Super Resolution for Automated Target Recognition

Maksim Levental

Abstract—Super resolution is the process of producing high-resolution images from low-resolution images while preserving ground truth about the subject matter of the images and potentially inferring more such truth. Algorithms that successfully carry out such a process are broadly useful in all circumstances where HR imagery is either difficult or impossible to obtain. In particular we look towards super resolving images collected using longwave infrared (LWIR) cameras since high resolution sensors for such cameras do not currently exist. We present an exposition of motivations and concepts of super resolution in general and current techniques, with a qualitative comparison of such techniques. Finally we suggest directions for future research, in particular with applications to LWIR images.

I. INTRODUCTION

Super-resolution (SR) is a collection of methods[1] that augment the resolving power of an imaging system. Here, and in the forthcoming, by resolving power we mean the ability of an imaging device to distinguish distinct but proximal objects in the scene. If such objects are modeled as point sources of light then the resolving power of the imaging system is defined by Rayleigh's criterion: two point sources are considered *resolved* when the first diffraction maximum[2] of one point source (at most) coincides with the first minimum of the other (see figure 1).

SR techniques yield high-resolution (HR) images from one or more observed low-resolution (LR) images by restoring lost fine details and reversing degradations produced by imperfect imaging systems. In the case when a single LR source image is used to construct the HR correspondent, the techniques are referred to as single-image-super-resolution (SISR) techniques. These techniques typically operate by either learning some mapping from low resolution chips (uniform partitions of the image, e.g. 3×3 pixels) to higher resolution chips that are highly similar (according to some metrics) and which obey regularity constraints (e.g. agreement at edges). In the case when multiple LR source images are used to construct the single HR correspondent, the techniques are referred to as multiple-image-super-resolution (MISR) techniques. MISR techniques rely on non-redundant and yet pertinent information in multiple images of the same scene (see figure 2). Note that for such information to exist there should be sub-
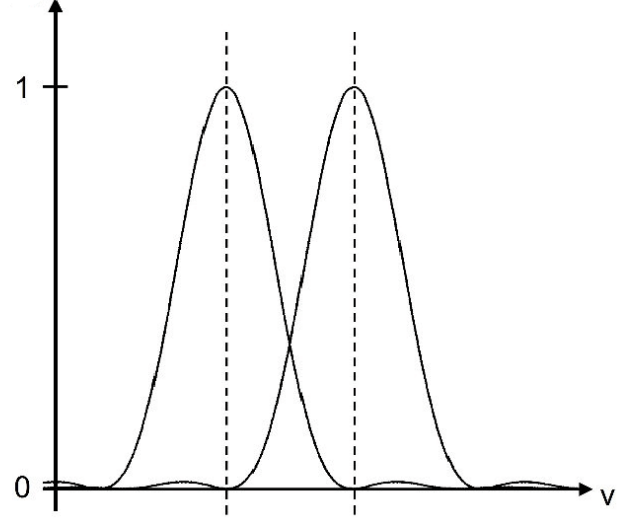
---

[1]We will often use the verb form "to super resolve" in order to denote the use of one or more such methods.

[2]The amplitude of the diffraction pattern (known as the Airy pattern) of a monochromatic point source through a circular aperture is given by

$$I(\theta) = I_0 \left[\frac{2J_1(ka \sin \theta)}{ka \sin \theta} \right]^2$$

where I_0 is peak intensity (at the center), $k = \frac{2\pi}{\lambda}$ is the wave number of the light, θ is the angle of observation, and J_1 is the Bessel function of the first kind of order one[1]. It is maxima/minima of this function that Rayleigh's criterion concerns.

Fig. 1. Rayleigh's criterion[2]

pixel[3] shifts in either the imaging system or the scene between consecutive images.

For typical imaging use-cases, high resolution images are preferable to low resolution images; higher resolutions are desirable in and of themselves and as inputs to later image processing transformations that can degrade image quality (e.g. by virtue of quantization or compression). In theory the resolving power of an imaging system is primarily determined by the number of independent sensor elements that comprise that imaging system (each of which collects a component of the ultimate image). Naturally then, a way to increase the resolution of such a system is to increase the density of such sensor elements per unit area. Unfortunately, and counter-intuitively, since the number of photons incident on each sensor decreases as the sensor shrinks, shot noise[4] thwarts that idea. Furthermore, while sensor density is primary, secondary effects due to optics limit resolution as well; the point spread of a lens (distortion of a point source due to diffraction), chromatic aberrations (distortion due to differing indices of refraction for differing wavelengths of light), and motion blur all function to obscure or erase details from the image.

In domains such as satellite/aerial photography, medical imaging, and facial recognition, high-resolution reconstruction of low-resolution samples is eminently useful since ab-initio acquisition of high-resolution images is either logistically difficult or impossible due to aforementioned imaging apparatus

