Information Fusion 42 (2018) 127–145

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Fusing synergistic information from multi-sensor images: An overview from implementation to performance assessment☆

Zheng Liu*,[a], Erik Blasch[b], Gaurav Bhatnagar[c], Vijay John[d], Wei Wu[e], Rick S. Blum[f]

[a] Faculty of Applied Science, University of British Columbia, BC, Canada
[b] US Air Force Research Laboratory, Rome, NY, USA
[c] Indian Institute of Technology Jodhpur, Rajasthan, India
[d] Toyota Technological Institute, Nagoya, Japan
[e] Sichuan University, Chengdu, China
[f] Lehigh University, Bethlehem, PA, USA

A R T I C L E I N F O

Keywords:
Image fusion
Fusion methods
Fusion implementation
Performance assessment
Significance test

2010 MSC:
00-01
99-00

A B S T R A C T

Image fusion is capable of processing multiple heterogeneous images acquired by single or multi-sensor imaging systems for an improved interpretation of the targeted object or scene. A diversity of applications have benefited from the fusion of multi-sensor images through a more reliable and comprehensive fused result. Likewise, numerous approaches to fuse multi-sensor images have been proposed and published in literature. However, due to a lack of benchmark resources and commonly accepted assessment measures, it is hard to identify the significance of new image fusion algorithms and implementations. This paper reviews and categorizes recent algorithms for image fusion and performance assessment based on reported comparative results. We recommend using non-parametric statistical tests to verify the performance of the pixel-level fusion algorithms. Furthermore, a comprehensive evaluation of 40 fusion algorithms from recently published results is conducted to demonstrate the significance of these algorithms in terms of statistical analyses within their respective applications. Although the results of these performance tests are limited by available data sets, baseline algorithms, and selected assessment metrics; it is a critical step for comparative image fusion research. This paper aims to advance image fusion development by creating a complete inventory of state-of-the-art image fusion techniques and advocating statistical comparison tests to avoid unnecessary duplication of development efforts. Establishing a benchmark study for image fusion is critical for performance comparisons of contemporary methods.

1. Introduction

Image fusion has benefited a diversity of applications, including medical diagnosis, security and surveillance, remote sensing, weather forecasting, industrial inspection, and biometrics, etc [1]. A fused image is characterized to provide more reliable information for understanding and perception of the scene or targeted object. Given the varied objectives of specific applications; image fusion, according to [2], is defined as “the process of combining information from two or more images of a scene into a single composite image that is more informative and is more suitable for visual perception or computer processing”. Image fusion algorithms can operate on a single frame or on a video sequence, as illustrated in Fig. 1. In terms of the types of image data, image fusion can be categorized into: 1) temporal image fusion, which fuses images in video time sequences into one meaningful image; 2) spatial image fusion, which stitches images together for a wider field of view; 3) volumetric fusion, which creates a 3D object from image slices; or 4) connotative image fusion, which fuses multi-sensor (e.g., multi-spectral or multi-modal) images to integrate complementary information from inputs into the fused result.

The inputs for *temporal image fusion* are a sequence of images from video or still photographs of different exposure lengths. Temporal image fusion manipulates the temporal dynamic range of the fused images through integrating the details and structures available in the inputs [3]. Pixel-level fusion of images over time was applied to moving object tracking [4]. A temporal series of images can also be used to generate a “clean” scene image of high quality [5,6]. *Spatial image fusion* stitches images into a sharpened panoramic image of the scene even with the blurry inputs [7]. The fused image can give a large field of view or combine the salient results of the input images. The purpose

☆ The R code and data sets used in this study will be available on IEEE DataPort™.

* Corresponding author.
E-mail address: zheng.liu@ieee.org (Z. Liu).

https://doi.org/10.1016/j.inffus.2017.10.010
Received 22 August 2017; Received in revised form 18 October 2017; Accepted 29 October 2017
Available online 04 November 2017

