Image fusion using FPGA in multi-sensor optoelectronic systems

Nguyen Quang Hiep[1*], Tran Dinh Hoa[1], Dang Van Tang[2]

[1]Military Technical Academy, 236 Hoang Quoc Viet, Bac Tu Liem, Hanoi, Vietnam;
[2]Photo Electric and Electronic One Member Limited Liability Company, Phu Dien, Bac Tu Liem, Hanoi, Vietnam.
*Corresponding author: quanghiepktq@gmail.com.

Received 2 Mar. 2025; Revised 20 Apr. 2025; Accepted 25 Apr. 2025; Published 25 Jun. 2025.
DOI: https://doi.org/10.54939/1859-1043.j.mst.104.2025.130-136

ABSTRACT

This study proposes a foundation for developing a real-time image fusion system for multi-sensor optoelectronic systems applied in defense and security, aiming to enhance target detection by combining images from thermal imaging and low-light channels. The VIP Board Big V3.8 with Cyclone IV FPGA, a custom FPGA development board, which integrates key peripherals such as SDRAM, VGA, HDMI, and image sensors, is utilized. The main hardware modules have been designed, including global clock management (PLL), UART communication, FIFO buffers, LCD display control, and image fusion. The image fusion module employs an algorithm that decomposes images into two layers, the base layer and the detail layer, enabling efficient data fusion from two input images to generate a high-quality output image. Experimental results demonstrate that the system operates stably, meeting speed and low-latency requirements, paving the way for future research on more powerful FPGA platforms.

Keywords: Multi-sensor optoelectronic system; FPGA; Image fusion.

1. INTRODUCTION

In modern optoelectronic systems, multi-sensor image fusion plays a critical role in enhancing situational awareness, target detection, and decision-making processes. Image fusion integrates data from multiple sensors, such as visible light, infrared (IR), and short-wave infrared (SWIR), to generate a single, more informative image. Each sensor provides unique advantages—visible light offers high-resolution details, thermal imaging captures temperature variations, and SWIR enhances visibility in low-light or obscured environments. By combining these modalities, image fusion improves contrast, reduces noise, and enhances object recognition, making it highly valuable in military, surveillance, medical imaging, and autonomous navigation applications. One of the issues that needs to be addressed in image fusion is the latency in real-time image processing [1-3].

To achieve real-time processing and high efficiency, Field-Programmable Gate Arrays (FPGAs) have become a preferred hardware platform for implementing image fusion algorithms [4, 5]. FPGAs offer parallel processing capabilities, low latency, and high-speed computation, making them well-suited for handling the vast amount of data generated by multiple sensors. Unlike traditional software-based implementations, FPGA-based image fusion enables real-time performance while maintaining flexibility for algorithm optimization and system integration.

This paper aims to implement a low-latency image fusion algorithm on FPGA in multi-sensor optoelectronic devices, highlighting its significance in real-world applications and advancements in hardware acceleration.

2. IMAGE FUSION METHOD

Image fusion aims to integrate information from multiple sensor images into a single image that retains the most relevant details from each source. A common and effective approach involves decomposing input images into base and detail layers, allowing for better separation of structural and textural information. This method improves contrast, preserves fine details, and enhances the overall quality of the fused image [6-8].

Step 1: Input image acquisition

Two or more images from different sensors (e.g., low-light (or visible) and infrared) are captured. These images may have different characteristics, with the low-light (or visible) image containing texture details and the infrared image highlighting thermal variations.

Step 2: Image decomposition into base and detail layers

Each input image I_i (where i = 1, 2 for two images) is decomposed into:

+ Base layer I_{i_b}: Represents the low-frequency components, capturing the overall intensity and structural information;

+ Detail layer I_{i_d}: Represents the high-frequency components, preserving fine textures and edges.

A common technique for decomposition is bilateral filtering, guided filtering, or wavelet decomposition. In this paper, guided filtering has been used for image decomposition.

Step 3: Fusion of base and detail layers

+ **Base Layer Fusion:** A weighted averaging or max-selection method is used to retain the most significant structural information

$$I_{\Sigma_\mathrm{b}} = \omega_1 I_{1_b} + \omega_2 I_{2_b} \tag{1}$$

where ω_1 and ω_2 are fusion weights that can be determined adaptively.

