1

Image super-resolution: Historical overview and future challenges

Jianchao Yang

University of Illinois at Urbana-Champaign

Thomas Huang

University of Illinois at Urbana-Champaign

CONTENTS

1.1 Introduction to super-resolution

In most digital imaging applications, high resolution images or videos are usually desired for later image processing and analysis. The desire for high image resolution stems from two principal application areas: improvement of pictorial information for human interpretation; and helping representation for automatic machine perception. Image resolution describes the details contained in an image, the higher the resolution, the more image details. The resolution of a digital image can be classified in many different ways: pixel resolution, spatial resolution, spectral resolution, temporal resolution, and radiometric resolution. In this context, we are mainly interested in spatial resolution.

Spatial resolution: *a digital image is made up of small picture elements called pixels. Spatial resolution refers to the pixel density in an image and measures in pixels per unit area. Fig. 1.1 shows a classic test target to determine the spatial resolution of an imaging system.*

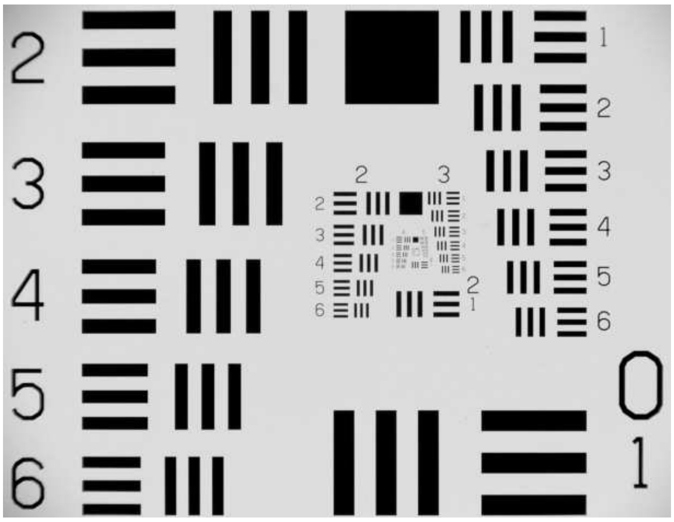

FIGURE 1.1
The 1951 USAF resolution test target, a classic test target used to determine spatial resolution of imaging sensors and imaging systems.

The image spatial resolution is firstly limited by the imaging sensors or the imaging acquisition device. Modern image sensor is typically a charge-coupled deveice (CCD) or a complementary metal-oxide-semiconductor (CMOS) active-pixel sensor. These sensors are typically arranged in a two-dimensional array to capture two-dimensional image signals. The sensor size or equivalently the number of sensor elements per unit area in the first place determines the spatial resolution of the image to capture. The higher density of the sensors, the higher spatial resolution possible of the imaging system. An imaging system with inadequate detectors will generate low resolution images with blocky effects, due to the aliasing from low spatial sampling frequency. In order to increase the spatial resolution of an imaging system, one straight forward way is to increase the sensor density by reducing the sensor size. However, as the sensor size decreases, the amount of light incident on each sensor also decreases, causing the so called shot noise. Also, the hardware cost of sensor increases with the increase of sensor density or correspondingly image pixel density. Therefore, the hardware limitation on the size of the sensor restricts the spatial resolution of an image that can be captured.

While the image sensors limit the spatial resolution of the image, the image details (high frequency bands) are also limited by the optics, due to lens blurs (associated with the sensor point spread function (PSF)), lens aberration effects, aperture diffractions and optical blurring due to motion. Constructing imaging chips and optical components to capture very high-resolution images is prohibitively expensive and not practical in most real applications, e.g., widely used surveillance cameras and cell phone built-in cameras. Besides the cost, the resolution of a surveillance camera is also limited in the camera speed and hardware storage. In some other scenarios such as satellite imagery, it is difficult to use high resolution sensors due to physical constraints. Another way to address this problem is to accept the image degradations and use signal processing to post process the captured images, to trade off computational cost with the hardware cost. These techniques are specifically referred as Super-Resolution (SR) reconstruction.

