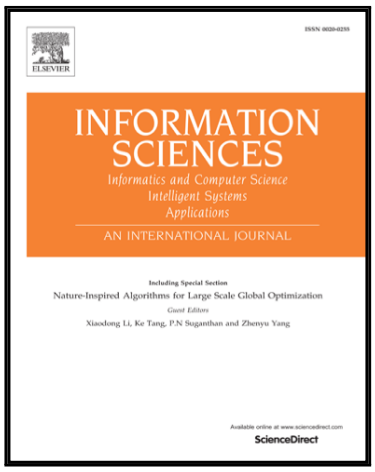

Pixel Convolutional Neural Network for Multi-Focus Image Fusion

Han Tang , Bin Xiao , Weisheng Li , Guoyin Wang

PII: S0020-0255(17)31164-7
DOI: 10.1016/j.ins.2017.12.043
Reference: INS 13337

To appear in: *Information Sciences*

Received date: 15 September 2017
Revised date: 21 December 2017
Accepted date: 25 December 2017

Please cite this article as: Han Tang , Bin Xiao , Weisheng Li , Guoyin Wang , Pixel Convolutional Neural Network for Multi-Focus Image Fusion, *Information Sciences* (2017), doi: 10.1016/j.ins.2017.12.043

Pixel Convolutional Neural Network for Multi-Focus Image Fusion

Han Tang, Bin Xiao*, Weisheng Li, and Guoyin Wang, Senior *Member, IEEE*
Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China;

Abstract: This paper proposes a pixel-wise convolutional neural network (p-CNN) that can recognize the focused and defocused pixels in source images from its neighbourhood information for multi-focus image fusion. The proposed p-CNN can be thought of as a learned focus measure (FM) and provides more efficiency than conventional handcrafted FMs. To enable the p-CNN with the strong capability to discriminate focused and defocused pixels, a comprehensive training image set based on a public image database is created. Furthermore, by setting precise labels according to different focus levels and adding various defocus masks, the p-CNN can accurately measure the focus level of each pixel in source images in which the artefacts in the fused image can be efficiently avoided. We also propose a method to implement the p-CNN with a conventional image convolutional neural network (image-wised CNN), which is almost 25 times faster than directly using the p-CNN in multi-focus image fusion. Experimental results demonstrate that the proposed method is competitive with or even outperforms the state-of-the-art methods in terms of both subjective visual perception and objective evaluation metrics.

Keywords: Multi-focus image fusion, Convolutional neural network, Deep learning, Focus measure

1 Introduction

It is difficult to capture an image in which all the objects are in focus under the condition of a finite depth-of-field. The main reason is that only the objects at a particular distance from the camera are in focus and sharply captured, whereas objects at other distances in front of or behind the focus plane are defocused and blurred [9] [30]. A common method to solve this problem is the multi-focus image fusion technique, which aims at obtaining an all-in-focus image by combining two or more images taken with diverse focal lengths. This all-in-focus image is more suitable for visual perception or computer processing. Thus, multi-focus image fusion can be described as a process that improves the quality of the information in a set of images [35, 43]. In the past few years, several multi-focus image fusion methods had been proposed. According to the fusion strategy, these methods can be broadly classified into two groups [30] [8]: transform domain-based and spatial domain-based methods.

A. Transform domain-based multi-focus image fusion methods

The transform domain-based methods [33] usually employ pyramid decomposition [29] [32], wavelet transform [18] [17] and multi-scale geometric analysis [31] [38] in multi-focus image fusion. These image fusion methods mainly include the Laplacian pyramid (LP) [6], the gradient pyramid (GP) [32], discrete wavelet transform (DWT) [18], dual tree complex

*Corresponding author at: Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China. Tel. /fax: +86 23 62471429. E-mail addresses: xiaobin@cqupt.edu.cn (B. Xiao)