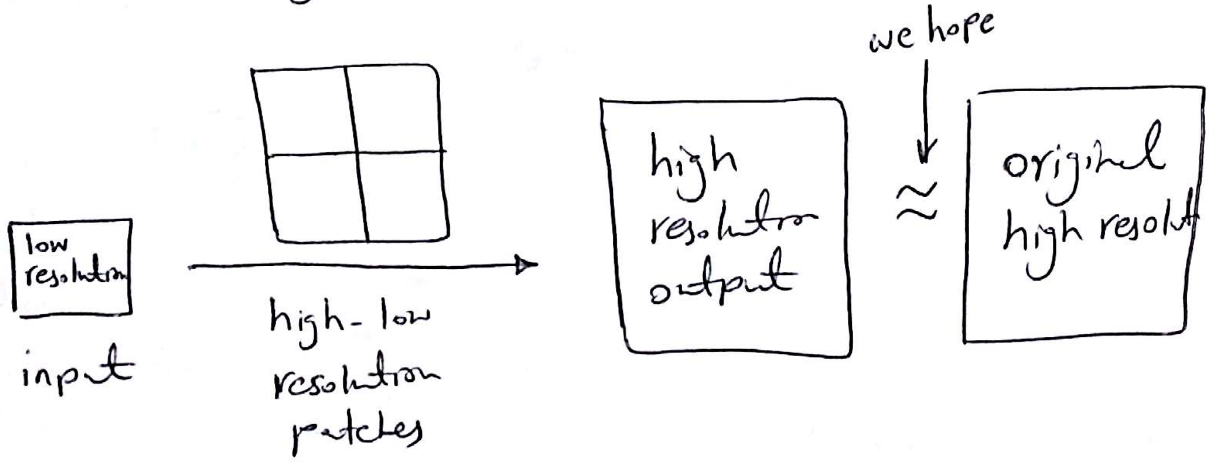
Image Super-resolution via Sparse representation

report by Mohsen Nabian

This paper presents a novel approach toward single image super-resolution based on sparse representation in terms of coupled dictionaries jointly trained from high and low resolution image patch pairs.

Problem: given a single low-resolution input, and a set of pairs (high-and-low-resolution) of training patches sampled from similar images, reconstruct a high resolution version of the input.



Previous works
- Markov Random Field
- Primal sketch prior
- Neighbor embedding
- Soft edge prior

Y

approach of the paper:

high resoulation patches have a sparse linear representation with respect to an overcomplete dictionary of patches randomly sampled from similar images.

$x \in \mathbb{R}^D$ ↳ output high res patch

$D_h \in \mathbb{R}^{D \times n}$ high-res dictionary

$x \approx D_h \alpha_0$ for some $\alpha_0 \in \mathbb{R}^n$ with $\|\alpha_0\|_0 \ll n$

We do not directly observe the high resolution patch but rather features of its low resolution version:

$D_\ell = L D_h \in \mathbb{R}^{d \times n}$

D_ℓ: dictionary of low resolution patches

L: down sampling or blurring operator

The input low resolution patches $y \in \mathbb{R}^d$ satisfies:

$$y = Lx$$
$$\approx L D_h \alpha_0 = D_\ell \alpha_0$$

d linear measurment of sparse coefficient vector α_0!

if we recover the sparse solution α_0 to the under determined system of linear equations $y = D_l \alpha$, we can construct x as $D_h \alpha_0$

Therefore, the problem is formulated as follows:

$$\left\{\begin{array}{l} \hat{\alpha}_0 = \operatorname{argmin} \|\alpha\|_0 \\ \text{s.t. } y = D_l \alpha \end{array}\right. \xrightarrow[\text{relaxation}]{\text{Convex}} \left\{\begin{array}{l} \hat{\alpha}_1 = \operatorname{argmin} \|\alpha\|_1 \\ \text{s.t. } y = D_l \alpha \end{array}\right.$$

$\Downarrow$

Can be solved by Linear programming.