+ Detail Layer Fusion: The detail layers are combined using an energy-based method (e.g., Laplacian-based weighting) or absolute maximum selection to preserve the strongest edge details:

$$I_{\Sigma_d} = \max\left(I_{1_d} + I_{2_d}\right) \tag{2}$$

Step 4: Reconstruction of the fused image

The final fused image I_Σ is obtained by summing the fused base and detail layers:

$$I_\Sigma = I_{\Sigma_b} + I_{\Sigma_d} \tag{3}$$

3. FPGA IMPLEMENTATION OF IMAGE FUSION

3.1. Description of the FPGA platform

The VIP Board Big V3.8 (figure 1) is an FPGA development platform featuring the Altera Cyclone IV FPGA, designed for high-performance applications such as image processing and fusion.

The Altera Cyclone IV FPGA series is known for its balance between performance and power efficiency, making it suitable for real-time image processing tasks. Its key features include [9]:

+ Logic Elements (LEs): Up to 114,480 LEs, providing ample resources for complex computations;

+ Embedded Memory: Up to 3.9 Mbits of RAM, facilitating efficient data storage and retrieval;

+ DSP Blocks: Up to 266 embedded multipliers for high-speed arithmetic operations, essential in image processing.

The VIP_Board Big provides all the essential features for FPGA development, including:

- CMOS sensors (such as OV7725, OV5640, MT9V034) that allow for high-quality image/video capture.
- VGA and HDMI connectivity, supporting video output with 1080p resolution.
- LCD1602 display and controllable LEDs for displaying information and signal testing.
- I/O ports and USB interfaces for peripheral device connections and programming communication.
- JTAG interface for FPGA programming and debugging.

Figure 1. *The **VIP Board Big V3.8** with Cyclone IV FPGA and peripheral interfaces.*

This board is particularly useful for projects requiring image and video processing, such as face recognition, real-time video processing, and high-performance embedded systems. With flexible connectivity and powerful features, the VIP_Board Big is an ideal tool for engineers and researchers in FPGA development.

3.2. The main modules of image fusion

a) The Phase-Locked Loop

In FPGA-based image fusion, a Phase-Locked Loop (PLL) plays a crucial role in clock generation, synchronization, and timing control. Since image fusion involves real-time processing of multiple sensor inputs (e.g., visible light, infrared), maintaining precise timing is essential for accurate data integration and smooth image processing.

Key Functions of PLL in Image Fusion on FPGA are Clock Generation & Frequency Synthesis, Synchronization of Multiple Sensor Inputs, Reducing Clock Jitter and Enhancing Stability, Pixel Clock and Frame Timing Control and Multiplexing Different Data Streams.

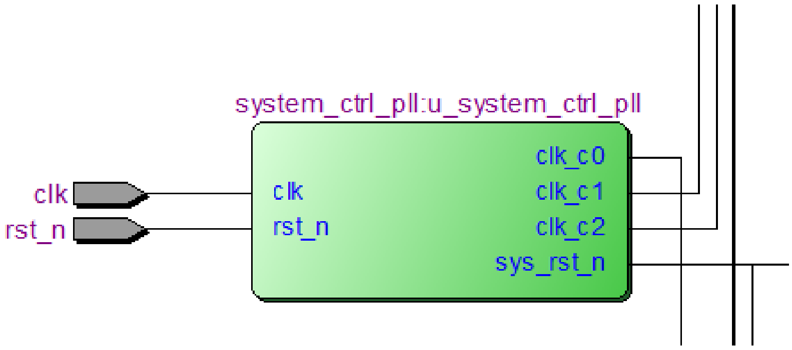

Figure 2. *The Phase-Locked Loop.*

A PLL (figure 2) consists of the following components:

- Phase Detector (PD): Compares the phase of the input and output signals.
- Low-Pass Filter (LPF): Filters out high-frequency noise from the phase detector output.
- Voltage-Controlled Oscillator (VCO): Generates an output frequency that adjusts based on the control signal.
- Frequency Divider (Optional): Used in frequency synthesis to scale the output frequency

In this paper, a 50 MHz clock signal, generated by a quartz crystal oscillator on the board, is used as the input signal for the PLL (Phase-Locked Loop) module. The output clock signals from the PLL module have the following frequencies: clk_c0 = 100 MHz, used for the write speed of the SDRAM; clk_c1 = 100 MHz (with a 90-degree phase shift), used as a synchronization clock for the SDRAM chip, assigned as the clock signal for the physical SDRAM chip; clk_c2 = 5 MHz, used as the input for the LCD-controlled module.

b) The FIFO module

In FPGA-based image fusion, a FIFO (First-In, First-Out) buffer plays a crucial role in data synchronization, temporary storage, and managing data flow between different processing blocks. Since image fusion involves multiple sensor inputs, each operating at different frame rates and resolutions, FIFO helps ensure smooth data handling and prevents data loss.

