Fundamental Limits of Reconstruction-Based Superresolution Algorithms under Local Translation

Zhouchen Lin, *Member*, *IEEE*, and Heung-Yeung Shum, *Senior Member*, *IEEE*

Abstract—Superresolution is a technique that can produce images of a higher resolution than that of the originally captured ones. Nevertheless, improvement in resolution using such a technique is very limited in practice. This makes it significant to study the problem: "*Do fundamental limits exist for superresolution*?" In this paper, we focus on a major class of superresolution algorithms, called the *reconstruction-based algorithms*, which compute high-resolution images by simulating the image formation process. Assuming local translation among low-resolution images, this paper is the first attempt to determine the *explicit* limits of reconstruction-based algorithms, under both real and synthetic conditions. Based on the perturbation theory of linear systems, we obtain the superresolution limits from the conditioning analysis of the coefficient matrix. Moreover, we determine the number of low-resolution images that are sufficient to achieve the limit. Both real and synthetic experiments are carried out to verify our analysis.

Index Terms—Superresolution, reconstruction-based algorithms, conditioning analysis, fundamental limits, magnification factor.

1 Introduction

Increasing image resolution is a fundamental problem in image processing. Instead of imposing higher requirements on hardware devices and sensors, it is often much more economical to take multiple images. This makes the technique of superresolution very useful as it combines low-resolution images to produce higher resolution ones.

Numerous algorithms ([1], [6], [8], [12], [14], [15], [21], [20], [10], [19], [7], [17], [4], and the references in [5]) have been proposed for superresolution. However, in practice, people have found that improvement in resolution is quite limited. When the magnification factor becomes a bit large, the performance of existing algorithms deteriorates. The images they produce either are overly smooth or contain undesirable details. This leads us to pose a question: "*Do fundamental limits exist for superresolution?*" The study on this problem is important because:

1. If such a limit exists, practitioners can stop trying those unduly large magnification factors, thus saving resources.
2. A better understanding of this problem can help people choose good magnification factors within the limits and capture just enough images so as to make better use of resources.

In this paper, we investigate this problem within reconstruction-based algorithms (RBAs, coined in [2]), which are by far the most commonly used superresolution algorithms, under the assumption that the movement between the low-resolution images is locally translational.

- *Z. Lin is with Microsoft Research, Asia, 5th floor, Sigma Building, Zhichun Road No. 49, Haidian District, Beijing, 100080, P.R. China. E-mail: zhoulin@microsoft.com.*
- *H.-Y. Shum is with Microsoft Research, Asia, 3rd floor, Sigma Building, Zhichun Road No. 49, Haidian District, Beijing, 100080, P.R. China. E-mail: hshum@microsoft.com.*

Manuscript received 8 July 2002; revised 27 Mar. 2003; accepted 2 Aug. 2003. Recommended for acceptance by M. Irani.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 116892.

1.1 Reconstruction-Based Algorithms

Reconstruction-based algorithms model the process of image formation to build the relation between the low-resolution images (LRIs) and the high-resolution image (HRI). They usually start from the continuous image formation equation:

$$L(\mathbf{y}) = \int B(\mathbf{y}, \mathbf{x}) H(\mathbf{x}) d\mathbf{x} + E(\mathbf{y}), \tag{1}$$

where $L(\mathbf{y})$ is the continuous low-resolution irradiance field on the image plane, $B(\mathbf{y}, \mathbf{x})$ is the blurring kernel (when there is motion blur, $B(\mathbf{y}, \mathbf{x})$ is supposed to have been integrated in time), $H(\mathbf{x})$ is the continuous high-resolution irradiance light-field that would have reached the image plane, and $E(\mathbf{y})$ is the noise. For the explanation of more abbreviations and notations, please refer to Table 1.

The discrete version of (1) is a linear system, where the sampling density of $L(\mathbf{y})$ is lower than that of $H(\mathbf{x})$. A smoothness prior, such as that involved in maximum a posteriori (MAP) [12], [8], [21] or regularized maximum likelihood (ML) [8], is often required in order to remove noise and solve the linear system stably.

Often, the blurring kernel is assumed to be shift invariant. This is true for local superresolution. In this case, it shows up as a point spread function (PSF). As mentioned by Baker and Kanade [2], the PSF can be decoupled into two parts:

$$B = PSF_{sensor} * PSF_{lens},$$

then

$$L(\mathbf{x}) = PSF_{sensor} * U(\mathbf{x}), \tag{2}$$

where $U(\mathbf{x}) = PSF_{lens} * H(\mathbf{x})$, PSF_{lens} and PSF_{sensor} are the PSFs of the lens and the sensor on the image plane, respectively. When there is motion blur, PSF_{lens} is also

TABLE 1
Abbreviations and Notations

MAP	**M**aximum **a** p*osteriori*	CCD	**C**harge-**C**oupled **D**evices
ML	**M**aximum **L**ikelihood	LSE	**L**east **S**quared **E**rror
POCS	**P**rojection **o**nto **C**onvex **S**ets	PSF	**P**oint **S**pread **F**unction
RHS	**R**ight **H**and **S**ide	ROI	**R**egion **o**f **I**nterest
RMSE	**R**oot **M**ean **S**quared **E**rror	PSNR	**P**eak **S**ignal to **N**oise **R**atio
RBA	**R**econstruction-**b**ased **A**lgorithm	LRI	**L**ow **R**esolution **I**mage
HRI	**H**igh **R**esolution **I**mage	LRP	**L**ow **R**esolution **P**ixel
HRP	**H**igh **R**esolution **P**ixel	EMF	**E**ffective **M**agnification **F**actor
RD	**R**elative **D**isplacement		
$L(\mathbf{y}), H(\mathbf{x})$	continuous low and high resolution irradiance field	$B(\mathbf{y}, \mathbf{x})$	blurring kernel
$E(\mathbf{y})$	noise	PSF_{lens}	PSF of the lens
PSF_{sensor}	PSF of the sensor	$U(\mathbf{x})$	continuous irradiance field before raster discretization
$\mathcal{M}$	magnification factor	$\delta\mathbf{H}$	$\tilde{\mathbf{H}} - \mathbf{H}$, deviation from the real solution
$\mathbf{H}$, $\mathbf{L}$	column vector of HRPs and LRPs, respectively	$\mathbf{P}$	exact coefficient matrix
$\mathbf{P}^+$	pseudo-inverse of $\mathbf{P}$	$\mathbf{E}$	column vector of noise
$\tilde{\mathbf{H}}$	column vector of estimated HRPs	$\tilde{\mathbf{P}}$	perturbed coefficient matrix
$\tilde{\mathbf{E}}$	equivalent noise	κ	condition number of $\mathbf{P}$
$\mathbf{r}$	residual error	$\delta\mathbf{E}$	$\tilde{\mathbf{E}} - \mathbf{E}$
$\delta\mathbf{P}$	$\tilde{\mathbf{P}} - \mathbf{P}$	$\Delta\mathbf{x}$	displacement of an LRP relative to an HRP
M, ε	integer and fractional part of $\mathcal{M}$	N_h, N_l	square root of the number of HRPs and LRPs, respectively
ω	parameter for the separable HRI used in norm estimation	Δ_r	maximum registration error, measured in the size of LRP
$g(\mathcal{M})$	a function related to $\|\|\mathbf{P}^+\|\|^{-1}$	δ_p	a function related to $\|\|\delta\mathbf{P}\|\|$
δ_e	RMS of $\delta\mathbf{E}$	δ_r	RMS of $\mathbf{r}$
σ_h	variation of $\mathbf{H}$	$\hat{\delta}_h$	a large threshold, chosen as 128

convolved in time. The PSF_{sensor} is a box function (see (10)) for CCD sensors [3].

1.2 What Is Superresolution and Its Limit?

As the readers may have different understanding about "superresolution," here we would like to make it more specific. In this paper, superresolution must have the effect of resolving LRIs into an HRI with *smaller* pixels. If the pixel size of HRI is identical to that of LRIs, we only refer to deblurring (or enhancement, etc.) to such techniques of sharpening images. It is obvious that PSF_{sensor} is related to the pixel size and accounts for fine-resolving, while PSF_{lens} is not and is for deblurring only. We will show later that the superresolution limits do not depend heavily on PSF_{lens}.

Although, in theory, the HRI at arbitrary magnification factor $\mathcal{M}$ can be computed by using smoothness regularization, we are in fact more interested in the "effective" magnification factors (EMFs), beyond which the resolution of HRI is more or less the same as that of those with smaller $\mathcal{M}$s. Here, two images are of the same resolution means that an image can be obtained by interpolating the other image or the salience of useful features in both images is identical. In this definition, the "resolution" of an image cannot simply be measured by the pixel size nor the highest frequency in the image (e.g., adding random noise does not improve resolution). On the other hand, the HRIs of interest are those images that are close to the ground truth. An HRI is useless if it does not reflect the ground truth, even if all its edges are sharp. Through the above analysis, the problem of finding the "limits of superresolution" can be converted to a more workable one: find the distribution of magnification factors beyond which the HRI deviates significantly from the ground truth. By finding such distribution, if $\mathcal{M}$ is not an EMF, people can save much computation and storage resources by trying a smaller EMF and then interpolating to this larger $\mathcal{M}$.

1.3 Previous Work and Main Results

The study on the problem of "limits of superresolution" is rare and incomplete. All existing studies focus on RBAs. Elad and Feuer [8] briefly discussed the choice of magnification factors, but they were considering the superresolution under different defocus and without relative motion and the coefficient matrix is block-Toeplitz. Moreover, their results appear to permit arbitrarily large magnification factors. Baker and Kanade [2] showed that both the condition number of the linear system and the volume of solutions grow fast with the increment of the magnification factor and an RBA can only generate an overly smoothed solution. However, they only pointed out the deteriorating tendency of the RBAs. No explicit limits are deduced. In fact, even the concept of "limits of superresolution" is not clarified in existing literature. We are the first ones to make this statement explicit.

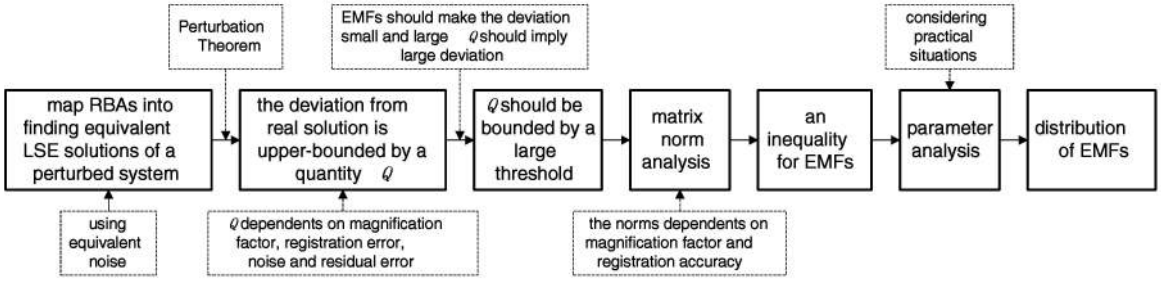

Fig. 1. The logic flow of our analysis on superresolution limits. The texts in dashed boxes are for supplementary information. We first formulate the RBAs as finding equivalent LSE solutions to the perturbed linear system due to noise and error. Using the perturbation theorem, we can have an upper-bound estimate Q on the deviation between the LSE solution and the ground truth. As the deviation and Q are correlated, Q should be small so that the deviation is also small. Then, the matrix norm analysis further gives an estimate on Q and an inequality that relates the EMF, registration error and noise level are derived. Considering practical situations, we finally derive the distribution of EMFs by assigning the parameters with their extreme values.

This paper is a breakthrough toward the problem of "do fundamental limits exist for superresolution?"[1] Also limited to RBAs and assuming local translation among LRIs, we give explicit bounds of the magnification factor and investigate sufficient number of LRIs.

We prove that fundamental limits do exist for RBAs under local translation. Specifically, we discuss two extreme cases and find that:

- The practical limit is 1.6, if the noise removal and registration is not good enough. Moreover, 2.5 is the first choice if one wants to try an $\mathcal{M}$ that is larger than 1.6.
- The theoretical limit is 5.7. Actually, the EMF can only distribute on some disjoint intervals.

Although we analyze by assuming box PSF, as we will shown later, the limits of superresolution is dominated by the ability of fine-resolving. Therefore, the above results are still general.