Super-resolution (SR) are techniques that construct high-resolution (HR) images from several observed low-resolution (LR) images, thereby increasing the high frequency components and removing the degradations caused by the imaging process of the low resolution camera. The basic idea behind SR is to combine the non-redundant information contained in multiple low-resolution frames to generate a high-resolution image. A closely related technique with SR is the single image interpolation approach, which can be also used to increase the image size. However, since there is no additional information provided, the quality of the single image interpolation is very much limited due to the ill-posed nature of the problem, and the lost frequency components cannot be recovered. In the SR setting, however, multiple low-resolution observations are available for reconstruction, making the problem better constrained. The non-redundant information contained in the these LR images is typically introduced by subpixel shifts between them. These subpixel shifts may occur due to uncontrolled motions between the imaging system and scene, e.g., movements of objects, or due to controlled motions, e.g., the satellite imaging system orbits the earth with predefined speed and path.

Each low-resolution frame is a decimated, aliased observation of the true scene. SR is possible only if there exists subpixel motions between these low resolution frames[1], and thus the ill-posed upsampling problem can be better conditioned. Fig. 1.2 shows a simplified diagram describing the basic idea of SR reconstruction. In the imaging process, the camera captures several LR frames, which are downsampled from the HR scene with subpixel shifts between each other. SR construction reverses this process by aligning the LR observations to subpixel accuracy and combining them into a HR image grid

[1]The mainstream SR techniques rely on motions, although there are some works using defocus as a cue.

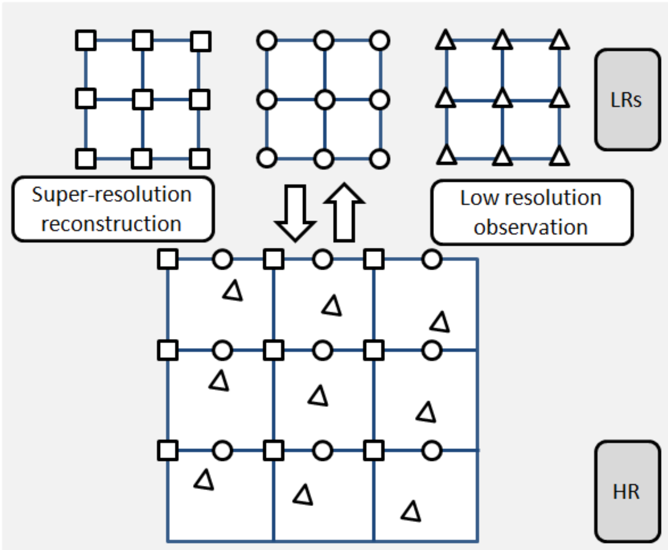

FIGURE 1.2
The basic idea for super-resolution reconstruction from multiple low-resolution frames. Subpixel motion provides the complementary information among the low-resolution frames that makes SR reconstruction possible.

(interpolation), thereby overcoming the imaging limitation of the camera. SR (some of which described in this book), arises in many areas such as:

1. Surveillance video [20, 55]: frame freeze and zoom region of interest (ROI) in video for human perception (e.g. look at the license plate in the video), resolution enhancement for automatic target recognition (e.g. try to recognize a criminal's face).

2. Remote sensing [29]: several images of the same area are provided, and an improved resolution image can be sought.

3. Medical imaging (CT, MRI, Ultrasound etc)[59, 70, 47, 60]: several images limited in resolution quality can be acquired, and SR technique can be applied to enhance the resolution.

4. Video standard conversion, e.g. from NTSC video signal to HDTV signal.

This chapter targets at an introduction to the SR research area, by explaining some basic techniques of SR, an overview of the literature, and discussions about some challenging issues for future research.

