A TWO-CHANNEL DETECTION AND TRACK SYSTEM DESIGN BASED ON FPGA AND DSP

HONG ZHANG¹, DONGDONG JIANG¹, DING YUAN¹, YUECHENG LI¹, MINGUI SUN²*
¹Image Processing Center, Beihang University, Beijing 100191, China
²Laboratory for Computational Neuroscience, University of Pittsburgh
*Tel: 412-802-6481, Fax: 412-802-6785, E-mail address: drsun@pitt.edu

ABSTRACT
A system based on FPGA and DSP is designed to detect and track targets, which receives real-time infrared and visible light video, and this paper also presents a new detecting algorithm. It introduces the operating principle, hardware architecture and algorithm flow. The experiment results show that the system is capable of processing 3 dB video data from two channels at the same time. It is designed in the way of modularity, so it’s convenient to update and modified.

KEY WORDS
Detection, Track, Two-channel, Modularity, Otsu

1 INTRODUCTION
The detection and track of infrared targets has been one of the difficulties in the field of image processing. It remains to be improved since the infrared image itself contains a lot of noise and the lack of texture information. So the visible light channel is added to this system. The two channels search the targets at the same time and the detection results are fitted according to the SNR of the image to increase the detection rate and reduce false alarm.

2 HARDWARE IMPLEMENT

![Fig. 1 hardware implement](image)

The system’s hardware consists of three modules: FPGA module, DSP module, video input&output module.

2.1 FPGA MODULE
The system is equipped with Xilinx’s XC5VSX-95T chip to pre-process the image and relieve the burden of DSP.

1) Image decoding:
Extract the valid video data base on the pixel clock, FVAL, LVAL and DVAL.

2) Image pre-processing:
Apply a highpass filter to the image to suppress the noise and highlight the target.
3) **Input&output buffer configuration:**

The input buffer is configured as a dual port RAM, a port is for preprocessed video-data and b port is connected with DSP via EMIF bus.

The output buffer is implemented by ISSI’s IS61WV102416 SRAMs. The video data superposed with the target first is written into the SRAMs and then read by the ADV7125 encoding chip.

4) **Communication with console and servo:**

The console selects the track types (auto or manual) and switches the IR channels (MW-IR or LW-IR). The servo receives the fitted location of the target and drives the motor to track it.

![Fig.2 FPGA module](image)

### 2. 2 DSP MODULE

The system chooses the TMS320C6455 DSP chips to meet the challenge of a real-time huge amount of video data.

The DSP module consists of two TMS320C6455 chips and its functions can be generalized as follows:

Receive the pre-processed image via EMIF, execute the detection and track algorithm, extend storage to read image from external DDR2, store the algorithm program in the extended Flash and communicate with FPGA.

### 2. 3 VIDEO INPUT&OUTPUT MODULE

The main functions can be generalized as follows:

The module is equipped with 3 DS90CR288As to convert 4 LVDS data streams captured by the camera into 28 bits of TTL data. ADV7125 encoding chips output the analog video signal to the LED monitors, with 140M max clock speed and 1440*1050 max resolutions.

The system applies the modularization to its hardware implement, which is flexible to update and modified.

### 3. ALGORITHMS FLOW

![Fig.3 the flow of algorithms](image)

1) **Pre-processing:** Apply a filter to the image to suppress the background and highlight the target.
2) **Segmentation:** Segment the pre-processed images to remove the noise, with the target left only in the image.

3) **Statistics:** Count up all the white spots in a 16*16 neighborhood with the target’s centroid as the center if the target is locked, otherwise, count up all the image.

4) **Clustering:** Calculate the centroid and size of the target after the statistics.

5) **Judgement:** Estimate the location of the target in the next frame based on the past five, if the target shows up, the targets is deemed to be locked, else, lost, and scan the whole image again.

6) **Output:** The target’s location and size are superposed on the target in the output video. At the same time; fit the target’s locations based on the SNR of the two channels. The equations are listed as

\[
LOC_{xy} = LOC_1 * A_1 + LOC_2 * A_2
\]

\[
A_1 = \frac{SNR_1}{SNR_1 + SNR_2}
\]

\[
A_2 = \frac{SNR_2}{SNR_1 + SNR_2}
\]

where \( LOC_1 \) and \( LOC_2 \) are the locations detected in IR and visible light channels; \( SNR_1 \) and \( SNR_2 \) are the SNRs of two channels. The \( LOC_{xy}, LOC_1 \) and \( LOC_2 \) are transmitted to the servo at the same time, and the servo will choose the best one to use.

4. THE IR DETECTION ALGORITHMS

In the IR vision system, it’s usual to use 14 bits or 16 bits to define the gray of a pixel rather than 8 bits. The system adopts a new method of processing 14 bits data: use the high 8 bits of the original 14 bits and then segment the image based on the threshold and the solution is analyzed as follows:

a) As PSNR is an effective criterion to measure the quality of images, we calculate the PNSR of the selected 7 images (from low bit to high bit) to find one that suits best. The results are shown in Tab.1.

<table>
<thead>
<tr>
<th>No.</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNSR (dB)</td>
<td>2.8</td>
<td>4.0</td>
<td>0.7</td>
<td>5.5</td>
<td>0.6</td>
<td>0.5</td>
<td>7.3</td>
</tr>
</tbody>
</table>

Image g has the best PNSR according to the results.

b) Analyze this from a theoretical point of view:

Selecting the high 8 bits equals shifting the original the gray 6 bits to the right or it’s divided by \(2^6\) and the target has a relatively high gray level. So the
discontinuous pixels are transformed to the same gray level as the surroundings and the target remains.

All in all, the system selects the high 8 bits for the further process.

The 8 bits image has a well-suppressed background, but the contrast is inevitably low due to the shifting, Fig.4. So the threshold segmentation should be done to the image.

An Otsu [5] based on gray-median gray 2D histogram threshold algorithm is adopted in this system. The gray-median gray histogram performs better in dealing with noise in IR images than the traditional gray-mean gray histogram. And it is used to segment the high 8 bits image in contrast with the traditional gray-mean gray 2D histogram[6]~[10]. The results are shown in Fig.5.

5. EXPERIMENT RESULTS
Fig.7 Detection results: (a) IR. (b) Visible light.

The experiment results are shown in Fig.7 when the system operates in real environment.

6 CONCLUSIONS

High-performance FPGA and DSP are applied in this system to detect and track targets in two-channels, with the location fitted based on SNR. And a new IR detection algorithm is adopted. What’s more, modularization is applied so it’s convenient to update and modify. The experiment results prove that, the system can meet all the technological specifications in real environment (IR: resolution: 640*512, frame-rate: 100; visible light: resolution: 1392*1040, frame-rate: 30).

ACKNOWLEDGMENT

This paper was supported by the National Natural Science Foundation of China under Grant No.61272351 and the US National Institutes of Health Grant U01 HL91736.

REFERENCES

[1] Alexander Samochin; Evgeny Artyomov; Orly Yadid-Pecht; Optimized hardware architecture for object recognition and tracking [J], 2010 49(10).