Image Super-Resolution via Sparse Representation

Stanley Chan

Purdue University

ECE/STAT 695 Spring 2018

Image Super-resolution

Fig. 1. Reconstruction of a raccoon face with magnification factor 2. Left: result by our method. Right: the original image. There is little noticeable difference visually even for such a complicated texture. The RMSE for the reconstructed image is 5.92 (only the local patch model is employed).

- Low resolution input
- Want high resolution output

Basic Setup

Model:

$$\boldsymbol{y} = \boldsymbol{SHD\alpha} \tag{1}$$

- $\boldsymbol{y} \in \mathbb{R}^{m\times 1}$: low resolution image
- $\boldsymbol{S} \in \mathbb{R}^{m\times n}$: Sampling matrix
- $\boldsymbol{H} \in \mathbb{R}^{n\times n}$: Blurring matrix
- $\boldsymbol{D} \in \mathbb{R}^{n\times p}$: Dictionary
- $\boldsymbol{\alpha} \in \mathbb{R}^{p\times 1}$: sparse coefficient vector

Questions:

- How to learn dictionary?
- Can we incorporate information from low resolution VS high resolution?
- Can we incorporate feature extraction methods?

Some Accessories 1

Image to Patch:

- Partition $\boldsymbol{y}$ into patches $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N$
- Patch extraction operator $\boldsymbol{R}_i$:

$$\boldsymbol{y}_i = \boldsymbol{R}_i \boldsymbol{y}, \qquad i = 1, \ldots, N$$

- $\boldsymbol{R}_i$ consists of rows of the identity matrix
- Overlap is allowed
- To extract overlap between $\boldsymbol{y}_i$ and $\boldsymbol{y}_{i-1}$:

$$\boldsymbol{w}_i = \boldsymbol{P}_i \boldsymbol{y}, \qquad i = 1, \ldots, N$$

Some Accessories 2

Feature Extraction:

- Extract gradient features of $\boldsymbol{y}_i$
- Gradient filter:

$$\boldsymbol{f}_1 = [-1,\ 0,\ 1], \qquad \boldsymbol{f}_2 = [-1,\ 0,\ 1]^T$$
$$\boldsymbol{f}_3 = [1,\ 0,\ -2,\ 0,\ 1], \qquad \boldsymbol{f}_4 = [1,\ 0,\ -2,\ 0,\ 1]^T.$$

- Define convolution matrices $\boldsymbol{F}_1, \boldsymbol{F}_2, \boldsymbol{F}_3, \boldsymbol{F}_4$.
- Concatenate to make $\boldsymbol{F} = [\boldsymbol{F}_1; \boldsymbol{F}_2; \boldsymbol{F}_3; \boldsymbol{F}_4]$.
- Extracted features:

$$\boldsymbol{F}\boldsymbol{y}_i = \boldsymbol{F}\boldsymbol{D}\boldsymbol{\alpha}_i$$

- $\boldsymbol{f}_j$ $(j = 1, \ldots, 4)$ is a gradient operator. Mean is removed.

Local Representation

- Construct two dictionaries: $\boldsymbol{D}_h, \boldsymbol{D}_l$.
- For $\boldsymbol{D}_l$, do

$$\underset{\boldsymbol{\alpha}_i}{\text{minimize}} \quad \|\boldsymbol{\alpha}_i\|_0 \quad \text{subject to} \quad \|\boldsymbol{F}\boldsymbol{D}_l\boldsymbol{\alpha}_i - \boldsymbol{F}\boldsymbol{y}_i\|_2^2 \leq \epsilon. \tag{2}$$

- $\boldsymbol{F}$ is a feature selection matrix.

