

Super-resolution in computational imaging

M. Bertero* P. Boccacci

INFN and DISI, University of Genova, Via Dodecaneso 35, Italy

Abstract

Super-resolution is a word used in different contexts but mainly in the case of methods aimed at improving the resolution of an optical instrument beyond the diffraction limit. Such a result may be achieved by means of specific instrumental techniques (such as, for instance, interferometry) or by means of a suitable processing of a digital image; in the latter case we will use the expression *computational super-resolution* (CSR). In this paper we describe the basic concepts underlying CSR without using the mathematics required for establishing its theoretical validity. The aim is to introduce a wide audience to this topic, to specify the situations where CSR is feasible and to emphasize the point that unlimited CSR is not possible.

Key words: Resolving power, super-resolution, out-of-band extrapolation

1 Introduction

According to geometrical optics the image of a point source is a perfectly sharp point. Such an approximation however does not apply to the most powerful optical instruments, such as microscopes and telescopes, where the effects due to the diffraction of the light by the boundaries of the pupil become important. As a result the image of a point source is no longer a point but rather a small bright patch with a well-defined shape, which is sometimes called the *diffraction pattern* of the source. In the cases where the isoplanatic approximation holds true, the shape of the pattern does not depend on the position of the point source and its intensity distribution as a function of the coordinates in the image plane is called the *point spread function* (PSF) of the optical instrument.

* Corresponding author. Tel: +39-010-3536733; fax:+39-010-3536699
Email address: bertero@disi.unige.it (M. Bertero).

Such a property has generated the concept of *resolving power* of an optical instrument which dates back to the classical works of Abbe (1873) and Rayleigh (1879) at the end of the XIX-th century and is related to the overlap of the images of two close point sources. For instance, the spatial resolution of a microscope is defined as the minimum distance between two point sources such that their images are distinguishable according to some specified criterion; analogously the angular resolution of a telescope is defined as the minimal angular separation between two distinguishable distant stars. A widely used criterion is the classical one due to Lord Rayleigh and the so-called Rayleigh limit, resulting from this criterion, is related to the width of the PSF of the optical instrument. In this context the term *super-resolution* is currently used for denoting instrumental or computational methods which allow to reach a resolution beyond the Rayleigh limit.

Since the advent of digital images and of powerful computers, the problem of improving the resolution by means of a suitable processing of the recorded images has been investigated. In this paper we focus on this kind of super-resolution and we do not consider that based on instrumental methods (such as apodization, interferometry or others). We call the corresponding topic *computational super-resolution* (CSR) since we intend it as a specific topic of computational imaging or, more precisely, of object restoration. The latter, also known as image restoration or image deconvolution, consists of all numerical methods designed to correct for the degradation of the image produced by known defects (described by the PSF) of the optical instrument, such as all possible kinds of aberrations. An introduction to this topic is given in Bertero & Boccacci (1998) (see also Boccacci & Bertero (2002)). Super-resolution is even more ambitious than that since it pretends to correct for the degradation due to diffraction.

The idea of overcoming the Rayleigh limit is quite old and has been discussed in papers by Toraldo di Francia (1955), Wolter (1961), Harris (1964), McCutchen (1967), Rushforth & Harris (1968), and others. The basic remark suggesting the possibility of CSR is clearly explained in the papers by Wolter and Harris mentioned above. However the fundamental limitations intrinsic to the suggested procedure were not completely understood so that the dream of almost unlimited CSR seemed to come true. Nowadays we know that CSR is feasible only to a limited extent and only in rather specific cases.

This paper is intended to provide an introduction to this topic in a way which should be accessible to a wide audience. To this purpose we have avoided as far as possible the use of advanced mathematical tools, even if it is obvious that mathematical concepts cannot be completely ignored in a topic which consists of suitable mathematical manipulations of digital images. The interested reader can find an account of the mathematical methods used for CSR in Bertero & De Mol (1996).

The paper is organized as follows. In Section 2 we introduce two complementary descriptions of an optical instrument (typically a microscope or a telescope), the so-called frequency-domain and space-domain description. In the frequency domain the instrument acts as a filter, as a consequence of diffraction, and this effect is described as a blurring in the space domain. The important concepts of band-width and resolution distance are introduced and their relationship is discussed. These concepts are basic for understanding the content of Section 3 where the feasibility of CSR is related to the possibility of extrapolating the spectrum of the object outside the band of the instrument. On such a basis the fundamental limitations in the amount of CSR are discussed. In addition it is clarified that the possibility of achieving CSR is related to the possibility of designing methods for object restoration with some specific properties. An example of these methods, very easy to implement, is described in Section 4 where the results of a few numerical simulations are also given. The conclusions derived from our experience in this topic are summarized in Section 5.

2 Filtering and resolution

Let us assume, for simplicity, a planar object placed in the focal plane of an optical instrument. The extension to 3D objects is, in principle, straightforward even if the discussion of such a case is obviously much more complex.

As a consequence of our assumption the object can be described by means of a function of two space variables x, y (the coordinates of a point in the focal plane) representing, for instance, the intensity of the light emitted by the points of the object. We denote this function by $O(x, y)$ and analogously we denote by $I(x, y)$ the function describing the image.

