The Twenty-Seventh International Symposium on Artificial Life and Robotics 2022 (AROB 27th 2022),
The Seventh International Symposium on BioComplexity 2022 (ISBC 7th 2022),
The Fifth International Symposium on Swarm Behavior and Bio-Inspired Robotics 2022 (SWARM 5th 2022)
ONLINE, January 25-27, 2022

Shortening calculation by Introducing Field Programmable Gate Array For 3D Space Sensing

Shiyu Wang† ,Jieyuan He[1] ,Takahiro Nida[1] ,Yuichiro Toda[1] and Mamoru Minami[1]

[1]Okayama University, Japan
Tel: +81-86-251-8233; E-mail: paw6968v@s.okayama-u.ac.jp

Abstract: With the development of intelligent robots, robots can be involved in more and more fields. Visual servoing system is widely used in target ' s pose recognition for it ' s adaption in all kinds of complex situations. However, the traditional visual system requires the size, color or vertex of the object to determine the spatial position of the object. In order to perform a given function, a robot needs to have a basic recognition of its environment and a processor that can handle huge amounts of computation. This paper mainly introduces the application of FPGA in 3D space sensing to accelerate recognition speed. We use FPGA to accelerate the computation speed of fitness and we let different location information be processed at the same time, and greatly accelerate the calculation speed. Lay a foundation for increasing the number of individuals and the number of evolution in the future. We describe our approach for arbitrary objects pose recognition system with FPGA technology, the proposed system is implemented on Zynq UltraScale+ MPSoC ZCU104.

Keywords: Visual servoing system, 3D space sensing, Fitness.

1. INTRODUCTION

According to statistics released by the Ministry of Internal Affairs and Communications, The population of Japan is 125.41 million (April 2021). With the development of Japan's economy and society, Japan is facing more and more serious population problems, which the most prominent is the aging of population and the Sub-replacement fertility. Japan has entered an aging society since the early 1970s, and the elderly over the age of 65 account for 7.1% of the total population. After more than 40 years of development, Japan has become the fastest-aging country in the world, with the proportion of the population over 65 rising to 23% in 2010, according to the United Nations' World Population Outlook. By 2060, nearly 40 percent of the total population will be aged 65 or older. The aging population increases the burden of social security in Japan. Labor shortage caused by aging has also become a major problem plaguing Japan's economic development. Thus, the development of intelligent robots to alleviate Japan's labor shortage became one of the most important development projects of the early 20th century. Intelligent robot technology is also distributed in different fields, such as industrial intelligent robot arm, autonomous driving robot, unmanned exploration robot and so on. In order to enable intelligent robots to achieve automatic driving and exploration, robot stereo vision has become the priority in the field of robot research. The robot binocular stereo vision is an important form of stereo vision, which is based on the parallax principle and a method to obtain three-dimensional geometric information of objects from multiple images. In the binocular stereo vision system, two digital images of the object are obtained by two cameras at the same time from different angles, or two digital images of the object are obtained by a single camera at different times from different angles. Based on the parallax principle, the three-dimensional geometric information of the object is recovered, and the three-dimensional outline and position of the object are reconstructed. Binocular stereo vision system has a wide application prospect in stereo vision field. Unlike the human brain, which automatically recognizes the complexity of a target, robot needs the color, vertex, size and other prerequisite information of the target, and in order to obtain the key information, the robot must carry out repeated and complex calculation, which causes a waste of time, so it is difficult to conduct real-time calculation and recognition. Field Programmable Gate Array (FPGA) is further developed on the basis of PAL, GAL and other Programmable devices. As a semi-custom circuit in the field of ASIC, it not only solves the deficiency of custom circuit, but also overcomes the shortcoming of limited number of gate circuits in original programmable devices. FPGA can be reprogrammed indefinitely, loading a new design in a few hundred milliseconds, and reconfiguration can reduce hardware overhead. Therefore, the recognition system can achieve high performance at a very early stage of development by using FPGA to accelerate the robot stereo vision algorithm.

† Shiyu Wang is the presenter of this paper.

2. BASIC PRINCIPLE OF MODEL-BASED METHOD

2.1. GA Algorithm

Genetic Algorithm (GA) was first proposed by John Holland in the United States in the 1970s, which was designed and proposed according to the laws of biological evolution in nature. It is a computational model of biological evolution process that simulates the natural selection and genetic mechanism of Darwin's biological evolution. It is a method to search the optimal solution by simulating the natural evolution process. By means of mathematics and computer simulation, the algorithm transforms the solving process of the problem into a process similar to the crossover and mutation of chromosome genes in biological evolution. When solving complex combinatorial optimization problems, compared with some conventional optimization algorithms, better optimization results can be obtained faster.

