Low-Complexity Single-Image Super-Resolution based on Nonnegative Neighbor Embedding

Marco Bevilacqua[1]
marco.bevilacqua@inria.fr

Aline Roumy[1]
aline.roumy@inria.fr

Christine Guillemot[1]
christine.guillemot@inria.fr

Marie-Line Alberi Morel[2]
marie_line.alberi-morel@alcatel-lucent.com

[1] INRIA
Campus universitaire de Beaulieu,
35042 Rennes Cedex, France

[2] Alcatel-Lucent, Bell Labs France
Route de Villejust,
91620 Nozay, France

Single-image super-resolution (SR) is the problem of generating a single high resolution (HR) image, given one low resolution (LR) image as input. In this paper we propose a low-complexity and yet efficient algorithm that reconstruct the HR image in one pass (instead, e.g. of 7 passes for a magnification factor of 4 in [3]). The proposed algorithm falls into the family of *example-based SR*. Taking inspiration from machine learning, it aims at learning the mapping from the LR image(s) to the HR image by using a dictionary: the learning process is performed locally, by trying to infer the HR details through the use of small "examples". For general SR purposes the examples used are patches (sub-windows of image); the dictionary is formed by pairs of LR and HR patches.

In particular, our method adopts the Neighbor Embedding (NE) approach [1, 2], that assumes a local similarities between the LR and HR spaces. For each LR input patch $\mathbf{x}_t^i$, we construct the corresponding HR output by following three steps.

1. *Nearest neighbor (NN) search*: find among the LR patches of the dictionary $\mathcal{X}_d = \{\mathbf{x}_d^j\}_{j=1}^{N_d}$ the K NN.
2. *LR patch estimation*: find a weighted combination of the selected neighbor that approximates $\mathbf{x}_t^i$.
3. *HR patch reconstruction*: keep the same weights with the corresponding HR patches to reconstruct the HR output patch $\mathbf{y}_t^i$.

The whole procedure is carried out in a feature space, i.e. each patch is represented by a vector of features computed on its pixels. In previous NE-based SR algorithms [1, 2], the weights of each linear combination (Step 2) are the result of the following constrained least squares (LS) minimization problem (*SUM1-LS*):

$$\mathbf{w}^i = \arg\min_{\mathbf{w}} \|\mathbf{x}_t^i - X_d^i \mathbf{w}\|^2 \quad \text{s.t.} \quad \mathbf{1}^T\mathbf{w} = 1\,. \qquad (1)$$

In the wake of the NE-based approach, we want to design a low-complexity and competitive algorithm for single-image SR. For this purpose, we analyze three key aspects:

1. the *features* used to represent the LR and HR patches
2. the method used to *compute the weights* of the patch combinations
3. the nature of the *dictionary* (external or "internal")

As for the feature representation, we propose to use centered luminance values (the straight luminance values of the pixels after subtracting the mean value of the patch) as unique features. This choice is in order to have consistency in the representation of the LR and HR patches, and to meet the low-complexity requirement (in fact, we have only one value per pixel). We show that, while keeping the usual *SUM1-LS* as a method to compute the weights, centered luminance values turn out to be good features, as they outperform gradient features (even more costly), if we observe the results for different values of the number of neighbors K.

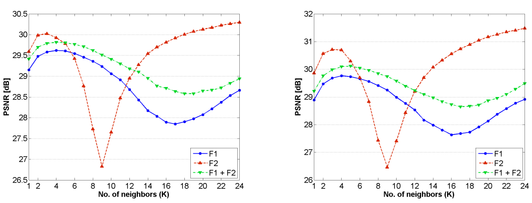

Figure 1: Comparison between LR features (*SUM1-LS*).

However, all the curves present a fall in the performance, that is even dramatic for our selected features. We explain this criticality with the fact that, for certain values of K, the LS solution is over fitted on the LR data and thus generates bad HR reconstruction. To overcome this problem, we propose another strategy to compute the weights, still based on the LS approximation of the input vectors, but with a more relaxed non-negativity constraint. The problem in (1) thus becomes the following non-negative LS problem (*NNLS*):

$$\mathbf{w}^i = \arg\min_{\mathbf{w}} \|\mathbf{x}_t^i - X_d^i \mathbf{w}\|^2 \quad \text{s.t.} \quad \mathbf{w} \geq 0\,. \qquad (2)$$

We show that the combination of centered luminance values and *NNLS* weights gives the best results and presents monotonically increasing PSNR values (Figure 2). As for the dictionary issue, as we want to realize a single-step upscaling without iterating the algorithm for small scale factors, the external training is shown to be the obvious solution. By deriving the patch correspondences from a "self-pyramid" in the way of [3], in fact, we run the risk of having extremely poor (in size) dictionaries. Moreover, the external solution offers the possibility to build a dictionary in advance, so reducing the online running time.

Our method presents much better results than other one-pass algorithms (the original NE-based algorithm of Chang et al. [2] and the Kernel Ridge Regression method of Tang et al. [4]), but it presents comparable results w.r.t. the multi-pass (and therefore more complex) algorithm of [3]. As future work, we plan to investigate other strategies for neighbor search (i.e. other metrics), in order to select "better neighbors" for the HR reconstructions and so improve the performance.

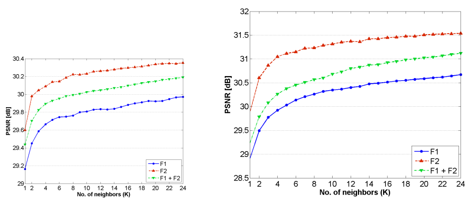

Figure 2: Comparison between LR features (*NNLS*).

		Our algorithm		Chang et al.		Glasner et al.	
Image	*Scale*	*PSNR*	Time	*PSNR*	Time	*PSNR*	Time
bird	2	34.69	18	32.94	110	34.42	406
head	2	32.88	18	32.34	145	32.68	367
woman	2	30.91	15	29.43	114	30.61	410
bird	3	31.37	9	29.71	47	32.16	281
head	3	31.46	12	30.82	68	31.69	370
woman	3	27.98	12	26.45	37	28.79	248
bird	4	28.99	6	27.37	21	30.07	475
head	4	30.26	6	29.57	26	30.86	379
woman	4	25.66	5	24.25	17	26.79	401

Table 1: Results (*PSNR* and running time in sec.) for different images.

Our algorithm   Chang et al. [2]   Glasner et al. [3]

[1] Tak-Ming Chan, Junping Zhang, Jian Pu, and Hua Huang. Neighbor embedding based super-resolution algorithm through edge detection and feature selection. *Pattern Recognition Letters*, 4 2009.

[2] Hong Chang, Dit-Yan Yeung, and Yimin Xiong. Super-Resolution Through Neighbor Embedding. In *2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, 2004.

[3] Daniel Glasner, Shai Bagon, and Michal Irani. Super-Resolution from a Single Image. In *2009 IEEE 12th International Conference on Computer Vision (ICCV)*, 10 2009.

[4] Yi Tang, Pingkun Yan, Yuan Yuan, and Xuelong Li. Single-image super-resolution via local learning. *International Journal of Machine Learning and Cybernetics*, 2011.