Quantitative control of the error bounds of a fast super-resolution technique for microscopy and astronomy

Pierre Chainais, Pierre Pfennig, Aymeric Leray

To cite this version:
Pierre Chainais, Pierre Pfennig, Aymeric Leray. Quantitative control of the error bounds of a fast super-resolution technique for microscopy and astronomy. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), May 2014, Florence, Italy. pp.2853 - 2857, 10.1109/ICASSP.2014.6854121. hal-01081402

HAL Id: hal-01081402
https://hal.archives-ouvertes.fr/hal-01081402
Submitted on 7 Nov 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
**ABSTRACT**

While the registration step is often problematic for super-resolution, many microscopes and telescopes are now equipped with a piezoelectric mechanical system which permits to accurately control their motion (down to nanometers). Therefore one can use such devices to acquire multiple images of the same scene at various controlled positions. Then a fast super-resolution algorithm [1] can be used for efficient super-resolution. However the minimal use of \( r^2 \) images for a resolution enhancement factor \( r \) is generally not sufficient to obtain good results. We propose to take several images at positions randomly distributed close to each reference position. We study the number of images necessary to control the error resulting from the super-resolution algorithm by [1] due to the uncertainty on positions. The main result is a lower bound on the number of images to respect a given error upper bound with probability higher than a desired confidence level.

**Index Terms**— microscopy; super-resolution; confidence level; error estimate

1. **INTRODUCTION**

Super-resolution algorithms which combine several low-resolution (LR) images to compute one high-resolution (HR) image have been the subject of a large amount of works over nearly 30 years now, see [2] for a review. Super-resolution methods are often either computationally expensive to access good performances in a general setting or fast at the price of simplifying assumptions (e.g. displacements of images are exactly known) which may limit their range of application. In the present work, we deal with a setting where such simplifying assumptions can be controlled so that a fast method [1] can be used for microscopy or astronomy imaging. While the registration step is often problematic for super-resolution, we propose to exploit the potential of a microscope or telescope equipped with a looped piezoelectric positioning system (precision goes down to 0.1 nm) combined with a fast super-resolution algorithm to propose a new cheap, fast and efficient super-resolution technique for microscopy or astronomy imaging. One can use such devices to acquire multiple images of the same scene at various controlled positions and assume that the registration problem is solved, in good approximation at least. The experimental protocol is much simpler than PALM/STORM [3, 4, 5]. We will consider pure translations only. Aiming at an integer resolution enhancement factor \( r \) (\( \times 2, \times 4, \ldots \)), we also consider that the \( r^2 \) possible translations multiple of \( 1/r \) low resolution pixel (1 LR pixel = \( r \) HR pixels) in the two horizontal directions are available. In practice, the positioning system is not perfect so that a targetted position is only approximately reached with some little random error. We show that the use of several acquisitions associated to the same targetted displacement with different positioning errors permits to obtain good results with controlled accuracy. Our main contribution is that this claim is supported by bounds and guarantees on the quality of the super-resolved image. To make sure that the technique is both fast and efficient, we study the relation between the number of images taken at each position and the error resulting from the use of the fast super-resolution algorithm [1] which assumes that translations are exactly known. The main result is a lower bound on the number of images that is necessary to respect a given error bound with probability higher than some desired confidence level. We emphasize that only few works, e.g. [6, 7] for iterative methods, deal with the mathematical analysis of super-resolution algorithms. The present study is performed in the Fourier domain so that the precise contribution of each aliasing term due to only approximately known translations is analyzed and quantitatively evaluated. The use of Bienaymé-Cebyshev inequality finally yields confidence intervals of practical use.

Section 2 presents the model and recalls about the fast algorithm in [1]. Section 3 presents our main theoretical results which predict the required number of image acquisitions at each position to ensure some given confidence level in the reconstructed image. Section 4 presents numerical results.
2. A FAST AND CONTROLLED SUPER-RESOLUTION TECHNIQUE