At the beginning, GA algorithm generates several model individuals with different positions through random numbers. And convert its position and posture information to binary code. Then, the binocular camera image information is read, and these randomly generated model individuals are put into the exploration space, and compared with the image information to calculate the fitness of each random model,choose the model individuals with highest fitness value, keep its position and posture information, individuals with low fitness are eliminated and allowed to crossover mutation and sudden mutation. Crossover mutation means that some genes of individuals with high fitness replace those of individuals with low fitness. Sudden mutation means that some genes of individuals with low fitness are randomized again. Finally, the fitness of each individual after evolution is calculated again and sorted, and the final step is selection. As a cycle, the position and posture of the individual gradually converge to the real position of the object.

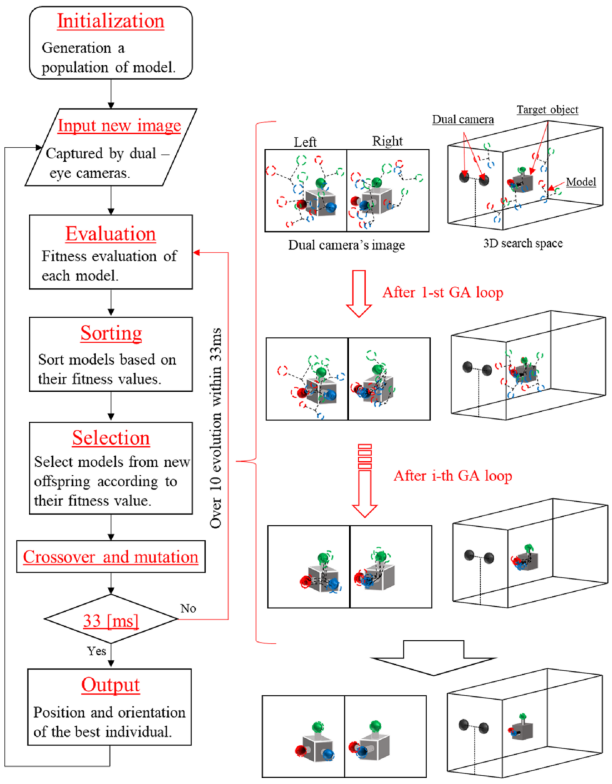

Fig. 1. GA algorithm flow

2.2. Model-based Matching

In this system, a model-based matching method is used to estimate the matching degree between the target model and the captured image. The pose estimation based on 2D-to-3D reconstruction using feature-based recognition is applied in other conventional methods.

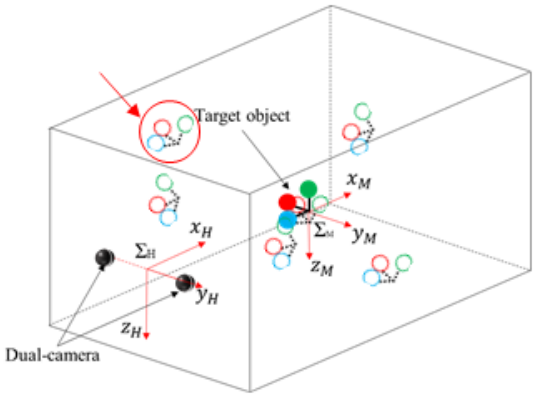

Fig. 2. Exploration space

The model is composed of point groups, and the hue value can be obtained by reading the color information of evaluation point coordinates. Evaluation points are awarded if the hue value of the evaluation point in the internal field is within a given range. Conversely, if the hue value of the evaluation point in the internal field is out of the given range, the score will be subtracted. The evaluation method of the background field is exactly opposite to that of the internal field. If the hue value of the evaluation point in the background field is within a given range, the score will be deducted. Conversely, evaluation points are awarded if the evaluation point hue value of the background field is outside the given range.

$$F_{L,HSV}(\phi) = \frac{\sum_{n=1}^{N_{in}}({}^{IR}r_{in}^{(n)}) - (\sum_{n=1}^{N_{out}}({}^{IR}r_{out}^{(n)})}{N_{in}+N_{out}} \quad (1)$$

$$F_{R,HSV}(\phi) = \frac{\sum_{n=1}^{N_{in}}({}^{IL}r_{in}^{(n)}) - (\sum_{n=1}^{N_{out}}({}^{IL}r_{out}^{(n)})}{N_{in}+N_{out}} \quad (2)$$

