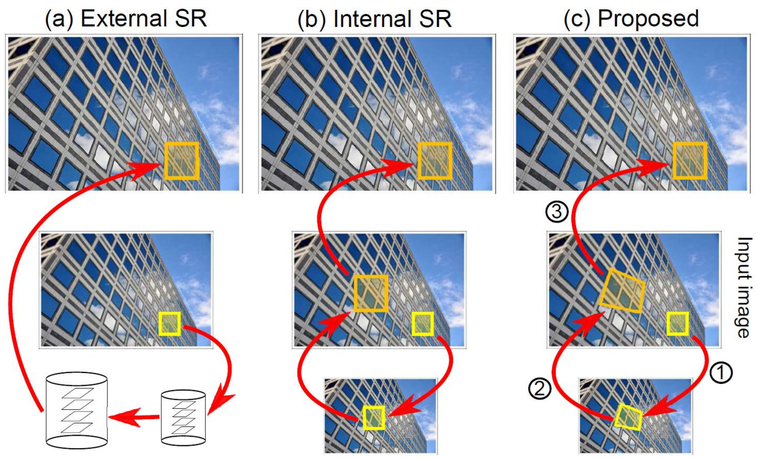
Single Image Super-Resolution using Transformed Self-Exemplars

Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja
University of Illinois, Urbana-Champaign

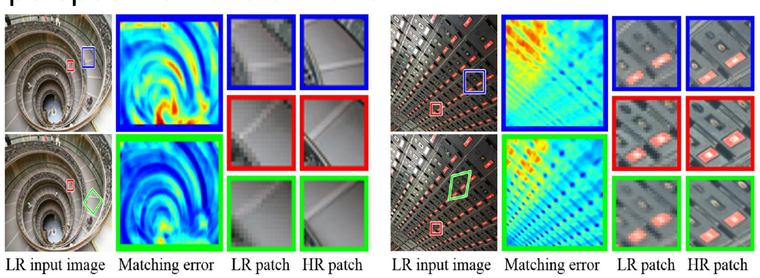
Goal:

Recover high-res image from low-res observation

Basic idea: exploit internal statistics by *transformed* self-exemplars



Searching internal repetition with affine and perspective transformation



Fractal structure of natural images

Project page: http://bit.ly/selfexemplarsr

Method

Nearest neighbor field (NNF) estimation

Objective function

$$\min_{\{\theta_i\}} \sum_{i \in \Omega} \mathbf{E}_{\text{app}}(\mathbf{t}_i, \theta_i) + \mathbf{E}_{\text{plane}}(\mathbf{t}_i, \theta_i) + \mathbf{E}_{\text{scale}}(\mathbf{t}_i, \theta_i)$$

Patch transformation

$$\mathbf{T}_i(\theta_i) = \mathbf{H}\left(\mathbf{t}_i, \mathbf{s}_i^x, \mathbf{s}_i^y, m_i\right) \mathbf{S}\left(\mathbf{s}_i^s, \mathbf{s}_i^\theta\right) \mathbf{A}\left(\mathbf{s}_i^\alpha, \mathbf{s}_i^\beta\right)$$

$\mathbf{H}\left(\mathbf{t}_i, \mathbf{s}_i^x, \mathbf{s}_i^y, m_i\right)$ Perspective transformation [Huang et al. SIGGRAPH 14]

$$\mathbf{S}\left(\mathbf{s}_i^s, \mathbf{s}_i^\theta\right) = \begin{bmatrix} \mathbf{s}_i^s \mathbf{R}(\mathbf{s}_i^\theta) & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} \qquad \mathbf{A}\left(\mathbf{s}_i^\alpha, \mathbf{s}_i^\beta\right) = \begin{bmatrix} 1 & \mathbf{s}_i^\alpha & 0 \\ \mathbf{s}_i^\beta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Inference: PatchMatch algorithm [Barnes et al. SIGGRAPH 09, ECCV 10]

Coarse-to-fine reconstruction

- Each level, perform NNF
- Reconstruct current level via voting
- Run iterative back-projection algorithm to ensure consistency of HR-LR

Quantitative evaluation

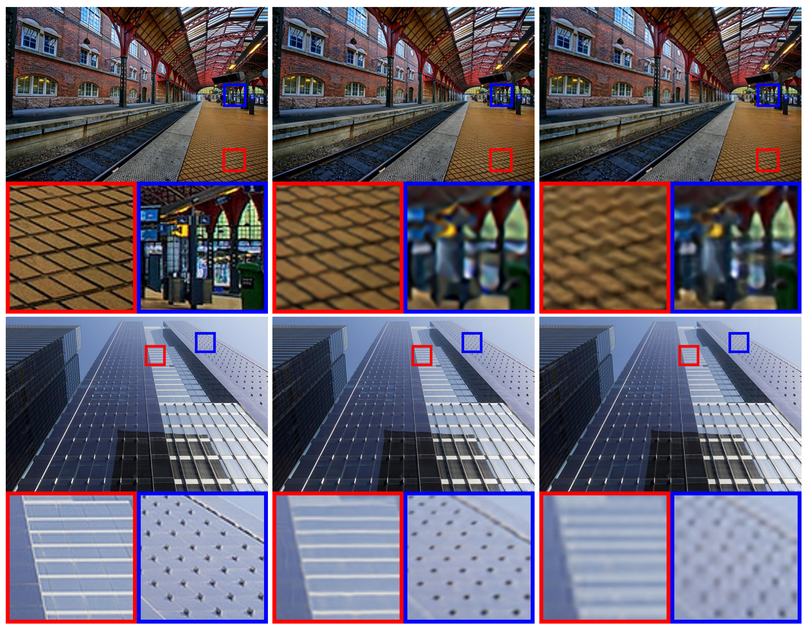
Table 1. Quantitative evaluation on *Urban 100* and *BSD 100* datasets. Red indicates the best and blue indicates the second best performance.

Metric	Scale	Bicubic	ScSR [40]	Kim [22]	SRCNN [9]	A+ [33]	Sub-band [27]	Glasner [15]	Ours
PSNR (*Urban*)	2x	26.66	28.26	28.74	28.65	28.87	28.34	28.15	29.05
	4x	23.14	24.02	24.20	24.14	24.34	24.21	23.79	24.67
SSIM (*Urban*)	2x	0.8408	0.8828	0.8940	0.8909	0.8957	0.8820	0.8743	0.8980
	4x	0.6573	0.7024	0.7104	0.7047	0.7195	0.7115	0.6838	0.7314
PSNR (*BSD*)	2x	29.55	30.77	31.11	31.11	31.22	30.73	30.56	31.12
	3x	27.20	27.72	28.17	28.20	28.30	27.88	27.36	28.20
	4x	25.96	26.61	26.71	26.70	26.82	26.60	26.38	26.80
SSIM (*BSD*)	2x	0.8425	0.8744	0.8840	0.8835	0.8862	0.8774	0.8675	0.8835
	3x	0.7382	0.7647	0.7788	0.7794	0.7836	0.7714	0.7490	0.7778
	4x	0.6672	0.6983	0.7027	0.7018	0.7089	0.7021	0.6842	0.7064

Contributions:

1. Increase the size of internal dictionary using transformed patches
2. Decomposition of perspective and affine transformation.
3. New dataset of urban images

Results on Urban 100 (4x)



Results on Sun-Hays 80 (8x)