IRGUN : Improved Residue based Gradual Up-Scaling Network for Single Image Super Resolution

Manoj Sharma, Rudrabha Mukhopadhyay, Avinash Upadhyay, Sriharsha Koundinya,
Ankit Shukla, Santanu Chaudhury.

CSIR- CEERI, Pilani
Rajasthan-333031, India
{mksnith, rudrabha, avinres, sriharsharaja, anktshkla.1307, schaudhury}@gmail.com

Abstract

Convolutional neural network based architectures have achieved decent perceptual quality super resolution on natural images for small scaling factors (2X and 4X). However, image super-resolution for large magnication factors (8X) is an extremely challenging problem for the computer vision community. In this paper, we propose a novel Improved Residual based Gradual Up-Scaling Network (IRGUN) to improve the quality of the super-resolved image for a large magnification factor. IRGUN has a Gradual Upsampling and Residue-based Enhancment Network (GUREN) which comprises of series of Up-scaling and Enhancement blocks (UEB) connected end-to-end and fine-tuned together to give a gradual magnification and enhancement. Due to the perceptual importance of the luminance in super-resolution, the model is trained on luminance (Y) channel of the YCbCr image. Whereas, the chrominance components (Cb and Cr) channel are up-scaled using bicubic interpolation and combined with super-resolved Y channel of the image, which is then converted to RGB. A cascaded 3D-RED architecture trained on RGB images is utilized to incorporate its inter-channel correlation. In addition to this, the training methodology is also presented in the paper. In the training procedure, the weights of the previous UEB are used in the next immediate UEB for faster and better convergence. Each UEB is trained on its respective scale by taking the output image of the previous UEB as input and corresponding HR image of the same scale as ground truth to the successive UEB. All the UEBs are then connected end-to-end and fine tuned. The IRGUN recovers fine details effectively at large (8X) magnification factors. The efficiency of IRGUN is presented on various benchmark datasets and at different magnification scales.

1. Introduction

Single image super-resolution (SISR) is a technique to construct a high-resolution (HR) image from its corresponding lower resolution version. Since a single low-resolution (LR) image has multiple possible higher resolution images, it is very difficult to reconstruct a high-quality HR image. The problem complicates when the scale of super-resolution increases. In recent years, deep learning based super-resolution architectures have performed reasonably well for lower scaling (2X and 4X) ratios. Most of the deep learning based super-resolution (SR) methods use bicubic interpolation to up-sample LR images as a pre-processing step [4, 5, 10–12, 21, 23, 24, 27, 32, 33, 35]. However, operation on these pre-processed images causes additional computational overhead and adds unwanted artifacts. To address these problems [1, 15, 18, 25, 28, 30] has incorporated convolutional transpose layer which reconstructs HR image in one up-scaling step. This sudden up-scaling to a higher magnification factor causes a lot of information loss, resulting in difficulty in training and the model is not able to reconstruct HR images efficiently. To overcome the issues due to sudden up-scaling, [14] has proposed Laplacian pyramid based super-resolution network(LapSRN) to gradually reconstruct the residuals of HR images at each pyramid level. LapSRN uses convolutional transpose layers for up-scaling between two SR scales. Zhao et.al. [38] has proposed a gradual up-scaling network(GUN) which uses multiple level up-sampling and convolutional layer to learn SR process for large scale. Though LapSRN [14] and GUN [38] were addressing the problem of large-scale SR by learning the process gradually in multiple levels, the results were not satisfactory due to which a need for an improved algorithm for 8X SR arises. In this paper, we propose an Improved Residual based Gradual Up-Scaling Network (IRGUN) to learn the mapping between LR and corresponding HR images. Y channel contains