The value calculated from the evaluation point can be obtained by the following calculation.

$$Fitness\ value = \tfrac{1}{2}(F_R + F_L) \quad (3)$$

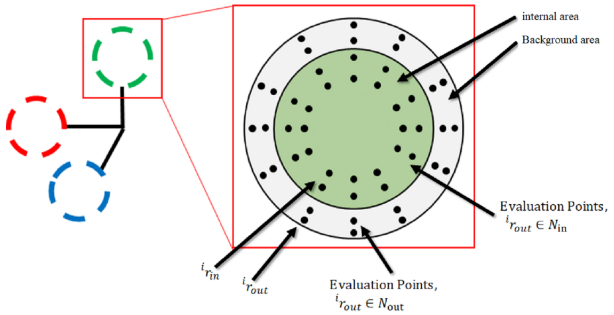

Fig. 3. Model-base Matching

3. RECOGNITION SYSTEM USING FPGA

3.1. OPENCL

OpenCL (Open Computing Language) is a unified programming environment for parallel programming of heterogeneous systems. It is widely applicable to multi-core processors (CPUS), graphics processing units (GPUS), Cell architectures, digital signal processors (DSP). FPGA and other parallel processors.

The main function of the CPU is to read the information of the picture, generate the genetic according to the GA algorithm, and then, according to the position and pose of model to evolution and mutation, finally synthesize the new image of the specified RGB information according to the evaluation point coordinate on the model. Then the creation of the FPGA channel is used, which is mainly used to use OPENCL. The main process is to read the FPGA device information, create a channel for data interaction with the FPGA, and the size of the predefined data buffer area, and then transfer the data to the FPGA buffer area.

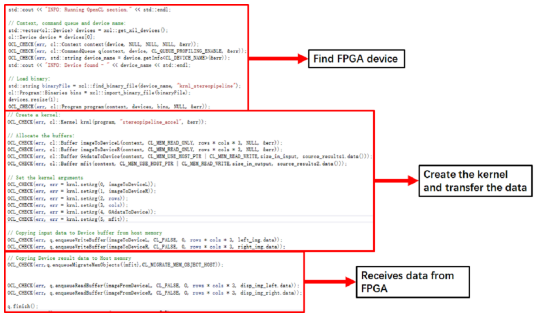

Fig. 4. OPENCL program

3.2. OPENCV

OpenCV is a cross-platform computer vision and machine learning software library. It is lightweight and efficient,it consists of a series of C functions and a small number of C++ classes. It also provides interfaces to many other languages and implements many common algorithms in image processing and computer vision. The optimized C code has greatly improved the execution speed, and the processing speed can be faster by using the Integrated Performance Primitives. Using OPENCV, CPU can read image information more quickly and conveniently, and transfer it to FPGA port.

Because FPGA cannot read the data of Mat type directly, it is necessary to convert the image of Mat data type into XF::Mat data type that FPGA can read .FPGA Development Environment Unlike CPU 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.

Table 1. XF::Mat data types

Option	Number of bits per Pixel	Type
XF_ 8UC1	8	Unsigned
XF_ 16UC1	16	Unsigned
XF_ 16SC1	8	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

Since the image we transferred is RGB image, we choose the three-channel type, and the RGB value range is 0-255, so we choose 8 bits and unsigned data type. So we select XF_ 8UC3 type.

3.3. DATAFLOW

As an optimization method of HLS, DATAFLOW is very effective for improving throughput and reducing Latency. The data processing process assembly line seems to have to mobilize, data a in accordance with the sequence of input, and then after processing, a continuous output, the processing process as far as possible not to have interruption。 HLS Dataflow allows user to schedule multiple task together to achieve higher throughput.The CPU does not proceed to the next function until one function is fully executed. Fpgas using Dataflow are different. Dataflow allows parallel processing of functions, passing the results of function A to function B while executing function A, and speeding up the computation by executing function B while function A is still running.

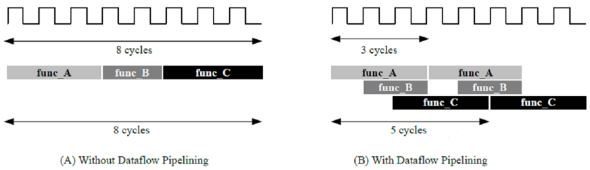

Fig. 5. DATAFLOW

3.4. The interaction flow between CPU and FPGA

FPGA adopts the concept of Logic Cell Array (LCA). The system provides a Configurable Logic Block (CLB), an Input/Output Block (IOB), and a different Interconnect module. FPGA is composed of hardware resources such as logic unit, RAM and multiplier. By organizing these hardware resources reasonably, hardware circuits such as multiplier, register and address generator can be realized.

