Optics and Laser Technology 109 (2019) 110–122

Contents lists available at ScienceDirect

Optics and Laser Technology

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

Full length article

Photo-realistic continuous digital zooming for an asymmetrical dual camera system

H. Ma, Q. Li, Z. Xu, H. Feng *, Y. Chen

State Key Laboratory of Modern Optical Instrumentation, College of Optical Engineering, Zhejiang University, Hangzhou 310027, China

ARTICLE INFO

Article history:
Received 29 January 2018
Received in revised form 22 May 2018
Accepted 19 July 2018

Keywords:
Asymmetrical dual camera system
Image inpainting
Image restoration
Super-resolution

ABSTRACT

Asymmetrical dual camera systems are becoming more and more popular among many commercial products and they are expected to provide digital zooming images. Existing related algorithms such as super-solution and image fusion fail to reconstruct photo-realistic results because of the sampling rate inconsistency, especially when the focal lengths of cameras vary largely. Nevertheless, magnification factor of super-resolution algorithms are fixed and inflexible as well. By looking into the inherent drawbacks of existing super-resolution algorithm, a new photo-realistic continuous digital zooming algorithm is then proposed in this paper on the basis of a novel super-resolution framework and a band-inpainting algorithm. Experiment result illustrates that proposed algorithm combines the advantages of both super-resolution algorithm and image inpainting algorithm: it reconstructs band information with super-resolution methods, and introduce image inpainting algorithm to reconstruct poorly super-resolved texture areas to produce state-of-the-art photo-realistic digital zooming images.

1. Introduction

Cameras have gradually become standard equipment in commercial applications. More and more sophisticated image processing applications based on cellphone cameras are being developed. In order to appeal the customers' needs, these applications are becoming more demanding on both processors and optical devices than ever before. However, the nature of cellphone largely limits the size or the design of the optical devices. One typical example is that very rare would cellphone use optical zooming camera lens. Optical zooming cameras are expensive and their sizes prevent them from being mounted on a cellphone. As an alternative, dual camera system are developed for many products. Cameras equipped with a short focal lens have large field of view, and cameras equipped with lone focal length can provide abundant high-frequency information. What's more, dual camera system apparently provides depth information which allows many intriguing functions. But such system also introduces a new problem: how to design an algorithm to generate a digital zooming image with both images acquired with a tele-view camera and the image acquired with a wide-view camera.

This paper designs a continuous digital zooming algorithm to generate a zooming image by exploiting both the information provided by a wide-view image and a tele-view image. A high-frequency inpainting algorithm is also designed to reconstruct poorly super-resolved texture regions and yield photo-realistic outputs. This paper is organized as follows: Section 2 introduces the image acquisition model of a dual camera system; Section 3 summarizes the related works done by other researcher; Section 4 describes our proposed algorithm in detail; the experiment results are demonstrated in Section 5 and the conclusion is summarized in the last section.

2. Image acquisition model

An asymmetrical dual-camera system refers to an optical system constituted by two cameras with different optical parameters such as focal lengths and exposure time. A typical dual camera system can be shown as Fig. 1. Such structure is getting more and more popular among commercial products because commercial applications usually emphasize greatly on the size and price of the optical devices. Two cameras with fixed focal length are preferred than a camera with optical zooming lens because fixed-focal cameras are small and economical. The typical structure of a dual camera system is shown in Fig. 1.

Under this circumstance, two fixed-focal cameras are expected to provide digital zooming images competitive with an optical zooming image. There are two problems that undermine the application: the first one is that two images acquired with different

* Corresponding author.
E-mail address: fenghj@zju.edu.cn (H. Feng).

