Efficient Image and Video Super-Resolution Using Deep Convolutional Neural Networks

ECE ILLINOIS
Department of Electrical and Computer Engineering

Ding Liu, Zhaowen Wang and Thomas Huang
Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign

ILLINOIS

Abstract

We propose a GPU-parallelized algorithm using deep convolutional neural network for single image super-resolution (SR), in which the restoration of high-resolution image from low-resolution image is accomplished by a deep learning model. . Unlike the traditional sparse coding based method, our deep learning approach parallelizes all the computation steps from end to end, without sacrificing the performance of the current state-of-the-art SR methods. The GPU parallelization accelerates our algorithm significantly, and enables us to build up a real-time video SR system for on-line applications.

Introduction

Image super-resolution is usually cast as an inverse problem of recovering the original high-resolution (HR) image from the low-resolution (LR) observation images. This technique can be utilized in the applications where high resolution is of importance, such as medical imaging, surveillance, satellite imaging and SDTV to HDTV conversion. The main difficulty resides in the loss of much information in the degradation process.

One representative of the current state-of-the-art SR methods is the sparse coding based method[1], which can be easily transformed into a deep convolutional neural network.

Related Work

Sparse-Coding-Based Method [1]

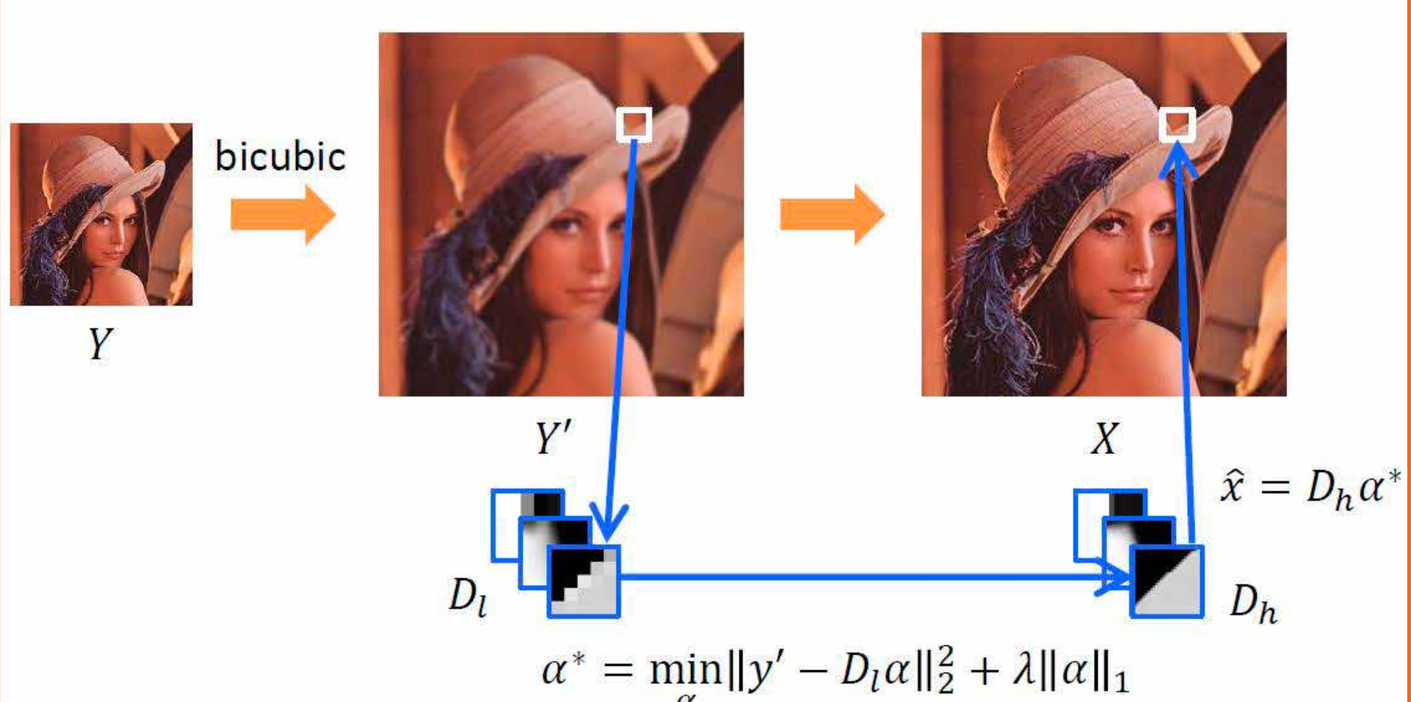
Coupled Dictionary Training

- Learn the transform from low-resolution (LR) feature space to high resolution (HR) feature space
- Find D_h, D_l such that the sparse code of y_i can well reconstruct x_i

$$\min_{\{D_l, D_h\}} \sum_i ||x_i - D_h z_i||_2^2 + \gamma ||y_i - D_l z_i||_2^2$$
$$s.t.\ z_i = \underset{\alpha_i}{\operatorname{argmin}} ||y_i - D_l \alpha_i||_2^2 + \lambda ||\alpha_i||_1$$