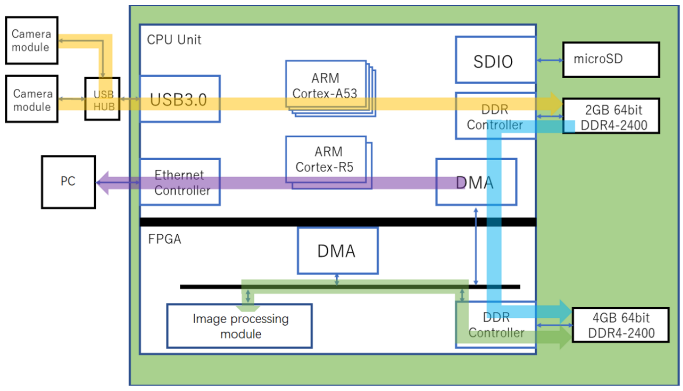

Fig. 6. The composition of FPGA

In the previous recognition program, all the steps were completed at the CPU side, which generated the first generation of random model individuals and read the image information, compared the position and posture of model individuals with the image, calculated the fitness, and finally sorted the evolution.

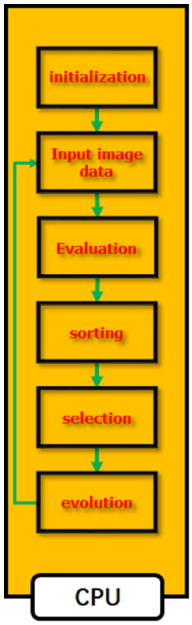

Fig. 7. CPU flow

The program flow using FPGA is different from that of CPU. First, CPU side generates the initial random model individual and reads the image information. Then, the gene information and image information are transmitted to the FPGA side through the kernel. Finally, the CPU uses the fitness value to sort and evolve, and transmits the new generation of gene information and image information to the FPGA.

3.5. Image input method for stream processing

Because the DATAFLOW processing method is used in the process of using FPGA, the traditional image input greatly increases the burden of stream processing, so we adopt a new image input method. The previous image input method is to input 640*480 image information to FPGA

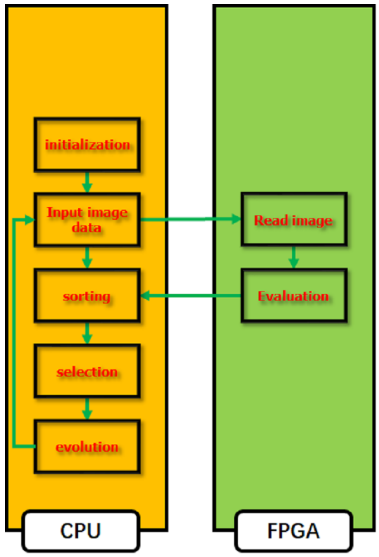

Fig. 8. FPGA flow

and then read the RGB information of corresponding coordinates. Due to DATAFLOW stream processing, FPGA can only read the image information from the beginning of the image (0,0) to the end of the image (640,480) in sequence. Not only does this take a lot of time, but information processed through the stream can only be read and written once. To this end, we introduce a new image input method.

