Resolution and its Enhancement in Imaging

Peyman Milanfar et al. *
Department of Electrical Engineering
University of California, Santa Cruz, CA 95064

Abstract

We present a definition and analysis of the concept of resolution and its enhancement in imaging based on statistical detection and estimation. We also present an overview of the problem of Super-resolution in imaging, which (similar to notions in MIMO communications) involves the reconstruction of high resolution images from a collection of "diverse" views of the same scene captured either in the presence of relative motion between the sensor and the scene (synthetic aperture), or with multiple nearby cameras (real aperture).

OCIS Codes:100.0100 (Image Processing); 100.6640 (Superresolution)

1 Statistical Analysis of Resolution

We know that the optics (lens, aperture, etc.) of an imaging system limit the amount of information – more specifically spatial frequency – that is received by the imager (CCD or CMOS device). Furthermore, the spatial sampling of the image at the focal plane by the imager also involves loss of information, particularly if it results in aliasing. While the above observation is nothing new, the same notions when discussed in the context of noisy measurements are less trivial. In particular, one may ask the question "What is the highest resolution one can measure with a camera in the presence of noise?" The popular measures provided to the public in terms of so many mega-pixels may sound wonderful in marketing copy, but are not very informative when we try to quantify the performance of these sensors in the presence of noise.

But what does resolution really mean? To address this question, we take the canonical case of a simple "pinhole" camera in 1-D, composed of a slit and a sampler. Reducing the problem to its basic essence, one can *define* resolution as the ability of the imaging system (or more generally an algorithm acting on an image) to distinguish the presence of two nearby point sources in the presence of noise. In particular, we can define two hypotheses – namely, whether one point source is being observed, or two, where the distance between the point sources is an unknown parameter d. The profile of two point sources imaged through such a system is an incoherent sum of two sinc functions. With proper normalization, in the classical view of resolution according to the Rayleigh limit, the point sources would not be "resolvable" if $d < 1$. This rule of thumb does not *define* the resolution limit of the imaging system. Instead, as we describe below, it is the SNR and the number of samples acquired by the sensor that dictate its ability to resolve.

To be more mathematically precise, we can write the ideal scene comprised of two point sources as:

$$\alpha\,\delta(x-\frac{d}{2})+\beta\,\delta(x+\frac{d}{2}), \tag{1}$$

where in general the distance d, and the amplitudes α and β are unknown. The measured signal is composed of discrete samples of the response of the imaging system to this input (measured at positions x_k for $k = 1, \cdots, N$ on the image plane), and corrupted with additive (readout) noise. That is to say, the measurement model can be described as:

$$g(x_k) = \alpha\,h\left(x_k-\frac{d}{2}\right)+\beta\,h\left(x_k-\frac{d}{2}\right)+w(x_k) \tag{2}$$

*This work is the culmination of efforts of the following members of the MDSP research group at UCSC: Morteza Shahram, Dirk Robinson, and Sina Farsiu, in collaboration with Dr. Michael Elad (CS department, the Technion). This work was supported by the NSF CAREER award CCR-9984246, US Air Force Grant F49620-03-1-0387, and the NSF Center for Adaptive Optics at UCSC. (e-mail contact: milanfar@ee.ucsc.edu) where $w(x_k)$ is assumed to be a zero-mean Gaussian white noise process[1] with variance σ^2. For our specific case of incoherent imaging, $h(x) = \text{sinc}^2(x) = \left[\frac{\sin(\pi x)}{\pi x}\right]^2$ is the point-spread function (PSF) of the system[2]. Posed in terms of the parameter d, the above hypothesis test is equivalent to one that asks whether d is equal to zero (null hypothesis $\mathcal{H}_0$) or not (alternate hypothesis $\mathcal{H}_1$). This "toy" problem is in fact a composite one-sided hypothesis test that leads to an optimum (uniformly most powerful) detector for small separations d– the only case of real interest. A question of great practical value, which is quite similar in nature to those asked by researchers studying communication systems, is *"What is the minimum distance d between two point sources that is detectable with very high probability of detection (say $P_d = 0.99$) and very low false alarm rate (say $P_f = 10^{-6}$) at a given SNR?"* As we have discovered, the answer turns out to be given by a power law [5]:

$$\text{SNR} \approx \frac{c}{N^2 d^4}, \tag{3}$$

where the constant c depends on the shape of the PSF (i.e. $h(x)$) and the selected P_f and P_d, and N^2 denotes the number of samples taken by the imager[3]. The above power law seems to be fundamental to the extent that even with a more complex imaging system model (e.g. inclusion of lenses, and more careful modelling of the spatial sampling process), the scaling law persists.

2 Overcoming Sensor Limitations by Post-processing

Given the above limitations we identified related to the sensor, one can proceed along a path of improved resolution by constructing a better sensor with more complex optics, higher number of pixels, and better SNR per pixel. This approach would likely not be scalable in the face of physical limitations, and economic constraints. An alternative approach is to collect a set of *moving* images of the scene of interest with the sensor at hand, and attempt to extract an enhanced image from this collection of frames.

The problem of resolution enhancement by fusion of multiple images (sometimes called super-resolution or SR) is one which has seen a flurry of recent activity [4]. In our ongoing work, we have addressed critical questions in the theoretical and practical aspects of resolution enhancement by developing a robust estimation framework for this problem [2, 3]. We have also extended this framework to treat the problem of color resolution enhancement [1] for sensors which measure incomplete color information (i.e. single-CCD Bayer-patterned images).

