Infrared Physics & Technology 82 (2017) 8–17

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Regular article

# Infrared and visible image fusion based on visual saliency map and weighted least square optimization

Jinlei Ma [a], Zhiqiang Zhou [a,*], Bo Wang [a], Hua Zong [b]

[a] School of Automation, Beijing Institute of Technology, Beijing 100081, China
[b] Beijing Aerospace Automatic Control Institute, 100854 Beijing, China

## HIGHLIGHTS

- Visual saliency map based rule is introduced to merge the base layers.
- The detail layers are fused through a weight least square optimization scheme.
- The fused image is visually more pleasing and suitable for human perception.
- Different features of IR and visible images are considered to fuse detail layers.

## ARTICLE INFO

*Article history:*
Received 28 October 2016
Revised 22 January 2017
Accepted 16 February 2017
Available online 20 February 2017

*Keywords:*
Multi-scale decomposition
Image fusion
Rolling guidance filter
Visual saliency map
Weighted least square optimization

## ABSTRACT

The goal of infrared (IR) and visible image fusion is to produce a more informative image for human observation or some other computer vision tasks. In this paper, we propose a novel multi-scale fusion method based on visual saliency map (VSM) and weighted least square (WLS) optimization, aiming to overcome some common deficiencies of conventional methods. Firstly, we introduce a multi-scale decomposition (MSD) using the rolling guidance filter (RGF) and Gaussian filter to decompose input images into base and detail layers. Compared with conventional MSDs, this MSD can achieve the unique property of preserving the information of specific scales and reducing halos near edges. Secondly, we argue that the base layers obtained by most MSDs would contain a certain amount of residual low-frequency information, which is important for controlling the contrast and overall visual appearance of the fused image, and the conventional "averaging" fusion scheme is unable to achieve desired effects. To address this problem, an improved VSM-based technique is proposed to fuse the base layers. Lastly, a novel WLS optimization scheme is proposed to fuse the detail layers. This optimization aims to transfer more visual details and less irrelevant IR details or noise into the fused image. As a result, the fused image details would appear more naturally and be suitable for human visual perception. Experimental results demonstrate that our method can achieve a superior performance compared with other fusion methods in both subjective and objective assessments.



## 1. Introduction

Infrared (IR) and visible image fusion plays an important role in military and civilian applications, such as target detection, surveillance and intelligence gathering. IR imaging sensors can capture thermal radiation emitted from objects, which is less affected by dark or adverse weather conditions. However, the acquired IR images usually lack enough background details of the scene. In contrast, visible images usually contain more detail and texture information, and also have higher spatial resolution than the corresponding IR images. IR and visible image fusion can produce a composite image that is more informative for human observation or computer vision tasks.

Multi-scale decomposition (MSD) based fusion methods have been widely studied in the last decades. Image structures of different scales are usually overlapped in the spatial domain. Small-scale structures in an image often correspond to the details and textures, while the large-scale ones may represent the objects in larger sizes. The essence of a multi-scale fusion method is to make the spatially-overlapped features separated in scales firstly with a MSD approach, then the separated features in different scales can be more effectively combined even if they occupy the same spatial area. In last decades, various MSD methods have been employed in

* Corresponding author.
*E-mail address:* zhzhzhou@bit.edu.cn (Z. Zhou).

http://dx.doi.org/10.1016/j.infrared.2017.02.005
1350-4495/

image fusion, such as Laplacian pyramid (LAP) [1], ratio of low pass pyramid (ROLP) [2], morphological pyramid (MOP) [3], discrete wavelet transform (DWT) [4,5], curvelet transform (CVT) [6,7] and nonsubsampled contour transform (NSCT) [8,9]. The great success of multi-scale fusion methods based on various MSDs has confirmed the importance of scale separation in image fusion process.