2.1. The super-resolution problem

For a given resolution enhancement factor \( r \), the most common linear formulation of the general super-resolution problem in the pixel domain is [1]:

\[
Y_k = D_k H_k F_k Y_{HR} + n_k \quad k = 1, \ldots, K, \tag{1}
\]

where \( Y_{HR} \) is the (desired) high resolution image one wants to estimate from the \( K \) low-resolution images \( \{Y_k, 1 \leq k \leq K\} \) and \( n_k \) is the noise, generally assumed to be Gaussian white noise so that \( E(n_k n_k^T) = \sigma^2 I \). Images \( Y_{HR} \), \( Y_k \) and \( n_k \) are rearranged in lexicographic ordered vectors. Each image \( Y_k \) is a low resolution observation of the same underlying scene translated by \( F_k \). The blur matrices \( H_k \) model the point spread function (PSF) of the acquisition system and matrix \( D_k \) is the decimation operator by a factor \( r \). If \( Y_{HR} \) is of size \( r^2 N^2 \times 1 \) and \( Y_k \) of size \( N \times N \), the matrices \( F_k \) and \( H_k \) are of size \( (rN)^2 \times (rN)^2 \) while \( D_k \) is \( N^2 \times (rN)^2 \). The least squares optimization problem can be formulated as:

\[
Y_{HR} = \text{argmin}_X \sum_{k=1}^K \|Y_k - D_k H_k F_k X\|_2^2. \tag{2}
\]

Other formulations based on the L1-norm [8] or adding some regularization have also been proposed. Many methods rely on iterative optimization methods which are often computationally expensive and time consuming. As far as applications to biological (alive) systems are concerned, a fast method is necessary. In the present setting, displacements are controlled thanks to a piezoelectric platform at a precision of 0.1 nm in every direction. As a consequence, the method in [1] appears as a good choice since it is fast (even though more recent algorithms may yield better empirical results) and above all we will be able to quantitatively analyze its performances and estimate error bounds and confidence level intervals as a function of the number of available images.

2.2. Super-resolution algorithm

The algorithm in [1] makes usual assumptions. The PSF of the acquisition system is known and spatially homogeneous so that \( \forall k, H_k = H \). Decimation is the same for all images so that \( \forall k, D_k = D \) in (1) & (2). Periodic boundary conditions make \( F_k \) and \( H \) circulant matrices. We will moreover assume that the \( r^2 \) possible translated images at integer multiples of the high resolution scale are available. As a consequence one can show following [1] that the intermediate blurred image \( Z \) can be computed as:

\[
Z := HY_{HR} = \sum_{k=1}^{r^2} F_k^T D^T Y_k \tag{3}
\]

Then the final HR image results from the deconvolution of \( Z \) only, which can be done using any state of the art algorithm. This algorithm [1] is fast thanks to one idealized assumption: displacements (matrices \( F_k \)) are assumed to be exact integer multiples of high resolution pixels. This assumption is obviously only approximately true in practice. In our setting, this is essentially due to the finite precision of the piezoelectric positioning system. However, we will stick to this assumption since it makes the algorithm very fast compared to iterative methods in particular. One way to compensate the errors resulting from the error on translation estimates is to take several images for each required position. For a given targeted position, the positioning system will be reset between each shooting so that positions are randomly distributed around the average position (which is assumed to be the right one). This averaging process is expected to enhance the super-resolution quality. We aim at quantifying how many images one should take to enter some given upper error bound of \( p\% \) (e.g. 10%) with probability (confidence) higher than \( P_{\min} \) (e.g. 0.90).

3. HOW MANY IMAGES TO ENSURE CONFIDENCE INTERVALS?

We consider the relation between the (blurred) HR image \( Z \) defined by (3) and the LR images \( Y_k = Y^d_k \) in the Fourier domain to detail the effects of aliasing. \( Y^d \) is the image of a scene \( Y \) translated by \( d \). Here \( d = (d_x, d_y) \) denotes the targeted displacement vector and \( d_e = d + b_d \) the real experimental displacement; \( b_d \) is the noise (error) on the platform position. Note that \( (F^d_t)^T F^d = I_{rN} \). For some integer \( n \), the interval \((-n : n)\) denotes the set of integers between \(-n \) and \( n \) (Matlab notations). When using the DFT, we denote by \( k \) the LR frequencies in \( D_{LR} = (-N/2 : N/2 - 1)^2 \) and \( k' \) the HR frequencies in \( D_{HR} = (-N/2 : rN/2 - 1)^2 \). Given some HR frequency \( k' \), we need to deal with corresponding aliased terms in the LR image. The integer vector \( \gamma \in (-r : r)^2 \) such that \( k' = k' - \gamma N \in D_{LR} \). We note \( \alpha \) integer vectors such that \( k' + \alpha N \in D_{HR} \). Sums over \( d \) are sums over all the considered displacements and sums over \( \alpha \) are sums over all possible HR frequencies \( k' + \alpha N \) (up to \( rN/2 \)). The Discrete Fourier Transform (DFT) of image \( Z \) is \( \hat{Z} \). Back to (3), note that when \( D \) is the decimation operator, \( D^T \) is an upsampling operation (inserting zeros between samples) which produces aliasing. Indeed, if \( F_{HR}^T \) is the HR DFT, for \( k' \in D_{HR} \):

\[
|F_{HR} D^T Y^d |(k') = \hat{Y}^d(k' + k' - \gamma N) \tag{4}
\]