This minimization says:

- Want $\boldsymbol{\alpha}_i$ to be sparse.
- Since $\boldsymbol{y}_i$ is low resolution, let us use $\boldsymbol{D}_l$.
- We also want to use the feature selection method. So we use $\boldsymbol{F}$.
- Since $\|\cdot\|_0$ is hard, we can switch to $\|\cdot\|_1$:

$$\underset{\boldsymbol{\alpha}_i}{\text{minimize}} \quad \|\boldsymbol{F}\boldsymbol{D}_l\boldsymbol{\alpha}_i - \boldsymbol{F}\boldsymbol{y}_i\|_2^2 + \lambda\|\boldsymbol{\alpha}_i\|_1 \tag{3}$$

High Resolution

For $i = 1, \ldots, N$, do

$$\widehat{\boldsymbol{\alpha}}_i = \begin{array}{ll} \arg\min\limits_{\boldsymbol{\alpha}} & \|\boldsymbol{\alpha}\|_1 \\ \text{subject to} & \|\boldsymbol{F}\boldsymbol{D}_l\boldsymbol{\alpha} - \boldsymbol{F}\boldsymbol{y}_i\|_2^2 \leq \epsilon_1 \\ & \|\boldsymbol{P}_i\boldsymbol{D}_h\boldsymbol{\alpha} - \boldsymbol{P}_i\boldsymbol{D}_h\widehat{\boldsymbol{\alpha}}_{i-1}\|_2^2 \leq \epsilon_2. \end{array} \tag{4}$$

- $\boldsymbol{D}_h\boldsymbol{\alpha}_i$ is the high-resolution patch
- $\|\boldsymbol{P}_i\boldsymbol{D}_h\boldsymbol{\alpha} - \boldsymbol{P}_i\boldsymbol{D}_h\widehat{\boldsymbol{\alpha}}_{i-1}\|_2^2$ ensures high resolution overlap region is consistent
- Raster scan image from left/right, top/down
- Final output: $\boldsymbol{x}_i = \boldsymbol{D}_h\widehat{\boldsymbol{\alpha}}_i + \mu_i$
- μ_i is the mean of $\boldsymbol{y}_i$

Enforce Global Consistency

First do the above sparse representation reconstruction

$$\widehat{\boldsymbol{x}}^0 = \text{Sparse Respresentation}(\boldsymbol{y})$$

Then do

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \ \|\boldsymbol{SHx} - \boldsymbol{y}\|^2 + \rho\|\boldsymbol{x} - \widehat{\boldsymbol{x}}^0\|^2. \tag{5}$$

- Ensures consistency with physical model
- Original paper: Use gradient descent
- Can be solved in closed-form using polyphase decomposition [Chan et al. 2017]

Face Hallucination

Can do better than natural image:

- Face is special
- Can be pre-trained to yield good initial guesses

Let $\widehat{\boldsymbol{x}}^0$ be initial guess (roughly super resolved image).
For patches $i = 1, \ldots, N$, do

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}_i = \quad & \underset{\boldsymbol{\alpha}}{\arg\min} && \|\boldsymbol{\alpha}\|_1 \\ & \text{subject to} && \|\boldsymbol{F}\boldsymbol{D}_l\boldsymbol{\alpha} - \boldsymbol{F}\widehat{\boldsymbol{x}}_i^0\|_2^2 \leq \epsilon_1 \\ & && \|\boldsymbol{P}_i\boldsymbol{D}_h\boldsymbol{\alpha} - \boldsymbol{P}_i\boldsymbol{D}_h\widehat{\boldsymbol{\alpha}}_{i-1}\|_2^2 \leq \epsilon_2. \end{aligned} \tag{6}$$

- Final output: $\boldsymbol{x}_i = \boldsymbol{D}_h\widehat{\boldsymbol{\alpha}}_i + \mu_i$
- μ_i is the mean of $\widehat{\boldsymbol{x}}_i^0$

Non-negative Matrix Factorization

Let $\boldsymbol{X}$ be a dataset; Each column is a sample patch.

$$(\boldsymbol{U}, \boldsymbol{V}) = \arg\min_{\boldsymbol{U}, \boldsymbol{V}} \|\boldsymbol{X} - \boldsymbol{U}\boldsymbol{V}\|_F^2, \qquad \text{s.t. } \boldsymbol{U} \geq 0, \ \boldsymbol{V} \geq 0. \tag{7}$$

Two standard approaches:

- Alternating minimization
- Multiplicative update

How to use NMF?

$$\boldsymbol{c}^* = \arg\min_{\boldsymbol{c} \geq 0} \|\boldsymbol{S}\boldsymbol{H}\boldsymbol{U}\boldsymbol{c} - \boldsymbol{Y}\|^2 + \eta \|\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{c}\|_2^2 \tag{8}$$

- $\boldsymbol{\Gamma}$ = "high-pass" matrix
- $\boldsymbol{c}$ = coefficient vector