Testing

- Find the sparse code α^* for the LR patch y'
- Obtain the estimated HR patch $\hat{x}$ by D_h and α^*



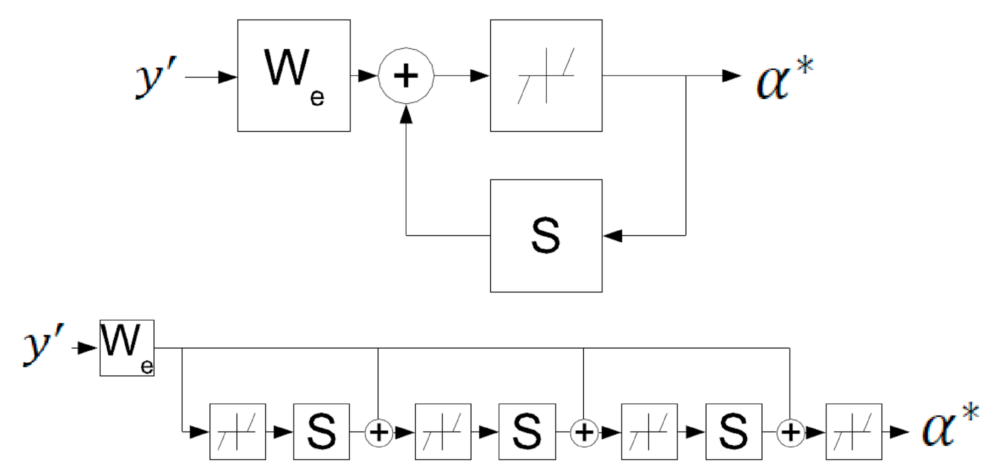
LISTA Algorithm [2]

- ISTA is used to find the solution to the l_1 norm optimization problem in testing phase. It iterates the following recursive equation to convergence:

$$\alpha(k+1) = h_\theta(W_e y' + S\alpha(k)), \alpha(0) = 0$$

where $W_e = \frac{1}{L} D_l^T, S = I - \frac{1}{L} D_l^T D_l$, $L >$ largest eigenvalue of $D_l^T D_l$,

$[h_\theta(V)]_i = \text{sign}(V_i)(|V_i| - \theta_i)_+$ and $\theta_i = \lambda / L$



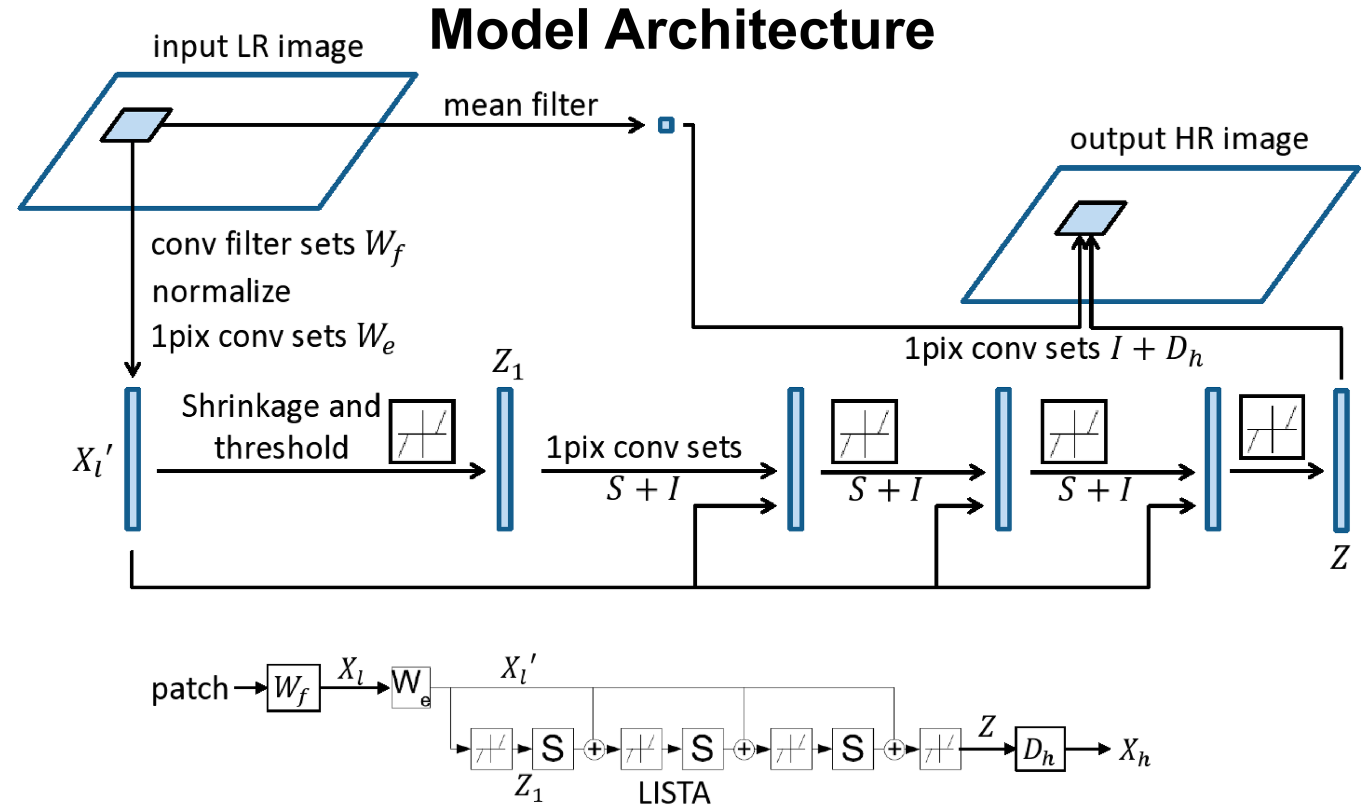
- LISTA takes a fixed number of iterations in ISTA, and W_θ, S, θ are learned in the process which is equivalent to a deep neural network

