Field programmable gate array based moving object tracking system for robot navigation

Hoang T. Tran¹, Dong LT. Tran¹, Quang N. Pham¹, Thanh C. Vo¹, Quan NA. Nguyen¹,
Thang K. Nguyen¹, Duyen M. Ha¹, Minh T. Nguyen²
¹Center of Electrical Engineering, Duy Tan University, Da Nang, Vietnam
²Thai Nguyen University of Technology, Thai Nguyen, Vietnam

ABSTRACT

This paper proposes a method in which an object tracking robot system is implemented on field programmable gate arrays (FPGAs). The OV7670 camera provides real-time object pictures to the system. To improve picture quality, images are put through the median filter phase. The item is distinguished from the backdrop based on color (red), after which it is subjected to a mathematical morphological approach of filtering to eliminate noise. To send the robot control signals, the object’s (new) coordinates are found. In this method, the median filter, color separation, hardware IP cores, and morphological filter are all part of the embedded system on FPGA. Through the direct memory access (DMA) controller, these cores may communicate and perform high-speed pipeline computing at higher data rates. The entire system is executed in real-time on Xilinx's spartan-6 FPGA KIT. The results show practical and promise.

Keywords: Embedded system, FPGA, Homomorphic filter, Median filter, MicroBlaze, Robot chasing objects

1. INTRODUCTION

The practice of tracking an object's movement over time using a camera is called object tracking [1]–[3]. Since the locations of the objects are always changing, tracking moving images is a crucial task in surveillance systems [4], [5]. Application areas for object recognition and tracking include autonomous robot navigation, surveillance, and vehicle navigation [6], [7]. Finding things throughout a series of frames is the process of object detection. Object tracking uses a camera to find the items that will gradually appear. A hard digital signal processing module, distributed memory, and a number of programmable logic blocks make up a field programmable gate array (FPGA), which can handle real-time objects [8], [9]. There are several real-world uses for object tracking, including security, surveillance, autonomous driving, automated traffic management, biological image analysis, and intelligent robot control [10], [11]. The goal of object tracking, like that of the majority of computer vision systems, is to identify and extract the target item from a stream of images that the camera continuously records [12]–[14]. Better object tracking is made possible by faster image processing computation. In actuality, when object tracking apps are deployed, they are often created on the open CV library [15]–[17] and operate on Windows or Linux. As a result, the graphics libraries used in image processing make the program's execution speed extremely reliant on the hardware setup, which raises the cost of design. Due to the current demand, tracking systems must be as affordable as feasible while still meeting the requirements of processing speed, well-handling, accuracy, and reaction time. Because of the quick
development time and low cost, while yet meeting the needs of reaction speed and precision, FPGA is the ideal solution for developing relativistic tracking algorithms.

In our study, the system makes use of a morphological noise-removal filter, a median filter, and red color to distinguish the target item from background objects. The robot motor is given a control signal by determining the object's location. In order to boost the system's speed, the system is constructed on an FPGA utilizing a combination of hardware cores and embedded microcontrollers from MicroBlaze. Different from previous work, the contributions of this work can be pointed out shortly as: i) an object tracking robot system based on FPGA is constructed, ii) the system captures real-time photos of objects from the OV7670 camera, iii) a mathematical morphological method is utilized to remove noise from around the object, and iv) the entire system is run in real-time on the Xilinx's spartan-6 FPGA KIT.

The rest of this paper is organized as follows. Section 2 provides the proposed object tracking system including an overview of the tracking system and a description of the algorithms to be used in the system. Section 3 describes the structure diagrams of the embedded system on the KIT FPGA and details the IP cores in the system. In section 4, some results of the implementation of the tracking system on FPGA and the experimental evaluation results are provided. Finally, conclusions and future work are addressed in section 5.

2. THE PROPOSED OBJECT TRACKING SYSTEM

2.1. System overview

First, the camera will provide the system one picture frame. The picture is then subjected to object separation image filtering to enhance image quality, red color separation, and eliminate noise-free zones. Following object separation, it will ascertain the item's shape and its location with respect to the image's center before generating a motor control signal that will cause the robot to move in that direction. A process of the object tracking system is shown in Figure 1.

![Figure 1. A process of object tracking system](image)

2.2. Deploying algorithm steps

2.2.1. Collecting object images

The OV7670 camera continually recorded photos of objects. The camera's data is delivered by the camera in parallel 8-bit frames, and the system configures the camera using the standard serial camera control bus (SCCB) [16]. The received picture is a 320 by 240 pixel RGB565 color image [18].