A basic point is that any function with arbitrary shape can be represented as a superposition of cosine and sine functions such as $\cos(\omega_x x + \omega_y y)$ and $\sin(\omega_x x + \omega_y y)$ with all possible values of the variables ω_x, ω_y . The latter are called *space frequencies* and, in the case of a microscope, are measured in μm^{-1} if x, y are measured in μm . The precise representation is given by the classical *Fourier integral* (see, for instance, Papoulis (1968)) while the coefficients appearing in the superposition of trigonometric functions constitute a function of the space frequencies ω_x, ω_y , which is called the *Fourier transform* (FT) of the function we are considering. If the function is $O(x, y)$ then its FT will be denoted by $\hat{O}(\omega_x, \omega_y)$. The characterization of the object by means of $O(x, y)$ is called its *representation in the space domain* while its characterization by means of $\hat{O}(\omega_x, \omega_y)$ is called its *representation in the frequency domain*.

In Figure 1 we give examples of 1D signals and of the corresponding FTs. It

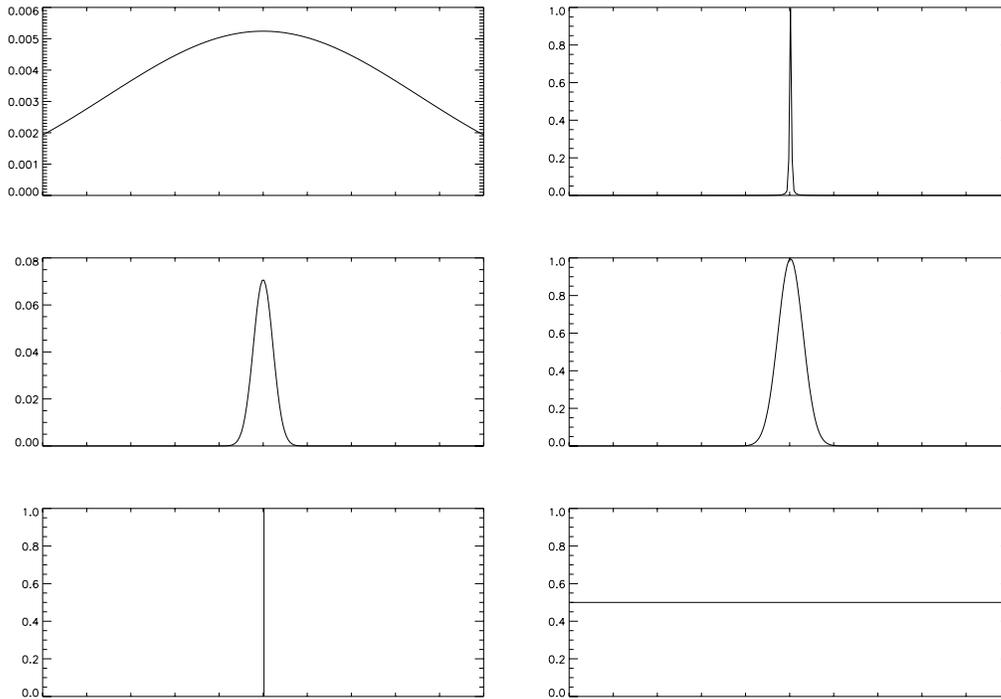


Fig. 1. Examples of Fourier transforms of 1D signals: on the left we plot the functions and on the right the corresponding Fourier transforms

is evident that the sharper is the signal, the broader is the corresponding FT. In the limiting case of an impulse signal the FT is constant, namely the signal contains all space frequencies with equal weights.

A first basic property of an optical system, at least in the isoplanatic approximation, is that it transforms a trigonometric signal into another trigonometric signal with the same space frequencies but a different amplitude and, possibly, a different phase. For simplicity we neglect the latter effect. Then the ratio between the amplitude of the input signal and that of its image is called the *modulation transfer function* (MTF) of the optical system. The MTF is a function of the space frequencies and characterizes the imaging properties of the system in the frequency domain. In this paper it is denoted by $\hat{T}(\omega_x, \omega_y)$ and therefore the relationship between the FT of the image and that of the object is given by

$$\hat{I}(\omega_x, \omega_y) = \hat{T}(\omega_x, \omega_y) \hat{O}(\omega_x, \omega_y). \quad (1)$$

This equation can be described by saying that any optical system is a *filter* in the frequency domain: it does not transmit exactly the input object but it provides an output which is obtained by modifying the FT of the input.