Method

End-to-end mapping between the LR patch and HR patch:

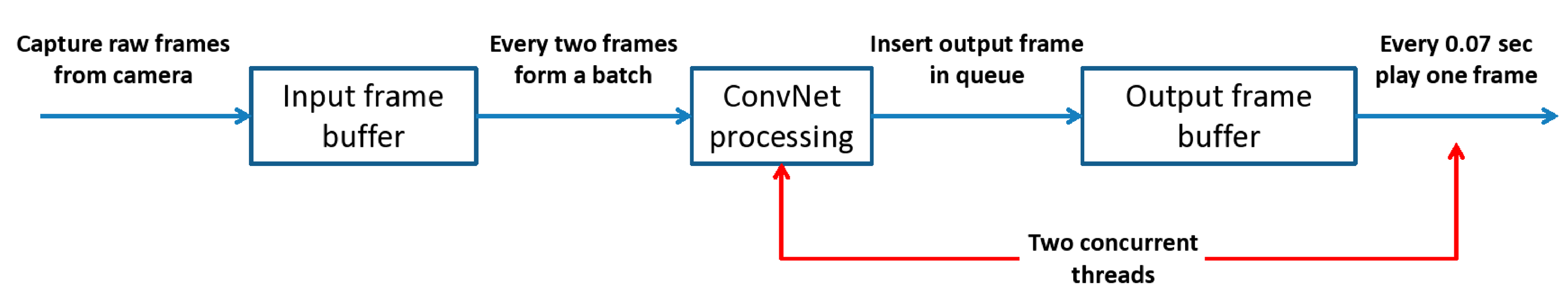
- Extract low-level features from each LR patch
- Utilize LISTA to find the sparse code and multiply it with D_h
- Scale it by the norm and add mean to it
- Average overlapping estimated patches to produce the HR output

Model Architecture

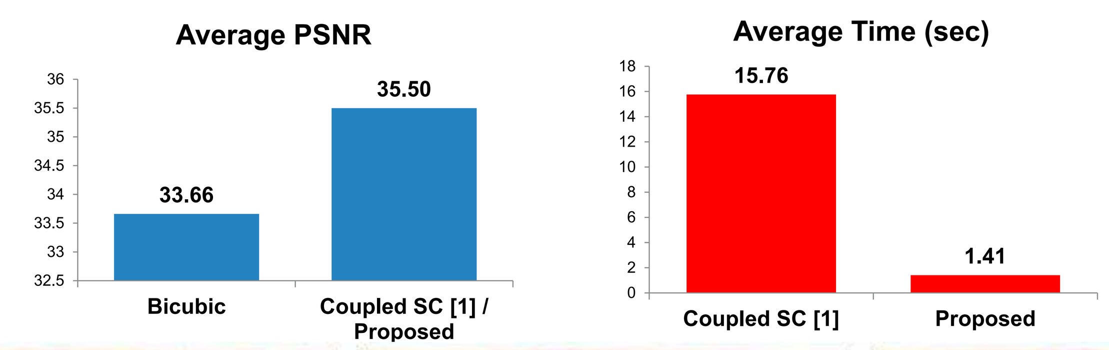


Real-time Video SR system

- Real-time video SR: frame by frame process
- Frame size: 200x250 pixels -> 400x500 pixels
- Frame rate: 0.07 sec / frame (≈14 frames per sec)



Experiments & Results



Ground Truth | Bicubic / 24.04 dB | Proposed / 30.76 dB

Ground Truth | Bicubic / 33.91 dB | Proposed / 37.90 dB

Conclusions & Future Work

A new single image SR method is proposed based on deep convolutional neural network, in which the restoration of HR image from LR image is achieved by a deep learning model. It is implemented at a much higher speed utilizing GPU computing. In the future, parameters in all layers can be trained in order to improve the restoration performance.

References

[1] Yang, Jianchao, et al. "Coupled dictionary training for image super-resolution."*Image Processing, IEEE Transactions on* 21.8 (2012): 3467-3478.
[2] Gregor, Karol, and Yann LeCun. "Learning fast approximations of sparse coding." *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. 2010.