Taking phase shift due to translations of \(-d \) associated to \( F_k^T \) into account in the DFT of (3) yields:

\[
\hat{Z}(k') = \sum_{d_e} \hat{Y}^d_e(k' - \gamma N) e^{\frac{2\pi i}{rN} d_e \cdot k'} \tag{5}
\]

Now we explicitly how the translated images \( \hat{Y}^d_e(k) \) for \( k \in D_{LR} \) relate to the high resolution image \( \hat{Y}_{HR}(k') \) for \( k' \in D_{HR} \).
in the Fourier domain. Since each observation is a decimated version of the blurred translated scene, one has:
\[ Y^d = DY^d_{sh} \]
where \( Y^d_{sh} = HF^d_{sh}Y_{HR} \)
in the spatial domain. In the Fourier domain:
\[ \hat{Y}^d = F_{LR}D_{HR}^* \hat{Y}^d_{sh} \]
and thanks to usual properties of the sum of roots of unity:
\[ \hat{Y}^d(k) = \frac{1}{r^2} \sum_{\alpha} \hat{Y}_{sh}(k + \alpha N) f^d(k + \alpha N) \]
where \( Y_{sh} = HY_{HR} \) and the translation phase factor \( f^d(k') \) is \( \exp\left( -\frac{2\pi i k'}{r^2} \right) \), except when \( k'_x \) or \( k'_y \) is \(-rN/2\). We have also used the fact that the homogeneous blur operator is diagonal in Fourier domain. One can explicitly see in (8) how the information at high frequencies \( k + \alpha N \) from the HR image is aliased at low frequency \( k \) in each LR image \( Y^d_{sh} \). By separating the desired contributions at \( k' = \tilde{k}' + \gamma N \) and aliasing terms at \( \tilde{k}' + \alpha N \) for \( \alpha \neq \gamma \), one gets from (5) and (8):
\[ \hat{Z}(k') = \hat{Y}_{sh}(k') g_{\alpha,b_d}(k') \]
\[ + \frac{1}{r^2} \sum_{\alpha \neq \gamma} \hat{Y}_{sh}(\tilde{k}' + \alpha N) g_{\alpha,b_d}(k') \]
where \( g_{\alpha,b_d}(k') = \sum_d e^{-\frac{2\pi i k'}{r^2} (\alpha - \gamma) N} d + (\tilde{k}' + \alpha N) \cdot b_d \),
(except when \( k'_x \) or \( k'_y \) is \(-rN/2\)). In the ideal case where \( b_d = 0 \) so that translations are exact multiples of the HR pixel scale, one retrieves (3) and \( \hat{Z} = \hat{Y}_{sh} = \hat{H}^{\gamma}Y_{HR} \). Note that high frequency terms at some \( k' \) are preserved only if the PSF \( \hat{H}(k') \) is not zero. In physical systems, the PSF has often a compact support so that its Fourier transform is non zero almost everywhere.

Now we study the errors on \( Z \) underlying the superresolved image \( \hat{Y}_{HR} \) due to inaccuracies of translation estimates. To compensate the inaccuracy of displacements which translates in \( g_{\alpha,b_d} \) terms above, we propose to use multiple acquisitions at the same targetted position with a different error \( b_{d,j} \) around the same expected value \( d \). Then we average \( n_d \) images per targetted position \( d \) so that (3) becomes:
\[ Z = \sum_d \frac{1}{n_d} \sum_{j=1}^{n_d} F^d_{d,j} D^d_{d,j} Y^d_{d,j} \]
\[ d_{d,j} = d + b_{d,j} \] is the effective experimental translation with error \( b_{d,j} \). Then (9) yields:
\[ \hat{Z}(k') = \hat{Y}_{sh}(k') G_{\alpha}(k') \]
\[ + \frac{1}{r^2} \sum_{\alpha} \hat{Y}_{sh}(\tilde{k}' + \alpha N) G_{\alpha}(k') \]
where \( G_{\alpha}(k') = \frac{1}{n_d} \sum_{j=1}^{n_d} g_{\alpha,b_{d,j}}(k') \). The first term will be called the **approximation term**, which should be as close as possible to \( Y_{sh}(k') \). The second term \( B(k') \) will be called the **aliasing term** and should be as small as possible compared to the main approximation term. Our purpose below is to establish conditions that guarantee such an approximation within quantitative bounds.