A second basic property is that, as a consequence of diffraction, the MTF be-

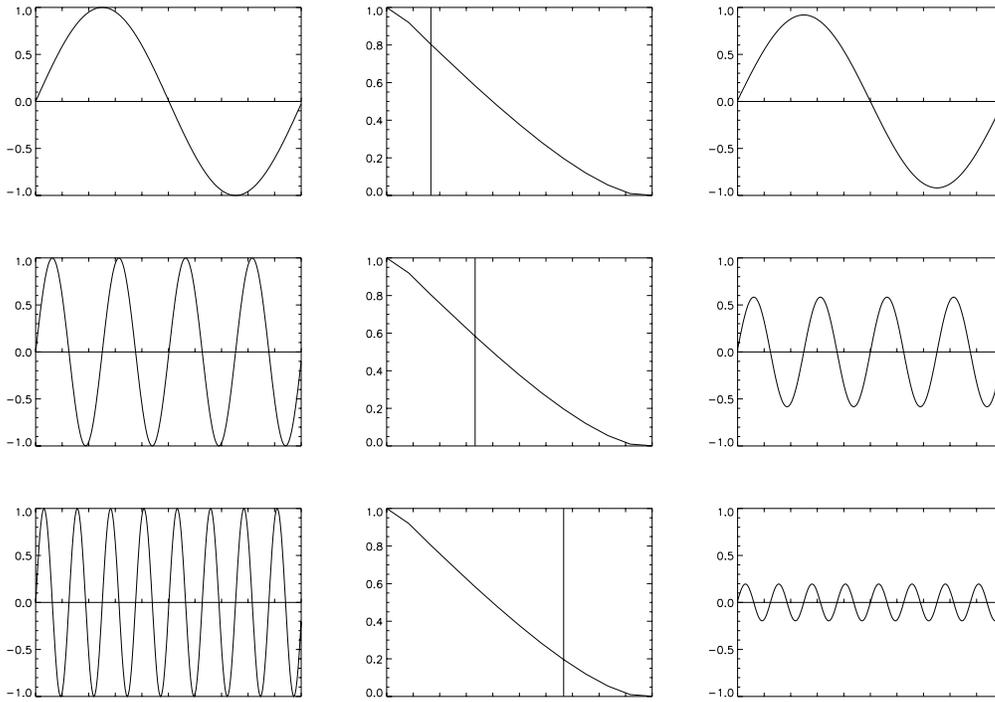


Fig. 2. Illustrating the filtering effect of an optical system; to the right we show sinusoidal input signals, in the middle the MTF with the indication of the frequency considered and to the left the corresponding output signals.

comes smaller and smaller when the space frequencies increase and eventually becomes exactly zero when they are higher than a certain limit which is called the *cut-off frequency* of the optical instrument. In Figure 2 we illustrate this effect, again in the case of 1D signals. In the 2D case, as a consequence of the circular symmetry of the pupils, the region where the MTF is different from zero is a disc whose radius Ω is just the cut-off frequency of the instrument. This disc is also called the *band* of the instrument and Ω , namely the cut-off frequency, is also called the *band-width* of the instrument.

This second property implies that the instrument is a *low-pass filter* (the low frequencies are almost perfectly transmitted while the high frequencies are attenuated or not transmitted at all) and has an important effect in the case of a point source. As we already observed such an object contains all frequencies with the same weights (the FT is constant, see Figure 1); but in its image the values corresponding to the higher frequencies are attenuated and set to zero when above the band-width. Therefore, as it can be easily understood from Figure 1, the image of a point source is not a point but a patch with a certain shape and width. This is precisely the PSF already mentioned in the Introduction and, mathematically, it is given by the inverse FT of the MTF, so that it will be denoted by $T(x, y)$.

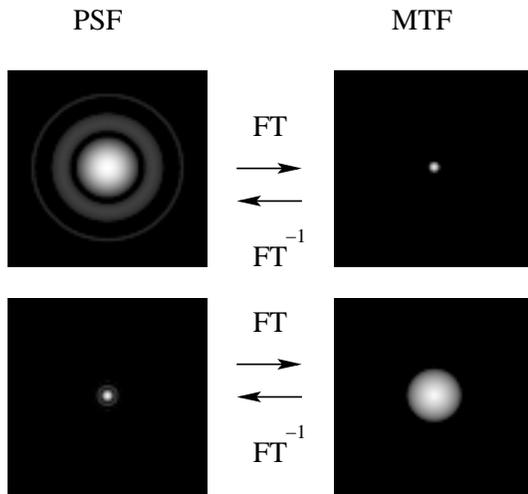


Fig. 3. Illustrating the relationship between the MTF and the PSF. We give two examples showing that when the band-width increases, the width of the PSF decreases.

For instance, in the case of a microscope and coherent illumination, the band-width Ω is given by

$$\Omega = 2\pi \frac{NA}{\lambda}, \quad (2)$$

where NA is the *numerical aperture* of the lens and λ is the wave-length of the illuminating light. If the lens is free of aberrations, then the MTF is constant (let us say 1, for simplicity) over the band and the PSF has the following expression

$$T(x, y) = \frac{\Omega}{2\pi} \frac{J_1(\Omega \rho)}{\rho}, \quad \rho = \sqrt{x^2 + y^2}, \quad (3)$$

where J_1 is the Bessel function of the first order. This function has a maximum at $\rho = 0$ and is zero over an infinity of circles, the radius of the first one being given by

$$\delta = 1.22 \frac{\pi}{\Omega} = 0.61 \frac{\lambda}{NA}. \quad (4)$$

This is precisely the well-known *Rayleigh resolution distance*, as we discuss in a moment.

However, in the case of fluorescence microscopy, since the emitted light is approximately incoherent, the PSF is proportional to the square of the coherent

PSF so that it takes the following expression

$$T(x, y) = \frac{\Omega^2}{\pi} \left(\frac{J_1(\Omega \rho)}{\Omega \rho} \right)^2, \quad (5)$$

which is also known as the *Airy function*. It is obvious that the positions of the zeros are not changed. However the band-width of this function is 2Ω and its FT, namely the MTF, is not constant over the band, the disc of radius 2Ω , but decreases from the value 1 at zero frequency up to 0 at the boundary of the band. The radial profile of this function is that given in Figure 2.

From Equation 4 it follows that *the broader is the band and the sharper is the PSF* and this relationship is illustrated in Figure 3 where we give two Airy functions with different band-widths and the corresponding MTF.