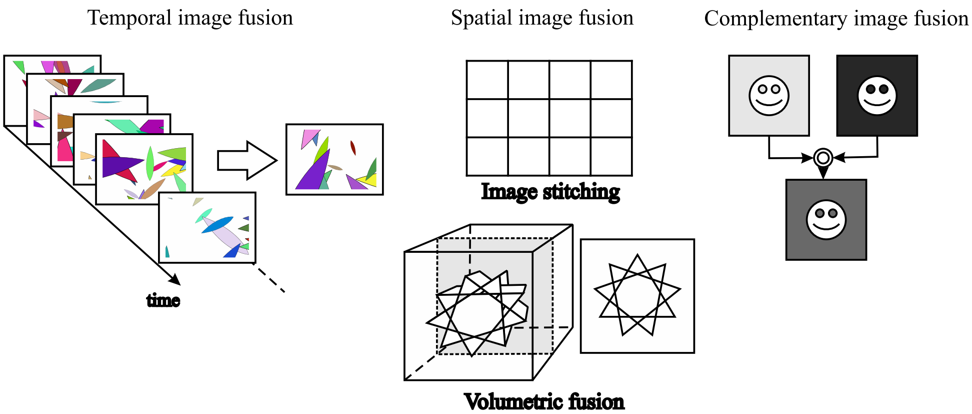


Fig. 1. Types of image fusion: (let) temporal, (middle top) spatial, (middle bottom) volumetric, and (right) complementary fusion.

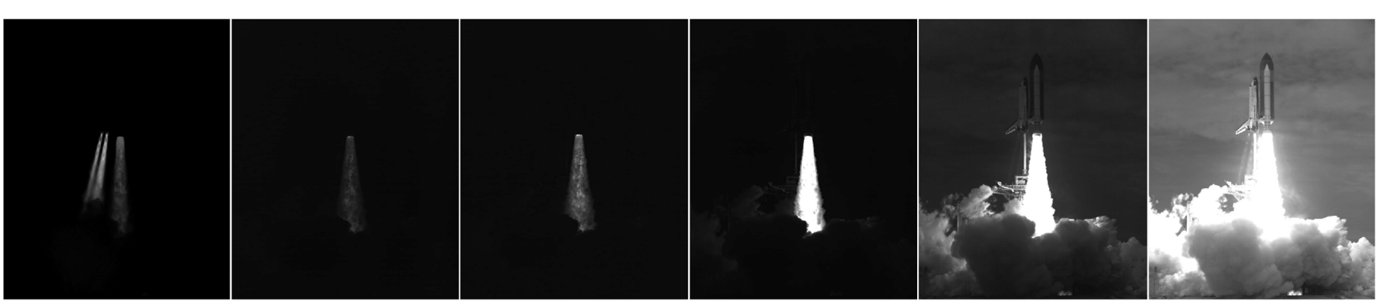
of temporospatial fusion is to achieve a high spatial and temporal resolution simultaneously. In the application of remote sensing, a high-quality temporospatial image analysis is useful for natural resource management and monitoring of land-use and land-cover changes as well as ecosystem dynamics [8]. Examples include sharp panoramas generated from motion-blurred videos through joint global motion estimation and multi-frame deblurring [9]. A closely related topic is image super-resolution, which is beyond the scope of this paper. An excellent summary of the image super-resolution techniques is available in a book [10]. Volumetric fusion targets the reconstruction of the 3D object from 2D image slices [11]. All methods require image registration for the success of reconstruction, where registration of a nonrigid body within an image remains a challenge [12].

Complementary fusion aims at integrating harmonious features or information from the input image. The focus of this paper is to review the methods for fusing images collected by multiple types of imaging sensors. Recording the space shuttle STS-135 launch by NASA provides an example of a highly successful application of image fusion [13]. The camera setup for STS-135 included seven cameras: five visible spectrum black-and-white, a high speed, a high resolution, and two thermal infrared cameras to capture temperature data (bottom left cameras) as shown in Fig. 2(a). One infrared camera did not function during the launch, so only six images in Fig. 2(b) are used to obtain the final fused image. According to [13], with the fused image, NASA researchers can better understand the structure of the plume when rockets fire, the motion of the flames flowing out of the rocket motor, and how to design optimal future motors.

A second example of fusing visual and thermal images is shown in Fig. 3 for context enhancement. The fused image can present both a clear foreground (human being) and background for easy scene interpretation. Other applications include medical imaging, remote sensing, and high dynamic range (HDR) image acquisition, which is achieved by multiple exposure fusion[14]. The most recent review of the state-of-the-art in medical image fusion can be found in [15]. Two reviews for remote sensing are published in literature [16] and [17] respectively. Readers are referred to these references for the details of the specific topics.

This paper focuses on general multi-sensor image fusion algorithms and performance assessment. Currently, the advantages of a fusion algorithm are judged by the improved values of selected fusion

(a)



(b)

Fig. 2. Image fusion for STS-135 space shuttle launch (image credit: NASA/Louise Walker/J.T. Heineck). (a) Left: NASA camera array “Walle”; Middle: STS-135 view without fusion; Right: fused image. (b) Six images taken for STS-135 space shuttle (The first one is taken by a thermal infrared camera).