We derive the above limits from an inequality that an EMF should satisfy. The cornerstone for proving the inequality is a theorem on the perturbation of the least-square-error (LSE) solutions[2] to overdetermined systems assembled by the RBAs. This theorem states how much the LSE solution deviates from the real one if noise and error exist. The limits are found by bounding the deviation. To estimate the deviation, we succeed in finely estimating the norms of the pseudoinverse of the coefficient matrix of the system and the perturbation coefficient matrix.

We further analyze the influence of increasing the number of LRIs. We find that:

- If a special set of low-resolution pixels (LRPs), called the vertices set, have been captured, then adding more LRIs will only improve the resolution marginally.
- If the fractional part of $\mathcal{M}$ is 0.5, then the corresponding sufficient number of LRIs is $4\mathcal{M}^2$. For an integer $\mathcal{M}$, the sufficient number of LRI is $\mathcal{M}^2$.

1.4 Guideline of Our Analysis

Since our analysis is a bit mathematical, we would like to give a rough guideline here in order to help the reader better understand the following proofs. Fig. 1 illustrates the logic flow of our analysis on superresolution limits.

As we analyze in Section 1.2, we need to estimate the deviation $\delta\mathbf{H}$ between the HRI and the ground truth. Therefore, we first formulate the RBAs as finding the equivalent LSE solutions to the linear systems used in RBAs. Due to noise and error, these linear systems are somewhat different from the real ones.

Then, we apply the perturbation theorem in matrix analysis to estimate $\delta\mathbf{H}$. It is upper-bounded by a quantity Q. Though, mathematically, a large Q does not necessarily lead to a large $\delta\mathbf{H}$, we justify that their magnitude should be correlated. At least for very large Q, $\delta\mathbf{H}$ should also exceed a relatively small threshold so that the computed HRI does not reflect the ground truth image well. Therefore, the EMFs should make Q smaller than a *large* threshold.

The upper-bound estimate Q depends on some matrix or vector norms. Our meticulous analysis gives their estimates, represented by the magnification factor, registration error, noise reduction level, etc. This is the critical part of our paper. By plugging the estimates in the inequality that bounds Q, we derive an inequality that the EMFs should satisfy. To obtain fundamental limits, we assign extreme values that are hard to break under practical situations to the parameters. Finally, we obtain the distribution of EMFs by solving the inequality numerically.

The remainder of this paper is organized as follows: We first present the mathematical formulation of the problem in Section 2. This is followed by an analysis of the norms of matrices in Section 3. Then, in Section 4, we present the fundamental limits of superresolution under both practical and synthetic situations. In Section 5, we study the impact of increasing the number of LRIs on the superresolution limits. In Section 6, both real and synthetic experiments are presented to verify our theory. Finally, we present the conclusion and suggest some directions of future work in Section 7.

2 Problem Formulation

2.1 Fundamental Linear System and Its Perturbation

The discrete version of (1) gives the relation between LRPs and high-resolution pixels (HRPs) via a linear system:

$$\mathbf{L} = \mathbf{P}\mathbf{H} + \mathbf{E}, \tag{3}$$

where $\mathbf{L}$ is the column vector of the irradiance of all LRPs considered, $\mathbf{H}$ is the vectorized (namely, concatenation of matrix rows) irradiance of the HRI, $\mathbf{P}$ gives the weights of the

1. Its early version appeared in the *Proceedings of Computer Vision and Pattern Recognition, CVPR '01* [16].

2. Those algorithms that do not seek explicit LSE solutions can be viewed to find the equivalent LSE solutions. See the justifications in Section 2.1.

HRPs in order to obtain the irradiance of the corresponding LRPs, and $\mathbf{E}$ is the noise. Usually, RBAs simply replace irradiance with gray level. In the following analysis, we also follow this convention.

An example of $\mathbf{P}$ is:

$$\mathbf{P} = \frac{1}{625}\begin{pmatrix} 196 & 154 & 0 & 154 & 121 & 0 & 0 & 0 & 0 \\ 0 & 196 & 154 & 0 & 154 & 121 & 0 & 0 & 0 \\ 0 & 0 & 0 & 196 & 154 & 0 & 154 & 121 & 0 \\ 0 & 0 & 0 & 0 & 196 & 154 & 0 & 154 & 121 \\ 72 & 192 & 36 & 78 & 208 & 39 & 0 & 0 & 0 \\ 0 & 0 & 0 & 72 & 192 & 36 & 78 & 208 & 39 \\ 72 & 78 & 0 & 192 & 208 & 0 & 36 & 39 & 0 \\ 0 & 72 & 78 & 0 & 192 & 208 & 0 & 36 & 39 \\ 16 & 64 & 20 & 64 & 256 & 80 & 20 & 80 & 25 \end{pmatrix}, \quad (4)$$

which relates the HRI and LRPs shown in Fig. 2 with a magnification factor $\mathcal{M} = 25/16$. For instance, the sixth LRI covers 72/625, 192/625, 36/625, 78/625, 208/625, and 39/625 of the fourth to ninth HRPs. Therefore, assuming a box-function PSF, the sixth row of $\mathbf{P}$ is $\frac{1}{625}(0, 0, 0, 72, 192, 36, 78, 208, 39)$. This example shows that the structure of $\mathbf{P}$ may not be block-Toeplitz, contrary to what some researchers have assumed [8], [9]. A block-Toeplitz coefficient matrix corresponds to a shift-invariant filter. The coefficient matrix is block-Toeplitz only when there is only one relative displacement (see Section 3.2) and the magnification factor is an integer so that the LRPs can form complete LRIs. Assuming a block-Toeplitz coefficient matrix can greatly facilitate the analysis (such as applying Fourier transform as done in [9]). However, as we will assume an unlimited number of LRIs, such an assumption is abandoned in our paper. Nevertheless, a block-Toeplitz coefficient matrix is a special case of our analysis.

No matter how the RBAs vary (such as MAP [7], [8], [12], [21], ML [8], POCS [19], Iterative Backprojection [15], [17], etc.), they can be viewed as denoising first and then solving (3) by finding its LSE solution. This can be justified by using the notion of "equivalent noise." Suppose an RBA finds a solution $\mathbf{H}_0$, then the equivalent noise estimated by the algorithm is

$$\mathbf{E}_0 = \mathbf{L} - \mathbf{P}\mathbf{H}_0.$$

One may check that the LSE solution to (3) is also $\mathbf{H}_0$ when $\mathbf{E}$ is estimated as $\mathbf{E}_0$. Note that the LSE solution to a system $\mathbf{A}\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = \mathbf{A}^+\mathbf{b}$, where $\mathbf{A}^+$ is the Moore-Penrose generalized inverse [11], [22] of $\mathbf{A}$. Similarly, the ground-truth solution is also an LSE solution with appropriate equivalent noise. With such a posteriori equivalent noise, we are freed from dealing with the details of RBAs.

The performance of RBAs is affected by several factors: the level of noise that exists in the LRIs, the accuracy of the blurring kernel estimation and the accuracy of registration, including the correction of geometric distortion. The higher the level of noise or the poorer the kernel estimation and registration, the less improvement in resolution. Except those considering motion blur [4], [6], [19] in the superresolution of video sequences, most existing RBAs throw all possible errors into $\mathbf{E}$ and the coefficient matrix $\mathbf{P}$ is deemed to be precise. We choose to differentiate the noise from misregistration and that from gray evel. Regarding the errors in $\mathbf{P}$, the system that an RBA assembles should be a perturbed version of (3):

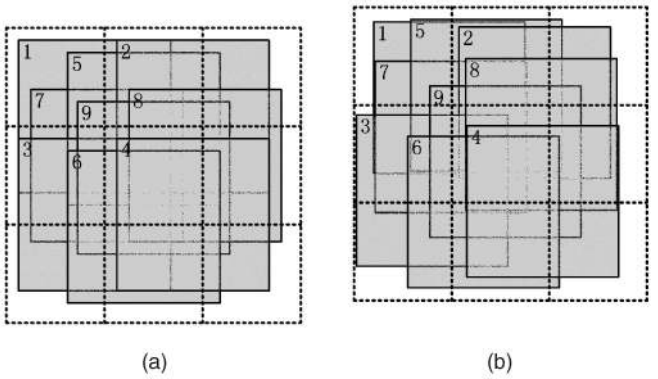

Fig. 2. The relation between low and high-resolution pixels. The dotted squares represent the HRPs, while the overlapped and shaded squares represent LRPs. The number on each LRP indicates the order of LRPs. The magnification factor is $\mathcal{M} = 25/16$. (a) The ground-truth registration. (b) The misregistration with registration error uniformly distributed between 0 and 0.08 pixel size of the LRP.

$$\mathbf{L} = \tilde{\mathbf{P}}\tilde{\mathbf{H}} + \tilde{\mathbf{E}}, \quad (5)$$

where $\tilde{\mathbf{P}} = \mathbf{P} + \delta\mathbf{P}$ is the estimated coefficient matrix, $\tilde{\mathbf{H}} = \mathbf{H} + \delta\mathbf{H}$ is the estimated solution, and $\tilde{\mathbf{E}}$ is the equivalent noise. Taking the coefficient matrix in (4), for example, the corresponding coefficient matrix due to misregistration is:

$$\tilde{\mathbf{P}} = \begin{pmatrix} 0.2847 & 0.2661 & 0 & 0.2322 & 0.2170 & 0 & 0 & 0 & 0 \\ 0 & 0.2918 & 0.2070 & 0 & 0.2932 & 0.2080 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.3603 & 0.2064 & 0 & .2755 & 0.1578 & 0 \\ 0 & 0 & 0 & 0 & 0.2657 & 0.2295 & 0 & 0.2709 & 0.2340 \\ 0.1502 & 0.3580 & 0.0512 & 0.1183 & 0.2820 & 0.0403 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1290 & 0.2766 & 0.0266 & 0.1695 & 0.3634 & 0.0349 \\ 0.1475 & 0.1483 & 0 & 0.3192 & 0.3208 & 0 & 0.0320 & 0.0322 & 0 \\ 0 & 0.1699 & 0.1364 & 0 & 0.3550 & 0.2850 & 0 & 0.0297 & 0.0239 \\ 0.0200 & 0.0838 & 0.0271 & 0.0976 & 0.4096 & 0.1328 & 0.0350 & 0.1466 & 0.0475 \end{pmatrix}, \quad (6)$$

where the LRPs are misregistered, as shown in Fig. 2b.

The error in estimating $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{E}}$ causes the deviation $\delta\mathbf{H}$ between $\tilde{\mathbf{H}}$ and $\mathbf{H}$. As we analyzed in Section 1.2, $\delta\mathbf{H}$ should not be too large.

2.2 The Perturbation Theorem

We should find out the relation between $\delta\mathbf{H}$ and the magnification factor. From the perturbation theory of linear systems, $\delta\mathbf{H}$ can be depicted by the following theorem [11], [13], [22]:

Theorem. *If both $\tilde{\mathbf{P}}$ and $\mathbf{P}$ are of full rank and $\kappa\varepsilon_{\mathbf{P}} < 1$, then*

$$||\delta\mathbf{H}|| \leq \frac{\kappa}{1 - \kappa\varepsilon_{\mathbf{P}}}\left[\varepsilon_{\mathbf{P}}\left(||\mathbf{H}|| + \kappa\frac{||\mathbf{r}||}{||\mathbf{P}||}\right) + \frac{||\delta\mathbf{E}||}{||\mathbf{P}||}\right], \quad (7)$$

where $\kappa = ||\mathbf{P}||||\mathbf{P}^+||$ is the condition number of the system, $\varepsilon_{\mathbf{P}} = ||\delta\mathbf{P}||/||\mathbf{P}||$, $\mathbf{r} = \mathbf{L} - \mathbf{P}\mathbf{H} - \mathbf{E}$ is the residual error[3] and $\delta\mathbf{E} = \tilde{\mathbf{E}} - \mathbf{E}$ is the error in estimating the noise. To be specific, the norms of a vector and a matrix are defined as: $||\mathbf{v}|| = \sqrt{\mathbf{v}^t \cdot \mathbf{v}}$, and $||\mathbf{A}|| = \max_{||\mathbf{x}||\neq 0}\frac{||\mathbf{A}\mathbf{x}||}{||\mathbf{x}||}$, respectively.

Inequality (7) is relatively sharp, i.e., the equality may hold in some circumstances [22].[4] This rules out abused versions of

3. As we deem (3) is exact, $\mathbf{r}$ is always $\mathbf{0}$. However, we choose to put it here in order to give the readers a complete picture of the perturbation theorem. This treatment does not affect our later analysis.