Until now we have considered the representation of an object or image in the frequency domain: in such a case the optical system is described by its MTF and the image is a filtered version of the object. Anyway an alternative representation is obtained by considering an object as a superposition of point-wise objects, each one with a different weight (its intensity); this is precisely the description in space domain. In such a case the behaviour of the optical system is described by its PSF and the image is a superposition of PSFs centred on the corresponding point objects, with weights given by their intensities. In mathematical terms the relationship is as follows:

$$I(x, y) = \int \int T(x - x', y - y') O(x', y') dx' dy' \quad (6)$$

and the r.h.s. of this equation is called the *convolution product* of PSF and object. It clarifies that the image is a *blurred* version of the object. Obviously this equation is equivalent to Equation 1 since one can be obtained from the other by means of the FT and this equivalence implies that *filtering and blurring are complementary descriptions of the same effect, namely light diffraction*.

Finally we relate the concept of resolution to the concepts we have just introduced. As we mentioned in the Introduction, according to Rayleigh the resolving power is a measure of the possibility of discriminating two close point sources. The problem is illustrated in Figure 4. Then the classical Rayleigh criterion consists in establishing that two point sources are resolved when the maximum of the image of the first one (i. e. a PSF centred on the position of the first source) falls on the first zero of the image of the second one (a PSF centred on the position of the second source). For a microscope this criterion provides a *resolution distance* δ which is given by Equation 4 and we conclude that *the resolution distance is inversely proportional to the band-width*. This result is true not only for a microscope but also for other optical instruments

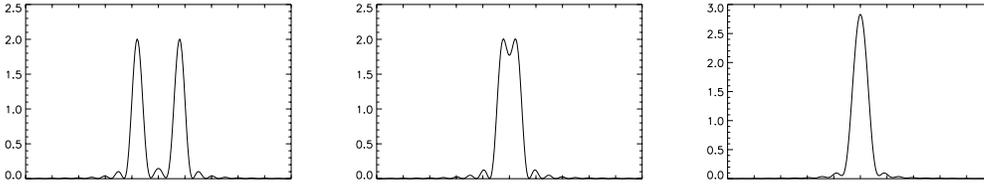


Fig. 4. Illustrating the problem of discriminating two point sources from the knowledge of their images as given by the PSF (1D case): from left to right we plot the images of two point sources with decreasing separation; the central one corresponds to the situation considered in the Rayleigh criterion.

(such as telescopes, etc.) even when the PSF is not the ideal one as that given, for instance, in Equation 5. Indeed the resolution distance can always be measured by the so-called FWHM (Full Width at Half Maximum) of the PSF and this parameter is also inversely proportional to the band-width of the instrument.

3 Computational super-resolution and out-of-band extrapolation

As follows from the analysis of the previous Section the limited resolution of an optical instrument is related to the limited information conveyed by the instrument in the frequency domain; indeed only the values of the FT of the object corresponding to frequencies inside the band are transmitted while the others are set to zero. Of course some of the in-band values are also attenuated but, in principle, they can be recovered if the MTF is known (see Equation 1). On the other hand the out-of-band values seem to be irretrievably lost, even if one can rise the question whether this fact is undoubtedly true.

In general the answer is yes. Indeed it is possible to imagine (and possibly also to manufacture) objects whose FT is exactly zero inside the band while takes any value outside the band. One can choose these values in order to get reasonable objects. Since their image is exactly zero, sometimes they are called *invisible objects*. Now it is obvious that, if we have recovered the in-band FT of the imaged object, we can take as out-of-band FT that of an invisible object and we obtain a new object which has exactly the same image of the true one. Such a procedure is clearly arbitrary because there exists an infinite set of invisible objects so that we can conclude that there exists an infinite set of objects compatible with the given image. Very simple and nice examples are given in Heintzmann & Sarafis (2000). In mathematical language this result is stated by saying that, if we attempt to extrapolate the FT of the object outside the band, then the solution of this problem, which is called *out-of-band extrapolation*, is not unique. In conclusion, since there exist an infinite set of

solutions, it is impossible to recover the true out-of-band FT of the imaged object.

However there is one case where the previous considerations are not true and precisely the case where the object has a finite extent. Indeed in such a case the FT of the object has a particular mathematical property, the so-called *analyticity*, which assures that *if the FT is known inside the band then it can be uniquely extrapolated outside the band*. In other words there exists a unique analytic function which coincides, inside the band, with the FT of the object (as recovered from the FT of the image - see Equation 1). This is precisely the basic remark contained in the papers by Wolter and Harris.

The implication of this result is evident: let us assume that it is possible to extrapolate the FT of the object outside the band in such a way that it can be estimated inside a disc with a radius twice the bandwidth of the instrument, for instance a microscope; then this procedure is equivalent to generate, by means of a processing of the data, a new imaging system, consisting of the microscope and of the computer (where the extrapolation method is implemented) in cascade, such that its band-width is twice that of the microscope; hence the corresponding resolution distance is one-half of the original one. We conclude that CSR has been achieved.

Since it is obvious that any object has a finite extent, CSR seems to be always obtainable by developing suitable methods and algorithms for out-of-band extrapolation. However there is a weak point in the remark of Wolter and Harris, namely the implicit assumption that the FT of the object is *exactly* known inside the band. This is never true in practice: detection and digitization of the image are always affected by noise and various kinds of random errors. In other words we only have at disposal a *noisy* version of the in-band FT of the object and, unfortunately, the noise destroys the analyticity of the FT.