---

[3]For example when a point source wholly captured by one sensor element shifts to distributing energy equally amongst the same element and a direct adjacent.
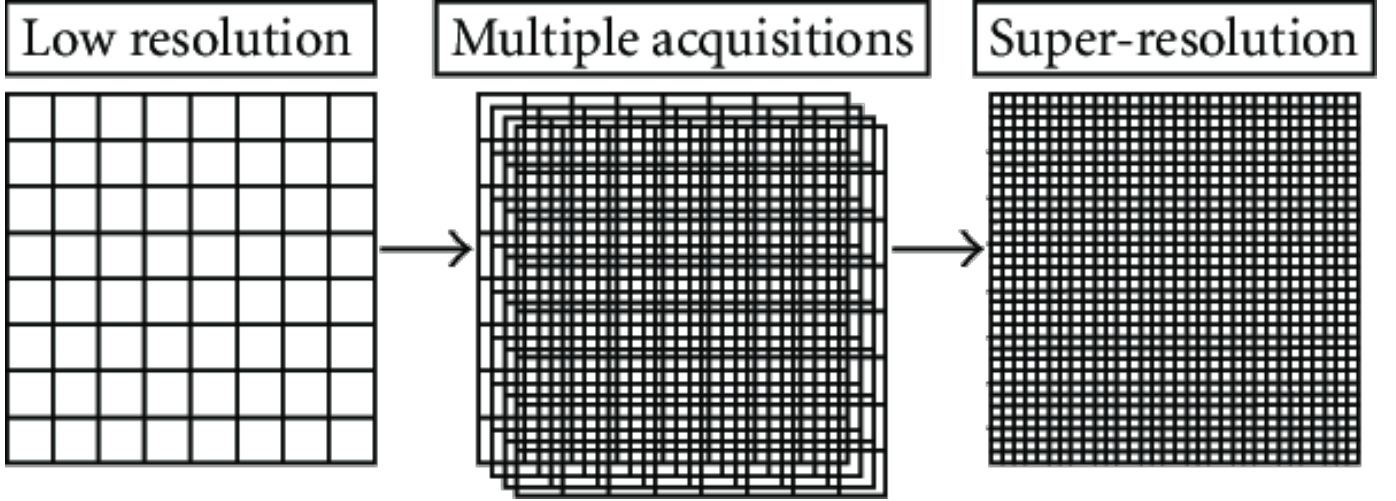
[4]TODO

Fig. 2. Multiple image super resolution[3]

limitations. For example in the instance of satellite imagery, acquisition of high-resolution imagery is primarily hampered by optics and physics[5]. In contrast, in the cases of medical imaging (where procedures are invasive and patient exposure time needs to be minimized[6]) and facial recognition (e.g. for purposes of surveillance) the primary challenge is logistics and access to repeat collection opportunities.

The benefits of enhancing images using SR techniques include not only more pleasing or more readily interpretable images for human consumption but higher quality inputs for automated learning systems as well. In particular object detection systems trained on super-resolved images outperform those trained on the low resolution originals[7]. Indeed this is our ultimate goal - not super-resolution per se but super-resolution in the service of improved object detection performance. Note that while practically speaking, there exist hardware and software solutions for increasing the resolution of an imaging system, we, owing to a "Ship of Theseus" consideration, discount such propositions. We instead take low resolution images as given and seek techniques that allow for ex post facto reconstruction or inference of precise details. This necessarily constrains techniques under consideration to be algorithmic in nature and software in practice.

The rest of this survey is outlined as follows: Section II introduces notation and the model of imaging that will be the mathematical framework for the proceeding sections, Section III surveys classical techniques (those that do not employ neural networks), Section IV surveys neural-network techniques with heavy emphasis on deep learning (i.e. deep networks), Section V discusses the scope and goals of the author's research program, and Section VI summarizes.

II. Background

III. Classical Algorithms

IV. Deep Learning Algorithms

V. Future Research

VI. Conclusion

VII. Appendix

TODO: work out diffraction circular aperture TODO: work-out poisson noise

---

[5]Rayleigh's criterion implies that the angular resolution R of a telescope with optical diameter $D = 2.4$m observing visible light ($\sim$500nm) is approximately[4]

$$R \approx 1.220\frac{\lambda}{D} = 1.220\frac{500\text{nm}}{2.4\text{m}} \approx 0.06\text{arcsec}$$

From an altitude of 250 km this corresponds to a ground sample distance of 6cm. This loss of resolving power is further exacerbated by refraction through turbulent atmosphere[5].

References

[1] J. W. Goodman, "Fresnel and fraunhofer diffraction," *Introduction to Fourier Optics, 3rd ed., by JW Goodman (pp. 76-78). Englewood, CO: Roberts & Co. Publishers, 2005*, vol. 1, 2005.

[2] P. Scholz, "Focused ion beam created refractive and diffractive lens techniques for the improvement of optical imaging through silicon," PhD thesis, Jul. 2012. DOI: 10.14279/depositonce-3270.

[3] J. Kennedy, O. Israel, A. Frenkel, R. bar shalom, and H. Azhari, "Improved image fusion in pet/ct using hybrid image reconstruction and super-resolution," *International journal of biomedical imaging*, vol. 2007, p. 46 846, Jan. 2007. DOI: 10.1155/2007/46846.

[4] L. R. F.R.S., "Xxxi. investigations in optics, with special reference to the spectroscope," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 8, no. 49, pp. 261–274, 1879. DOI: 10.1080/14786447908639684. eprint: https://doi.org/10.1080/14786447908639684. [Online]. Available: https://doi.org/10.1080/14786447908639684.

[5] D. L. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," *J. Opt. Soc. Am.*, vol. 56, no. 10, pp. 1372–1379, 1966. DOI: 10.1364/JOSA.56.001372. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=josa-56-10-1372.

[6] E. Van Reeth, I. W. K. Tham, C. H. Tan, and C. L. Poh, "Super-resolution in magnetic resonance imaging: A review," *Concepts in Magnetic Resonance Part A*, vol. 40A, no. 6, pp. 306–325, 2012. DOI: 10.1002/cmr.a.21249. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cmr.a.21249. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cmr.a.21249.

[7] J. Shermeyer and A. V. Etten, "The effects of super-resolution on object detection performance in satellite imagery," *CoRR*, vol. abs/1812.04098, 2018. arXiv: 1812.04098. [Online]. Available: http://arxiv.org/abs/1812.04098.