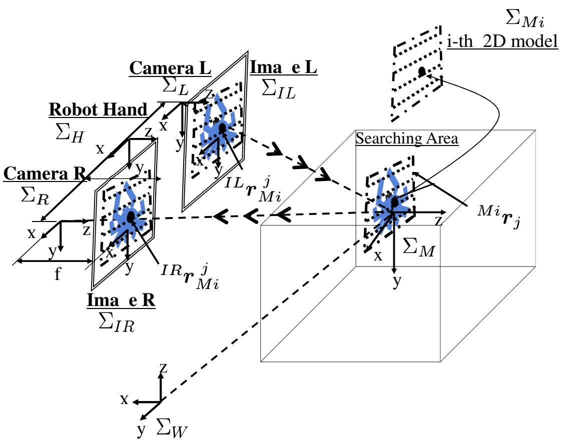

Fig. 1. Principle of Projection-Based Method

2.2 GA and fitness function

As noted above, the assumed pose generates by GA (genetic algorithm), which is commonly used to generate high-quality solutions to optimization and search problem and the evaluation method is used to evaluate the recognition result.

3 CONSTRUCTION OF FPGA DEVELOPMENT ENVIRONMENT

3.1 Environment preparation

To implement FPGA technology, previous computer environment preparation is necessary. Ubuntu Linux 16.04 LTS(64 bit) is installed. And the Vitis unified software platform enables the development of FPGAs. The new Vivado Design Suite supply design teams with the tools and methodology needed to leverage C-based design. The PetaLinux Tools offers everything necessary to customize, build and deploy Embedded Linux solutions on Xilinx processing systems.

3.2 OPEN CL and OPEN CV

3.2.1 OPEN CL

Open Cl (Open Computing Language) is a framework for writing programs that execute across heterogeneous platforms consisting of central processing units (CPUs) field-programmable gate arrays (FPGAs) and other processors or hardware accelerators.

3.2.2 OPEN CV and xf::open cv

Open CV (Open Source Computer Vision Library) is a library of programming function aimed at real-time computer vision. Picture data is read by open cv from the CPU and stored in a class named mat. Mat is like an image container which save information and attributes of pictures. xf::open cv is intended for FPGA development supplied by Xilinx.

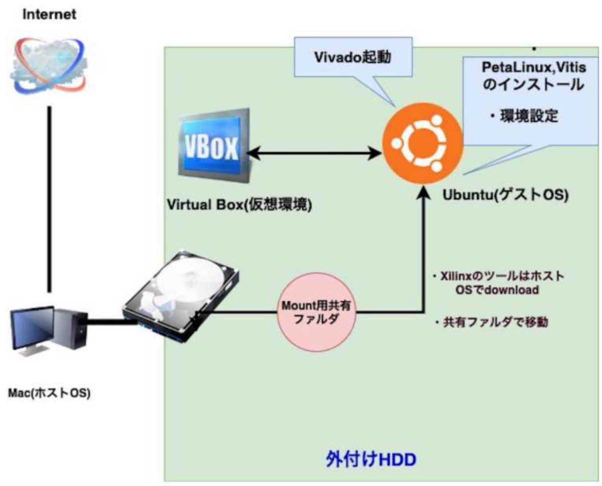

Fig. 2. FPGA Development Environment

Unlike cv::mat which stores picture information as int type, it’s easy to read r,g,b. The data type of xf:: mat differs based on the number of pixels to process per clock cycle and the type parameter, there are different possible data types.

Option	Number of bits per Pixel	Type
XF_8UC1	8	Unsigned
XF_16UC1	16	Unsigned
XF_16SC1	16	Signed
XF_32UC1	32	Unsigned
XF_32FC1	32	Float
XF_32SC1	32	Signed
XF_8UC2	8	Unsigned
XF_8UC4	8	Unsigned
XF_8UC3	8	Unsigned
XF_2UC1	2	Unsigned

Table 1. xf::Mat Class - Available Data Types

$$wordwidth = pixeldepth \times numberofchannels \times numberofpixelstoprocesspercycle(NPPC) \quad (1)$$

For example, this time we use picture of 3 channels and process one pixel per clock cycle, so the picture data would be $ap\_uint < 24 >$。

$$Wordwidth = 8 \times 3 \times 1 = 24 \quad (2)$$

The first eight bit represents B, then the second eight bit represent G, and the final one represent R. we could use program below to read the data from pixel。

$$Pixel1.range(0,7) = mat\_L.data \times (y \times mat\_L.cols + x)$$
$$Pixel1.range(7,15) = mat\_L.data \times (y \times mat\_L.cols + x)$$
$$Pixel1.range(16,23) = mat\_L.data \times (y \times mat\_L.cols + x) \quad (3)$$

4 SCHEME OF FITNESS CALCULATION WITH FPGA IMPLEMENTATION

The idea to calculate fitness value is : firstly get image data from CPU and give a target pose to simulate model including outer area and the inner part. Secondly, read r, g, b value of model sample points and use it to transfer them into the brightness and hue. Finally, calculate the fitness function of left and right images. Based on the idea of fitness calculation, there are 8 schemes of program structures prepared previously. And the target pose is set as xsft =22.66, ysft =19.24, zsft = 339.99,n1 = 0.0052, n2 = 0.0354, n3 = 0.0234.