As it was stressed for the first time in Viano (1977), the problem of out-of-band extrapolation belongs to a class of problems which are called *ill-posed* because they do not satisfy at least one of the three classical conditions for *well-posedness* as stated by the French mathematician Jacques Hadamard at the beginning of the last century: existence, uniqueness and stability of the solution. The last condition essentially means that a small variation of the data must produce a small variation of the solution. In the case of out-of-band extrapolation the solution, when it exists, is certainly unique but the difficulty is that, in general, it does not exist for noisy data and, in any case, it is not stable with respect to small variations of the data.

In the case of an ill-posed problem special techniques, known as *regularization methods* and introduced by the Russian mathematician A. N. Tikhonov, must be used for obtaining approximate and sensible solutions. Even a brief account

of the basic ideas of regularization theory is beyond the scope of this short note. An introduction with applications to imaging can be found in Bertero & Boccacci (1998).

The use of regularization methods allows to perform out-of-band extrapolation but only to a limited extent. The crucial question is: *is it possible to estimate how much extrapolation is feasible and therefore how much CSR can be achieved?* A first answer to this question is given in Bertero & Pike (1982). As shown in that paper, the amount of CSR is basically controlled by two parameters: *the space-bandwidth product* (SBP) and *the signal-to-noise ratio* (SNR).

As concerns the first parameter, it can be defined as follows: let us assume that our object is zero outside a disc of radius X while the band-width for coherent illumination is Ω ; then the SBP is given by $c = X\Omega/\pi$. If we use Equation 4, we find that, in the case of a microscope

$$c = 1.22 \frac{X}{\delta} \tag{7}$$

and therefore *SBP is a measure of the extent of the object in units given by the resolution distance of the instrument*. In Bertero et al. (1982) it is shown that the same parameter can be used for quantifying CSR also in the case of incoherent light (remember that, in such a case, the band-width is twice the coherent one).

The main result contained in Bertero & Pike (1982) is that a significant amount of out-of-band extrapolation, hence a significant improvement of resolution, is possible only if SBP is not much greater than one, the precise amount of extrapolation depending of course on the value of SNR and increasing when the SNR increases. Quantitative estimates are given in the paper quoted above. For instance for $c = 2$ and $SNR = 100$ an improvement of resolution by a factor 2.3 is possible. It is not a tremendous improvement but it may be very important in applications to microscopy or astronomy.

The physical meaning of the condition on SBP is clarified by Equation 7. Indeed it implies that the spatial extent of the object must be not much greater than the Rayleigh resolution distance of the optical instrument. It is obvious that the result applies also to the case where the object consists of separate unresolved objects over a black background, all with dimensions not much greater than the resolution distance. In these cases it may be possible to resolve details of the objects if a suitable algorithm for out-of-band extrapolation can be designed.

In conclusion we must also say that, as follows from numerical experiments as well as from theoretical results, a further improvement in out-of-band extrap-

olation can be obtained if one is searching for solutions which are non-negative in the space domain, as it is required in the case of incoherent light.

4 An algorithm for super-resolution

If we know that the SBP of the object satisfies the conditions stated in the previous Section, then a super-resolving method must be a method able to produce an estimate of the object which is compatible with the detected image and, in addition, is non-negative and zero outside the *domain* of the object, namely the region of the focal plane where the object is different from zero. Of course it is necessary to know such a domain and we discuss in a moment how it can be determined, at least approximately. It is also important to emphasize that it is not necessary to evaluate an object which satisfies exactly Equation 1 because, as we know, the detected image is corrupted by errors and noise.

Several methods with the properties stated above have been designed. All are suitable for CSR. Here we give a method for image restoration which is very popular in microscopy as well as in astronomy, is very easy to implement and can be modified in such a way that it becomes a super-resolving method. We are talking about a maximum likelihood method, which is known in astronomy as Richardson-Lucy method (RL) (Richardson (1972), Lucy (1974)) and in a wider audience as Maximum-Likelihood Expectation-Maximization (ML-EM) method (Shepp & Vardi (1982)). We do not discuss the foundations of this method; we only describe the corresponding algorithm which is iterative, so that it must be initialized with some guess $O^{(0)}(x, y)$ of the object which is updated at each iteration. If we denote by $O^{(k)}(x, y)$ the result of the k-th iteration, then the scheme is as follows:

- store $I(x, y)$, $\hat{T}(\omega_x, \omega_y)$;
- choose $O^{(0)}(x, y)$;
- for $k = 0, 1, 2, \dots$,
 - given $O^{(k)}(x, y)$ compute $\hat{O}^{(k)}(\omega_x, \omega_y)$;
 - $\hat{I}^{(k)}(\omega_x, \omega_y) = \hat{T}(\omega_x, \omega_y) \hat{O}^{(k)}(\omega_x, \omega_y)$; (8)
 - compute $I^{(k)}(x, y)$, the inverse FT of $\hat{I}^{(k)}(\omega_x, \omega_y)$;
 - $H^{(k)}(x, y) = \frac{I(x, y)}{I^{(k)}(x, y)}$; (9)
 - compute $\hat{H}^{(k)}(\omega_x, \omega_y)$, the FT of $H^{(k)}(x, y)$;
 - $\hat{G}^{(k)}(\omega_x, \omega_y) = \hat{T}(\omega_x, \omega_y) \hat{H}^{(k)}(\omega_x, \omega_y)$; (10)
 - compute $G^{(k)}(x, y)$, the inverse FT of $\hat{G}^{(k)}(\omega_x, \omega_y)$;
 - $O^{(k+1)}(x, y) = O^{(k)}(x, y) G^{(k)}(x, y)$. (11)

In Equation 10 we have assumed, for simplicity, that the MTF is real valued; this is in agreement with the assumption of Section 2 where the change of phase

of a trigonometric signal, as produced by the instrument, is neglected. In the general case the MTF takes complex values and one must use in Equation 10 the complex conjugate of the MTF.