2.2.2. Apply filters to separate objects

The color on red, green, and blue (RGB) images collected from the camera will be applied a median filter of size 3x3 to reduce noise. Then, the image will be color separated into a binary image in which only the black background and white pixel blocks represent the object to be processed. The binary image is then applied a morphological filter to enhance the quality to make it easier to find the contour around the object.

a. Filter the median

Boyle's median filter is a popular choice for mild noise reduction (impulse noise). The median filter minimizes noise by replacing a pixel's value with the median of the gray levels of nearby pixels because pulse noise frequently appears unique, and its gray level value differs significantly from that of its neighbors. The core principle of the median filter technique is to employ a mask that scans every pixel of the input picture sequence; typically, a mask of sizes 3x3, 5x5, and 7x7 is employed. A 3x3 median filter is shown in Figure 2. Take the value of the associated pixels in the mask area at each pixel location and arrange them in ascending or descending order. After sorting the range of pixels for the pixel value being taken into consideration for the final picture, assign the pixel to the middle (median) of the range [19].

b. Color separation of objects

The color separation block, which is a block that recognizes the object based on the color of the item compared to the picture backdrop, is used to filter the noise after obtaining the RGB image. It is important to examine the RGB values at several locations across the image in order to distinguish the color of the item. The red ball used in this article has an RGB color space value of (255, 0, 0), however the picture captured by the
camera has a low resolution and is influenced by light. Light causes the obtained R, G, and B picture area's red color to vary around the values 255, 0, 0. The object's color value is equal to 1, the backdrop color returns a value of 0, and the thresholds for the color channels R, G, and B are experimentally chosen to separate the color. Figure 3 shows the binary output picture that is the end result.

Figure 2. An illustration of the median filter: the median value of its neighboring pixels (3x3) replaced the current pixel value

Figure 3. Image after filtering and color separation

c. Morphological filtering

Due to the aforementioned circumstances, the object that was removed from the backdrop becomes the subject of the survey as well. This includes noise and any interior spaces that were left empty due to background confusion. Therefore, the object should be refined by eliminating the noise regions that are not the object and filling the vacant spaces inside the object in order to guarantee that the best information is delivered for the following blocks of the system. Using mathematical morphology (MM) is one of the techniques used to filter items after removing them from the background [20], [21].

MM is a set theory-based method for treating geometrical structures. The basic morphological procedures used by this approach, which is based on structure and shape, allow the picture to be made simpler while preserving the key elements seen in Figure 4 of the original. In order to assess if a given basic block, or structural element, fits or misses the form in the picture, the fundamental goal of MM is to find images that contain that block. In the case of Figure 4(a), applying dilation helps to connect the dashed points of the image that increases the details of the image. Next, in Figure 4(b), the erosion removes groups of pixels that are much smaller than the size of the object in the image to remove noisy areas for more accurate object for identification. There are 4 basic morphological operations [22] in Figure 5; i) dilation: used to expand or thicken the object in the frame, ii) erosion (shrink): used to shrink or thin the object in the frame, iii) opening: opening combines an erosion and a dilation with the same structuring element, and iv) closing: closing combines a dilation and an erosion with the same structure element.

Figure 4. An example of applying two MM's on block A, (a) block A after performing the dilation and (b) block A after performing the erosion
2.2.3. Determine object coordinates

In Figure 6, the first thing we can find a way around the object. The contour edges are constructed based on the minimum distance from the subject to the corresponding sides of the frame. The coordinates of the object are determined based on the center of the contour relative to the center of the image frame. From there, determine the direction of movement for the robot so that the object returns to the center of gravity.

3. FPGA EMBEDDED SYSTEM STRUCTURE

3.1. Embedded system structure diagram

Figure 7 shows a detailed built-in diagram of the embedded system on the spartan-6 FPGA SP605 evaluation KIT including:

- One 32-bit MicroBlaze processor core [23] running at 100 MHz with 32 K of data and instruction memory is connected to high-speed computer and peripheral cores through the AXI interface.
- A single UART controller with a baud rate of 128,000 that can transport from the computer to the board the picture to be processed and receive the completed image for display on the computer.
- The picture to be filtered is stored in external RAM with a maximum memory capacity of 128 MB, which is connected to a single core SDRAM controller.
- IP core for a median filter with two memory FIFOs and one median filter.
- The IP core does the dilation process, which includes the morphological filtering of 1 math filter and 2 FIFO memory. The IP core performs erosion math (performs morphological filtering), includes 2 FIFO memories and 1 math filter core erosion.
- A DMA controller core that uses Xilinx to increase data processing performance while moving data between hard-core IP cores and external RAM [24].
- The PWM core regulates pulses to operate the robot's motor.
- Two Xilinx-powered clock sources, one with a 200 Mhz oscillator (both positive and negative side) and the other with a 27 Mhz oscillator (single rib) [25].
- The system employs the AXI interface, which includes AXI4, AXI4-Lite, and AXI stream [26], to connect with MicroBlaze.
3.2. Structure of hardware cores
3.2.1. Filter median and color separation of objects

a. Median filter element

Figure 8 shows the block diagram of the median and color separation filters including:

- The median filter and color separation comprise two FIFOs for synchronizing data between two clock domains; one data domain is obtained from an AXI stream with a frequency of 100 MHz and the other from the internal clock domain of the median filter, which operates at 30 MHz.
- The mechanism pipeline is responsible for filtering the median, and a stiff core does this.
- A control_bi block to manage synchronizing the writing of data from the median filter output to the FIFO OUT.
- The work of binaryizing each R, G, and B channel with the necessary thresholds is essentially what makes up a color separation block.
- Consider a 3x3 mask with the pixels sorted in ascending order for each row, then in descending order for each row, and lastly sorted diagonally. The median of the diagonal equals the 3x3 mask's median [27].

Figure 9 shows the hardware that was constructed using that approach. To compare two 8-bit A and B inputs and output the bigger number \( H \) and the smaller number \( L \), use the basic node block. Based on the fundamental nodes in Figure 10, the aforementioned procedure is used to sort a block and determine the median of a 3x3 mask.

![Figure 7. The embedded system structure diagram on spartan-6 FPGA SP605 evaluation KIT from Xilinx](image)

![Figure 8. Block diagram of median filter and color separation](image)
3.2.2. Morphological filtering

We construct 2 IP cores, including dilation and erosion, based on the homomorphic filtering theory. The hardware architecture and I/O data flow for these two IP cores are identical. In Figure 13, the entire picture is scanned using a 9x9 structural element. The structural element will travel one pixel at a time throughout the whole image, from left to right and top to bottom. A new picture pixel will be generated for each 9x9 block of the image that corresponds to the identified 9x9 structural element. As a result, it is clear that, with an image size of 320x240 pixels, translating the structure pixel-by-pixel from top to bottom will take a very long time. Because of this, the article uses a pipeline computation with the input data stream in each 9x9 picture block in Figure 14 to shorten the execution time. After the first 9 cycles since the first data is pushed in, we will get the
exact 9 bits of data from Out1 to Out9 show at Figure 15. Then every cycle we will have 9 bits of data after homomorphic filtering. Therefore, base on the pipeline architecture, the calculation speed of the system is very high.

3.2.3. Engine control
The robot follows the object based on the coordinates and contour of the object as shown in Figure 15. Controlling the robot forward or backward is based on the parameters, \( Y_{\text{min}} \) and \( Y_{\text{max}} \): i) \( \beta < Y_{\text{min}} - Y_{\text{max}} \leq \alpha \): the robot is stationary, ii) \( \beta \leq Y_{\text{max}} - Y_{\text{min}} \): the robot is moving backwards, and iii) \( Y_{\text{max}} - Y_{\text{min}} \): the robot is moving forwards. Controlling the robot to turn left or right is based on the coordinates of the center of gravity of the object \( x_{\text{ob}}, y_{\text{ob}} \): i) \((x_{\text{ob}}, y_{\text{ob}}) \in \{B \cup C\}\): the robot turns left, and ii) \((x_{\text{ob}}, y_{\text{ob}}) \in \{A \cup D\}\): the robot turns right.
Figure 16 depicts the IP core for motor control. The motor control core consists of: FSM module and PWM module. The FSM module receives signals (cometo, backward, right, left) from the DSP module to decide to control the moving robot to follow the object. The output data of the FSM module (driver_1, driver_2) is sent to the PWM module to control pulses (signal_1, signal_2) for the two robot motors to rotate at the right speed and in the right direction.

Figure 16. Motor control IP-core

4. RESULTS OF IMPLEMENTATION AND ASSESSMENT
4.1. Hardware synthesis results
The embedded system implemented on the spartan-6 KIT. It’s used about 80% of the memory elements including RAM blocks, LUTs and about 20% of other logic of the spartan-6 KIT. This result shows that the design is suitable for resource-constrained systems.

