ARTICLE IN PRESS
JID: NEUCOM [m5G;September 28, 2017;8:30]

Neurocomputing 000 (2017) 1–14

Contents lists available at ScienceDirect

Neurocomputing

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

Deep learning based image Super-resolution for nonlinear lens distortions

Qinglong Chang, Kwok-Wai Hung*, Jianmin Jiang

College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China

ARTICLE INFO

Article history:
Received 1 March 2017
Revised 6 September 2017
Accepted 13 September 2017
Available online xxx

Communicated by Jun Yu.

Keywords:
Nonlinear lens distortion
Radial distortion
Super-resolution
Up-sampling

ABSTRACT

Recent developments of virtual reality applications have accelerated the usage of cameras with wide-angle and telephoto/macro lens, which produce nonlinear radial lens distortions, such as barrel distortion and pincushion distortion. However, due to many reasons, the resolution of images with nonlinear lens distortions is often limited. In this paper, we address the image super-resolution (SR) for images with nonlinear lens distortions through the deep convolutional neutral network with residual learning, which can significantly improve the image quality before and after the camera calibration. The proposed deep learning network was trained using hundreds of simulated images and tested on real cameras with fish-eye and macro lens. Experimental results show that the proposed image SR method outperforms state-of-the-art SR methods for various degrees of radial-based barrel and pincushion distortions.

1. Introduction

Image super-resolution has been a popular research topic due to its wide applications [1–12], such as 2D/3D video coding [1,2], scalable video coding [3], inpainting [4], surveillance [5], remote sensing [6–8], multiview synthesis [9,10], etc. For example, due to limited network bandwidth and cost of hardware, online video streaming are often transmitted in 720P or 1080P resolutions, which are required to be super-resolved for displaying on recent 4k displays using state-of-the-art super-resolution algorithms [11,12].

Recently, due to popularity of virtual reality (VR) applications, the panoramic system often makes use of multiple cameras with different field-of-views (FOVs) for stitching/fusing the captured multi-view images to form a panoramic picture [13]. However, due to many reasons, such as the cost of cameras, limited network bandwidth, etc., the resolutions of images captured by fish-eye or telephoto/macro lens with different FOVs are often limited to provide a high resolution panoramic picture after image stitching/fusion. To the best knowledge of authors, there are very limited researches for directly improving the resolutions of images captured by fisheye or telephoto/macro lens [14,16], which inherently generate nonlinear lens distortions, as shown in Fig. 1.

However, the existing super-resolution methods for fisheye cameras [14–16] were designed for one particular lens [14] and required multi-frame image registration (video sequences) for super-resolution [14–16]. On the contrary, there are some researches on super-resolution for images captured by omnidirectional camera, which captures the panorama with one standalone camera [17–20]. Moreover, instead of software post-processing, hardware design of super-resolution system for wide-angle capture was proposed [21]. Super-resolution for wide-angle corneal images captured from human eyes was also investigated [22]. Moreover, to improve the image quality after camera calibration [23,24], the research community suggested a more accurate camera model [25,26], a more accurate rectification process [27], a local planarity and orthogonality constrain [28], etc.

In the literature, there are still a lack of super-resolution algorithms for images captured by both fisheye lens and telephoto/macro lens, which inherently generate nonlinear lens distortions. The difficulty of image super-resolution caused by nonlinear lens distortions is the complex image characteristics due to spatially-varying image resolutions along the radius from the image center. To address this complicated problem for both fisheye and telephoto/macro lens, it is essential to apply the deep learning techniques such as convolutional neural networks with thousands of filter parameters through learning from abundant training datasets.

In this paper, we propose a learning-based single-frame super-resolution method for increasing the resolutions of images with nonlinear radial lens distortions, as shown in Fig. 1. More specifically, we analyze the image formulation model of cameras with nonlinear lens and propose a deep convolutional neural network to learn the explicit end-to-end relationship of original high-resolution images and observed low-resolution images. To simulate

* Corresponding author.
E-mail address: kwhung@szu.edu.cn (K.-W. Hung).

https://doi.org/10.1016/j.neucom.2017.09.035
0925-2312/

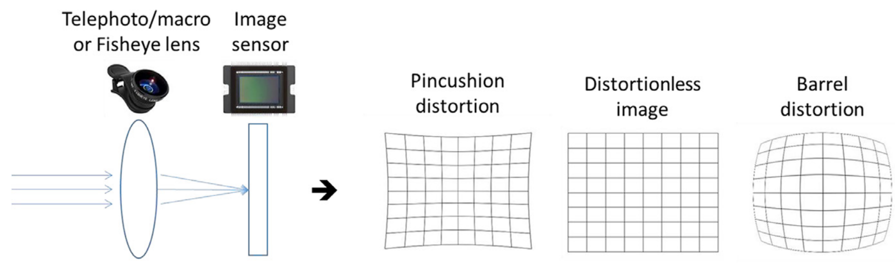

Fig. 1. Cameras with nonlinear lens distortions.