Figure 1 demonstrates a typical example of the aforementioned technology as applied to a "real-world" measured sequence. The success of these techniques in fact prompted us to develop a software package[4]. The SR problem can be modelled as a nonlinear estimation problem as

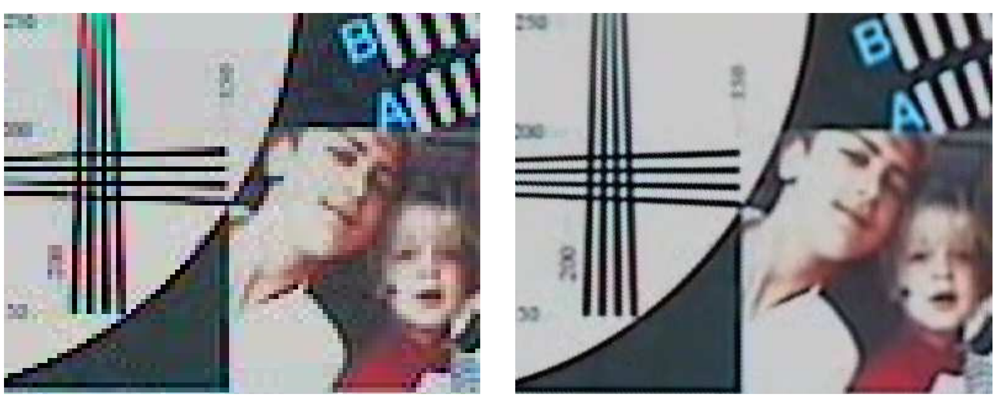

Figure 1: Multi-frame color super-resolution implemented on a real-world data sequence.

$$\mathbf{y}_k = \mathbf{DHF}(\mathbf{v}_k)\mathbf{x} + \mathbf{w}_k \quad \Longrightarrow \quad \mathbf{y} = \mathbf{A}(\mathbf{v})\mathbf{x} + \mathbf{w}, \tag{4}$$

[1]This idealization implies an imaging system that is not photon-limited. This assumption may be relaxed to allow a more general analysis within the same mathematical framework.

[2]The same approach can be applied for other choices of PSF's as well.

[3]Note that the sensor noise variance σ^2 is absorbed into the definition of SNR.

[4]See http://www.soe.ucsc.edu/∼milanfar/SR-software.htm

where $k = 1, \cdots, M$ index the number of available frames. The vectors $\mathbf{y}_k$ here represent the measured images scanned to form column vectors. Likewise, $\mathbf{x}$ represents the unknown high resolution image to be reconstructed (similarly scanned). The matrix $\mathbf{D}$ captures the down-sampling operation, $\mathbf{H}$ the blurring operation due to the imaging system point spread function (PSF), and $\mathbf{F}(\mathbf{v}_k)$ the (translational) motion operator with $\mathbf{v}_k = [v_{1,k},\ v_{2,k}]^T$ being the translation vector. Finally, $\mathbf{w}_k$ represents the vector of additive measurement noise.

It is well-known in practice that even small errors in the values of the motion vector set $\mathbf{v}$ can cause severe and visually disturbing artifacts in the reconstructions. These observations led us to consider robust statistical estimation procedures for the solution of the SR problem. More specifically, we have employed the following penalty function

$$C(\mathbf{x}, \mathbf{v}) = \|\mathbf{y} - \mathbf{A}(\mathbf{v})\mathbf{x}\|_1 + \lambda \sum_{l=0}^{P} \sum_{m=0}^{P} \alpha^{m+l} \|\mathbf{x} - S_x^l S_y^m \mathbf{x}\|_1 \tag{5}$$

where the first term is relating the measurements to the desired image $\mathbf{x}$ through the model we described earlier, while in the second term, S_x^l and S_y^m are operators corresponding to shifting the image by l pixels in horizontal direction and m pixels in vertical direction, respectively. This last operation corresponds roughly to computing "derivatives" across multiple scales. The scalar weight α, $0 < \alpha < 1$, is applied to give a spatially decaying effect to differences of pixels that are farther away.

3 Some Future Directions

Along with the continuing work on identifying the performance limits of sensors and related algorithmic approaches, we have endeavored to use the results of such analysis to design improved imaging systems. We have also worked toward integration of sensing and processing mechanism that will overcome the limitations of the parts by appropriate exchange of information between the sensor and the processor. One such direction we have recently embarked upon is the study of optimal image/video capture from a custom-designed sensor which will adapt its (spatial and temporal) sampling rate to the scene being observed. The aim is to produce a data stream that contains the most information (in terms of details in space and time), and which is most amenable to post-processing in the manner such as the resolution enhancement algorithms we described above. This is ongoing work.

4 References

[1] S. Farsiu, M. Elad, and P. Milanfar. Multi-frame demosaicing and super-resolution of color images. *Submitted to IEEE Tran. on Image Proc.*, 2004.

[2] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar. Advances and challenges in superresolution. *International Journal of Imaging Systems and Technology*, 14(2):47–57, August 2004.

[3] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar. Fast and Robust Multi-frame Super-resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344, October 2004.

[4] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20(3):21–36, May 2003.

[5] M. Shahram and P. Milanfar. Imaging Below the Diffraction Limit: A Statistical Analysis. *IEEE Transactions on Image Processing*, 13(5):677–689, May 2004.