4. As it is inequality, the equality cannot hold for all cases.

the argument in this paper, namely, enlarge the right-hand side (RHS) of (7) in order to obtain smaller superresolution limits. Unfortunately, it is not guaranteed that $||\delta\mathbf{H}||$ will reach the magnitude of the RHS. For example, in the extreme case, it is theoretically possible that $\delta\mathbf{H}$ happens to be $\mathbf{0}$ or, equivalently, $\mathbf{H}$ is not only the solution to both (3) and (5), but also the solution to $\delta\mathbf{E} = -\delta\mathbf{P}\mathbf{H}$. However, usually we are not that lucky because $\delta\mathbf{P}$ and $\delta\mathbf{E}$ are independent random noise. In practice, they come from completely different sources: $\delta\mathbf{P}$ comes from registration error, while $\mathbf{E}$ and $\tilde{\mathbf{E}}$ come mainly from sensor noise. As a result, $\delta\mathbf{P}$ and $\delta\mathbf{E}$ cannot be correlated and, hence, the possibility that the RHS is large while $||\delta\mathbf{H}||$ is small seems to be slight. Therefore, we should view $||\delta\mathbf{H}||$ as being quite correlated with the RHS. If the RHS is very large, it is expected that $||\delta\mathbf{H}||$ will be above a threshold T such that there is no further resolution improvement in $\tilde{\mathbf{H}}$. As we will show, such a threshold T is, in general, quite small.

Then, it is apparent that the following should hold:

$$\text{RHS} \leq \hat{\delta}_h N_h, \tag{8}$$

where $\hat{\delta}_h$ is a number, say 128, that is much larger than T and N_h is the square root of the number of HRPs. Hence, estimating the RHS is of central role. Note that the RHS of (7) can be rewritten as:

$$\frac{||\mathbf{P}^+||}{1 - ||\mathbf{P}^+||||\delta\mathbf{P}||}[||\delta\mathbf{P}||(||\mathbf{H}|| + ||\mathbf{P}^+||||\mathbf{r}||) + ||\delta\mathbf{E}||],$$

where $||\mathbf{P}||$ is canceled. Then, (8) can be written as:

$$||\mathbf{P}^+||[||\delta\mathbf{P}||(||\mathbf{H}|| + ||\mathbf{P}^+||||\mathbf{r}|| + \hat{\delta}_h N_h) + ||\delta\mathbf{E}||] < \hat{\delta}_h N_h. \tag{9}$$

Consequently, estimating $||\mathbf{P}^+||$ and $||\delta\mathbf{P}||$ is critical.

3 Estimating the Norms of $\mathbf{P}^+$ and $\delta\mathbf{P}$

Several researchers have recognized the importance of the condition number of the coefficient matrix in such ill-posed problems of deblurring and superresolution. In [18], the growth rate of the condition number of Toeplitz matrix is studied. In [9], the conditioning of block-Toeplitz matrix that corresponds to a symmetric 2D PSF is investigated using a new conditioning measure. Both studies assume fixed weights and rely on root finding. Therefore, their problems are different from ours and the relation between condition number and magnification factor cannot be derived. Baker and Kanade [2] may be the first to directly tackle the relationship between condition number and magnification factor. In comparison, our estimation is much finer than Baker and Kanade's.

Before presenting the details of the norm estimation in this section, we first make some assumptions and introduce some notations.

3.1 Assumptions

In order to make the theoretical analysis possible, we make the following assumptions:

1. The superresolution is done locally on small regions of interest (ROI). If a large area is required, one may partition the area into relatively small regions and do superresolution region by region. This approach not only speeds up the computation and lowers the memory requirement, but also simplifies the problem since both the PSF and geometric distortion are uniform within a small region.
2. The number of LRIs is unlimited, but there are no multiple images taken at the same place. As a consequence, one cannot remove the noise simply by averaging the multiple images at the same place. The problem of sufficient number of LRIs will be investigated in Section 5.
3. The camera movement is locally translational.
4. The PSF is a box function, namely,

$$PSF(\mathbf{x}) = \begin{cases} S^{-2}, & |x| \leq S/2, |y| \leq S/2, \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

where $\mathbf{x} = (x, y)$ and S is the size of the LRP. Such PSF weights HRPs according to their portions of area inside the LRPs. For most applications, recovering $U(\mathbf{x}) = PSF_{lens} * H(\mathbf{x})$ is more fundamental since estimating PSF_{lens} is not an easy task. Moreover, we may compute $U(\mathbf{x})$ first and then deconvolve it to obtain $H(\mathbf{x})$. Most importantly, usually the PSF_{lens} at high resolution is simply the interpolation of that at low resolution. In this case, the deconvolved image at high resolution are also nearly the interpolation of that at low resolution. The proof is sketched in Appendix D. Therefore, the limits of superresolution is dominated by the limits of fine-resolving *unless* the PSF_{lens} at high resolution is not the interpolation of that at low resolution. As a result, we need not take PSF_{lens} into consideration.

3.2 Notations

Each LRP is a blending among its neighboring HRPs. We call the set of HRPs that contribute to an LRP the "influence region" of the LRP (Fig. 3a). In order to avoid the nasty boundary problem, which needs to extend the pixel values outside the ROI, we only utilize those LRPs whose influence regions are completely inside the ROI.

Note that, for a matrix $\mathbf{A}$, its norm $||\mathbf{A}|| = \max_{||\mathbf{x}|| \neq 0} \frac{||\mathbf{A}\mathbf{x}||}{||\mathbf{x}||}$, where $\mathbf{x}$ need not be a real vector. Therefore, given an $\mathbf{H}$ and the corresponding $\mathbf{L}$, $||\mathbf{L}||/||\mathbf{H}||$ is a lower-bound estimate of the norm of the coefficient matrix, where $\mathbf{H}$ need not be real. For an arbitrary $\mathbf{H}$, the corresponding $||\mathbf{L}||$ usually does not have a simple closed-form solution. In order to facilitate our norm estimation of both $\mathbf{L}$ and $\mathbf{H}$, we choose separable HRIs (Fig. 3b): $\mathbf{I}_h = \mathbf{W} \cdot \mathbf{W}^t$, where $\mathbf{W} = (1 \ \omega \ \omega^2 \ \cdots \ \omega^{N_h - 1})^t$. The vectorization of $\mathbf{I}_h$ is denoted by $\mathbf{H}_\omega$ and the corresponding vector of LRPs is denoted by $\mathbf{L}_\omega$. The choice of ω (may not be real) is dependent on our purpose. We have tried other choices of $\mathbf{I}_h$, but have found that the norm estimation is not as convenient.

Since the number of LRIs is unlimited, we may assume that whenever the top-left corner of an LRP is at a relative displacement (RD, Fig. 3a) $\Delta\mathbf{x} = (\Delta x, \Delta y)$ $(0 \leq \Delta x, \Delta y < 1)$ inside an HRP (i, j), then (3) includes all other LRPs with the same RD inside all other HRPs. In particular, for every RD, there is a "base LRP" whose top-left corner is inside the HRP $(0, 0)$ (Fig. 3b). With our separable HRI, if the pixel value of this base LRP is $L_{0,0}$, then the pixel value of other LRP at HRP (i, j) is $\omega^{i+j} L_{0,0}$. Therefore, both the norms of $\mathbf{H}_\omega$ and $\mathbf{L}_\omega$ can easily be written down and we can concentrate on investigating the base LRP.

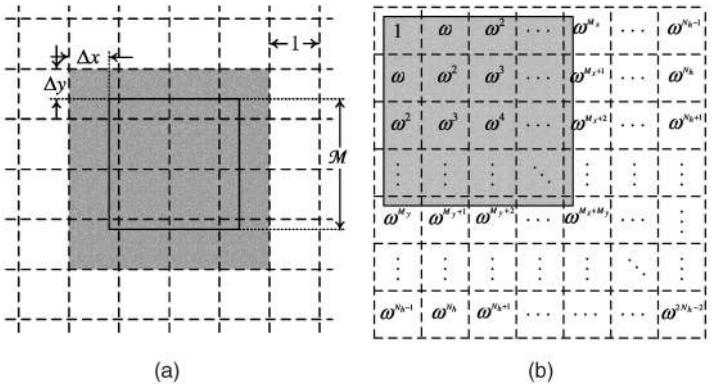

Fig. 3. The notations and graphical illustrations. (a) The influence region (shaded square) of an LRP (square in solid lines) and its relative displacement. (b) The separable HRI used for norm estimation and a base LRP (shaded square).

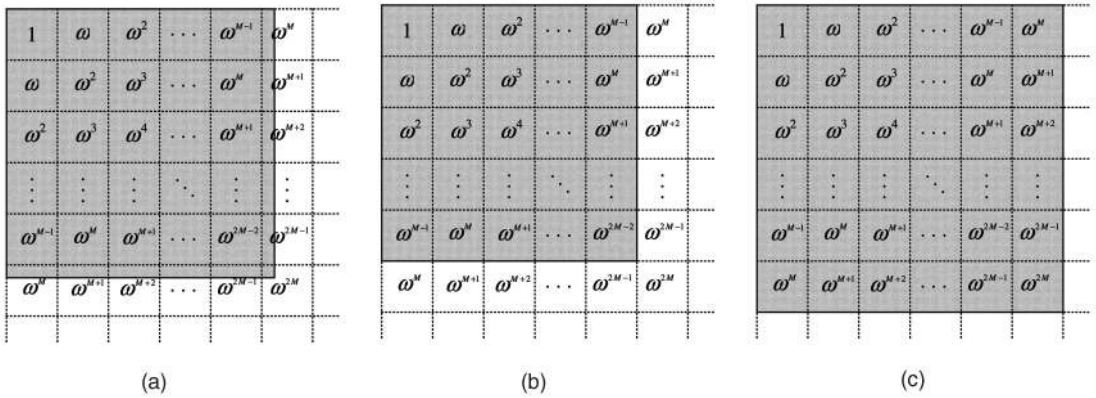

Fig. 4. The choice of ω_i $(1 = 1, 2, 3)$ is to make the integration on the shaded square zero, where the size of the shaded square is (a) $\mathcal{M} \times \mathcal{M}$, (b) $M \times M$, and (c) $(M + 1) \times (M + 1)$, respectively.

Finally, suppose that we want to raise the number of pixels in the ROI to $N_h \times N_h$ and the magnification factor is $\mathcal{M} > 1$, which can be decomposed into integer and fractional parts: $\mathcal{M} = M + \varepsilon$ $(0 \le \varepsilon < 1)$. We may normalize the size of HRPs to be unit 1, then the size of LRPs is $\mathcal{M}$. Now, we set out to estimate $||\mathbf{P}^+||$ and $||\delta\mathbf{P}||$.

3.3 Estimating $||\mathbf{P}^+||$ and $||\delta\mathbf{P}||$

$||\mathbf{P}^+||$ can be computed via [11], [22]: $||\mathbf{P}^+||^{-1} = \min_{||\mathbf{x}|| \neq 0} ||\mathbf{Px}||/||\mathbf{x}||$. For our problem, $\mathbf{x}$ is any HRI and $\mathbf{Px}$ is the corresponding vector of LRPs. Therefore, given any ω, $||\mathbf{P}^+||^{-1} \le ||\mathbf{L}_\omega||/||\mathbf{H}_\omega||$, where $\mathbf{L}_\omega$ and $\mathbf{H}_\omega$ have been defined in last section. To have a good estimate on $||\mathbf{P}^+||^{-1}$, ω had better be chosen such that every element of $\mathbf{L}_\omega$, i.e., the integral within every LRP, is as close to 0 as possible so that $||\mathbf{L}_\omega||$ is small. Obviously, if an ω can make the integral 0 in either the shaded squares shown in Fig. 4,[5] then other LRPs at other places will be close to 0. The corresponding ω satisfies:

$$\sum_{n=0}^{M-1} \omega_1^n + \varepsilon\omega_1^M = 0 \quad \text{or} \quad \sum_{n=0}^{M-1} \omega_2^n = 0 \quad \text{or} \quad \sum_{n=0}^{M} \omega_3^n = 0, \quad (11)$$

respectively (Figs. 4a, 4b, and 4c). With such ωs, we have an estimate:

5. Another choice of ω is permitted, but does not help our norm estimation.

$$||\mathbf{P}^+||^{-1} \le \mathcal{M}^{-2} g^2(\mathcal{M}) N_l N_h^{-1}, \text{ where}$$
$$g(\mathcal{M}) = \begin{cases} 1 - \varepsilon, & \text{if } M = 1, \\ \min(\varepsilon|1 - \omega_1^{M+1}||\omega_1|^{-\frac{M}{2}}, \varepsilon, 1 - \varepsilon), & \text{if } M > 1, \end{cases} \quad (12)$$

in which N_l is the square root of the number of LRPs and ω_1 should be chosen among the solutions so that it minimizes $|1 - \omega_1^{M+1}||\omega_1|^{-\frac{M}{2}}$. The detailed proof can be found in Appendix A. The curve of $g(\mathcal{M})$ is shown in Fig. 6a. We see that, when $1 \le \mathcal{M} \le 2$, $g(\mathcal{M})$ decreases linearly from 1 to zero. When $2 < \mathcal{M} < 4$, the curve of $g(\mathcal{M})$ consists of two saw-teeth. When $\mathcal{M} > 4$, the local maxima of $g(\mathcal{M})$ gradually decrease to 0. Moreover, $g(\mathcal{M}) = 0$ for all integer $\mathcal{M} > 1$. When $\mathcal{M} > 2$, the local maxima appear when $\varepsilon = 0.5$. This indicates that choosing such an $\mathcal{M}$ with $\varepsilon = 0.5$ can make the system most stable and error-resistant.