The algorithm described above provides a restoration method which is usually initialized with a uniform object, namely $O^{(0)}(x, y) = \text{constant}$. It is known that it provides, in general, satisfactory results if the iterations are not pushed too far because when the number of iterations k is too large the result is degraded by the amplification of the noise corrupting the image. However the choice of the number of iterations is not very critical because the convergence is slow and the result remains stable for a considerable number of iterations. As easily follows from an analysis of the algorithm, it provides restored objects which are non-negative when both the image and the PSF are non-negative. Since one of the requirements for a super-resolving method is already satisfied, we have only to show that the algorithm can also implement the constraint on the domain of the object.

As follows from Equation 11, if $O^{(k)}(x, y)$ is zero in a point x_0, y_0 , then also $O^{(k+1)}(x, y)$ will be zero in the same point as well as all the subsequent iterates. Therefore, if we know the domain of the object, we can use this property by initializing the algorithm with an object $O^{(0)}(x, y)$ which is constant over the domain and zero elsewhere. As a consequence all the iterates $O^{(k)}(x, y)$ will be zero outside the domain of the object and therefore the constraint on the domain is automatically satisfied.

Another nice property of the algorithm is that, if the MTF is normalized in such a way that its value at zero-frequency is one, then all the iterates have the same total intensity as the original image $I(x, y)$. It follows that, if we initialize the algorithm in the way indicated above, then all the energy of the image is concentrated inside the domain of the object and is not spread out in the restoration procedure. Such a remark makes plausible the super-resolving effect obtainable in such a way.

In practical applications we have the problem of estimating the domain of the object. As it was clarified in the previous Section, the typical situation we are considering is that of a bright object, not much broader than the resolution distance of the instrument, over a black background. Therefore a possible approach can be provided by the following two-steps procedure: a first restoration of the object can be obtained by means of ML-EM with a standard initialization (a uniform object); next one estimates the region where the values of the restored object are greater than some threshold value (for instance, a few percent of the maximum value of the restored object); such a region is taken as an estimate of the domain of the object; finally ML-EM is reused with a new initialization provided by the mask function which is constant inside the estimated domain and zero elsewhere.

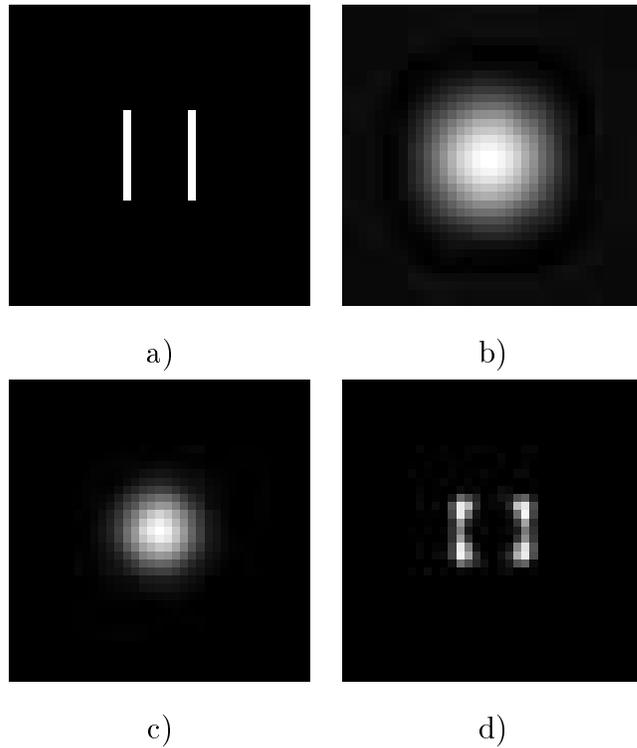


Fig. 5. Restoration of two vertical bars, 1 pixel wide, spaced by 9 pixels: a) the original object; b) the image obtained by convolving the object with an Airy function (Rayleigh resolution distance = 11 pixels) and corrupting the result with background and Poisson noise; c) the restoration provided by ML-EM without the constraint on the domain; d) the restoration provided by ML-EM with the constraint on the domain.

The previous analysis is based, for simplicity, on the assumption that the background is zero. Such an assumption, however, is not really necessary since it is known that the ML-EM method can be easily modified in order to take into account the presence of a constant and positive background on the image (van Kempen et al. (1997), Boccacci & Bertero (2002)). The modification consists uniquely in adding the constant background at the denominator of the r. h. s. of Equation 9.

We illustrate the previous method with a few numerical simulations. We consider the case of incoherent light with $\lambda = 0.58 \mu m$ and an ideal microscope with $NA = 1$ so that the PSF is the Airy function given in Equation 5 with $\Omega = 10.83 \mu m^{-1}$; the Rayleigh resolution distance is $\delta = 0.35 \mu m$. In addition we assume digital images 256×256 corresponding to physical images which are $8 \mu m$ wide; hence the pixel size is $0.031 \mu m$. It follows that the resolution distance corresponds to about 11 pixels and that we are considering images which are definitely oversampled; but this is necessary in the case of CSR. We also point out that the pixel size in the frequency domain is $0.78 \mu m^{-1}$, so that the band-width of the Airy function is about 28 pixels (we recall again that the incoherent band-width is 2Ω).