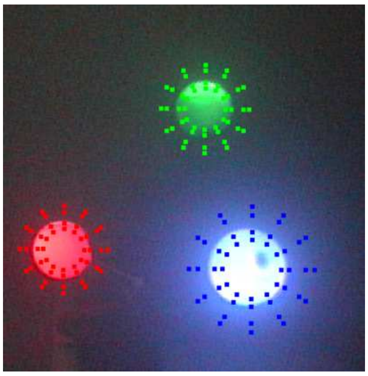

Fig. 3. 3-ball model in left image

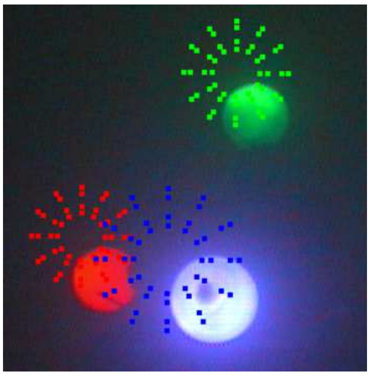

Fig. 4. 3-ball model in right image

The first and second one is that the image data from CPU is saved as stream mode and stable mode. In the stream, the data of the image could be read once because variables of stream are implemented as RAM. While in the stable mode, the data of image could be read repeatedly, but the memory consuming would be further huge.

The third scheme is that get the image data in stable mode and calculate fitness of left image only.

The forth and fifth scheme is that get the image data in stable mode and then transfer the BGR image into HSV image or HSL image. So it could save the memory wasted in calculation of brightness and hue.

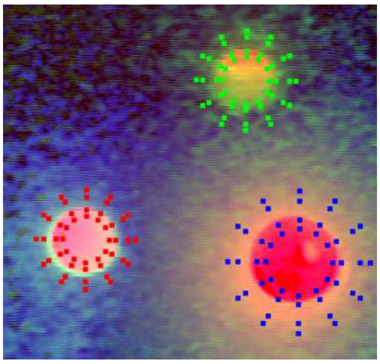

Fig. 5. 3-ball Model in HSV Image

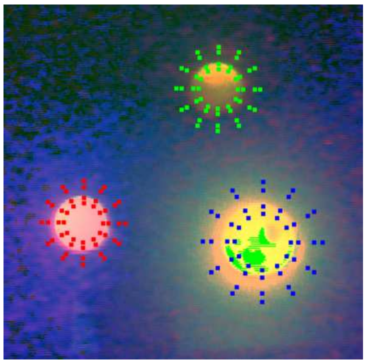

Fig. 6. 3-ball Model in HSL Image

The sixth and seventh scheme get image data in stream mode. Considering to the possible huge amount of memeory consuming, stream mode would be a better way to calculate fitness function. However, under that mode, ther will only one time to read the picture data, so there are two loops used to scan all the pixel in picture to identify if the pixel is what we need ,if so, save data. And the only difference between sixth and seventh one is that the seventh one will save all pixel data in the scope of blue square and then verify if the points in the square is the sample points. The four vertice of the blue square is from the maxmum and minmum of x and y coordinate of the out area of model. So no loop is need to verify if the point is in the square. Supposing points of inner part and outer part of model is 24. There would be $640 \times 480 \times 24$ loops needed to make sure if the point is what we need in the the sixth scheme, while $640 \times 480 + 80 \times 80 \times 20$ loops are needed in the seventh. So it could save memory at some extent.

The ninth scheme and eighth scheme is that all the coordi-

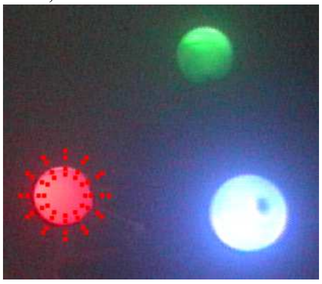

Fig. 7. Red Ball Model in Left Image

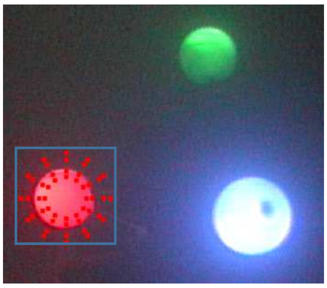

Fig. 8. Out Scope of Red Ball Model

nates of sample point and its r, g, b would be transferred from the CPU, so it could alleviate the computing load of FPGA to the most extent.