Take the coefficient matrix $\mathbf{P}$ in (4), for example, $||\mathbf{P}^+||^{-1} = 0.0166$, while the estimated value is 0.0784. Although the estimation does not seem very tight, the reader should be reminded that we need to derive an estimate that depends *only* on the magnification factor and the size of the matrix, i.e., it should be fulfilled by *any* coefficient matrix at a given magnification factor and size, in order to arrive at fundamental limits.

Our estimate is much better than that in [2], where the value of pixel (p, q) of the HRI is chosen as $(-1)^{p+q}$ and, thus, $g(\mathcal{M})$ is roughly estimated as 1.

Also, by choosing special ω, we have:

$$||\delta\mathbf{P}|| \geq \delta_p N_l/N_h, \text{where}$$
$$\delta_p = \Delta_r^2/4\left[M(1+M)^{-(1+\frac{1}{M})}\right]^2, \tag{13}$$

where $\mathcal{M}\Delta_r$ is the maximum registration error. The detailed proof can be found in Appendix B. Taking the coefficient matrices in (4) and (6), for example (where $\Delta_r = 0.08$ and $\mathcal{M} = 25/16$), the exact $||\delta\mathbf{P}||$ is 0.1243, while the estimated norm is 10^{-4}. Although the estimation seems quite pessimistic, it suffices to help us in achieving our results.

4 Limits of Reconstruction-Based Superresolution

As we have clarified in Section 1.2, a large deviation can imply that the details in the real solution are totally lost so that the computed solution is overly smooth or the computed solution contains details that are quite different from those in the real solution or even the profile of the superresolved image is quite different from the real one. In fact, the resolution is quite sensitive to the amount of deviation. As shown in Fig. 5, the average deviation between Figs. 5a and 5b is only 18.55 gray levels, but the details in Fig. 5a is completely lost in Fig. 5b.

To obtain fundamental limits, we set the average deviation $\hat{\delta}_h$ at 128, which is much larger than the reference threshold T=18.55, in order to ensure the correlation between the RHS of (7) and $||\delta\mathbf{H}||$.

Suppose

$$||\delta\mathbf{E}|| = \delta_e N_l,\ ||\mathbf{r}|| = \delta_r N_l,\ ||\mathbf{H}|| = \sigma_h N_h.$$

δ_e is nothing but the root mean square error (RMSE) of LRPs after denoising. Since globally shifting the intensity does not affect $||\delta\mathbf{H}||$, σ_h should be estimated by the variation of the superresolved image. Therefore, σ_h stands for the richness of details in the image. As for δ_r, it is always 0 (see the footnote in Section 2.2). Then, from (9), (12), and (13), we obtain:

$$\mathcal{M} < f(\mathcal{M}) \equiv g(\mathcal{M})\sqrt{\frac{\hat{\delta}_h}{\delta_e + (\sigma_h + \hat{\delta}_h)\delta_p}}. \tag{14}$$

The above inequality demarcates the existence intervals for EMFs.[6] Those that fall outside these intervals can still produce resolutions higher than that of low-resolution images, but may not be efficient in resolution improvement because they require stronger smoothness regularization and, hence, waste the effort of increasing pixel numbers. At this time, using smaller magnification factors is more economical.

Now, we may examine (14) to find the fundamental limits. In the ideal case, the data may be noise free except for the rounding error and the registration is perfect ($\Delta_r = 0$). We may view the rounding error as uniformly distributed between 0 and 0.5; therefore, we may assume δ_e to be its expectation value 0.25. With these ideal parameters, $\mathcal{M}$ must be in the disjoint intervals where the curve of $f(\mathcal{M})$ is above the line $\mathcal{M} = f$ in Fig. 6b, namely, $\mathcal{M} \in (1, 1.93) \cup (2.10, 2.89) \cup (3.15, 3.84) \cup (4.20, 4.80) \cup (5.27, 5.68)$.

6. Note that we never claim that any magnification factor inside these intervals is effective. We just claim that all EMF should be inside these intervals.

(a) (b)

Fig. 5. The sensitivity of resolution to deviation. (a) A high-resolution image. (b) The blurred image of (a). The average deviation between (a) and (b) is only 18.55 gray levels.

People may care more about the fundamental limits under practical situations. Indeed, as shown in (14), the limits depend on many factors. When the registration error is small (then $\delta_p \ll 1$), denoising becomes crucial. Moreover, the more details (i.e., the larger σ_h) in the image, the harder to recover them. One may check that, when $\sigma_h \geq 15$, $\Delta_r \geq 0.125$, and $\delta_e \geq 5$, the EMF can only be the interval where the curve of $f(\mathcal{M})$ is above the line $\mathcal{M} = f$ in Fig. 6c, namely, $\mathcal{M} < 1.59$.[7] This means that, unless the image is "appropriate" ($\sigma_h \leq 15$) and one can solve the problems of registration and denoising very well ($\Delta_r \leq 0.125$ and $\delta_e \leq 5$), the limit 1.6 is unbreakable. Unfortunately, these requirements are quite demanding because:

- $\sigma_h \leq 15$ may correspond to an image without any detail. Therefore, its resolution cannot be improved. For example, the variation of Fig. 5a is 56.25. Even an almost textureless image, Fig. 5b, has a variation of 44.40 gray levels.
- $\delta_e \leq 5.0$ is equivalent to a peak signal to noise ratio (PSNR) that is higher than 34.15dB. Such a PSNR is relatively high in image processing community. Although much higher performance was reported in the literature, only synthetic images with known noise type were tested in order to obtain the ground truth. In the practice of superresolution, neither the ground truth nor the noise type is known.
- With a relatively high level of noise, noise removal cannot be achieved very effectively without strong a priori knowledge[8] of the image content and the noise type. Moreover, image registration cannot be very accurate.
- The nonlinearity of camera response makes the linear relation (3) invalid unless the computation is totally done on radiometric domain, namely, recovering the irradiance at every pixel. Unfortunately, this is seldom considered in the practice of superresolution. This makes the denoising even harder.

As a result, we may consider 1.6 as the fundamental limit under practical situations. This means that if one wants to try a magnification factor larger than 1.6, then, instead of computing at such a large magnification factor directly, there

7. The peak at $\mathcal{M} = 2.5$ just touches the line. However, a bit more careful treatment in estimating $||\mathbf{P}^+||$ can easily invalidate $\mathcal{M} = 2.5$. We choose not to bother the reader with such minute details.

8. Otherwise, recognition techniques will replace superresolution.

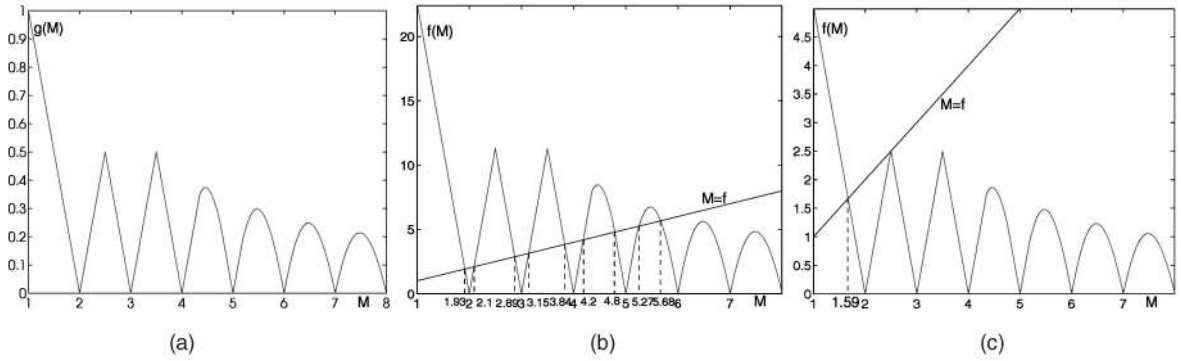

Fig. 6. The plot of $g(\mathcal{M})$ and the intervals that effective magnification factors distribute on. In (b) and (c), EMF must be in the intervals where the curve of $f(\mathcal{M})$ is above the line $\mathcal{M} = f$. (a) The curve of $g(\mathcal{M})$. (b) For synthetic data, $\mathcal{M} \in (1, 1.93) \cup (2.10, 2.89) \cup (3.15, 3.84) \cup (4.20, 4.80) \cup (5.27, 5.68)$. (c) Under practical situations, $1 < \mathcal{M} < 1.59$.

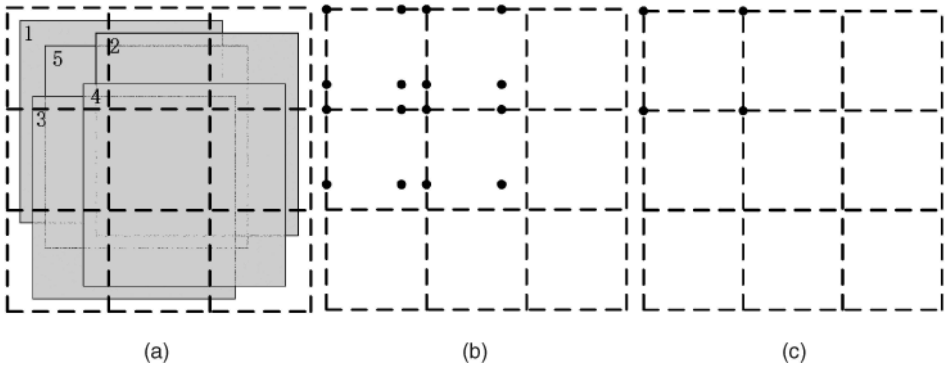

Fig. 7. The linear dependency among LRPs and the vertices sets. The white squares in dotted lines represent the HRPs. (a) The fifth LRP is a convex linear combination of numbers 1 to 4 LRPs. The overlapped and shaded large squares represent the LRPs. (b) The simplest vertices set when $\mathcal{M} = 5/4$ and $N_h = 3$. The dots indicate the top-left corners of the LRPs in the vertices set. (c) The simplest vertices set when $\mathcal{M} = 2$ and $N_h = 3$.

is a better way, namely, computing at a relatively small magnification factor, say 1.6, first, then interpolating the resulting image to the desired magnification factor. Nevertheless, $\mathcal{M} = 2.5$ is the best choice that is worth trying if one is very optimistic about his noise reduction and is eager to try a magnification factor that is larger than 1.6. Unfortunately, people have never realized this and usually try 2.0.

5 The Sufficient Number of LRIs

It is intuitive that increasing the number of LRPs can help resolution improvement. In this section, we will show that such improvement quickly becomes marginal. Consequently, people need not collect too many LRIs.