Algorithm Detail:

Combine local estimates:

- Sample 3x3 low resolution patches y on a regular grid.
- Allow 1 pixel overlap between adjucent patches.
- Enforce agreement between overlapping high resolution reconstruction.

So: Simultancy Solution for $\{\alpha\}$ for all patches.
$\hookrightarrow$ large but sparse Convex program $\rightarrow$ still too slow :(

3/

Fast approximation:

Comput α for each patch in raster scan order, enforce consistency with previously computed patch solutions:

$$\begin{cases} \hat{\alpha} = \arg\min \|\alpha\|_1 \\ \text{s.t. } \|Fy - FD_l \alpha\|_2^2 \leq \varepsilon_1 \\ \text{s.t. } \sum_{\alpha'} \|T'(FD_h \alpha') - T(FD_h \alpha)\|_2^2 \leq \varepsilon_2 \end{cases}$$

→ reconstruction constraint

T, T' select overlap between patches

F linear feature extraction operator

overlapping agreement

or $\{ \|PD_h\alpha - w\|_2^2 \leq \varepsilon_2$

- Here, F concatenate first and second image partial derivatives, computed from a low bicubic interpolation of the low-resolution input.

Complete feature vector for each low-resolution patch is 384 dimensional.

4

Super resolution via sparsity Algorithm

Input : training dictionaries D_h and D_ℓ
low resolution image Y.

for each 3x3 patch y of Y, taken in raster scan order with 1 pixel overlap,

- $\hat{\alpha} = \operatorname{argmin} \lambda \|\alpha\|_1 + \frac{1}{2} \|\tilde{D}\alpha - \tilde{y}\|_2^2$

  where $\tilde{D} = \begin{bmatrix} F D_\ell \\ \beta P D_h \end{bmatrix}$ and $\tilde{y} = \begin{bmatrix} F y \\ \beta w \end{bmatrix}$ → reconstruction → overlapping pixels

- place high resolution patch $x = D_h \hat{\alpha}$ in X_0.

end

Using back-projection, find the closest image to X_0 satisfying the reconstruction constraint :

$$\begin{cases} X^* = \underset{X}{\operatorname{argmin}} \|X - X_0\| \\ \text{s.t. } LX = Y \end{cases}$$

output : Super resolution image X^*

S/

Experimental results:

3x3 resolution patches over lap = 1 pixel ← generic images

5x5 " " " = 1 pixel ← face images

A) Single image super-resolution:

D_h and D_L patches are trained from 100,000 patches pairs randomly chosen from natural images from internet.

fix D size to 1024

fix $\lambda = 0.1$

flower image

input Y

Bicubic interpolation

RMSE = 4.066

NE

RMSE = 4.888

our method

RMSE = 3.761

original

orig

girl image

input Y

Bicubic

RMSE = 6.84

NE

RMSE = 7.74

our method

RMSE = 6.525

original

6/

2) Face super-resolution :