1.2 Notations

Before talking about SR techniques, we introduce the notations we use in this chapter. Upper case bold letters $\boldsymbol{X}$ and $\boldsymbol{Y}$ denote the vector form in lexicographical order for HR and LR images respectively. Lower case bold letters $\boldsymbol{x}$ and $\boldsymbol{y}$ denote the vector form in lexicographical order for HR and LR image *patches* respectively. Underlined upper case bold letters are used to denote a vector concatenation of multiple vectors, e.g., $\underline{\boldsymbol{Y}}$ is a vector concatenation of $\boldsymbol{Y}_k$ $(k = 1, 2, ..., K)$. We use plain upper case symbols to denote matrices, and plain lower case symbols to denote scalars.

1.3 Techniques for super-resolution

SR reconstruction has been one of the most active research areas since the seminal work by Tsai and Huang [99] in 1984. Many techniques have been proposed over the last two decades [4, 65] representing approaches from frequency domain to spatial domain, and from signal processing perspective to machine learning perspective. Early works on super-resolution mainly followed the theory of [99] by exploring the shift and aliasing properties of the Fourier transform. However, these frequency domain approaches are very restricted in the image observation model they can handle, and real problems are much more complicated. Researchers nowadays most commonly address the problem mainly in the spatial domain, for its flexibility to model all kinds of image degradations. This section talks about these techniques, starting from the image observation model.

1.3.1 Image observation model

The digital imaging system is not perfect due to hardware limitations, acquiring images with various kinds of degradations. For example, the finite aperture size causes optical blur, modeled by Point Spread Function (PSF). The finite aperture time results in motion blur, which is very common in videos. The finite sensor size leads to sensor blur; the image pixel is generated by integration over the sensor area instead of impulse sampling. The limited sensor density leads to aliasing effects, limiting the spatial resolution of the achieved image. These degradations are modeled fully or partially in different SR techniques.

Fig. 1.3 shows a typical observation model relating the HR image with LR video frames, as introduced in the literature [65, 82]. The input of the imaging system is continuous natural scenes, well approximated as band-limited signals. These signals may be contaminated by atmospheric turbulence be-

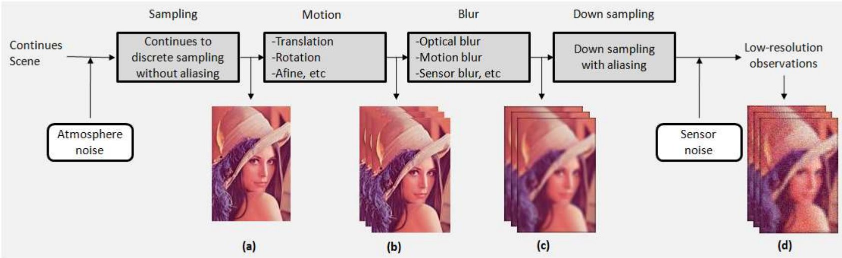

FIGURE 1.3
The observation model of a real imaging system relating a high resolution image to the low resolution observation frames with motion between the scene and the camera.

fore reaching the imaging system. Sampling the continues signal beyond the Nyquist rate generates the high resolution digital image (a) we desire. In our SR setting, usually there exists some kind of motion between the camera and scene to capture. The inputs to the camera are multiple frames of the scene, connected by possibly local or global shifts, leading to image (b). Going through the camera, these motion related high resolution frames will incur different kinds of blurring effects, such as optical blur and motion blur. These blurred images (c) are then downsampled at the image sensors (e.g. CCD detectors) into pixels, by an integral of the image falling into each sensor area. These downsampled images are further affected by the sensor noise and color filtering noise. Finally the frames captured by the low resolution imaging system are blurred, decimated, and noisy versions of the underlying true scene.

Let $\boldsymbol{X}$ denote the HR image desired, i.e., the digital image sampled above Nyquist sampling rate from the bandlimited continuous scene, and $\boldsymbol{Y}_k$ be the k-th LR observation from the camera. $\boldsymbol{X}$ and $\boldsymbol{Y}'_k s$ are represented in lexicographical order. Assume the camera captures K LR frames of $\boldsymbol{X}$, where the LR observations are related with the HR scene $\boldsymbol{X}$ by

$$\boldsymbol{Y}_k = D_k H_k F_k \boldsymbol{X} + V_k, k = 1, 2, ..., K, \tag{1.1}$$

where F_k encodes the motion information for the k-th frame, H_k models the blurring effects, D_k is the down-sampling operator, and V_k is the noise term. These linear equations can be rearranged into a large linear system