Consider five LRPs with gray level L_i $(i = 1, 2, \cdots, 5)$:

$$L_i = \mathbf{P}_i^t \mathbf{H} + E_i,$$

where $\mathbf{P}_i^t$ are the corresponding row-vectors of $\mathbf{P}$ and E_i are the corresponding entries of $\mathbf{E}$, respectively. Note that as long as their influence regions are the same (Fig. 7a), the coefficient vectors $\mathbf{P}_i$ $(i = 1, 2, \cdots, 5)$ are linearly dependent. For brevity, in the sequel, we simply say that the five LRPs are linearly dependent. Suppose the fifth LRP is a convex linear combination of the other four:

$$\mathbf{P}_5 = (\mathbf{P}_1\ \mathbf{P}_2\ \mathbf{P}_3\ \mathbf{P}_4)\mathbf{b},$$

where the entries of $\mathbf{b}$ are all nonnegative, then $L_5' = (L_1\ L_2\ L_3\ L_4)\mathbf{b}$ is usually more accurate than L_5 for the fifth LRP, as it is well known that a convex linear combination is a low-pass filtering process and suppresses noise. On the other hand, it is well known that the standard deviation of the average of N independent and identically distributed random variables reduces by a factor of $1/\sqrt{N}$. The reduction of this factor is stagnant when $N > 4$. Therefore, the improvement from the fifth LRP is marginal. In this viewpoint, if we have already captured a set V of LRPs such that every LRP is a convex linear combination of some LRPs in V, then the improvement from extra LRPs is marginal. Such V is the vertices set of all LRPs.

It is easy to see that, when $\mathcal{M}$ is not an integer, then the LRPs with RDs at $(0, 0)$, $(0, 1 - \varepsilon)$, $(1 - \varepsilon, 0)$ and $(1 - \varepsilon, 1 - \varepsilon)$ form the vertices set (Fig. 7b). Therefore, the sufficient number of LRPs is $[2 \times \text{ceil}(N_h - \mathcal{M})]^2$. Unfortunately, in practice, we cannot capture pixel by pixel. Instead, a set of LRPs that forms an LRI is acquired simultaneously and, in general, the RDs of these LRPs are quite different except for some special $\mathcal{M}$. Therefore, we cannot obtain the vertices set conveniently and economically for an arbitrary $\mathcal{M}$. However, things become simple when the fractional part of $\mathcal{M}$ is 0.5: To capture the vertices set, the camera only needs to be shifted $(2\mathcal{M})^2$ times, at a stepsize that equals half the size of the HRP, to make it run through an area of LRP. Hence the sufficient number of LRIs is $4\mathcal{M}^2$.

For an integer $\mathcal{M}$, the LRPs with RDs at $(0, 0)$ form the vertices set (Fig. 7c). One only needs to shift the camera $\mathcal{M}^2$ times, using a stepsize that equals the size of the HRP. The sufficient number of LRIs is thus $\mathcal{M}^2$.

6 Experiments

In the synthetic experiments, for the sake of visual comparison and in order to produce sharp HRIs, we only use the box function to filter a large image (Fig. 8a) to obtain 36 LRIs (Figs. 8b and 8c). The translation between nearby LRIs is only $1/7$ pixels. Each LRI is of size 23×23 pixels. Moreover, no noise except the rounding error is introduced to the LRIs and the registration is exact.

We compute superresolution results via MAP or ML. The MAP algorithm follows [8] and [12], in which the solution minimizes:

$$\Phi(\mathbf{H}) = (\mathbf{L} - \mathbf{PH})^t(\mathbf{L} - \mathbf{PH}) + \lambda \mathbf{H}^t \mathbf{\Lambda}^{-1} \mathbf{H}, \qquad (15)$$

where $\mathbf{\Lambda}^{-1}$ corresponds to a Laplacian operator that penalizes discontinuities:

$$\mathbf{\Lambda}^{-1} = \sum_{m=1}^{N_h^2} \mathbf{d}_m \mathbf{d}_m^t,$$

and $\mathbf{d}_m$ is the coefficient vector that corresponds to computing the Laplacian at the mth HRP:[9]

$$\mathbf{d}_m(n) = \begin{cases} 1, & \text{if } n = m, \\ -1/4, & \text{if pixel } n \text{ is one of the 4-neighbors of pixel } m, \\ 0, & \text{otherwise.} \end{cases}$$

The MAP solution $\mathbf{H}_{MAP}$ satisfies:

$$(\lambda \mathbf{\Lambda}^{-1} + \mathbf{P}^t \mathbf{P}) \mathbf{H}_{MAP} = \mathbf{P}^t \mathbf{L}. \qquad (16)$$

When $\lambda = 0$, the MAP solution reduces to the ML solution.

We first use all 36 LRIs for superresolution at various magnification factors, from $\mathcal{M} = 1.5$ to $\mathcal{M} = 3.0$. The results are shown in the second and fifth rows of Fig. 9, where all are ML solutions except Figs. 9e and 9l because the corresponding ML solutions exhibit completely random patterns due to the degeneracy of their coefficient matrices. Instead, the MAP solutions are shown there. The corresponding ground-truth images are shown in the first and fourth rows of Fig. 9. They are all scaled by bicubic interpolation to the size of Fig. 9l for better visual comparison.

We see that the resolution increases from $\mathcal{M} = 1.5$ (Fig. 9a) to $\mathcal{M} = 1.9$ (Fig. 9c). Then, the noise becomes unacceptable when $\mathcal{M} = 1.98$ (Fig. 9d) and $\mathcal{M} = 2.02$ (Fig. 9f). Therefore, these two magnification factors need smoothness regularization. The MAP solution at $\mathcal{M} = 2.0$ ($\lambda = 5 \times 10^{-14}$) does not exhibit a resolution that is higher than that of the ML solution at $\mathcal{M} = 1.9$. The criterion of choosing λ is that the resultant noise is approximately the same as that in the corresponding image in Fig. 9c. Smaller λs could result in sharper images, but the noise will also increase and, hence, more human knowledge is required to differentiate useful features from noise. Starting from $\mathcal{M} = 2.15$ (Fig. 9g), the resolution increases again until $\mathcal{M} = 2.8$ (Fig. 9j). Again, the noise in Fig. 9k is unacceptable and, hence, needs regularization. Moreover, the resolution of the MAP solution at $\mathcal{M} = 3.0$ ($\lambda = 5 \times 10^{-2}$) is not higher than that of the ML solution at $\mathcal{M} = 2.8$ either. Note that 1.98, 2.0, 2.02, 2.94, and 3.0 all fall outside the existence intervals of EMFs and, when $\mathcal{M}$ is inside those intervals, a larger $\mathcal{M}$ does produce higher resolution. The statistics in Table 2 also show that, when $\mathcal{M}$ is inside those regions, the RMSE from the ground truth is well-controlled and $||\mathbf{P}^+||$ is relatively small; when $\mathcal{M}$ is outside those regions, the RMSE and $||\mathbf{P}^+||$ increases drastically. Moreover, all the exact $||\mathbf{P}^+||^{-1}$s are well bounded by our estimation.

9. Minor adjustment is required when pixel m is on the image border.

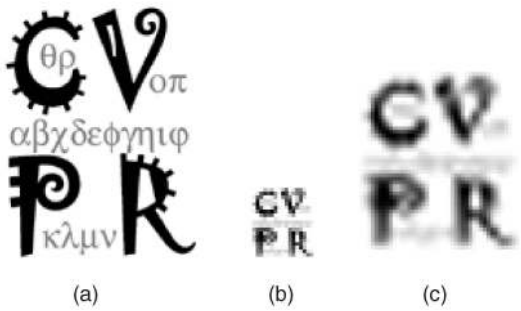

Fig. 8. The ideal image and its low-resolution images. (a) The ideal image used for generating LRIs. (b) One of the LRI. (c) Bicubic interpolation of (b) to the size of Fig. 9l for better visual comparison.

We also perform superresolution with the vertices sets only, where the members of the vertices set are found pixel by pixel. The results are shown in the third and sixth rows of Fig. 9. We see that the resolution difference from those from 36 full LRIs is negligible.[10] The RMSEs in Table 3 show that, in general, these results are closer to the ground truth. Note that the number of pixels (N_l^2) in the vertices sets are much less than the total number of LRPs in the 36 LRIs. This testifies to the conclusion in Section 5 that the vertices set can suppress noise better and using LRPs beyond the vertices set can only improve the resolution marginally. Again, the exact $||\mathbf{P}^+||^{-1}$s are well-bounded by our estimation.

To carry out real experiments, we utilize a computer-controllable vertical XY-table shown in Fig. 10a, where a monochromic CCD camera (white box in Fig. 10a) is attached vertically to the XY-plane. The camera can be moved at a stepsize of 0.025 mm and the error is below 0.2 percent. In our experiments, we place a piece of paper, printed with texts (Fig. 10b), about 2 meters away from the camera. When the camera moves 10,000 steps in both horizontal and vertical directions, the disparities are 190 ± 1 and 186 ± 1 pixels, respectively. Therefore, we can register the images at a higher accuracy than any existing registration algorithms, such as [15]. On the other hand, multiple images (20 images in this experiment) can be captured at the same place in order to suppress the random noise by simple averaging. Hence, we waive the vision and image processing techniques for registration and denoising.

We set the gamma of the CCD camera to be 1 and captured 12×12 images (Fig. 10c) of the texts, where the horizontal and vertical disparities between successive images are 0.0950 and 0.0934 pixels, respectively. The superresolution results on the region of interest are shown in Fig. 11, where all are MAP solutions and the $\mathcal{M}$s are 1.5, 2, and 2.5, respectively. The second to fourth rows are the results with box PSF_{sensor} only. The first and third columns

10. The noise in the superresolution result Fig. 9k with the vertices set seems acceptable. This is because the rounding error happens to be quite small. We have tested that, when there is no rounding error, even the results of $\mathcal{M} = 1.98$ and 2.02 are very good. Note that the limits are found by assuming that the rounding error is uniformly distributed between 0 and 0.5.