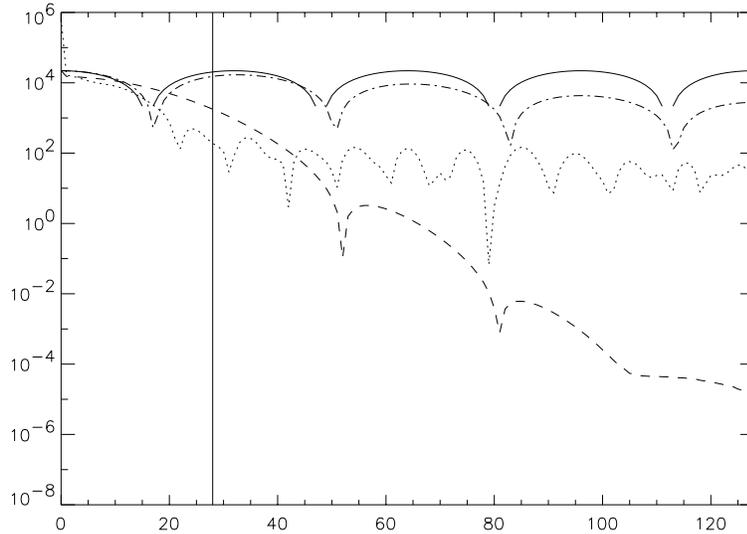


Fig. 6. Semi-logarithmic plot of the spectrum of the original object of Figure 5 (full line), of the corresponding image (dotted line), of the restoration provided by ML-EM (dashed line) and, finally, of the restoration provided by ML-EM with the constraint on the object domain (dash-dotted line). The spectrum is obtained by computing the square modulus of the FT on the line corresponding to $\omega_y = 0$ and is given as a function of the number of pixels; the position of the vertical line corresponds to the band-width of the Airy function. The larger value of the spectrum of the image at zero frequency is due to the background.

Our first object consists of two vertical bars, 1 pixel wide, and located at a distance of 9 pixels, hence non-resolved by our PSF. The image is generated by convolving the two bars with the PSF, by adding a suitable background, and corrupting the result with Poisson noise. In Figure 5 we show the original object, the noisy image, the restoration provided by ML-EM when initialized with a constant image and that provided by ML-EM when initialized with a rectangular mask a bit broader than the object. The two bars are clearly resolved in the last numerical experiment. Moreover, in order to make evident that the result is a consequence of out-of-band extrapolation, in Figure 6 we compare the profiles of the spectra (square modulus of the FT) of the original object, of the image, of the unresolved restoration and of the resolved one, when taken along the line corresponding to $\omega_y = 0$. In the Figure the vertical line indicates the band-width of the microscope. We remark that, in this particular example, the algorithm achieves a considerable amount of out-of-band extrapolation.

The second object consists of two clusters of bright spots consisting of 2×2 pixels. The distance between the centres of mass of the two clusters is greater than the resolution distance of the microscope so that the two clusters are resolved; however, in each cluster, the average distance between the spots is about 9 pixels, so that the spots are not resolved. Again an image is generated

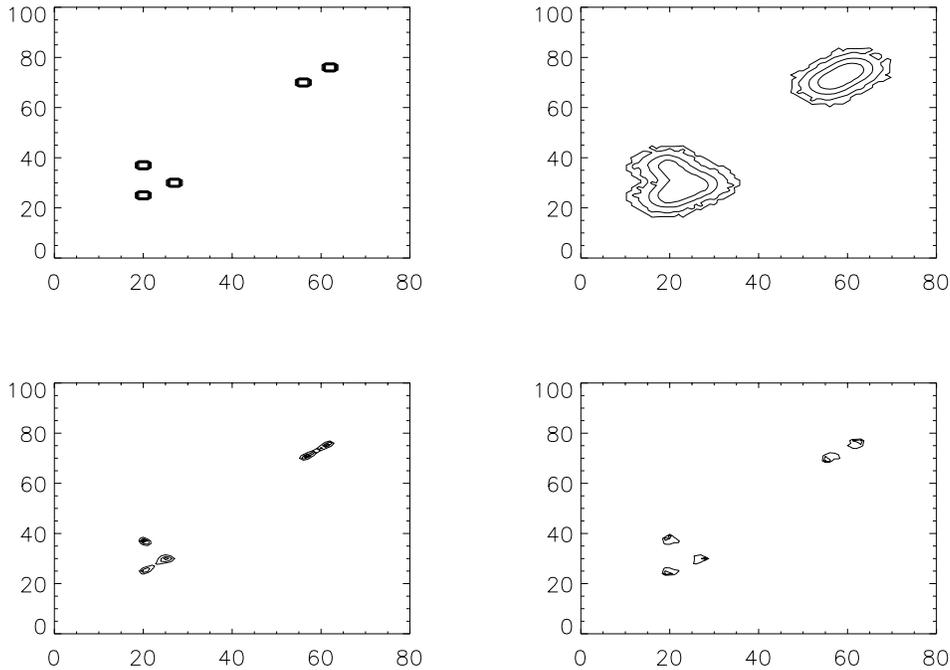


Fig. 7. Contour plot of the restoration of two clusters of bright spots: the original object (top left); the image obtained by convolving the object with the same PSF used in the first example and corrupting the result with background and Poisson noise (top right); the restoration provided by ML-EM (bottom left); the restoration provided by ML-EM initialized with a mask obtained from the previous result (bottom right).