In this paper, two FIFO buffers, fifo_8to24, which facilitates data transfer from the uart_receive block to the SDRAM block, and fifo_24to8, which facilitates data transfer from the SDRAM block to the uart_transfer block, are designed. Since the data output format of the uart_receiver module is 8-bit, while the data format written to the SDRAM is 24-bit, a buffer module fifo_8to24 is required. Similarly, as the data output format of the image_fusion module is 24-bit, and the input data format of the uart_transfer module is 8-bit, a buffer module fifo_24to8 is needed. Since both FIFO blocks have identical functionality, we will focus on explaining the fifo_8to24 buffer module (figure 3) in detail.

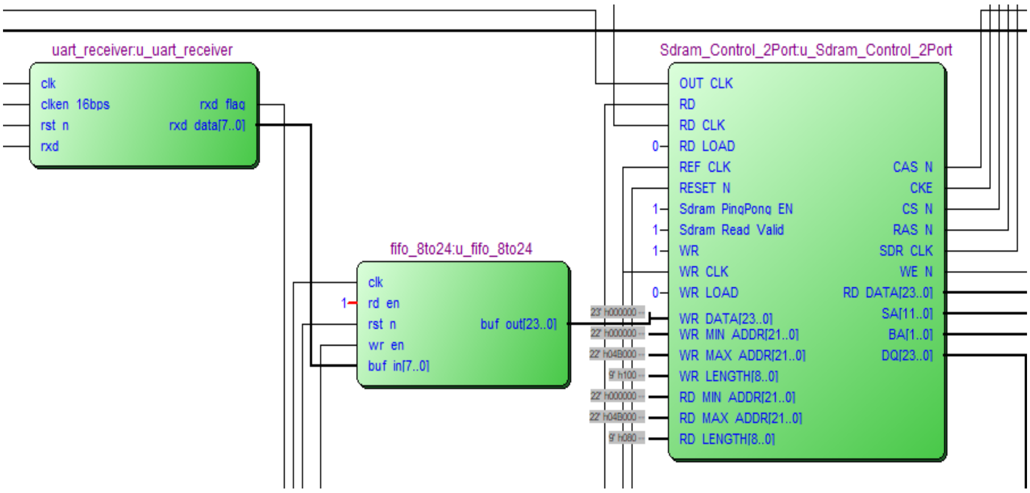

Figure 3. *Communication schema of FIFO module.*

The fifo_8to24 module is a FIFO buffer with an 8-bit input and a 24-bit output. It is designed to handle incoming data with a length of 8 bits and output data with a length of 24 bits. This module utilizes a memory with 64 locations, where each location can store 8-bit data, enabling the conversion from 8-bit to 24-bit when reading data.

c) LCD-controlled module

This module is designed to control the GM7123 chip and the VGA port to ensure that the image is displayed on the screen as intended (with pixels appearing in the correct position and at the right time) (figure 4).

The module is responsible for generating control signals for the LCD screen, including:

- Horizontal synchronization (lcd_hs)
- Vertical synchronization (lcd_vs)
- Display enable (lcd_en)
- Pixel data (lcd_rgb)

The input clock (clk) regulates the synchronization process, and the LCD data is transmitted in the form of a 24-bit RGB value (lcd_data).

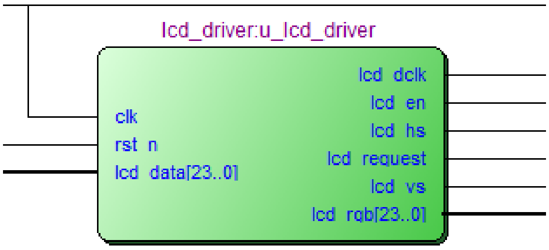

Figure 4. *LCD-controlled module.*

Figure 5. *Image fusion module.*

d) Image fusion module

After the image data is transmitted to the FPGA via the UART module and stored in SDRAM, the image fusion module retrieves two image data signals from two different address regions in SDRAM and fuses their pixels using the mentioned above method (figure 5).