Dictionary Learning

- Given training dataset $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ for $i = 1, \ldots, N$.
- $\boldsymbol{x}_i \in \mathbb{R}^q$, $\boldsymbol{y}_i \in \mathbb{R}^p$.
- Solve

$$\widehat{\boldsymbol{D}}_h = \underset{\boldsymbol{D}_h, \{\boldsymbol{\alpha}_i\}}{\arg\min} \sum_{i=1}^{N} \left\{ \|\boldsymbol{x}_i - \boldsymbol{D}_h \boldsymbol{\alpha}_i\|^2 + \lambda \|\boldsymbol{\alpha}_i\|_1 \right\}$$

$$\widehat{\boldsymbol{D}}_l = \underset{\boldsymbol{D}_l, \{\boldsymbol{\alpha}_i\}}{\arg\min} \sum_{i=1}^{N} \left\{ \|\boldsymbol{y}_i - \boldsymbol{D}_l \boldsymbol{\alpha}_i\|^2 + \lambda \|\boldsymbol{\alpha}_i\|_1 \right\}$$

- How to ensure common sparsity?

$$\underset{\boldsymbol{D}_h, \boldsymbol{D}_l, \boldsymbol{A}}{\text{minimize}} \quad \frac{1}{p}\|\boldsymbol{Y} - \boldsymbol{D}_l \boldsymbol{A}\|^2 + \frac{1}{q}\|\boldsymbol{X} - \boldsymbol{D}_h \boldsymbol{A}\|^2 + \lambda \left(\frac{1}{p} + \frac{1}{q} \right) \|\boldsymbol{A}\|_1.$$

Results

Fig. 5. Results on an image of the Parthenon with magnification factor 3 and corresponding RMSEs. Top row: low-resolution input, Bicubic interpolation (RMSE: 12.724), BP (RMSE: 12.131). Bottom row: NE (RMSE: 13.556), SE [7] (RMSE: 12.228), and our method (RMSE: **11.817**).

Results

Fig. 3. Results of the flower image magnified by a factor of 3 and the corresponding RMSEs. Left to right: input, Bicubic interpolation (RMSE: 4.066), NE [11] (RMSE: 4.888), our method (RMSE: **3.761**), and the original.

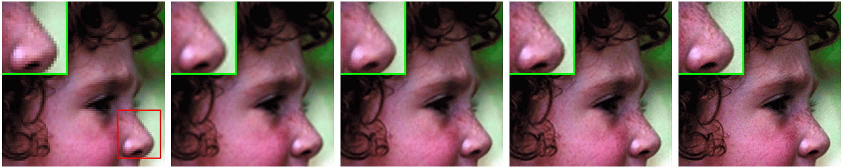

Fig. 4. Results of the girl image magnified by a factor of 3 and the corresponding RMSEs. Left to right: input, Bicubic interpolation (RMSE: 6.843), NE [11] (RMSE: 7.740), our method (RMSE: **6.525**), and the original.

Results

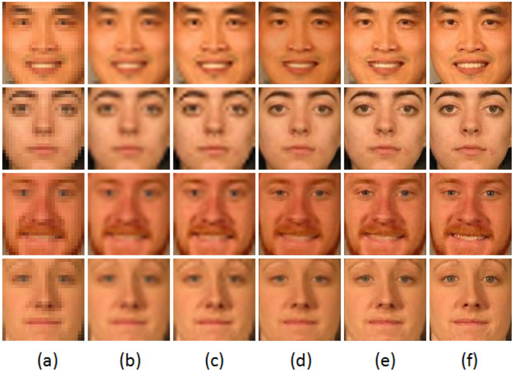

Fig. 8. Results of our algorithm compared to other methods and the corresponding average RMSEs. From left to right columns: (a) low resolution input; (b) Bicubic interpolation (RMSE: 8.024); (c) back projection (RMSE: 7.474); (d) global NMF modeling followed by bilateral filtering (RMSE: 10.738); (e) global NMF modeling and Sparse Representation (RMSE: **6.891**); (f) Original.

Reference

- J. Yang, J. Wright, Y. Huang and Y. Ma, Image Super-resolution via Sparse Representation, IEEE Trans. Image Process. pp.2861-2873, vol. 19, no. 11, Nov. 2010.