4.1 Conclusion

From the result above, it ’s easy to tell that the ninth one is the most practicable one and it could be speculated that model creating would be the one consuming the most memory. So in the future, we would continue optimize the grogram of fitness calculation and try as hard as we could to fulfill the three ball model fitness calculation. Besides, due to the characteristic of FPGA, once we complete the fitness calculation with a FPGA implementation, it's possible to accelorate the speed of recognition so it would pave the base stone for recognizing objects with fast speed.

REFERENCES

[1] Yu Wei,Xiong Bing,C. Chareonsak, “ FPGA implementation of AdaBoost algorithm for detection of face biometrics,” *IEEE International Workshop on Biomedical Circuits and Systems*, 2004.

[2] Lwin K N, Myint M, Mukada N, Yamada D, Matsuno T, Saitou K, Godou W, Sakamoto T, Minami M, “ Sea Docking by Dual-eye Pose Estimation with Optimized Genetic Algorithm Parameters,”*Jouranl of of Intelligent & Robotic Systems*, Vol.92, Issue 1, pp.159-186, DOI (identifier) 10.1007/s10846-018-0970-x, 2018.

[3] 寇　ギョウ郡, 田　宏志, 李　想, 見浪　護, ”Projection-based ビジュアルサーボシステム”, 第 *13* 回コンピューテーショナル・インテリジェンス研究会,CI13-02, 2018.6.16-17.

[4] Vivado Design Suite ユーザー ガイド: 高位合成 (UG902).

[5] Vivado Design Suite チュートリアル: 高位合成 (UG871).

[6] ザイリンクス OpenCV ユーザー ガイド.

		Result of FPGA	Result of CPU Type
1	Stream left and right	Fitness=0	Fitness_all=0.4
2	Stable left and right	1700 cell units needed but 642 available	Fitness_all=0.4
3	Stable left	945 cell units needed but 642 available	Fitness_L=0.8
4	Stable HSV left	842 cell units needed but 642 available	Fitness_L=0.8
5	Stable HLS left	842 cell units needed but 642 available	Fitness_L=0.8
6	Stream scan all picture left	-1749124	Fitness_L=0.8
7	Stream square scope left	-1749124	Fitness_L=0.8
8	Directly red ball of left image	Fitness=0.6	Fitness_L_red=0.6

