NOTES: Computational analysis

Thea Quistgaard

Master Thesis 2020/2021

1 NONPARAMETRIC BLIND SUPER-RESOLUTION T. Michaeli & M. Irani

Super-resolution (SR) is the recovery of a high-resolution image from one or several low-resolution versions. It is recovering information through a thorough analysis of one or more low-res images to piece together the bigger high-res picture.
Usually, SR-methods use a known blur kernel to recover the high-res image, by i.e. knowing the Point-Spread Function (PSF) of the camera that took the picture. But SR will only work if this blur kernel is actually the right kernel to be used - otherwise the new picture will be just as bad or worse than the low-res image.
With blind SR, the blur kernel is assumed unknown often a parametric model is assumed for the kernel to approximate. But blind SR can also assume the kernel is unknown and nonparametric. This means that recovering the high-res image is not the sole objective, because it can not be attempted before the kernel is estimated. This is what nonparametric blind SR focuses on.
Luckily, the frequent recurrence of natural image patches (pixel clusters) is a universal property, both within the same image and in other images, e.g. in an external natural image database. Most important findings:

- A way of estimating the blur kernel k relating high- and low-res images. k is often narrower than the PSF, more oscillatory and can assume negative values.
- k can be estimated from patch recurrence across scales of a single input image. Optimal SR kernel is the oen which maximizes similarity of recurring patches.
- k can also be estimated through coupling with an external database with high-res images, examining recurring patches in both high- and low-res image.
- The algorithm computes the maximum a posterior estimate of the kernel.

1.1 Correct SR kernel

Super-resolution (SR) can be seen as attempting to take a new photograph of the same scene but now with an optical zoom. The zoom-in factor α allow for finer details than that of a single pixel size in the low-res image l. The low-resolution image l is generated from a continuous space image f(a continuous scene (real-life)), and a continuous space PSF b_L (which can be understood as the camera) as:

$$l[n] = \int f(x) b_L(n - x)\, dx \tag{1}$$

Square brackets refer to discrete space and parentheses to continuous.
The, unknown, higher resolution (though not continuous) image h corresponds to a finer grid(more pixels), given by the scaling factor α:

$$h[n] = \int f(x) b_H(\frac{n}{\alpha} - x)\, dx \tag{2}$$

where b_H is the high-resolution PSF. In case of the previously mentioned optical zoom, b_H is a scaled down version of b_L as:

$$b_H(x) = \alpha b_L(x) \tag{3}$$

and the high- and low-resolution images are related a blur and a resampling due to the scaling as:

$$l = (h * k) \downarrow_\alpha \tag{4}$$

By knowing the relation between the discrete high- and low-resolution images exists we can write a discrete description of the relation as

$$l[n] = \sum_m h[m]k[\alpha n - m] \tag{5}$$

where we are on the look-out for the discrete blur kernel k. From equations 1, 2 and 5 we get:

$$\int f(x) b_L(n-x)\,dx = \sum_m \int f(x) b_H(\frac{m}{\alpha} - x)\,dx\,k[\alpha n - m] \tag{6}$$

$$b_L(n-x) = \sum_m b_H(\frac{m}{\alpha} - x) k[\alpha n - m] \tag{7}$$

$$b_L(x) = \sum_m k[m] b_H(x - \frac{m}{\alpha}) \tag{8}$$

From 8 it can be seen that $b_L(x)$ is a linear combination of translated versions of b_H where the linear coefficients constitutes the blur kernel k. Note: If 8 cannot by directly satisfied (i.e. solved), the optimal blur kernel can be found as the projection of b_L onto the space spanned by $\left\{b_H(x - \frac{m}{\alpha})\right\}$ and can be obtained by a least squares solution of 8.

In Figure 2 a one-dimensional visual representation of a naive and a more optimal blur kernel can be seen. The naive blur kernel k_{naive} is a discretization of the one-dimensional rectangular PSF $b_L(x)$, and the optimal blur kernel k_{optimal} may look very different from this - even with negative values and a more oscillatory nature.