In recent years, edge-preserving filters have been successfully applied in image fusion. For example, Li et al. [10] propose a fast and effective image fusion method, named guided filtering based fusion method (GFF), in which a novel guided filtering based weight construction method is introduced to combine pixel saliency and spatial consistency. Zhou et al. [11] present a Gaussian and bilateral filter based hybrid multi-scale decomposition method (HMSD) to achieve better fusion results for human visual perception. However, most of the conventional edge-preserving filters tend to smooth details according to their contrasts (i.e., remove low-contrast details first), and take less consideration of their spatial scales in image decomposition, for example, bilateral filter [12], guided filter [13], $L_0$ gradient minimization [14], RTV [15] and WLS [16]. Thus, the corresponding MSD can not achieve a good scale separation for image features, whereas this property would give great benefit to image fusion as we have known from the nature of multi-scale fusion methods. In order to preserve edges while smoothing image structures according to their scales, Zhang et al. [17] propose a scale-aware edge-preserving filter, named rolling guidance filter (RGF). Owing to the useful scale-aware and edge-preserving properties for image fusion, RGF is used to perform the MSD in our method.

The source image can be decomposed into a base layer and several detail layers through the MSD. The base layer mainly contains low-frequency information, which can control the overall look and contrast of the fused image [11]. However, the conventional "averaging" fusion rule for the base layer can not take full advantage of these low-frequency information effectively, usually leading to losing of contrast in the fused image. In order to address this problem, we take use of the visual saliency map (VSM) [18] to merge base layers. The VSM can extract salient structures, regions and objects of an image. The proposed VSM-based fusion rule for the base layer is able to effectively avoid losing contrast and get a better overall appearance for the fused image. The detail layers are usually merged with the conventional "max-absolute" fusion rule, due to the fact that the larger absolute values of the detail layer coefficients correspond to the more salient features. However, the characteristics of IR and visible images are quite different, visible images often contain fine-scale visual detail information, while IR images usually present coarse-scale structures or many incompatible IR details and noise. Simply applying the "max-absolute" rule may transfer less fine-scale details from the visible image and more irrelevant IR details or noise into the fused image. In order to solve this problem, we propose a novel weighted least square (WLS) optimization scheme for the fusion of detail layers. It is based on the different characteristics of IR and visible image features and can selectively fuse detail information from source images.

The rest of this paper is organized as follows. In Section 2, the RGF and Gaussian filter based MSD is presented. We show that this MSD can reduce halos and preserve the information of specific scales. Section 3 describes our IR and visible image fusion algorithm. We propose a VSM-based weighted average technique to fuse base layers and a WLS optimization scheme to fuse detail layers. Experimental results and comparisons are given in Section 4. In Section 5, we conclude this paper.

## 2. MSD based on RGF and Gaussian filter

The Laplacian pyramid, which is constructed with Gaussian filter, is widely used in the field of multi-scale image processing. However, Gaussian filter blurs all edges no matter whether they are strong or weak (Fig. 1 shows an example to illustrate this fact). This property may result in annoying halos near edges in many applications including image fusion. Edge-preserving filters can successfully solve this problem. When smoothing an image, an edge-preserving filter can retain the boundaries of image structures to reduce the halos [10] and maintain the spatial consistency of structures. Thus, the fusion methods based on edge-preserving filters usually achieve better performance [10,11,19]. However, most of the edge-preserving filters preserve edges of image contents only according to their contrasts without considering scales. As a result, the MSDs based on these filters may not be very suitable for the application of a multi-scale image fusion, since they can not well achieve scale separation of spatially-overlapped features. Recently, Zhang et al. [17] propose a state-of-the-art filtering technique named RGF, which has both scale-aware and edge-preserving properties. RGF includes two main steps: small structure removal and edge recovery.

The first step is small structure removal using Gaussian filter. The filtered image $G$ from the input image $I$ can be denoted by

$$G = Gaussian(I, \sigma_s) \tag{1}$$

where $Gaussian(I, \sigma_s)$ denotes the Gaussian filtering with the standard deviation $\sigma_s$ served as the scale parameter. This filter can remove structures whose scales are smaller than $\sigma_s$ as claimed in scale-space theory [20].

The second step is iterative edge recovery using some types of joint filters, such as joint bilateral filter [21], guided filter [13], domain transform [22] and recursive bilateral filter [23]. Here we choose the guided filter as the joint filter because of its high computational efficiency and good edge-preserving property. This step is an iterative process in which the recovered image $J^t$ is iteratively updated and the initial image $J^1$ is the Gaussian smoothed image $G$. The $t$-th iteration can be expressed as

(a) (b) (c) (d) (e)

**Fig. 1.** The source image is repeatedly smoothed (from left to right) using Gaussian filter.