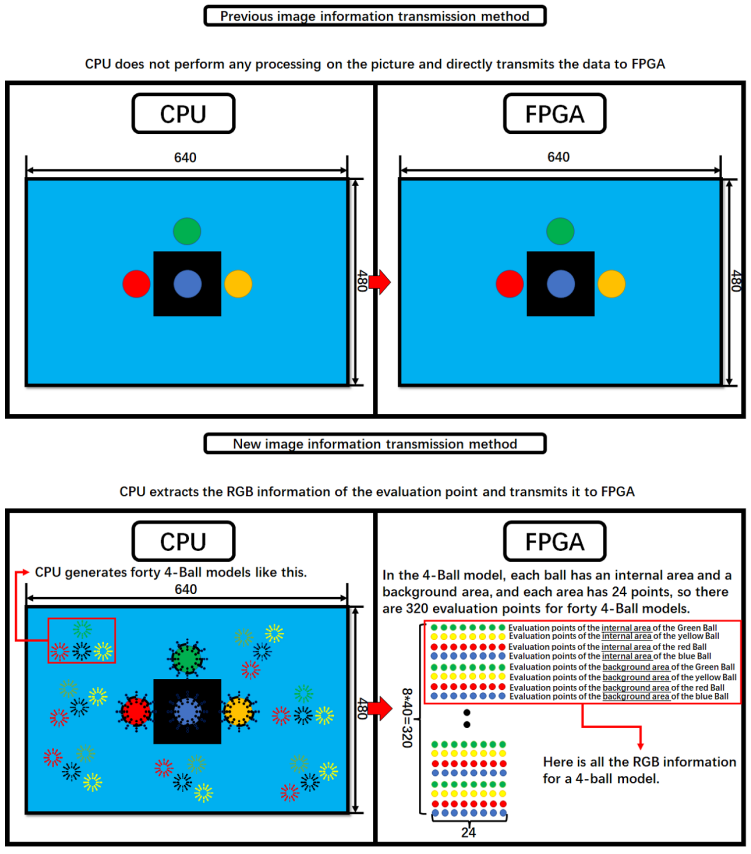

Fig. 9. Image input method for stream processing

Now, instead of transmitting the entire 640*480 image, we extract the image information of the evaluation points and compose a new image, which is sent to FPGA. 4 ball model has a total of four balls, the evaluation of every ball area is composed of internal areas and background areas respectively, that is to say, a single camera has a total of eight evaluation areas, and any evaluation area has 24 points, that is, a single camera requires Fitness calculation points is 8 * 24, And 40 individuals are 40 * 8 * 24, that is, 320*24 size image information. The advantage of this method is that the unnecessary image information is discarded, the amount of data is reduced, and the reading speed is accelerated.This method can reduce the program burden (no longer need IF function to determine whether a coordinate point is consistent with the evaluation point coordinates).

4. PROGRAM STRUCTURE AND EXPERIMENTAL RESULTS

4.1. Serial port debugging tool

We need to use PUTTY serial port debugging tool, through this tool we can achieve serial port debugging and FPGA operation. Enter the USB port position manually and set the serial port speed to 115200.

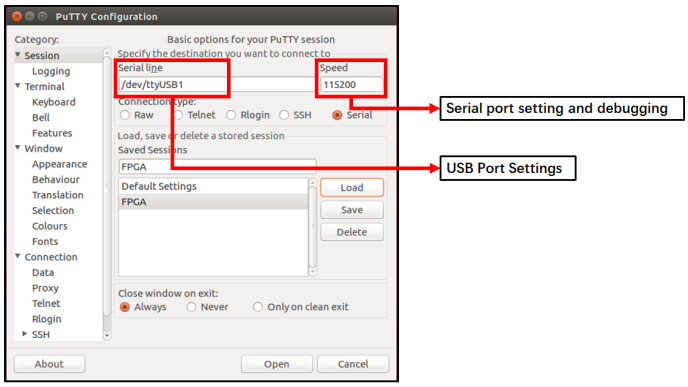

Fig. 10. Serial port debugging tool

4.2. Experimental results

We placed the camera in the center of the model 700mm on the Z axis.

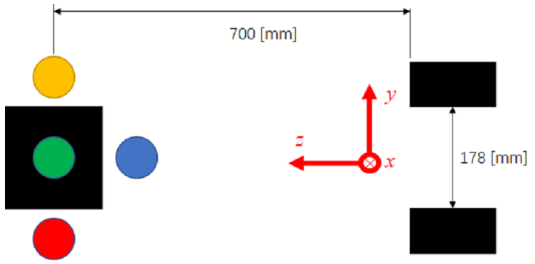

Fig. 11. Experimental environment

Compared with CPU, FPGA adopts DATAFLOW data processing method to realize parallel processing, which greatly reduces the calculation time. Originally,calculating a genetic CPU needs about 125ms, but FPGA can complete it with only about 0.1ms. Nearly 100 times faster than a CPU.

Table 2. XF::Mat data types

Device	Calculation time
CPU	125.00012 ms
FPGA	0.15872 ms

This is the result of the knowledge of the system, and the data is gradually receiving the value. Fitness and position information are shown below.As can be seen from the figure, the randomly generated individuals of the first generation did not fit the model at all. With the increase of evolution times, the positions of individuals gradually moved closer to the model and finally fit.