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_K \end{bmatrix} = \begin{bmatrix} D_1 H_1 F_1 \\ D_2 H_2 F_2 \\ \cdot \\ \cdot \\ D_K H_K F_K \end{bmatrix} \boldsymbol{X} + \underline{V} \tag{1.2}$$

or equivalently

$$\underline{\boldsymbol{Y}} = M\boldsymbol{X} + \underline{V} \tag{1.3}$$

The involved matrices D_k, H_k, F_k or M are very sparse, and this linear system is typically ill-posed. Furthermore, in real imaging systems, these matrices are unknown and need to be estimated from the available LR observations, leaving the problem even more ill-conditioned. Thus proper prior regularization for the high resolution image is always desirable and often even crucial. In the following, we will introduce some basic super-resolution techniques proposed in the literature and give an overview of the recent developments.

1.3.2 Super-resolution in the frequency domain

The pioneering work for super-resolution traces back to Tsai and Huang [99], in which the authors related the high resolution image with multiple shifted low-resolution images by a frequency domain formulation based on the shift and aliasing properties of the Continuous and Discrete Fourier Transforms. Let $\boldsymbol{x}(t_1, t_2)$ denote a continuous high resolution scene. The global translations yield K shifted images, $\boldsymbol{x}_k(t_1, t_2) = \boldsymbol{x}(t_1 + \triangle_{k_1}, t_2 + \triangle_{k_2})$, with $k = 1, 2, ..., K$, where $\triangle_{k_1}$ and $\triangle_{k_2}$ are arbitrary but known shifts. The continuous Fourier transform (CFT) of the scene is given by $\mathcal{X}(u_1, u_2)$ and those of the translated scenes by $\mathcal{X}_k(u_1, u_2)$. Then by the shifting properties of the CFT, the CFT of the shifted images can be written as

$$\mathcal{X}_k(u_1, u_2) = \exp\left[j2\pi(\triangle_{k_1} u_1 + \triangle_{k_2} u_2)\right] \mathcal{X}(u_1, u_2). \tag{1.4}$$

The shifted images are impulse sampled with the sampling period T_1 and T_2 to yield observed low resolution image $\boldsymbol{y}_k[n_1, n_2] = \boldsymbol{x}_k(n_1 T_1 + \triangle_{k_1}, n_2 T_2 + \triangle_{k_2})$ with $n_1 = 0, 1, 2, ..., N_1 - 1$ and $n_2 = 0, 1, 2, ..., N_2 - 1$. Denote the discrete Fourier transforms (DFTs) of these low resolution images by $\mathcal{Y}_k[r_1, r_2]$. The CFTs of the shifted images are related with their DFTs by the aliasing property:

$$\mathcal{Y}_k[r_1, r_2] = \frac{1}{T_1 T_2} \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \mathcal{X}_k\left(\frac{2\pi}{T_1}\left(\frac{r_1}{N_1} - m_1\right), \frac{2\pi}{T_2}\left(\frac{r_2}{N_2} - m_2\right)\right). \tag{1.5}$$

Assuming $\mathcal{X}(u_1, u_2)$ is band-limited, $|\mathcal{X}(u_1, u_2)| = 0$ for $|u_1| \geq (N_1\pi)/T_1$, $|u_2| \geq (N_2\pi)/T_2$, combining Eqn. 1.4 and Eqn. 1.5 we relate the DFT coefficients of $\mathcal{Y}_k[r_1, r_2]$ with the samples of the unknown CFT of $x(t_1, t_2)$ in matrix form as [2]

$$\underline{\mathcal{Y}} = \boldsymbol{\Phi}\underline{\mathcal{X}}, \tag{1.6}$$

where $\underline{\mathcal{Y}}$ is a $K \times 1$ column vector with the k^{th} element being the DFT coefficient $\mathcal{Y}_k[r_1, r_2]$, $\underline{\mathcal{X}}$ is a $N_1 N_2 \times 1$ column vector containing the samples of the