with the same procedure used in the first example. A first restoration is obtained by means of ML-EM initialized with a constant image and is used for generating a mask which is obtained by taking the regions where the intensity is greater than 3% of the maximum value. Finally ML-EM is initialized with this mask. In Figure 7 we show the original object, the image and the two restorations. We remark that in this particular example the spots are already resolved in the first experiment. This is a demonstration of the beneficial effect of the positivity constraint but it is also the effect of a particular property of ML-EM since this method has the tendency to concentrate the energy of the image into localized regions. However the separation of the spots provided by the restoration initialized with the mask is even better.

We point out again that the method proposed in this paper is not necessarily the best one for CSR; its main advantage is that it is very simple and can be very easily implemented since it only requires the use of a FFT routine. Therefore it can be used as a flexible tool for performing numerical experiments and investigate the amount of CSR which can be reached in specific problems.

5 Concluding remarks

The main aim of this paper was to show that CSR is intimately related to the problem of extrapolating the spectrum of the object outside the band of the optical instrument. Obviously if the instrument is affected by strong aberrations a simple deconvolution can provide an improvement of the quality of the image, allowing to detect details which are not clearly visible in the original image. However this is not a genuine CSR if the method is only able to recover the in-band values of the FT of the object. Indeed out-of-band extrapolation is a difficult problem and it can be successfully handled only if we have additional information on the object. In this paper we have indicated a crucial one, namely the information that the size of the domain of the object is not much greater than the resolution distance of the instrument. Another important kind of information is the non-negativity of the object, at least in cases where such a property applies (for instance, incoherent light). However, as clearly shown by our first example non-negativity alone is not always sufficient for obtaining a significant CSR. We also point out that in our examples we require much more than the simple condition that the spots or the lines are resolved according to the Rayleigh criterion; indeed in our restorations the small features (spots, bars) are *visually* resolved. Obviously this is possible because we only search for a relatively small amount of super-resolution.

Better results could be obtained by means of further additional information on the object (for instance, statistical properties) and more sophisticated restoration methods implementing this additional information. However, it must be clear that no miracle is possible and that it is not possible to go much beyond the limitations indicated in this paper.

Our discussion was basically restricted to the 2D case. However it applies also to the 3D case and, in particular, to confocal microscopy. In 3D the limitations on the domain of the object must be satisfied in the three directions. However, since the axial resolution distance is greater than the lateral one, the limitations in the axial direction are less stringent than those in the lateral ones. In particular cases, namely thin objects with a depth not much greater than the axial resolution distance, it should be possible, by means of a super-resolving method acting only in the axial direction, to obtain restored images such that the resolution is approximately uniform in the three directions.

Acknowledgments

We thank Vassilios Sarafis for suggesting and stimulating this work. We also thank Christine De Mol for a critical reading of the manuscript.

References

- Abbe, E., 1873, *Archiv. Mikrosk. Anat. Entwicklungsmech.* 9, 413
- Bertero, M., and Boccacci, P., 1998, *Introduction to Inverse Problems in Imaging* (IoP Publishing, Bristol)
- Bertero M., and De Mol, C., 1996, *Progress in Optics* (ed. E. Wolf) Vol. XXXVI, ch. III (Elsevier, Amsterdam)
- Bertero, M., and Pike, E. R., 1982, *Opt. Acta* 29, 727
- Bertero, M., Boccacci, P., and Pike, E. R., 1982, *Opt. Acta* 29, 1599
- Boccacci, P., and Bertero, M., 2002, *Confocal and Two-Photon Microscopy: Foundations, Applications, and Advances* (ed. A. Diaspro), ch. 12 (Wiley-Liss, New York)
- Harris, J. L., 1964, *J. Opt. Soc. Am.* 54, 931
- Heintzmann, R., and Sarafis, V., 2000, *Optik* 112, 114
- van Kempen, G. M. P., van Vliet, L. J., Verveer, P. J., and van der Voort, H. T. M., 1997, *J. Microsc.* 185, 354
- Lucy, L. B., 1974, *AJ* 79, 745
- McCutchen, C. W., 1967, *J. Opt. Soc. Am.* 57, 1190
- Papoulis, A., 1968, *Systems and Transforms with Applications in Optics* (McGraw-Hill, New York)
- Rayleigh, J. W. S., 1879, *Phil. Mag.* 8, 261
- Richardson, W. H., 1972, *J. Opt. Soc. Am.* 62, 55
- Rushforth, C. K., and Harris, R. W., 1968, *J. Opt. Soc. Am.* 58, 539
- Shepp, L. A., and Vardi, Y., 1982, *IEEE Trans. Med. Imaging* MI-1, 113
- Toraldo di Francia, G., 1955, *J. Opt. Soc. Am.* 45,497
- Viano, G. A., 1977, *J. Math. Phys.* 17, 1160
- Wolter, H., 1961, *Progress in Optics* (ed. E. Wolf) Vol. I, ch. V (Elsevier, Amsterdam)