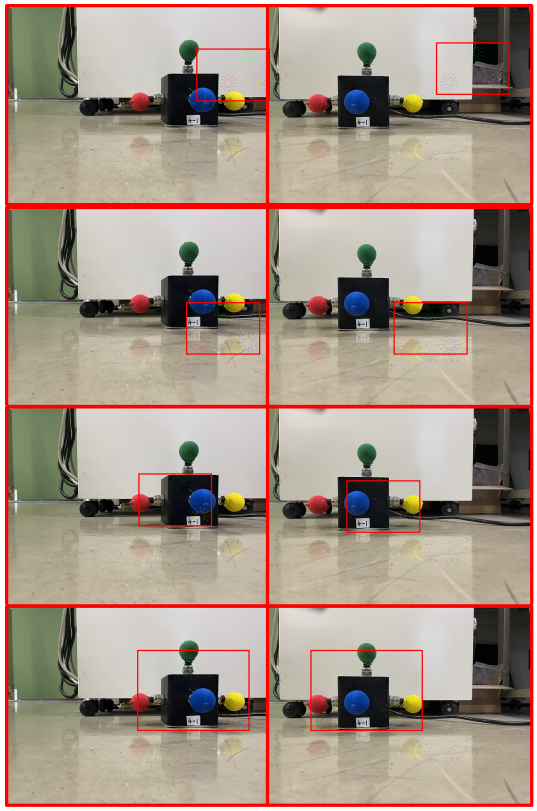

Fig. 12. Cognition result

The results of this experiment were demonstrated in only four of the 50 times evolutions. With the evolution, it can be seen that Fitness value gradually increases, and the peak value appears at the 50th time, and Fitness reaches 0.428175. At this time, it can be clearly seen from the picture that the individual's position and posture at this time are almost the same as the real position and posture.

we can see that the error between the initial individual and the individual after five times of evolution is large in the X and Y directions. With the evolution, the value of the X and Y coordinate position gradually decreases, and the Z coordinate also tends to be close to the real distance of 700mm.

Table 3. Fitness

Number of evolution	Fitness
0 time	0.010417
5 times	0.125000
25 times	0.203125
50 times	0.428175

Table 4. position

Number of evolution	X	Y	Z
0 time	240.62	-14.84	963.04
5 times	309.18	-147.56	839.50
25 times	-52.34	-17.68	927.39
50 times	-17.38	-34.38	701.32

Table 5. posture

Number of evolution	N1	N2	N3
0 time	0.0529	0.1744	0.0255
5 times	0.0023	-0.0097	0.0265
25 times	-0.0062	-0.1920	0.0264
50 times	0.0442	-0.0376	-0.0238

5. CONCLUSION

Through this research, we successfully combined FPGA and CPU to realize the parallelization of Fitness computing, and the computing time of Fitness was shortened from the original 120ms to 0.1ms. It also verifies the accuracy of cognition.

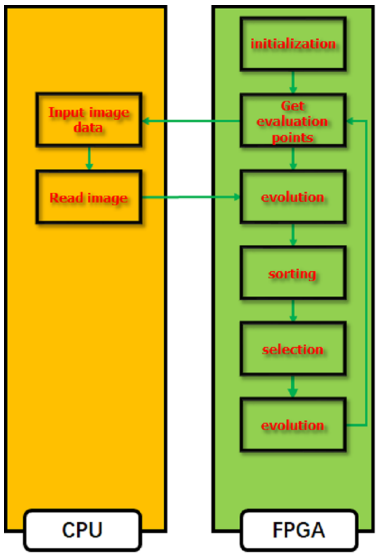

Fig. 13. New system flow

However, at this time, FPGA memory occupancy rate is only 20%, in order to maximize the use of FPGA computing unit and buffer space, we plan to gradually move GA algorithm to FPGA, so as to realize the initial and evolution of parallel processing cognitive system. As shown in the figure.13 CPU only needs to read the picture information and extract the RGB information of the specified coordinate in the picture according to the genetic information completed in the initial stage of FPGA. Finally, the RGB information of the specified coordinate points is synthesized into a MAT-type picture and its data is sent to FPGA for the calculation, evolution and variation of Fitness through FPGA. The new generation of genetic information is passed to THE CPU and the RGB information of the specified coordinates in the picture is extracted and transmitted to FPGA again to form a cycle.

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*.

[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] Shiqian Luo, Yuichiro Toda, Mamoru Minami," Fitness Calculation with a FPGA Implementation", *26th International Symposium on Artificial Life and Robotics, pp.705-710, January 21-23, 2021*.

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