[2]Strictly, subscripts $\{r_1, r_2\}$ should be used in the following equation. We omit those for uncluttered presentation.

unknown CFT coefficients of $x(t_1, t_2)$, and $\mathbf{\Phi}$ is a $K \times N_1N_2$ matrix relating $\underline{\mathcal{Y}}$ and $\underline{\mathcal{X}}$. Eqn. 1.6 defines a set of linear equations from which we intend to solve $\underline{\mathcal{X}}$ and then use the inverse DFT to obtain the reconstructed image.

The above formulation for SR reconstruction assumes a noise-free and global translation model with known parameters. The downsampling process is assumed to be impulse sampling, with no sensor blurring effects modeled. Along this line of work, many extensions have been proposed to handle more complicated observation models. Kim et al. [49] extended [99] by taking into account the observation noise as well as spatial blurring. Their later work in [5] extend the work further by introducing Tikohonov regularization [95]. In [89], a local motion model is considered by dividing the images into overlapping blocks and estimating motions for each local block individually. In [98], the restoration and motion estimation are done simultaneously using an EM algorithm. However, the frequency domain SR theory of these works did not go beyond as what was initially proposed. These approaches are computationally efficient, but limited in their abilities to handle more complicated image degradation models and include various image priors as proper regularization. Later works on super-resolution reconstruction have been almost exclusively in the spatial domain.

1.3.3 Interpolation-restoration: non-iterative approaches

Many spatial domain approaches [4, 82, 65, 2] have been proposed over the years to overcome the difficulties of the frequency domain methods. As the HR image and the LR frames are related in a sparse linear system 1.3, similar to the traditional single image restoration problem [26], many flexible estimators can be applied to the SR reconstruction. These include Maximum likelihood (ML), Maximum *a Posteriori* (MAP)[84, 35], and Projection Onto Convex Sets (POCS)[88]. In this section, we start with the simplest and a non-iterative forward model for SR reconstruction in the spatial domain, in analogy to the frequency domain approach.

Assume H_k is Linearly Spatial Invariant (LSI) and is the same for all K frames, and we denote it as H. Suppose F_k considers only simple motion models such as translation and rotation, then H and F_k commute [27, 30] and we get

$$\boldsymbol{Y}_k = D_k F_k H \boldsymbol{X} + V_k = D_k F_k \boldsymbol{Z}, k = 1, 2, ..., K, \tag{1.7}$$

which motivates a forward non-iterative approach based on interpolation and restoration. There are three stages for this approach 1) low resolution image registration; 2) nonuniform interpolation to get $\boldsymbol{Z}$ and 3) deblurring and noise removal to get $\boldsymbol{X}$. Fig. 1.4 shows the procedure of such an approach. The low resolution frames are first aligned by some image registration algorithm [77] to subpixel accuracy. These aligned low resolution frames are then put on a high resolution image grid, where nonuniform interpolation methods are used to fill in those missing pixels on the HR image grid to get $\boldsymbol{Z}$. At last, $\boldsymbol{Z}$ is deblurred by any classical deconvolutional algorithm with noise removal to achieve $\boldsymbol{X}$. Keren *et. al* [48] proposed an early two step approach to SR reconstruction based on a global translation and rotation model. Gross *et. al*[101] proposed a nonuniform interpolation of a set of spatially shifted low resolution images by utilizing the generalized multi-channel sampling theorem by Yen [109] and later Papulis [64], followed by deblurring. Nguyen and Milanfar [62] proposed an efficient wavelet-based interpolation SR reconstruction algorithm by exploiting the interlacing sampling structure in the low resolution data. Alam [1] presented an efficient interpolation scheme based on weighted nearest neighbors, followed by Wiener filtering for deblurring. Focusing on the special case of SR reconstruction where the observation is composed of pure translation, space invariant blur, and additive Gaussian noise, Elad and Hel-Or [27] presented a very computationally efficient algorithm. [52] proposed a triangulation-based method for interpolating irregularly sampled data. The triangulation method, however, is not robust to noise commonly present in real applications. Based on the normalized convolution [50], Pham *et al.* [71] proposed a robust certainty and a structure-adaptive applicability function to the polynomial facet model and applied it to fusion of irregularly sampled data. Recently, Takeda *et. al* [91] proposed an adaptive steering kernel regression for interpolation on the high resolution image grid where the low resolution images are registered and mapped on.