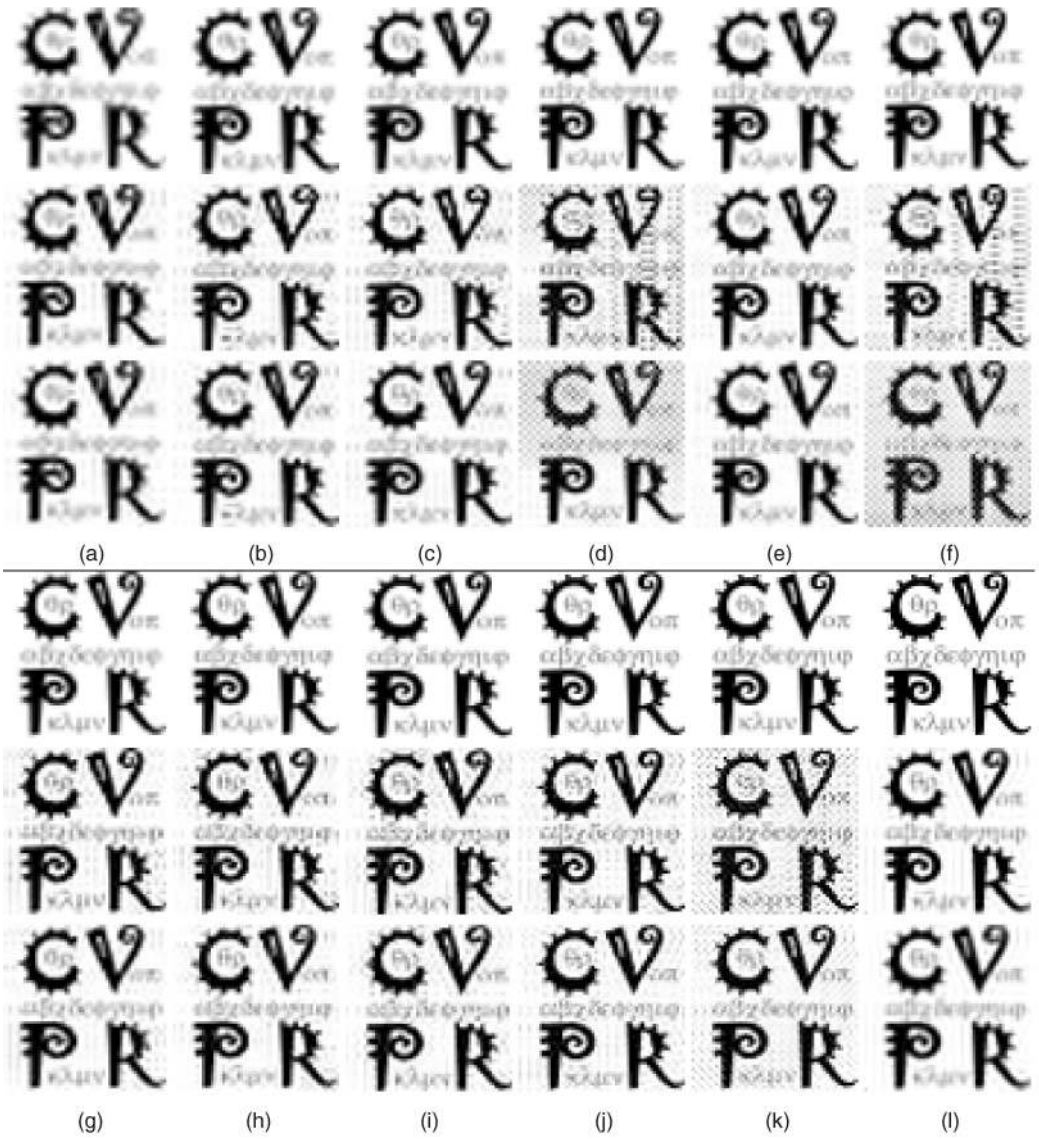

Fig. 9. The ground truth and superresolution results. The first and fourth rows are the ground truth images. The second and fifth rows are superresolution results using 36 LRIs. The third and sixth rows are superresolution results using the vertices sets only. They are all ML solutions except the second and third images of (e) and (l), which are MAP solutions with λs: 5×10^{-14}, 5×10^{-15}, 5×10^{-2}, and 5×10^{-2}, respectively. The criterion of choosing λs is that the resultant noise is approximately the same as that in the corresponding images in (c) or (j) in order to compare the resolution in a fair manner. Their corresponding ML solutions are random patterns and, hence, not shown here. The magnification factors are: (a) $\mathcal{M} = 1.5$. (b) $\mathcal{M} = 1.8$. (c) $\mathcal{M} = 1.9$. (d) $\mathcal{M} = 1.98$. (e) $\mathcal{M} = 2.0$. (f) $\mathcal{M} = 2.02$. (g) $\mathcal{M} = 2.15$. (h) $\mathcal{M} = 2.3$. (i) $\mathcal{M} = 2.5$. (j) $\mathcal{M} = 2.8$. (k) $\mathcal{M} = 2.94$. (l) $\mathcal{M} = 3.0$.

use those 36 images whose horizontal and vertical indices are both even, while the second and fourth columns use those images that are closest to the corresponding vertices sets, namely, Figs. 11d and 11f use images whose horizontal and vertical indices are both 0, 3, and 6. Figs. 11h and 11j use images whose horizontal and vertical indices are both 0 and 5, while Figs. 11l and 11n use images whose horizontal and vertical indices are both 0, 2, 4, 6, and 8. We see that, when $\mathcal{M}$ increases, the resolution remains the same.[11] Moreover, using 36 full images does not produce resolution that is higher than that using vertices sets.[12]

To verify the resolution dominance of fine-resolving, we also take PSF_{lens} into consideration. The results are shown in the fifth to seventh rows of Fig. 11, where the configurations are the same as those of previous three rows. The PSF_{lens} is assumed to be a Gaussian and its σ is estimated as 0.7 by examining the "i" dots in the image. When PSF_{lens} is involved, the coefficient matrix $\mathbf{P}$ in (15) and (16) should both be replaced by $\mathbf{PB}$, where $\mathbf{B}$ corresponds to the blur introduced by PSF_{lens}.

We see that, when $\mathcal{M}$ increases, the resolution still does not improve. Therefore, PSF_{lens} does not influence the limits of superresolution. Again, using more images than the vertices set only gives marginal improvement.

The relevant data are also shown in Table 4. In both cases, our experiments do not break the limit 1.6.

11. Someone may think that the resolution at $\mathcal{M} = 2.5$ is the highest. Actually, there are block artifacts.

12. The results of using vertices set are not as good as those in Fig. 9 partly because the LRPs used are not exactly the vertices set due to finite capturing resolution.

TABLE 2
Some Information of Superresolution with 36 Full LRIs

$\mathcal{M}$	1.5	1.8	1.9	1.98	2.0	2.02
N_h^2	841	1156	1296	1444	1444	1521
N_l^2	12996	12996	12996	12996	12996	12996
RMSE	19.9739	23.3299	24.7337	39.0442	170.1411	34.8174
$\|\|P^+\|\|^{-1}_{exact}$	0.1847	0.0307	0.0068	0.0003	0.0	0.0002
$\|\|P^+\|\|^{-1}_{estimate}$	0.4234	0.0392	0.0083	0.0003	0.0	0.0003
$\mathcal{M}$	2.15	2.3	2.5	2.8	2.94	3.0
N_h^2	1681	1936	2209	2809	3136	3249
N_l^2	12996	12996	12996	12996	12996	12996
RMSE	23.3550	24.6588	26.4026	24.8635	35.1881	173.9753
$\|\|P^+\|\|^{-1}_{exact}$	0.0100	0.0216	0.0159	0.0036	0.0003	0.0
$\|\|P^+\|\|^{-1}_{estimate}$	0.0129	0.0421	0.0911	0.0104	0.0008	0.0

TABLE 3
Some Information of Superresolution with the Vertices Sets

$\mathcal{M}$	1.5	1.8	1.9	1.98	2.0	2.02
N_h^2	841	1156	1296	1444	1444	1521
N_l^2	3136	4356	4900	5476	1369	5476
RMSE	13.0102	12.6872	10.8545	34.3737	108.3541	41.2814
$\|\|P^+\|\|^{-1}_{exact}$	0.1999	0.0239	0.0054	0.0002	0.0	0.0002
$\|\|P^+\|\|^{-1}_{estimate}$	0.2146	0.0240	0.0054	0.0002	0.0	0.0002
$\mathcal{M}$	2.15	2.3	2.5	2.8	2.94	3.0
N_h^2	1681	1936	2209	2809	3136	3249
N_l^2	6084	7056	8100	10404	11664	3025
RMSE	12.2171	14.1028	16.6417	13.7887	15.6689	171.9798
$\|\|P^+\|\|^{-1}_{exact}$	0.0089	0.0251	0.0296	0.0066	0.0006	0.0
$\|\|P^+\|\|^{-1}_{estimate}$	0.0093	0.0325	0.0766	0.0098	0.0008	0.0

7 Conclusions and Future Work

We analyze reconstruction-based algorithms for superresolution under local translation and give explicit limits under both practical and synthetic conditions. Under practical conditions, the limit is found to be 1.6 if the denoising and registration is not good enough, while, under synthetic conditions, it is 5.7 and the effective magnification factors distribute on disjoint intervals. We also find that if the vertices set has been captured, then it is hard to further improve the resolution by adding more LRIs. Furthermore, when the fractional part of $\mathcal{M}$ is 0.5, the sufficient number of LRIs is $4\mathcal{M}^2$. When the magnification factor is an integer, the sufficient number of LRIs is $\mathcal{M}^2$. Our experiments have verified all these conclusions. Our analysis indicates that to achieve superresolution at large magnification factors, RBAs are not favorable. One should try other kinds of superresolution algorithms, such as recognition-based algorithms [1].

Our analysis is based on the assumption of local translation. We believe that allowing higher order motion, such as rotation and perspective transformation, will make the limits larger, but they are still bounded. Moreover, in recent years, component-wise perturbation theory [13] has been developed for more accurate estimation. We hope this novel theory could improve our results. In addition, we also want to extend our analysis to the case that PSF_{sensor} is not a box function. Finally, the analysis on other kinds of superresolution algorithms, such as frequency-domain methods [14] or learning-based methods [1], [10], or motionless superresolution, such as superresolution from defocus [20], is also attractive.

Appendix A
Proof of (12)

Proof. Since the PSF is also separable, we may discuss the 1D case first and then derive in 2D. For each RD, there are exactly $N_x N_y$ LRPs that are inside the ROI, where

$$N_x = N_h - \text{ceil}(\mathcal{M} + \Delta x) + 1,$$
$$N_y = N_h - \text{ceil}(\mathcal{M} + \Delta y) + 1,$$

Suppose there are K different RDs, then $\mathbf{H}$ and $\mathbf{L}$ are column vectors of sizes $N_h^2 \times 1$ and $N_l^2 \times 1$, respectively, where

$$N_l = \sqrt{\sum_{k=1}^{K} N_x^{(k)} N_y^{(k)}}.$$

For the kth RD $\Delta\mathbf{x}^{(k)} = (\Delta x^{(k)}, \Delta y^{(k)})$, let $M_x^{(k)}$+1 be the width of the influence region of an LRP (Fig. 12), then $M_x^{(k)} = \text{ceil}(\mathcal{M} + \Delta x^{(k)}) - 1 = N_h - N_x^{(k)}$ and the horizontal portions of the HRPs inside the LRP are $\mathcal{M}^{-1} x_n^{(k)},\ (n = 0, \cdots, M_x^{(k)})$, respectively, where (Fig. 12)

$$x_n^{(k)} = \begin{cases} 1 - \Delta x^{(k)}, & \text{if } n = 0, \\ 1, & \text{if } 0 < n < M_x^{(k)}, \\ \mathcal{M} - (M_x^{(k)} - \Delta x^{(k)}), & \text{if } n = M_x^{(k)}. \end{cases}$$

Then, the value of the base LRP is

$$L_{0,0}^{(k)} = (\mathcal{M}^{-1} L_x^{(k)})(\mathcal{M}^{-1} L_y^{(k)}) = \mathcal{M}^{-2} L_x^{(k)} L_y^{(k)},$$

Fig. 10. The device and images in the real experiment. (a) The vertical XY-table and the CCD camera (white box) in our experiments. (b) The content of the LRIs. (c) One of the LRIs captured and the region of interest (white box).

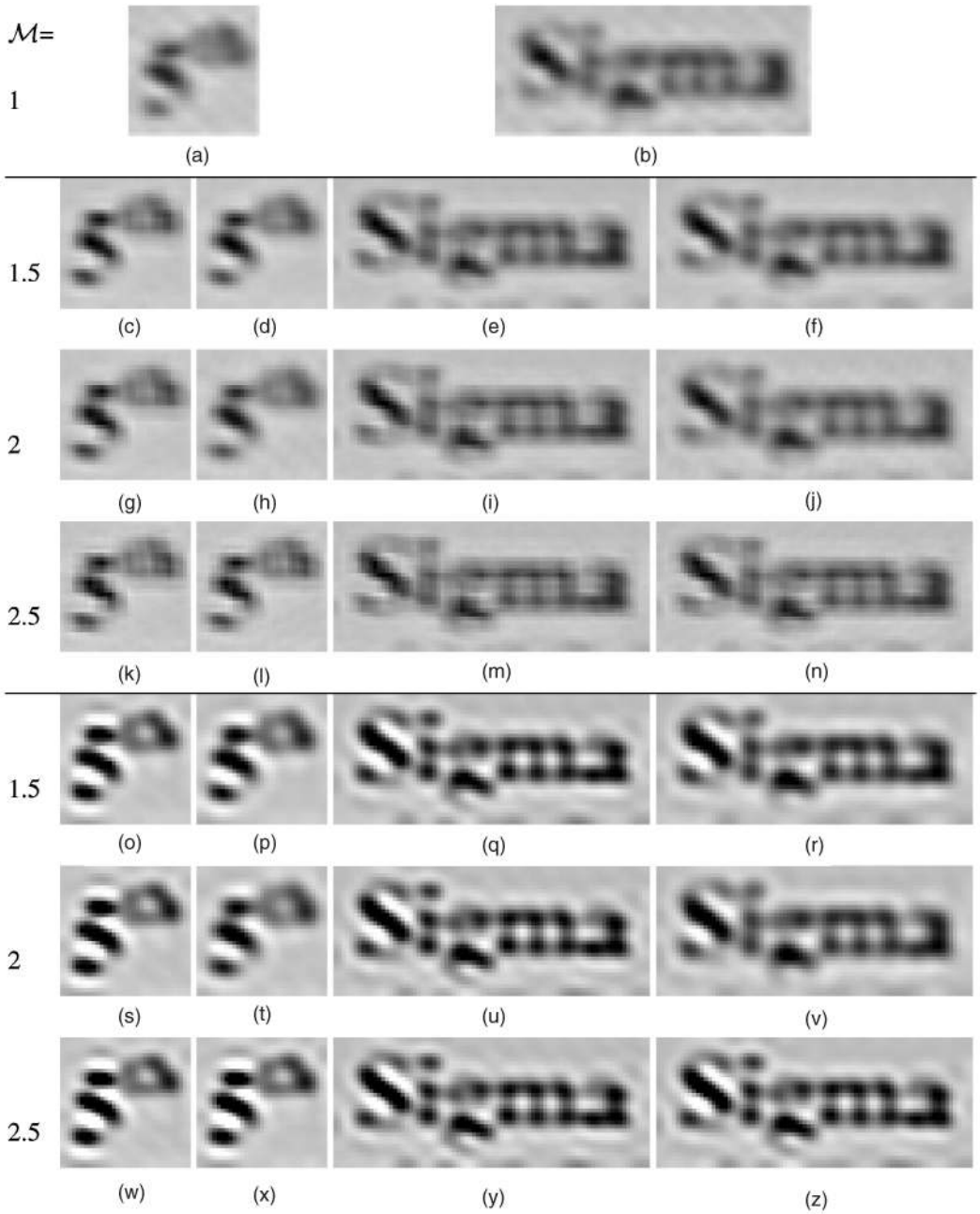

Fig. 11. Superresolution of the ROI by MAP with $\lambda = 0.5$. (a) and (b) Blowup of the ROI. (c)-(n) Superresolution with only the box PSF_{sensor} being considered. The magnification factor is shown on the left. The first and third columns use 36 full images, (d) and (f) use nine full images, (h) and (j) use four full images only, and (l) and (n) use 25 full images. The configurations of (o)-(z) are the same as those of (c)-(n), except that PSF_{lens} is also considered. All images are enlarged to the size of (z) by bicubic interpolation.