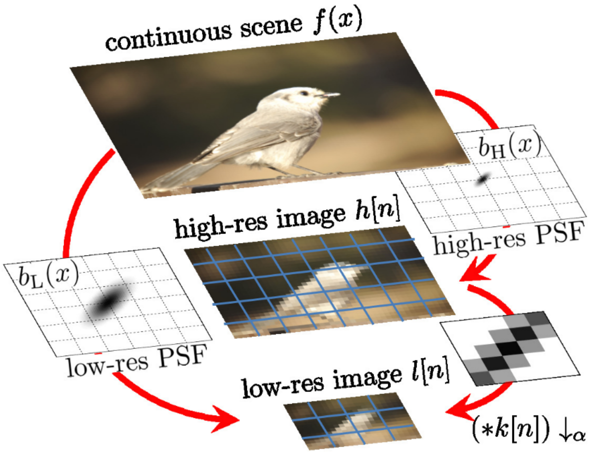

Figure 1

The actual blur kernel k is composed of two operations. A deconvolution with the high-resolution PSF $b_H(x)$ followed by a convolution with the low-resolution PSF b_L. k should thus have the same effect as a continuous-domain filter k_c, whose Fourier transform is:

$$K_c(\omega) = \frac{B_L(\omega)}{B_H(\omega)} = \frac{B_L(\omega)}{B_L(\frac{\omega}{\alpha})} \tag{9}$$

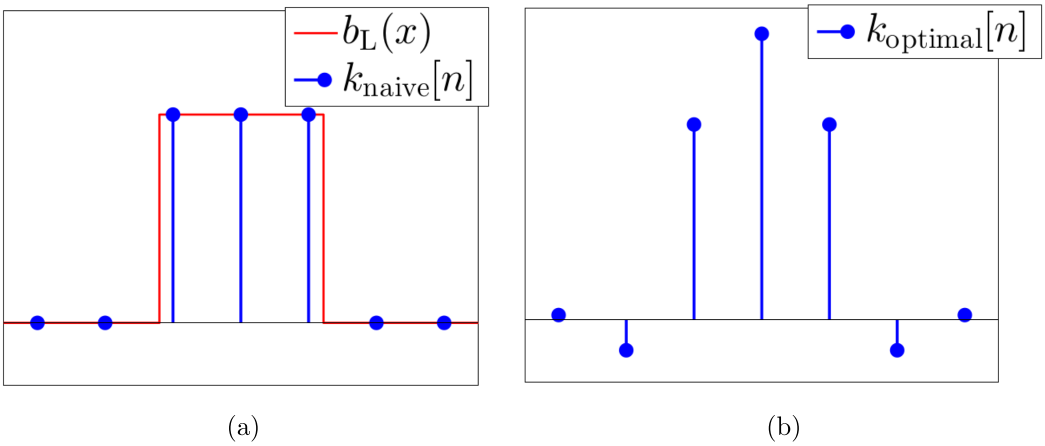

Figure 2: $k_{\mathrm{naive}}[n]$ is a discretization of the low-resolution PSF $b_L(x)$, (2a), but the actual optimal blur kernel $k_{\mathrm{optimal}}[n]$ may be very different, (2b).

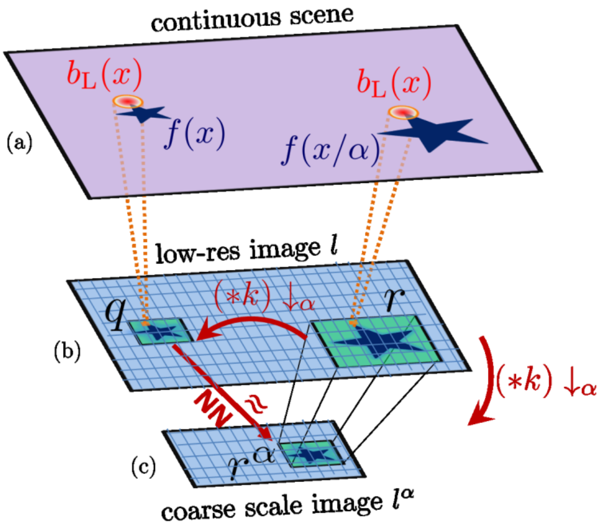

Figure 3