These interpolation-restoration forward approaches are intuitive, simple and computationally efficient [30], [18], assuming simple observation models. However, the step-by-step forward approach does not guarantee optimality of the estimation. The registration error can easily propagate to the later processing. Also, the interpolation step is suboptimal without considering the noise and blurring effects. Moreover, without HR image prior as proper regularization, the interpolation based approaches need special treatment of limited observations in order to reduce aliasing.

1.3.4 Statistical approaches

Unlike the interpolation-restoration approaches, statistical approaches relate the SR reconstruction steps stochastically toward optimal reconstruction. The HR image and motions among low-resolution inputs can be both regarded as stochastic variables. Let $M(\nu, h)$ denote the degradation matrix defined by the motion vector ν and blurring kernel h, the SR reconstruction can be cast

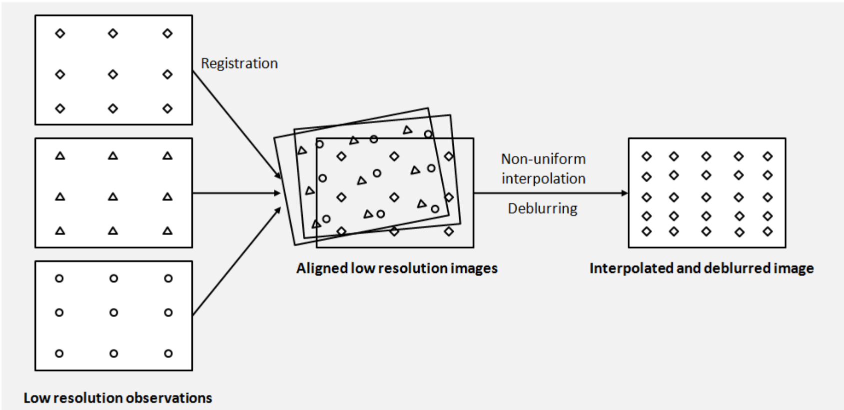

FIGURE 1.4
The interpolation SR approach based on alignment and post processing of deblurring.

into a full Bayesian framework:

$$\begin{aligned} \boldsymbol{X} &= \arg\max_{\boldsymbol{X}} Pr(\boldsymbol{X}|\underline{\boldsymbol{Y}}) \\ &= \arg\max_{\boldsymbol{X}} \int_{\nu,h} Pr(\boldsymbol{X}, M(\nu,h)|\underline{\boldsymbol{Y}})d\nu \\ &= \arg\max_{\boldsymbol{X}} \int_{\nu,h} \frac{Pr(\underline{\boldsymbol{Y}}|\boldsymbol{X}, M(\nu,h))Pr(\boldsymbol{X}, M(\nu,h))}{Pr(\underline{\boldsymbol{Y}})} d\nu \\ &= \arg\max_{\boldsymbol{X}} \int_{\nu,h} Pr(\underline{\boldsymbol{Y}}|\boldsymbol{X}, M(\nu,h))Pr(\boldsymbol{X})Pr(M(\nu,h))d\nu. \end{aligned} \tag{1.8}$$

Note that $\boldsymbol{X}$ and $M(\nu,h)$ are statistically independent [35]. Here $Pr(\underline{\boldsymbol{Y}}|\boldsymbol{X}, M(\nu,h))$ is the data likelihood, $Pr(\boldsymbol{X})$ is the prior term on the desired high resolution image and $Pr(M(\nu,h))$ is a prior term on the motion estimation. $\underline{V}$ in Eqn. 1.3 usually stands for additive noise, assumed to be a zero-mean and white Gaussian random vector. Therefore,

$$Pr(\underline{\boldsymbol{Y}}|\boldsymbol{X}, M(\nu,h)) \propto \exp\left\{-\frac{1}{2\sigma^2}\|\underline{\boldsymbol{Y}} - M(\nu,h)\boldsymbol{X}\|^2\right\}. \tag{1.9}$$