where $L_x^{(k)} = \sum_{n=0}^{M_x^{(k)}} x_n^{(k)} \omega^n$ (Fig. 12) and $L_y^{(k)}$ is defined similarly. As the value of other LRP with the same RD inside an HRP (l, m) is $L_{l,m}^{(k)} = \omega^{l+m} L_{0,0}^{(k)}$, the norms of low and high-resolution vectors are

$$||\mathbf{L}_\omega|| = \left(\sum_{k=1}^{K} \sum_{l=0}^{N_y^{(k)}-1} \sum_{m=0}^{N_x^{(k)}-1} |\mathcal{M}^{-2} L_x^{(k)} L_y^{(k)}|^2 |\omega|^{2l+2m} \right)^{\frac{1}{2}}$$

$$= \begin{cases} \mathcal{M}^{-2} |1-|\omega|^2|^{-1} \left(\sum_{k=1}^{K} |L_x^{(k)} L_y^{(k)}|^2 (|\omega|^{2N_x^{(k)}} - 1)(|\omega|^{2N_y^{(k)}} - 1) \right)^{\frac{1}{2}}, & \text{if } |\omega| \neq 1, \\ N_l \mathcal{M}^{-2} |L_x^{(k)} L_y^{(k)}|, & \text{if } |\omega| = 1, \end{cases}$$

and

$$||\mathbf{H}_\omega|| = \left(\sum_{l=0}^{N_h-1} \sum_{m=0}^{N_h-1} |\omega|^{2l+2m} \right)^{\frac{1}{2}}$$

$$= \begin{cases} |1 - |\omega|^2|^{-1} ||\omega|^{2N_h} - 1|, & \text{if } |\omega| \neq 1, \\ N_h, & \text{if } |\omega| = 1, \end{cases}$$

respectively.

To estimate $||\mathbf{P}^+||$, as explained in Section 3.3, ω can be chosen to be ω_1, ω_2, and ω_3 that satisfy one of the equations in (11), respectively. For ω_1, which is one of the solutions to:

TABLE 4
Some Information of Real Experiments

	Fig. 11(c)	Fig. 11(d)	Fig. 11(e)	Fig. 11(f)
$\mathcal{M}$	1.5	1.5	1.5	1.5
N_h^2	400	400	960	960
N_l^2	5550	1406	14025	3478
$\|\|P^+\|\|^{-1}_{exact}$	0.157484	0.066341	0.210126	0.071710
$\|\|P^+\|\|^{-1}_{estimate}$	0.413880	0.208315	0.424691	0.211488
	Fig. 11(g)	Fig. 11(h)	Fig. 11(i)	Fig. 11(j)
$\mathcal{M}$	2	2	2	2
N_h^2	676	676	1664	1664
N_l^2	5329	625	13651	1575
$\|\|P^+\|\|^{-1}_{exact}$	0.0	0.0	0.0	0.0
$\|\|P^+\|\|^{-1}_{estimate}$	0.0	0.0	0.0	0.0
	Fig. 11(k)	Fig. 11(l)	Fig. 11(m)	Fig. 11(n)
$\mathcal{M}$	2.5	2.5	2.5	2.5
N_h^2	1089	1089	2640	2640
N_l^2	5476	3782	13838	9516
$\|\|P^+\|\|^{-1}_{exact}$	0.020720	0.015595	0.020745	0.015736
$\|\|P^+\|\|^{-1}_{estimate}$	0.089697	0.074543	0.091579	0.075943

$$\sum_{n=0}^{M-1} \omega^n + \varepsilon\omega^M = 0, \tag{17}$$

we can prove that $|\omega_1| > 1$ and

1. If $x_0^{(k)} \geq \varepsilon$, then $M_x^{(k)} = M$ (Fig. 13a), hence

$$\begin{aligned} L_x^{(k)} &= x_0^{(k)} + \sum_{n=1}^{M-1} \omega_1^n + x_M^{(k)}\omega_1^M \\ &= x_0^{(k)} + (-\varepsilon\omega_1^M - 1) + (1 + \varepsilon - x_0^{(k)})\omega_1^M \\ &= (x_0^{(k)} - 1)(1 - \omega_1^M). \end{aligned}$$

Therefore,

$$\begin{aligned} |L_x^{(k)}| &= |x_0^{(k)} - 1||1 - \omega_1^M| \leq (1 - \varepsilon)|1 - \omega_1^M| \\ &= |(1 - \varepsilon)(1 - \omega_1^M)|. \end{aligned}$$

From (17), we have

$$\frac{1 - \omega_1^M}{1 - \omega_1} + \varepsilon\omega_1^M = 0.$$

or, equivalently,

$$1 - \omega_1^M + \varepsilon\omega_1^M - \varepsilon\omega_1^{M+1} = 0.$$

Hence,

$$\begin{aligned} |L_x^{(k)}| &\leq |(1 - \varepsilon)(1 - \omega_1^M)| \\ &= |-\varepsilon + \varepsilon\omega_1^{M+1}| = \varepsilon|1 - \omega_1^{M+1}|. \end{aligned}$$

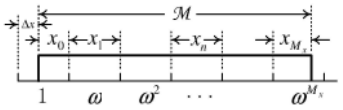

Fig. 12. The notations and graphical illustrations for matrix norm estimation. Δ_x is the relative displacement of the LRP. $M_x + 1$ is the width of the influence region. x_i $(i = 0, \cdots, M_x)$ are the portion of an HRP inside the LRP.

2. Else $M_x^{(k)} = M + 1$ (Fig. 13b), then

$$\begin{aligned} L_x^{(k)} &= x_0^{(k)} + \sum_{n=1}^{M} \omega_1^n + x_{M+1}^{(k)}\omega_1^{M+1} \\ &= x_0^{(k)} - \varepsilon\omega_1^{M+1} + (\varepsilon - x_0^{(k)})\omega_1^{M+1} \\ &= x_0^{(k)}(1 - \omega_1^{M+1}). \end{aligned}$$

Therefore, we always have $|L_x^{(k)}| \leq \varepsilon|1 - \omega_1^{M+1}|$. Similarly, $|L_y^{(k)}| \leq \varepsilon|1 - \omega_1^{M+1}|$. Consequently,

$$\begin{aligned} ||\mathbf{P}^+||^{-1} &\leq ||\mathbf{L}_{\omega_1}||/||\mathbf{H}_{\omega_1}|| \\ &= \frac{\mathcal{M}^{-2}|1-|\omega_1|^2|^{-1}\left(\sum_{k=1}^{K}|L_x^{(k)}L_y^{(k)}|^2(|\omega_1|^{2N_x^{(k)}}-1)(|\omega_1|^{2N_y^{(k)}}-1)\right)^{\frac{1}{2}}}{|1-|\omega_1|^2|^{-1}||\omega_1|^{2N_h}-1|} \\ &\leq \mathcal{M}^{-2}\varepsilon^2|1-\omega_1^{M+1}|^2\left(\sum_{k=1}^{K}\frac{|\omega_1|^{2N_x^{(k)}}-1}{|\omega_1|^{2N_h}-1}\frac{|\omega_1|^{2N_y^{(k)}}-1}{|\omega_1|^{2N_h}-1}\right)^{\frac{1}{2}} \\ &\leq \varepsilon^2\mathcal{M}^{-2}|1-\omega_1^{M+1}|^2|\omega_1|^{-M}N_lN_h^{-1}, \end{aligned}$$

where we have used the inequality

$$\frac{a^{2N_x} - 1}{a^{2N_h} - 1} \leq \frac{N_x}{N_h}a^{N_x - N_h} \leq \frac{N_x}{N_h}a^{-M}, \ \forall\ a > 1.$$

To best estimate $||\mathbf{P}^+||^{-1}$, ω_1 should be chosen such that it satisfies (17) and, on the other hand, minimizes $|1 - \omega_1^{M+1}|^2|\omega_1|^{-M}$.

In the second and third steps, we choose ω_2 and ω_3 to be any one of the solutions to $\sum_{n=0}^{M-1}\omega^n = 0$ and $\sum_{n=0}^{M}\omega^n = 0$, respectively. Mimicking the previous proof, we have:

$$\begin{aligned} &||\mathbf{P}^+||^{-1} \leq ||\mathbf{L}_{\omega_2}||/||\mathbf{H}_{\omega_2}|| \leq \varepsilon^2\mathcal{M}^{-2}N_lN_h^{-1}, \\ &||\mathbf{P}^+||^{-1} \leq ||\mathbf{L}_{\omega_3}||/||\mathbf{H}_{\omega_3}|| \leq (1 - \varepsilon)^2\mathcal{M}^{-2}N_lN_h^{-1}. \end{aligned}$$

Summing up, we have (12). □

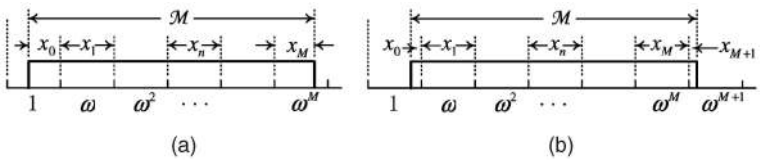

Fig. 13. The graphical illustrations of different cases in estimating $||\mathbf{P}^+||^{-1}$. (a) When $x_0 \geq \varepsilon$, the LRP covers $M + 1$ HRPs. (b) When $x_0 < \varepsilon$, the LRP covers $M + 2$ HRPs.

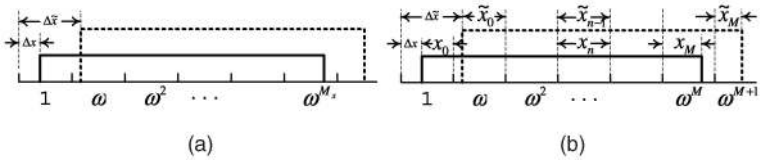

Fig. 14. Registration error of an LRP. The rectangle within solid lines is the correctly registered LRP, while the rectangle enclosed in dotted lines is the misregistered LRP. (a) An LRP at relative displacement Δx is misregistered at $\Delta\tilde{x}$. (b) The case that the correct and wrong registrations are in nearby HRPs and both cover $M+1$ HRPs.