The approximation term is rather easy to control by noticing that
\[ G_{\alpha}(k') = \sum_d \frac{1}{n_d} \sum_{j=1}^{n_d} e^{-\frac{2\pi i k'}{r^2} \cdot \alpha \gamma} b_{d,j} \]
can be seen as an estimate of the characteristic function of \( b_d \). It appears reasonable to empirically assume that the \( b_d \) are uniformly distributed within some uncertainty interval \((-\epsilon, \epsilon)\) in LR pixel unit or \((-\epsilon_r, \epsilon_r)\) where \( \epsilon_r = r \epsilon \) in HR pixel unit. The generalization to distributions with bounded support is the subject of ongoing work. Therefore we define
\[ A(q, \epsilon_r) = \frac{1}{r^2} E_{b_d} |G_{\gamma}(k') - B(k')| \]
\[ = \text{sinc}(q \epsilon_r) \text{sinc}(q \epsilon) \]
(13)
where \( q = \frac{2\pi}{r} k' \). Since \( |A(q, \epsilon_r) - 1| \leq \frac{2^2}{r^2} \), it is easy to control that \( A \approx 1 \) in very good approximation for all \( k' \) as soon as \( r^2 \epsilon^2 / 3 \ll 1 \). Now we need to control the error \( \frac{1}{r^2} G_{\alpha}(k') - A(q, \epsilon_r) \) to control \( \frac{1}{r^2} G_{\gamma}(k') - 1 \) in the end. Using third order approximations of the real and imaginary parts of \( \frac{1}{r^2} G_{\gamma}(k') - A(q, \epsilon_r) \) one can bound this difference by quantities which depend only on powers of \( ||q||_1 \epsilon_r \) and \( ||q||_1 \epsilon \). The variances of powers of \( q \cdot b_{d,j} \) can be evaluated and bounded from above so that we can use Bienaymé-Cebycev inequality recalled below for i.i.d. random variables \( x_j \):
\[ P\left( \left| \frac{1}{n} \sum_{j=1}^{n} x_j - E[x] \right| \geq \delta \right) \leq \frac{\text{var}[x]}{n \delta^2} \]
(14)
By choosing \( \delta = c ||q||_1 \epsilon_r \) where \( c \) is some proportionality parameter, one can show the following important inequality:
\[ P\left( \left| \frac{1}{r^2} G_{\alpha}(k') - A(q, \epsilon_r) \right| \geq m(q, \epsilon_r, c) \right) \]
\[ \leq \frac{||q||_2^2}{3 ||q||_1^2 n_d c^2} \left( 1 + f(\epsilon_r) \right) \]
(15)
where \( f(\epsilon_r) \) remains negligible as soon as \( \epsilon_r \ll 0.4 \) and
\[ m(q, \epsilon_r, c) = c ||q||_1 \epsilon_r \cdot s(||q||_1, \epsilon_r) \]
(16)
\[ s(||q||_1, \epsilon_r) = \left( 2 + \frac{||q||_1^2 \epsilon_r^2}{10 c} \left( 1 + \frac{||q||_1 \epsilon_r}{6} \right)^{1/2} \right) \]
(17)
Our aim is to guarantee a good approximation up to at most \( p\% \), e.g. 10\%, of relative error with probability greater than \( P_{\text{min}} \), e.g. 0.9. To this aim we need to estimate for
Fig. 1. Required minimum number \( n_d(k') \) of images per position for \( r = 2, \epsilon = 0.01 \) and \( r = 4, \epsilon = 0.001 \) to ensure a relative error \(< 10\%\) with probability \( \geq P_{\text{min}} = 0.90; N = 32 \); (a-b) approximation; (c-d) aliasing term.

Each frequency \( k' \) the maximum value \( c_{\text{max}}(k') \) such that
\[
\min_{k' \in D_{HR}} \|q_{d}(k') \|^{2} [3(1 - P_{\text{min}})^{2}c_{\text{max}}(k')^{2}q_{d}^{2}]^{-1}
\]
\[ (18) \]

Turning to the aliasing term \( B(k') \) in (10), let \( q_{\alpha} = \frac{2n}{rN}(k' + \alpha N) \). Using similar arguments one can show that:
\[
P \left( |B(k')| \leq \sum_{\alpha \neq \gamma} |\mathcal{Y}_{\alpha}(k')|\sqrt{2}\|q_{\alpha}\|_{2} \epsilon_{r} \left( c + \frac{\|q_{\alpha}\|_{2}\epsilon_{r}}{2} \right) \right) 
\geq \prod_{\alpha \neq \gamma} \left( 1 - \frac{1}{3n_{d}c^{2}} \right) \quad (19)
\]