Table 2. Result of Each scehmes

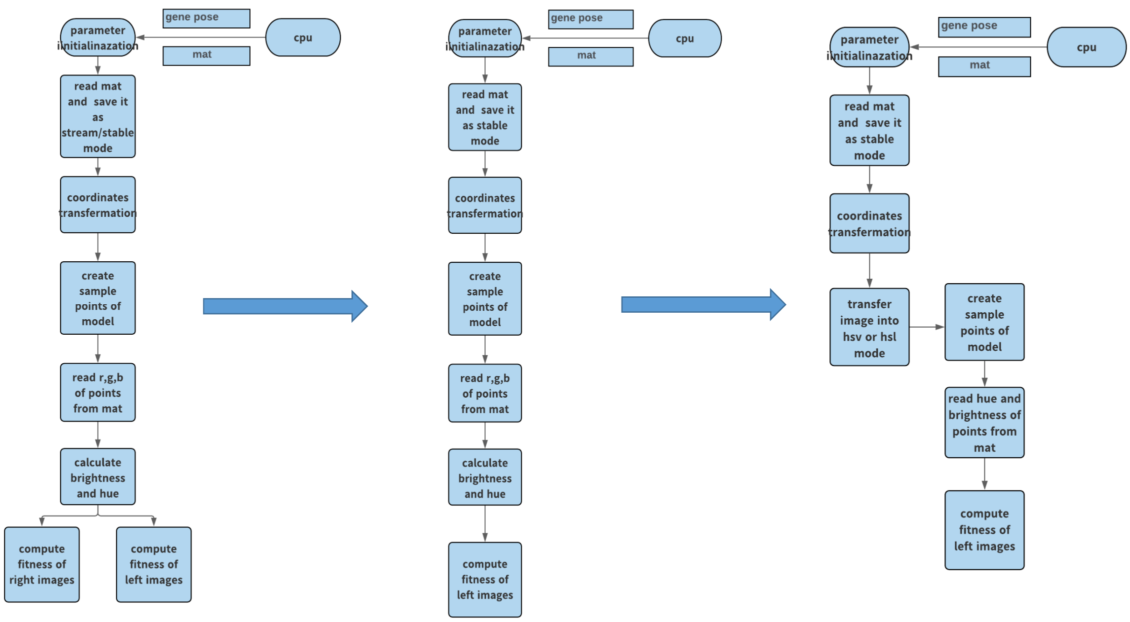

Fig. 9. Scheme 1 to Scheme 5

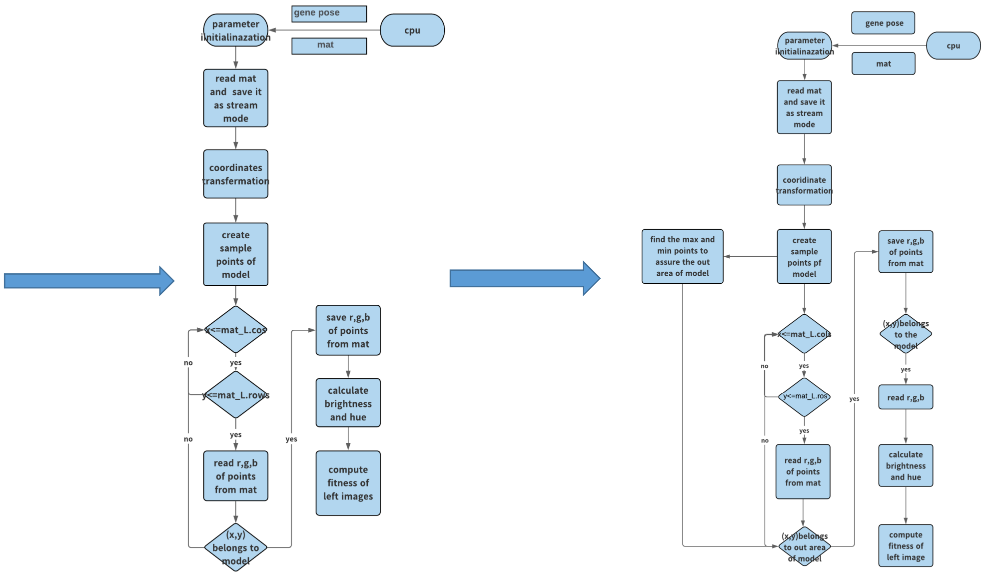

Fig. 10. Scheme 6 and Scheme 7

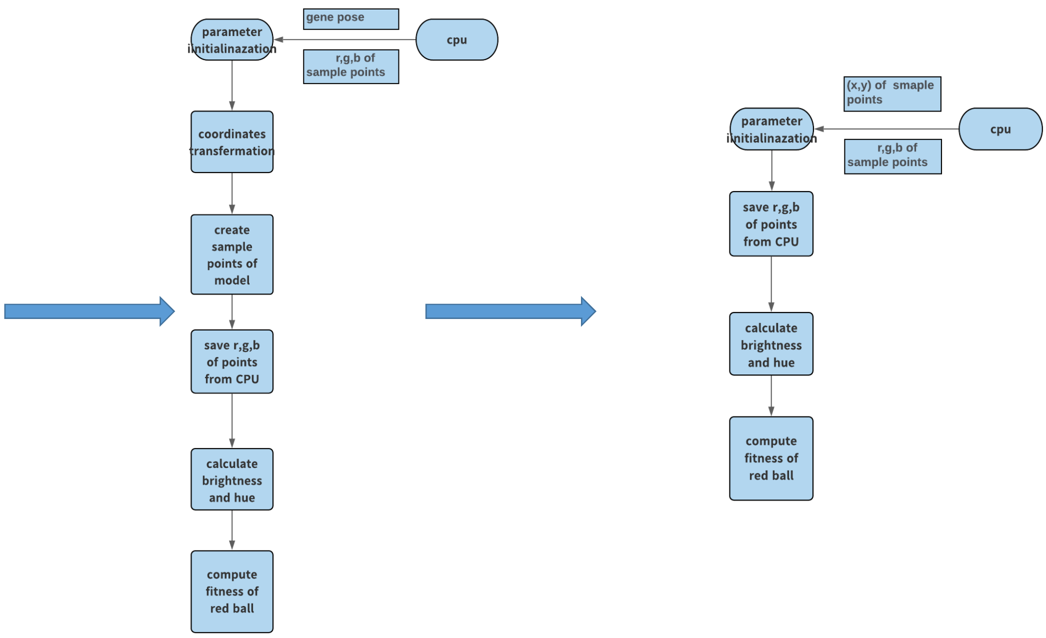

Fig. 11. Scheme 8 and Scheme 9