face d-b : FRGC Ver 1.0

faces are aligned → croped to 100x100 pxl

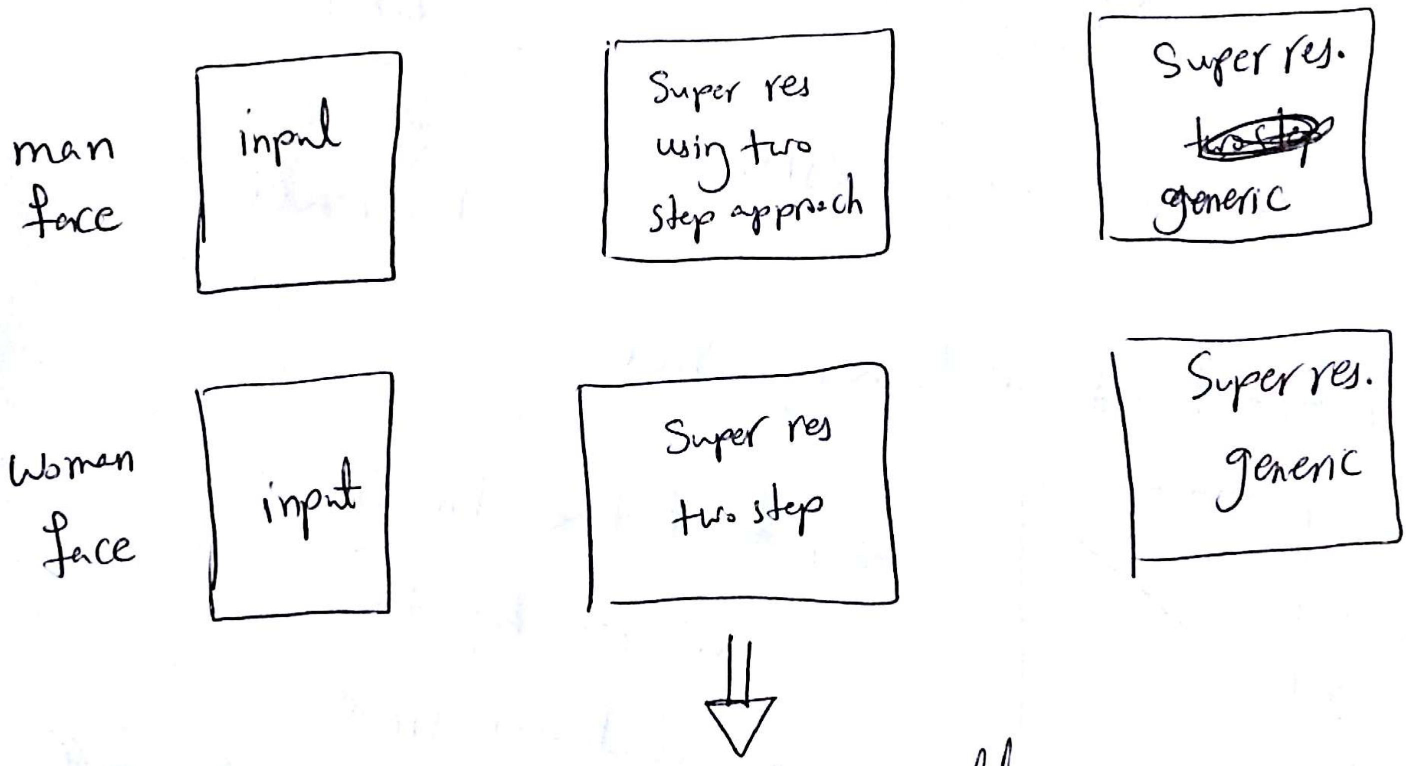
540 images for D_h & D_L

↳ blured D_h

→ 100,000 patches pairs

30 new people (no in training) used for test:

result:



better result with two-step approach

7/

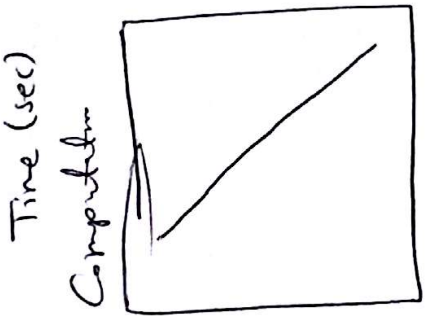
in Fig 8: we compare other methods:

low res input	Bicubic interpolat	back projection	global NMF BLF	global NMF SR	original
	RSME = 8.024	RSME = 7.47	RSME = 10.7	RSME = 6.891	

⇓ best result

we use average RSME.

B) Effect of dictionary size



Computational time linearly increases by the dictionary size

Lena image	D= 256	D= 512	D= 1024	D= 2048	D= whole patch

$\underbrace{\text{D = 2048} \quad \text{D = whole patch}}$ → almost equally well

8/

C) Robustness to noise

different noise levels (σ) = $\{0, 4, 6, 8\}$

Compare $\{$Bicubic, Neighbor embedding, our method$\}$

our method resulted in lowest RMSE

D) Effect of λ

baby image

$\lambda=0.01$ $\lambda=0.05$ $\lambda=0.1$ $\lambda=0.2$ $\lambda=0.3$

$\Downarrow$

Smoother results

$\lambda\uparrow$ Smoother results