Once the fused pixel values are calculated, the module stores the results in SDRAM at the appropriate addresses and then outputs the image to the LCD screen. Specifically, each pixel is processed based on the image width and height, its fusion value is computed, and the result is stored in memory. Table 1 presents the FPGA resources utilized in this image fusion module.

Table 1. *FPGA resources utilized in the image fusion module.*

Resource	Used	Available	% of All
Logic Elements (LEs)	2.789	15.408	18%
Pins (I/O)	96	166	58%
Memory Bits (on-chip) BRAM M9K	82.328	516.096	16%
DSPs	21	112	19%
PLLs	1	4	25%

In addition to displaying the fused image on the LCD screen, the processed image data can also be sent back to a PC via the UART module.

Thus, the interface diagram of all the modules can be represented in figure 6.

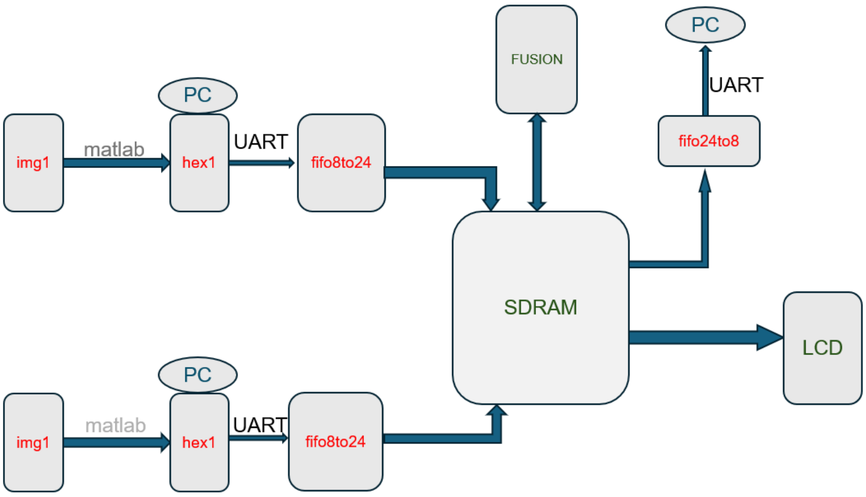

Figure 6. *Image fusion schema.*

4. EXPERIMENTAL RESULT AND DISCUSSION

The images from the low-light and thermal imaging channels are transferred from the PC and stored in SDRAM of VIP Board Big V3.8 (figure 7 and figure 8).

Figure 7. *The image from the low-light channel.*

Figure 8. *The image from the thermal imaging channel.*

Figure 9 shows the fused images obtained on FPGA. It is observed that the fused image shows a significant improvement in information content, specifically with better contrast, clearer details, and enhanced target distinction.

Figure 9. *The fused image obtained on FPGA.*

After implementing the image fusion system on the VIP Board Big V3.8, the following results were achieved:

\+ The system successfully combined two different image sources (thermal imaging and low-light imaging) using an image fusion algorithm implemented on the FPGA.

\+ The resulting fused image retains the same resolution as the original images while incorporating useful information from both channels. It exhibited better contrast, clearer details, and improved target differentiation compared to the other fusion methods (table 2). All of these aspects enhance target recognition, especially in low-light and night-time conditions.

Table 2. *Objective analysis for different fusion methods.*

	Max Fusion	Mean Fusion	Guided Filter
SSIM	6,282	6,358	6,529
MI	2,702	2,88	2,945
RMSE	25,828	25,229	23,086
AG	9,142	6,93	12,137
EN	6,964	6,936	7,228

\+ The FPGA demonstrated its parallel processing capability, enabling real-time image fusion (essentially instant) without significant latency. Compared to the processing speed implemented on a PC, the speed of processing on FPGA is roughly 19 times faster, leading to almost real-time output of the fused image (table 3).

Table 3. *Speed performance compared to PC implementation.*

FPGA Maximum Clock Frequency	100 Mhz
Time consumption on FPGA	0,01s
Time consumption on PC (i7-1165G7 @2.8GHz)	0,19s

+ The designed hardware modules operated synchronously and accurately, ensuring smooth system functionality.

