International Journal on Semantic Web and Information Systems
Volume 21 • Issue 1 • January-December 2025

Semantic-Driven Ghosting-Free Image Fusion Using Dual Gain Video Stream With FPGA

Yang Xu
Zhejiang Sci-Tech University, China

Longhua Xie
Zhejiang Sci-Tech University, China

Hongchuan Huang
Zhejiang Sci-Tech University, China

Feihong Yu
Zhejiang University, China

Tingyu Zhao
https://orcid.org/0000-0001-5073-7623
Zhejiang Sci-Tech University, China

ABSTRACT

Conventional methods that merge multiple images with different exposure levels often suffer from blur and ghosting due to object movement. Existing ghosting removal algorithms are usually complex and slow, making them unsuitable for real-time video applications. To address this challenge, this study proposes a novel ghosting-free image fusion method using a dual gain video stream setup on an FPGA. IMX662 image sensor is employed, which simultaneously captures both HCG and LCG images with the same exposure time, enabling efficient HDR image synthesis. The proposed method directly addresses the source of the problem, eliminating the need for post-processing steps, thereby preserving algorithmic simplicity. Experimental results reveal that the proposed method not only removes ghosting by 100% but also processes data on an FPGA 98.79% faster than traditional software-based HDR fusion techniques, enabling real-time video stream processing. This dual gain, ghosting-free fusion algorithm demonstrates promising potential for use in high-speed photography and surveillance.

KEYWORDS

Image Fusion, Ghosting-Free, High Conversion Gain, Low Conversion Gain, Field Programmable Gate Array

INTRODUCTION

Real-world scenes encompass a vast and diverse range of image information, characterized by variations in brightness and color. However, conventional digital image acquisition technologies are often inadequate in fully capturing and representing the comprehensive details inherent in such scenes (Chang & Zhang, 2023; Chopra et al., 2022; Yao et al., 2024; Yu et al., 2018). These technologies frequently face challenges in simultaneously preserving extreme brightness levels, such as those observed in bright sunlight and deep shadows, within a single image (Xu et al., 2020; Yuan et al.,

DOI: 10.4018/IJSWIS.367281

2023; Zhu et al., 2022). This limitation has driven the advancement of high dynamic range (HDR) image fusion technology, which integrates under-exposed images (highlighting bright details) with over-exposed images (highlighting dark details) to improve overall image quality (Tan et al., 2023; Wei et al., 2024).

To facilitate this process, basic quality metrics, such as saturation and contrast, guide the fusion of multiple exposures, eliminating the need for camera response curve calibration (Mertens et al., 2009). However, real-world scenes are often dynamic, and any changes during the shooting process, such as object movement, can introduce ghosting artifacts into the fused image (Hu et al., 2022; Li et al., 2021; Li et al., 2020; Wu & Wang, 2023). A multi-exposure fusion technique utilizing dense scale-invariant feature transform (SIFT) has been proposed to address this issue. This method generates a well-exposed image by integrating information from multiple low-dynamic-range images captured with varying exposure settings (Liu & Wang, 2015). Dense SIFT descriptors serve as activity level measurements to extract local details from source images and reduce ghosting effects. While effective in mitigating ghosting artifacts in small regions, this approach is less successful in addressing artifacts in larger areas.

To enhance versatility, Ulucan et al. (2023) introduced a method that extracts feature maps using adaptive exposure and saliency map weighting, achieving fusion through guided filtering and pyramid decomposition. This approach is applicable to both static and dynamic scenes. However, the elimination of ghosting in dynamic images still relies on supplementary algorithms, which increase processing complexity and prolong image synthesis time. Additionally, these algorithms risk misidentifying and filtering useful information as ghosting artifacts, thereby reducing image clarity and compromising the realism and completeness of the final output (Li et al., 2013).

To address the outlined challenges, this paper investigates the root causes of ghosting (Kang et al., 2024) and introduces a novel ghosting-free image fusion method based on dual-gain video streams implemented with field-programmable gate array (FPGA) technology. Traditional methods rely on multiple frames with varying exposure times for HDR image fusion (Bavirisetti et al., 2019), which frequently results in inconsistencies in captured image information, especially in dynamic scenes. Such inconsistencies hinder the production of ghosting-free HDR outcomes (Xiao et al., 2022). Recognizing that ghosting primarily stems from object motion and frame misalignments, this study advocates the use of dual-gain images (Rasheed et al., 2022) for HDR fusion. In dual-gain mode, exposure time remains constant across dynamic and static scenes, ensuring uniformity in captured scene information. This method effectively prevents ghosting during the fusion process while streamlining the synthesis algorithm.

Traditional methods often rely on personal computers for HDR image fusion (Singh et al., 2021). However, this approach greatly reduces processing efficiency, particularly when managing high-speed video stream data. These methods are typically insufficient for capturing dynamic semantic content within images, such as moving objects and evolving scenes. By contrast, utilizing FPGA chips for video image data processing significantly enhances efficiency, making them well-suited for real-time analysis of image content and semantic understanding.

By eliminating the need for additional artifact-removal algorithms and leveraging the high-speed parallel processing capabilities of FPGA, the proposed HDR fusion algorithm achieves a frame rate of 192.9 FPS, enabling full real-time processing. In contrast, existing algorithms achieve a maximum frame rate of only 2.3 FPS, which falls far short of real-time processing requirements.

The potential applications of this technology are vast and are anticipated to significantly impact the fields of image processing and analysis. For instance, intelligent surveillance systems can enhance the accurate identification of suspicious activities, thereby improving security. Similarly, in medical imaging analysis, it enables the production of clearer images, aiding physicians in making more precise diagnoses. Table 1 lists the professional terms used in this framework.

Related Content

Enterprise Tomography: Maintenance and Root-Cause-Analysis of Federated ERP in Enterprise Clouds

Jan Aalmink, Timo von der Dovenmühleand Jorge Marx Gómez (2013). *Advancing Information Management through Semantic Web Concepts and Ontologies (pp. 133-153).*

www.irma-international.org/chapter/enterprise-tomography-maintenance-root-cause/71853

On the Social Shaping of the Semantic Web

Paul T. Kidd (2009). *Handbook of Research on Social Dimensions of Semantic Technologies and Web Services (pp. 341-356).*

www.irma-international.org/chapter/social-shaping-semantic-web/35736

A Semantic Tree-Based Fast-Moving Object Trajectory Tracking Algorithm for Table Tennis

Zechen Jin, Tianjian Zou, Dazhuang Sun, Yu Yangand Jun Liu (2024). *International Journal on Semantic Web and Information Systems (pp. 1-17).*

www.irma-international.org/article/a-semantic-tree-based-fast-moving-object-trajectory-tracking-algorithm-for-table-tennis/337320

Data Linking for the Semantic Web

Alfio Ferraram, Andriy Nikolovand François Scharffe (2013). *Semantic Web: Ontology and Knowledge Base Enabled Tools, Services, and Applications (pp. 169-200).*

www.irma-international.org/chapter/data-linking-semantic-web/76176

The Berlin SPARQL Benchmark

Christian Bizerand Andreas Schultz (2009). *International Journal on Semantic Web and Information Systems (pp. 1-24).*

www.irma-international.org/article/berlin-sparql-benchmark/4112