To guarantee an error smaller than \( p\% \) we need an assumption on the spectrum of the image. We will use the usual property of the natural images as a worst case with an energy spectrum \( \propto 1/\|k'\|_{2}^{(1+\eta)} \) where usually \( |\eta| \ll 1 \). Eventually we numerically estimate \( c_{\text{max}}(k') \) such that for any \( c \leq c_{\text{max}}(k') \):
\[
\sum_{\alpha \neq \gamma} \left( \frac{\|q\|_{2}^{2}}{\|q_{\alpha}\|_{2}^{2}} \right)^{(1+\eta)/2} \sqrt{2}\|q_{\alpha}\|_{2} \epsilon_{r} \left( c + \frac{\|q_{\alpha}\|_{2}\epsilon_{r}}{2} \right) \leq \frac{p}{100}
\]
\[ (20) \]

Then remains to estimate the minimum \( n_d \) such that:
\[
n_d(k') \geq \left( 3(1 - P_{\text{min}})^{2}c_{\text{max}}(k')^{2} \right)^{-1}
\]
\[ (21) \]

The main results of this work are the two lower bounds on \( n_d \) given by (18) and above all (21) which control the minimum number of images \( n_d \) such that some relative error bound of \( p\% \), e.g. \( 10\% \), can be guaranteed with probability \( \geq 90\% \). The related constraints on \( n_d(k') \) can be computed numerically.

<table>
<thead>
<tr>
<th>( r )</th>
<th>( n_d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>6 / 151</td>
</tr>
<tr>
<td>4</td>
<td>22 / unreach.</td>
</tr>
<tr>
<td>6</td>
<td>50 / unreach.</td>
</tr>
</tbody>
</table>

Table 1. Required minimum number \( n_d \) of images per position to ensure a relative error \(< 10\%\) with probability \( > 0.9 \) on approximation/aliasing term as a function of the uncertainty on the translations expressed in LR pixel unit.

4. NUMERICAL RESULTS

Fig. 1 shows the numerical results from (18) and (21) in Fourier domain for two resolution enhancement factors \( r = 2 \) and \( r = 4 \). As expected, the accurate recovery of higher frequencies calls for more LR images. The results do not depend on the size \( N \) of images. The larger \( r \), the larger the need for multiple images. The control of aliasing effects is the most constraining. In our microscopy setting, 1 LR pixel \( \approx 100 \) nm. The random bias on the platform positioning system is between 0.1 and 1 nm that is \( \epsilon \approx 0.001 - 0.01 \) LR pixel. The acquition of 1 LR image usually takes about 0.5 s. For \( r = 2 \) and \( \epsilon = 0.001 \), resp. \( \epsilon = 0.01 \), an accuracy better than \( p = 10\% \) on the restored image can be guaranteed with probability \( \geq 0.90 \) by using at least \( n_d = 2 \), resp. \( n_d = 151 \), images/position. In practice, \( r^2 = 4 \) displacements are necessary which means an acquisition time between \( 4 \times 2 \times 0.5 = 4 \) s and \( 4 \times 151 \times 0.5 = 302 \) s which remains reasonable to observe many biological systems. Tab. 1 gathers the constraints for various values of \( r \) and \( \epsilon \). For \( r = 6 \) and \( \epsilon = 0.01 \) there is no way to guarantee the quality of reconstruction ("unreachable" in Tab. 1). However, for sufficiently accurate positioning with \( \epsilon = 0.0001 \), \( 6^2 \times 189 = 6804 \) images are sufficient (\( \approx 57 \) min at 0.5 s/image).

5. CONCLUSION

We propose a controlled and fast super-resolution technique which takes benefit from (1) the piezoelectric positioning systems of microscopes (or telescopes) platforms to realize accurate translations, (2) the analysis of the fast algorithm proposed in [1] so that error confidence intervals can be computed as a function of the number of available images. This is made possible by the algorithm itself and by exploiting the averaging of LR images taken at positions that are randomly distributed around the same reference position. This technique is cheap and realistic to enhance the resolution of many devices. It may be adapted to many applications ranging from biology to astronomy where the need for guarantees on the restored information is crucial. The present bounds yield sufficient conditions but may not be optimally tight. Tighter bounds would also be more optimistic since less images would be necessary. This would be useful for applications. This is the subject of ongoing work.
6. REFERENCES