5. CONCLUSIONS

In summary, this study proposes a hardware-efficient infrared and low-light image fusion method based on guided filtering, wherein the input images are decomposed into base and detail layers. The approach is implemented on an FPGA platform, offering a practical solution for real-time applications such as video surveillance and military reconnaissance. The resulting fused image exhibits better contrast, clearer details, and improved target differentiation compared to the original images. Experimental results demonstrate that the system operates stably, meeting speed and low-latency requirements, paving the way for future research on more powerful FPGA platforms.

REFERENCES

[1]. Sicong Zheng, "*Pixel-level Image Fusion Algorithms for Multi-camera Imaging System*", Masters Theses, University of Tennessee, Knoxville, (2010).

[2]. Lei Yan, Qun Hao, Jie Cao , Rizvi Saad, Kun Li, Zhengang Yan & Zhimin Wu, "*Infrared and visible fusion via octave Gaussian Pyramid framework*", Nature Research, (2021).

[3]. Ling Zhang, Xuefei Yang, Zhenlong Wan, Dingxin Cao, and Yingcheng Lin, "*A Real-Time FPGA Implementation of Infrared and Visible Image Fusion Using Guided Filter and Saliency Detection*" Sensors, 22, 8487, (2022). https://doi.org/10.3390/s22218487.

[4]. A. Boutros and V. Betz, "*FPGA architecture: Principles and progression*," IEEE Circuits and Systems Magazine, vol. 21, no. 2, pp. 4-29, (2021). DOI: 10.1109/MCAS.2021.3071607

[5]. T. Saegusa, T. Maruyama, and Y. Yamaguchi, "*How fast is an FPGA in image processing*?," in 2008 International Conference on Field Programmable Logic and Applications, (2008): IEEE, pp. 77-82. DOI: 10.1109/FPL.2008.4629911

[6]. Li, S.; Kang, X.; Hu, J *"Image fusion with guided filtering",* IEEE Trans. Image Process. 22, 2864–2875, (2013). DOI: 10.1109/TIP.2013.2244222

[7]. Li, H.; Wu, X.J. "*DenseFuse: A Fusion Approach to Infrared and Visible Images",* IEEE Trans. Image Process. 28, 2614–2623, (2019). DOI: 10.1109/TIP.2018.2887342

[8]. Li, Q.; Han, G.; Liu, P.; Yang, H.; Wu, J.; Liu, D. "*An Infrared and Visible Image Fusion Method Guided by Saliency and Gradient Information*", *IEEE Access*, *9*, 108942–108958, (2021).

[9]. A. F. Djomo, A. Tiedeu, and J. Fotsing, "*Design and implementation in an Altera's cyclone IV EP4CE6E22C8 FPGA board of a fast and robust cipher using combined 1D maps*," IET Image Processing, vol. 18, no. 7, pp. 1823-1843, (2024). doi.org/10.1049/ipr2.13066

TÓM TẮT

Nghiên cứu trộn ảnh bằng cách sử dụng FPGA trong hệ thống quang điện tử đa cảm biến

Nghiên cứu là cơ sở để phát triển hệ thống xử lý hình ảnh thời gian thực cho các hệ thống quang điện tử đa kênh ứng dụng trong quốc phòng và an ninh, nhằm tăng cường khả năng phát hiện mục tiêu bằng cách kết hợp hình ảnh từ 2 kênh: ánh sáng yếu và ảnh nhiệt. Bo mạch VIP Board Big V3.8, dựa trên FPGA Cyclone IV, tích hợp các thiết bị ngoại vi chính như SDRAM, VGA, HDMI và cảm biến hình ảnh, được sử dụng. Các mô-đun phần cứng chính đã được thiết kế, bao gồm: quản lý xung nhịp toàn cục (PLL), giao tiếp UART, bộ đệm FIFO, mô-đun trộn ảnh, mô-đun điều khiển hiển thị LCD. Mô-đun trộn ảnh sử dụng thuật toán phân tích hình ảnh thành hai lớp: lớp cơ sở và lớp chi tiết, cho phép trộn dữ liệu từ hai hình ảnh đầu vào để tạo ra hình ảnh đầu ra chất lượng tốt hơn. Kết quả thử nghiệm chứng minh rằng hệ thống hoạt động ổn định, đáp ứng các yêu cầu về tốc độ và độ trễ thấp, mở đường cho các nghiên cứu trong tương lai trên các nền tảng FPGA mạnh mẽ hơn.

Từ khoá: Hệ thống quang điện tử đa kênh; FPGA; Trộn ảnh.