4.2. Execution time results
Table 1 shows the execution time results. Following all the algorithm steps, the execution times are shown corresponding to each block name.

<table>
<thead>
<tr>
<th>No.</th>
<th>Block name</th>
<th>Execution time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Collect photo frames</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>Median filter</td>
<td>1.5</td>
</tr>
<tr>
<td>3</td>
<td>Color separation</td>
<td>0.5</td>
</tr>
<tr>
<td>4</td>
<td>Morphological filter</td>
<td>2.0</td>
</tr>
</tbody>
</table>

4.3. Performance evaluation results
To evaluate the system’s tracking performance, the team evaluated the system based on the good light environment and the robot’s ability to follow in the right direction. Table 2 shows the performance evaluation of each movement direction.

<table>
<thead>
<tr>
<th>Light conditions</th>
<th>Movement direction</th>
<th>Result</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>Forward (25 times)</td>
<td>23/25</td>
<td>88%</td>
</tr>
<tr>
<td></td>
<td>Back (25 times)</td>
<td>22/25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Left (25 times)</td>
<td>22/25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right (25 times)</td>
<td>21/25</td>
<td></td>
</tr>
</tbody>
</table>

5. CONCLUSION AND FUTURE NETWORK
The real-time object tracking robot control system on Spartan®-6 FPGA SP605 evaluation KIT is proposed in this study as being low-cost and low-power. The technology is being tested on the KIT and is capable of precisely directing the robot to pursue red objects under various lighting conditions. The system is constructed using high-speed pipelined hardware cores. In order to enable the system’s high speed operation, DMA is also used to transfer data in bursts back and forth between external DDR3 memory and hardware IP cores. The system's response time to the movement of the item is adequate. Our upcoming study will focus on
implementing robot collaboration using FPGA. Fast real-time robot embedded systems will be crucial in assuring the effectiveness of robot collaborative work as robotics adoption in manufacturing rises via work alongside humans.

ACKNOWLEDGEMENTS
The authors would like to thank Thai Nguyen University of Technology (TNUT), Vietnam for the support.

REFERENCES
BIOGRAPHIES OF AUTHORS

Hoang T. Tran received an engineering degree in system measurement and control from Danang University of Science and Technology in 1998, received a master's degree in network and electrical systems in 2009 from the University of Danang; Defending the PhD thesis majoring in Electronics and Communication Technology in 2016 at the University of Technology, Vietnam National University, Hanoi. A guide to synthesizing data sensors, wireless network transformation and the fields of control automation. He is currently the Director of The Center of Electrical Engineering, Duy Tan University, Da Nang, Vietnam, and the Lecture of the Faculty of Electronic Engineering, Duy Tan University, Da Nang, Vietnam. He can be contacted at email: tranthuanhoang@duytan.edu.vn.

Dong LT. Tran received the B.S. degree from the Duy Tan University, Da Nang, Vietnam, in 2009, the M.S. degree from the Duy Tan University, Da Nang, Vietnam, in 2012, and became a PhD student at Thai Nguyen University from 2022 to the present. He is currently the Vice Director of The Center of Electrical Engineering, Duy Tan University, Da Nang, Vietnam, and the Lecture of the Faculty of Electronic Engineering, Duy Tan University, Da Nang, Vietnam. He can be contacted at email: tranthangdong@duytan.edu.vn.

Quang N. Pham had a mechatronic engineering degree in 2016 from Danang University of Technology, Da Nang, Vietnam and this is also the place where he obtained his master's degree in Mechanical Engineering in 2018. At present, Mr. Quang is a researcher in mechanical design and automation products, especially in implementing robot in many practical applications. He can be contacted at email: phamngocquang3@dtu.edu.vn.

Thanh C. Vo received a degree in Mechatronic Engineering from Danang University of Technology, Vietnam in 2013. Engineer Thanh is currently an expert at Center for Electrical and Electronic Engineering, Duy Tan University, Vietnam (CEE). He has interest and expertise in research topics in the field of embedded programming, automatic control. He can be contacted at email: vochithanh.cdt@gmail.com.

Quan NA. Nguyen received a degree in automatic electrical engineering from Duy Tan University in 2019, Da Nang, Vietnam. Eng. Quan is currently an expert at Center for Electrical and Electronic Engineering, Duy Tan University, Vietnam (CEE). He has interest and expertise in research topics in the field of automatic control. He can be contacted at email: nguyenquan97@hotmail.com.
Field programmable gate array based moving object tracking system for robot navigation (Hoang T. Tran)