APPENDIX B
PROOF OF (13)

Proof. Inheriting previous notations, suppose the base LRP is at $\Delta\mathbf{x}^{(k)}$ in the first HRP and it is registered at $\Delta\tilde{\mathbf{x}}^{(\mathbf{k})}$, relative to the same HRP (Fig. 14a). Now, $\Delta\tilde{\mathbf{x}}^{(\mathbf{k})}$ is not constrained to satisfy $0 \le \Delta\tilde{x}^{(k)}, \Delta\tilde{y}^{(k)} < 1$.). We can prove (see Appendix C) that the difference of pixel values at these two registrations is:

$$\begin{aligned} |\delta L_x^{(k)}| &= |L_x^{(k)} - \tilde{L}_x^{(k)}| \ge |\Delta x^{(k)} - \Delta\tilde{x}^{(k)}|\omega_4(1-\omega_4^M) \\ &\equiv \varepsilon_x^{(k)}\omega_4(1-\omega_4^M), \end{aligned} \tag{18}$$

where $0 \le \varepsilon_x^{(k)} \le \mathcal{M}\Delta_r$ and $\mathcal{M}\Delta_r$ is the maximum registration error. Similarly, $|\delta L_y^{(k)}| \ge \varepsilon_y^{(k)}\omega_4(1-\omega_4^M)$, where $0 \le \varepsilon_y^{(k)} \le \mathcal{M}\Delta_r$. Both $\varepsilon_x^{(k)}$ and $\varepsilon_y^{(k)}$ can be viewed as a random variable uniformly distributed between 0 and $\mathcal{M}\Delta_r$. Therefore, expectations of $\varepsilon_x^{(k)}$ and $\varepsilon_y^{(k)}$ are both $\mathcal{M}\Delta_r/2$ and $||\delta\mathbf{P}||$ can be estimated as:

$$\begin{aligned} &||\delta\mathbf{P}|| \ge ||\delta\mathbf{L}_{\omega_4}||/||\mathbf{H}_{\omega_4}|| \\ &\ge \omega_4^2(1-\omega_4^M)^2\Delta_r^2/4\left(\sum_{k=1}^{K}\frac{1-\omega_4^{2N_x^{(k)}}}{1-\omega_4^{2N_h}}\frac{1-\omega_4^{2N_y^{(k)}}}{1-\omega_4^{2N_h}}\right)^{\frac{1}{2}} \\ &\ge \omega_4^2(1-\omega_4^M)^2\Delta_r^2/4N_lN_h^{-1}, \end{aligned}$$

where we have used the inequality

$$\frac{1-a^{2N_x}}{1-a^{2N_h}} \ge \frac{N_x}{N_h},\ \forall\ 0 \le a < 1.$$

To best estimate $||\delta\mathbf{P}||$, ω_4 should be chosen such that it maximizes $\omega(1-\omega^M)$, or $\omega_4 = (1+M)^{-\frac{1}{M}}$. As a result,

$$||\delta\mathbf{P}|| \ge \Delta_r^2/4\left[M(1+M)^{-(1+\frac{1}{M})}\right]^2 N_lN_h^{-1}.$$

Thus, we have (13). □

APPENDIX C
PROOF OF (18)

Proof. Since the weights for an LRP change with its influence region, we should differentiate various combinations of $\Delta x^{(k)}$ and $\Delta\tilde{x}^{(k)}$. If we constrain the registration error to be less than one pixel size of the HRP, i.e., $\mathcal{M}\Delta_r < 1$, then there are *10* possible cases in total. So, we only take one nontrivial case for example. The other cases can be analyzed similarly. For $\mathcal{M}\Delta_r \ge 1$, $|\delta L_x^{(k)}|$ will be larger.

Suppose that $\Delta x^{(k)} < 1-\varepsilon$ and $\Delta\tilde{x}^{(k)} > 1$ (Fig. 14b), then $x_0^{(k)} \ge \varepsilon$ and from $\mathcal{M}\Delta_r < 1$, we have $\varepsilon < \tilde{x}_0^{(k)} < 1$. Therefore,

$$\begin{aligned} |L_x^{(k)} - \tilde{L}_x^{(k)}| &= |[x_0^{(k)} + \sum_{n=1}^{M-1}\omega_4^n + (1+\varepsilon-x_0^{(k)})\omega_4^M] \\ &\quad - [\tilde{x}_0^{(k)}\omega_4 + \sum_{n=2}^{M}\omega_4^n + (1+\varepsilon-\tilde{x}_0^{(k)})\omega_4^{M+1}]| \\ &= (1+x_0^{(k)}-\tilde{x}_0^{(k)})\omega_4(1-\omega_4^M) \\ &\quad + (1-\omega_4)[x_0^{(k)}(1-\omega_4^M) + \varepsilon\omega_4^M] \\ &> (1+x_0^{(k)}-\tilde{x}_0^{(k)})\omega_4(1-\omega_4^M) \\ &= (\Delta\tilde{x}^{(k)} - \Delta x^{(k)})\omega_4(1-\omega_4^M). \end{aligned}$$

□

APPENDIX D
PROOF OF THE RESOLUTION DOMINANCE OF FINE-RESOLVING

Proof. In this appendix, we will show that, if two blurred images are of the same resolution and so are their blurring kernels, then, after deblurring, the two images are still of the same resolution. With such a conclusion, we are freed from considering PSF_{lens}.

Let $\mathbf{U}_i$ $(i=1,2)$ be the images fine-resolved using smaller pixel sizes, $\mathbf{B}_i$ be the blurring kernels that correspond to PSF_{lens} and $\mathbf{H}_i$ be the high-resolution images after deblurring, namely,

$$\begin{aligned} \mathbf{U}_1 &= \mathbf{B}_1\mathbf{H}_1, \\ \mathbf{U}_2 &= \mathbf{B}_2\mathbf{H}_2. \end{aligned} \tag{19}$$

Suppose that the magnification factor of $\mathbf{H}_2$ is larger than that of $\mathbf{H}_1$ and the resolution of $\mathbf{U}_2$ is the same as that of $\mathbf{U}_1$, namely, there exists an interpolation matrix $\mathbf{T}$ such that $\mathbf{U}_2 = \mathbf{T}\mathbf{U}_1$. In order not to drown the reader with complex notations, in the sequel we prove the 1D case only. Then, we may write down the entries of the ideal interpolation matrix:

$$\mathbf{T}_{k,l} = \mathrm{sinc}(x_k - l),$$

where x_ks are the coordinate of pixels of $\mathbf{U}_2$ and the pixel size of $\mathbf{U}_1$ is assumed to be 1. Then, because:

$$\begin{aligned} &\sum_k \mathrm{sinc}(x_k-l)\mathrm{sinc}(x_k-n) \\ &\approx \int \mathrm{sinc}(x-l)\mathrm{sinc}(x-n)dx = \mathrm{sinc}(l-n) = \begin{cases} 1, & \text{if } l=n, \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

we have:

$$\mathbf{T}^t\mathbf{T} \approx \mathbf{I},$$

where $\mathbf{I}$ is the identity matrix. If the blurring kernel $\mathbf{B}_2$ is also the interpolation of $\mathbf{B}_1$, namely, $\mathbf{B}_2 = \mathbf{T}\mathbf{B}_1\mathbf{T}^t$, then $\mathbf{U}_2 = \mathbf{T}\mathbf{U}_1 = \mathbf{T}\mathbf{B}_1\mathbf{H}_1 \approx \mathbf{B}_2(\mathbf{T}\mathbf{H}_1)$. This implies that the solution $\mathbf{H}_2$ to (19) is close to $\mathbf{T}\mathbf{H}_1$, or the interpolation of $\mathbf{H}_1$, due to the smoothness regularization used in deblurring algorithms. Therefore, there is little resolution improvement. □

REFERENCES

[1] S. Baker and T. Kanade, "Hallucinating Faces," Technical Report CMU-RI-TR-99-32, The Robotics Inst., Carnegie Mellon Univ. 1999.

[2] S. Baker and T. Kanade, "Limits on Super-Resolution and How to Break Them," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 372-379, 2000.

[3] D.F. Barbe, *Charge-Coupled Devices.* Springer-Verlag, 1980.

[4] B. Bascle, A. Blake, and A. Zisserman, "Motion Deblurring and Super-Resolution from an Image Sequence," *Proc. European Conf. Computer Vision,* pp. 573-581, 1996.

[5] S. Borman and R.L. Stevenson, "Spatial Resolution Enhancement of Low-Resolution Image Sequences: A Comprehensive Review with Directions for Future Research," technical report, Univ. of Notre Dame, 1998.

[6] N.K. Bose, H.C. Kim, and H.M. Valenzuela, "Recursive Implementation of Total Least Squares Algorithm for Image Reconstruction from Noisy, Undersampled Multiframes," *Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing,* vol. 5, pp. 269-272, 1993.

[7] P. Cheeseman, B. Kanefsky, R. Kraft, and J. Stutz, "Super-Resolved Surface Reconstruction from Multiple Images," Technical Report FIA-94-12, NASA, 1994.

[8] M. Elad and A. Feuer, "Restoration of Single Super-Resolution Image from Several Blurred, Noisy and Down-Sampled Measured Images," *IEEE Trans. Image Processing,* vol. 6, no. 12, pp. 1646-1658, 1997.

[9] K. Forbes and V.V. Anh, "Condition of System Matrices in Image Restoration," *J. Optical Soc. Am.,* vol. 11, no. 6, pp. 1727-1735, 1994.

[10] W.T. Freeman and E.C. Pasztor, "Learning Low-Level Vision," *Proc. Int'l Conf. Computer Vision,* pp. 1182-1189, 1999.

[11] G.H. Golub and C.F. Van Loan, *Matrix Computations,* third ed. Baltimore: The John Hopkins Univ. Press, 1996.

[12] R.C. Hardie, K.J. Barnard, and E.E. Amstrong, "Joint Map Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images," *IEEE Trans. Image Processing,* vol. 6, no. 12, pp. 1621-1633, 1997.

[13] N.J. Higham, "A Survey of Componentwise Perturbation Theory in Numerical Linear Algebra," *Math. Computation 1943-1993: Proc. Symposia in Applied Math.,* vol. 48, pp. 49-77. 1994.

[14] T.S. Huang and R. Tsai, "Multi-Frame Image Restoration and Registration," *Advances in Computer Vision and Image Processing,* vol. 1, pp. 317-339, 1984.

[15] M. Irani and S. Peleg, "Improving Resolution by Image Restoration," *Computer Vision, Graphics, and Image Processing,* vol. 53, pp. 231-239, 1991.

[16] Z. Lin and H.-Y. Shum, "On the Fundamental Limits of Reconstruction-Based Super-Resolution Algorithms," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 1171-1176, Dec. 2001.

[17] S. Mann and R.W. Picard, "Virtual Bellows: Constructing High Quality Stills from Video," *Proc. IEEE Conf. Image Processing,* pp. 363-367, Nov. 1994.

[18] F. Milinazzo, C. Zala, and I. Barrodale, "On The Rate of Growth of Condition Numbers for Convolution Matrices," *IEEE Trans. Acoustics, Speech, and Signal Processing,* vol. 35, no. 4, pp. 471-475, 1987.

[19] A.J. Patti, M.I. Sezan, and A.M. Tekalp, "Superresolution Video Reconstruction with Arbitrary Sampling Lattices and Nonzero Aperture Time," *IEEE Trans. Image Processing,* vol. 6, no. 8, pp. 1064-1076, 1997.

[20] D. Rajan and S. Chauhuri, "Simultaneous Estimation of Super-Resolved Intensity and Depth Maps from Low Resolution Defocused Observations of a Scene," *Proc. Int'l Conf. Computer Vision,* pp. 113-118, July 2001.

[21] R. Schultz and R. Stevenson, "Extraction of High-Resolution Frames from Video Sequences," *IEEE Trans. Image Processing,* vol. 5, no. 6, pp. 996-1011, 1996.

[22] S. Xu, *The Theory and Methods of Matrix Computation* (in Chinese). Beijing, China: Peking Univ. Press, 1995.

Zhouchen Lin received the Bachelor of Science degree in pure mathematics from Nankai University in 1989, the Master of Science degree in applied mathematics from Peking University in 1996, the Master of Philosophy degree in applied mathematics from Hong Kong Polytechnic University in 1998, and the PhD degree in applied mathematics from Peking University in 2000, respectively. He is currently a researcher in the Multi-Modal User Interface Group, Microsoft Research Asia. His research interests include: computer vision, computer graphics, pattern recognition, statistical learning, document processing, and human computer interaction. He is a member of the IEEE.

Heung-Yeung Shum received the PhD degree in robotics from the School of Computer Science, Carnegie Mellon University in 1996. He worked as a researcher for three years in the Vision Technology Group at Microsoft Research, Redmond, Washington. In 1999, he moved to Microsoft Research Asia, where he is currently a senior researcher and the assistant managing director. His research interests include computer vision, computer graphics, human computer interaction, pattern recognition, statistical learning, and robotics. He is on the editorial boards for the *IEEE Transactions on Circuit System Video Technology* (*CSVT*), *IEEE Transactions on Pattern Analysis and Machine Intelligence* (*PAMI*), and *Graphical Models*. He is a senior member of the IEEE and is the general cochair of the 10th International Conference on Computer Vision (ICCV 2005 Beijing).