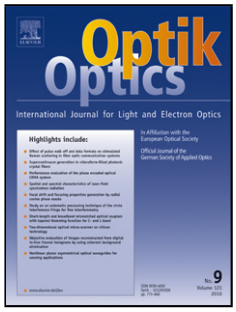

Title: Multi-focus image fusion using deep support value convolutional neural network

Authors: ChaoBen Du, SheSheng Gao, Ying Liu, BingBing Gao

PII: S0030-4026(18)31392-5
DOI: https://doi.org/10.1016/j.ijleo.2018.09.089
Reference: IJLEO 61528

To appear in:

Received date: 24-2-2018
Revised date: 3-6-2018
Accepted date: 17-9-2018

Please cite this article as: Du C, Gao S, Liu Y, Gao B, Multi-focus image fusion using deep support value convolutional neural network, *Optik* (2018), https://doi.org/10.1016/j.ijleo.2018.09.089

Multi-Focus Image Fusion Using Deep Support Value Convolutional Neural Network

ChaoBen Du [a,*], SheSheng Gao [a], Ying Liu [b, c], BingBing Gao [a]

[a] School of Automation, Northwestern Polytechnical University, Xi'an 710129, China

[b] Center for Image and Information Processing, Xi'an University of Posts and Telecommunications, Xi'an, 710121, China,

[c] Key Laboratory of Electronic Information Application Technology for Scene Investigation, Ministry of Public Security, Xi'an, 710121, China

[*] **Corresponding author.**
E-mail address: dcbxjdaxue@163.com (C. Du).

Abstract

A novel multi-focus image fusion algorithm based on deep support value convolutional neural network (DSVCNN) is proposed for multi-focus image fusion. First, a deep support value training network is presented by replacing the empirical risk minimization-based loss function by a loss function based on structural risk minimization during the training of convolutional neural network (CNN). Then, to avoid the loss of information, max-pooling/subsampling of the feature mapping layer of a conventional convolutional neural network, which is employed in all conventional CNN frameworks to reduce the dimensionality of the feature map, is replaced by standard convolutional layers with a stride of two. The experimental results demonstrate that the suggested DSVCNN-based method is competitive with current state-of-the-art approaches and superior to those that use traditional CNN methods.

Keywords: multi-focus image, convolutional neural network, image fusion, decision map

1. INTRODUCTION

In natural images, the imaging equipment usually captures a target image, which includes all the image objects that are effectively captured in focus. In general, by setting the focal length of the optical lens, only the objects in the depth-of-field (DOF) area are clearly visible in the picture, while others are blurred [1]. Consequently, due to the shortcomings of the depth-of-focus (DOF) of optical lenses in charge-coupled device devices, it is difficult to obtain an image where all the relevant objects are effectively captured in focus. To overcome this issue, an image fusion algorithm is introduced in this paper in which multiple source images of the same scene are combined to form a fused image where all the targets of interest are fully focused [2]. One of the basic requirements for image fusion is that all the details should be extracted from multiple source images and preserved in the final fusion image. To some extent, for multi-focus image fusion, only the focused regions in the multi-focus source images need to be preserved perfectly in the final fused image, while all the defocused regions should be completely removed [3].

A variety of multi-focus image fusion algorithms have been proposed over the last decade [2]–[8]. Overall, these methods can be classified into two categories: transform domain and spatial domain methods [2]. In the literature, multi-scale transform (MST) is one of the most popular transform domain methods [4]. Conventional MST image fusion methods include pyramid-based [6], wavelet transform-based [9], curvelet transform-based [10], shearlet transform-based [11] and non-subsampled contourlet transform-based (NSCT) [3], [12] algorithms. Because representation approaches of these image are consistent with the human visual system (HVS); the transform domain methods are it generally considered highly

* Corresponding author.
E-mail address: dcbxjdaxue@163.com (C. Du).