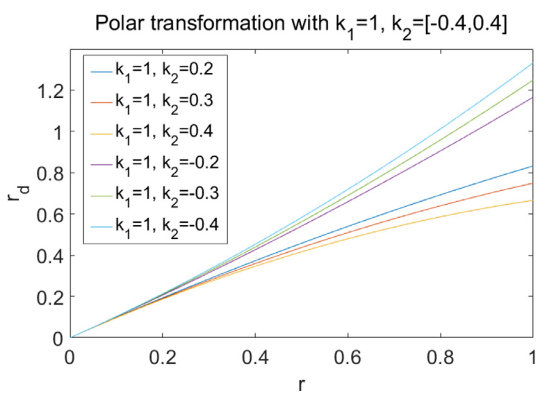

Fig. 2. Relationship between radius r and distorted radius r_d for distortion coefficients $k_1 = 1$, $k_2 = [-0.4,0.4]$.

the nonlinear lens distortions, we adopt a common lens distortion model using the polar transformation [29,30], in order to generate the training samples for network learning. Inspired by the state-of-the-art deep convolutional neural networks [31–33], we analyze the adjustable settings, such as convolutional filter sizes, channels per layer, number of network layers, activation functions, and batch normalization, in order to obtain the suitable network for the single-frame super-resolution for images with nonlinear lens distortions.

For the state-of-the-art super-resolution methods using deep neural networks, such as SRCNN [31], FSRCNN [32], and VDSR [33], these methods were not optimized for images with nonlinear lens distortions. Due to the fact that the degrees of distortion vary from lens to lens, the proposed method further embeds various degrees of nonlinear distortions using a single model, in order to improve the ease of use for different lens. Experimental results were done using both simulated data and real data (camera with fisheye and macro lens) to verify the performance of the proposed super-resolution method before and after camera calibration. Specifically, the proposed method outperforms the state-of-the-art super-resolution methods [31,33,34] in terms of PSNR (0.6–1.03 dB) and SSIM values on average for the simulated data.

The contributions of this paper are summarized as follows:

- The proposed method is the first single-frame image super-resolution algorithm for images with nonlinear lens distortions.
- Convolutional neural networks with various network architectures are analyzed to use a single model for different kinds of lens distortions.
- The resolution of images with nonlinear lens distortions is significantly improved before and after camera calibration for both simulated data and real data.

The rests of the organization of this paper are as follows. Section 2 explains the state-of-the-art deep convolutional networks for generic image super-resolution. Section 3 describes the image formulation model for cameras with nonlinear lens and gives the detail descriptions of the proposed deep convolutional network for images with nonlinear lens distortions. Section 4 shows the experimental results of simulated data and real data. Section 5 gives the conclusion and further discussions of this work.

2. Deep convolutional neural networks for image super-resolution

Deep neural networks have been widely adopted for many image processing applications [35–37], such as image super-resolution [31–33], human pose recovery [38,39], image ranking [40], image privacy protection [41], image recognition [42], image restoration [43], etc. In this section, let us review the state-of-the-art deep convolutional neural networks for generic image super-resolution of images without distortion effects, which aim for a fundamentally different objective from the proposed work.

2.1. SRCNN [31]

SRCNN is the first deep convolutional network for image super-resolution to obtain successful results. Let us denote the convolution layer as $Conv(f_i,\ n_i,\ c_i)$, where the variables f_i, n_i, c_i represent the filter size, the number of filters and the number of channels, respectively. SRCNN uses 3 layers, including *Conv*(9,64,1), *Conv*(5,32,64), *Conv*(5,1,32), as the network structure. For the training process, no adjustable gradient clipping was applied for fast training. For the scales, it uses a standalone model for each super-resolution scale.

2.2. FSRCNN [32]

FSRCNN is an improved and accelerated version of the SRCNN image super-resolution, which is targeted for real-time applications. For the network structure, it uses 8 layers, including *Conv*(5,56,1), *Conv*(1,12,56), 4×*Conv*(3,12,12), *Conv*(1,56,12), *DeConv*(9,1,56), where the de-convolutional layer is denoted as $DeConv(f_i,\ n_i,\ c_i)$. For the training process, no adjustable gradient or residual learning was applied for fast training. For the scales, it shares the first seven convolutional layers for different super-resolution scales.

2.3. VDSR [33]

VDSR is a state-of-the-art image super-resolution method using deep convolutional network. For the network structure, it consists of 20 convolution layers, including *Conv*(3,64,1), 18×*Conv*(3,64,64) and *Conv*(3,1,64). For the training process, the residual learning and adjustable gradient clipping were utilized for fast training. For the scales, it uses a single model for multiple SR scales by embedding training samples with different scales.

Please cite this article as: Q. Chang et al., Deep learning based image super-resolution for nonlinear lens distortions, Neurocomputing (2017), https://doi.org/10.1016/j.neucom.2017.09.035