https://doi.org/10.1016/j.optlastec.2018.07.056
0030-3992/

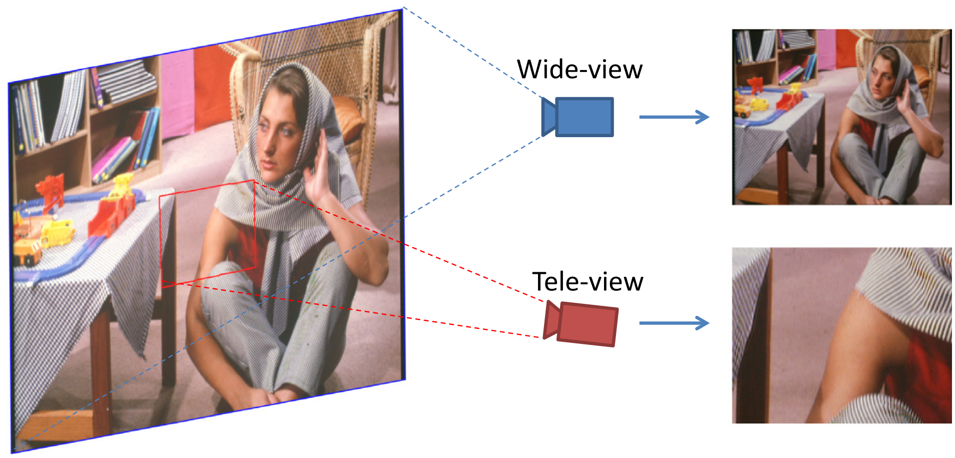

Fig. 1. A typical structure of an asymmetrical dual-camera system. The long focal length camera provides a small field of view with rich details (rich fabric textures), while short focal length camera provides large field of view with poor details (missing fabric textures).

cameras need to be registered properly, and normalized from different lighting conditions; the second one is that lots of high-frequency information is lost in the wide-view image because of its low spatial sampling rate, and an algorithm need to be invented to compensate for its loss in high-frequency component to make it comparable to the tele-view image. In order to solve the first problem, algorithm in [1] can be used to register the images if disparity caused by focal length can be negligible. Otherwise, optical-flow method [2] can be used as well. Difference in lighting, color, distortion, etc. can be compensated with distortion correction algorithm [3] and brightness correction correctness methods [4]. Existing registration methods can be used to register two images. Therefore we can focus on the second problem in the following sections.

3. Related works

The most intuitive thought to tackle the problem is to use super-resolution algorithm to reconstruct the wide-view image. However, despite current super-resolution algorithms have fixed zooming factor, most of them have cost functions of minimizing mean square error, which limits their capability to generate rich high-frequency details. Image inpainting algorithm, on the contrary, can create high-frequency components from nothing, and might create enough details to the image. But naïve image inpainting algorithms cannot exploit the information attained with the wide-view camera.

3.1. Image super-resolution

Most super-resolution algorithms can be classified into multi-frame methods and single-frame methods. Multi-frame super-resolution algorithm takes multiple low-resolution images as input. Input images have small displacements with each other, and once they are registered, the latent unknown high-resolution image can be reconstructed by simple interpolation via registration parameters. The registration process can be executed in frequency domain or spatial domain [5]. However, the performance of multi-frame super-resolution relies heavily on the registration accuracy of different frames. But registration process is very sensitive to noise and different exposures in real application. Hence we focus on the single-frame super-resolution algorithm in this paper.

In many handcrafted state-of-the-art super-resolution algorithms, single frame super-resolution is done in a standard pipeline: features are extracted from the image patches, and regressors are trained to reconstruct the unknown high-resolution image. In the feature extraction process high-pass filters and dimension reduction methods are used first to extract the features from each low-resolution sample [6–9]. Then different regressors, such as L2 regularized regression [6], random forest [7], KSVD [8], or hybrid method [9] are designed to calculate the estimated high-resolution counterparts by minimizing the Mean Square Error (MSE) between reconstructed image and original images. Single-frame super-resolution method aims to utilize the priori information provided by the offline data set, and sharpen the image more naturally with fewer artifacts comparing to traditional image enhancing method. It is observed in the experiment that most MSE-based super-resolution has very good performance in reconstructing strong edges, while provides over-smoothed reconstruction in texture areas. Xu et al. [10] provides an insight view of this phenomenon: strong edges usually have an explicit correspondence of their down-sampled version. Based on statistical learning, the super-resolution algorithm can find this correspondence easily. However, degraded texture regions usually have multiple correspondences and the model thus tends to provide the averaged version of all possible solutions to minimize MSE, which is usually an over-smoothed version. CNN-based methods are also designed to use fully convolutional neural network to produce image with arbitrary size, and reach state-of-the-art performance in terms of Peak Signal Noise Ratio (PSNR) [11,12]. In general, the perceptual performances of top MSE-based algorithms are quite close to each other: they sharpen strong edges and produce over-smoothed texture areas. Algorithm on the basis of Generative Adversarial Network (GAN) [13] also provides a viable solution to the problem rather than minimizing mean square error. Instead of minimizing the Euclidean distance between super-resolved image and the original one, it tries to minimize the distance between the manifold of generated image and the manifold of natural images [14]. This distance is depicted with three terms, and its produced result yields photo-realistic effect. It is also notable that, although results are photo-realistic, the high-frequency component generated by SRGAN is not always faithful. As a result, images produced by SRGAN usually have lower PSNR than MSE-based super-resolution algorithm. Nevertheless, GAN-based algorithms have very high demand on experimental instruments.

3.2. Image inpainting

Image inpainting algorithms are designed to inpaint missing areas, fill in gaps, or remove unwanted objects of an image, etc.