$Pr(\boldsymbol{X})$ is typically defined using the Gibbs distribution in an exponential form

$$Pr(\boldsymbol{X}) = \frac{1}{Z}\exp\{-\alpha A(\boldsymbol{X})\}, \tag{1.10}$$

where $A(\boldsymbol{X})$ is a non-negative potential function, and Z is just a normalization factor. The Bayesian formulation in Eq. 1.8 is complicated and difficult to evaluate due to the integration over motion estimates. If $M(\nu, h)$ is given or estimated beforehand (denote as M), Eqn. 1.8 can be simplified as

$$\begin{aligned} \boldsymbol{X} &= \arg\max_{\boldsymbol{X}} Pr(\underline{\boldsymbol{Y}}|\boldsymbol{X}, M)Pr(\boldsymbol{X}) \\ &= arg\min_{\boldsymbol{X}}\{\|\underline{\boldsymbol{Y}} - M\boldsymbol{X}\|^2 + \lambda A(\boldsymbol{X})\}, \end{aligned} \quad (1.11)$$

where λ absorbs the variance of the noise and α in Eqn. 1.10, balancing the data consistence and the HR image prior strength. Eqn. 1.11 is the popular *Maximum a Posteriori* (MAP) formulation for SR, where M is assumed to be known. The statistical approaches discussed below vary in the ways of treating degradation matrix $M(\nu, h)$, prior term $Pr(\boldsymbol{X})$ and statistical inference methods toward Eqn. 1.8.

1.3.4.1 Maximum Likelihood

If we assume uniform prior over $\boldsymbol{X}$, Eqn. 1.11 reduces to the simplest maximum likelihood (ML) estimator (motion estimation is assumed as a prior). The ML estimator relies on the observations only, seeking the most likely solution for the observations to take place by maximizing $p(\underline{\boldsymbol{Y}}|\boldsymbol{X})$, giving

$$\hat{\boldsymbol{X}}_{ML} = \arg\min_{\boldsymbol{X}} \|\underline{\boldsymbol{Y}} - M\boldsymbol{X}\|^2. \quad (1.12)$$

Differentiating Eqn. 1.12 with respect to $\boldsymbol{X}$ and setting the derivative to be zero gives the classical pseudo inverse result

$$\hat{\boldsymbol{X}}_{ML} = (M^T M)^{-1} M^T \underline{\boldsymbol{Y}}. \quad (1.13)$$

If $M^T M$ is singular, the problem is ill-posed and there are infinite many possible solutions due to the null space of M. This naturally leads to the term of regularization for the sake of unique solution from purely the algebraic point of view, which although can be interpreted in the MAP framework. For computation, direct inverse of matrix as $M^T M$ is usually prohibitive in practice due to the high dimensionality problem. For example, if the low resolution images are of size 100×100 and are to be zoomed to a single high resolution frame $\boldsymbol{X}$ of 300×300, M is of the size 90000×90000, requiring inverse of a matrix of size 90000×90000. Therefore, many iterative methods for practical ways to solve this large set of sparse linear equations have been suggested in the literature [111].

Irani and Peleg proposed a simple but very popular method, based on error back-projection scheme inspired by computer-aided tomography, in [39, 40, 41]. The algorithm iteratively updates the current estimation by adding back the warped simulation error convolved with a back-projection function (BPF):

$$\boldsymbol{X}^{i+1} = \boldsymbol{X}^i + c\sum_k F_k^{-1}[h_{bpf} * S\uparrow(\widehat{\boldsymbol{Y}}_k - \boldsymbol{Y}_k)], \quad (1.14)$$

where c is a constant, h_{bpf} is the back-projection kernel, $S \uparrow$ is the upsampling operator, and $\widehat{\boldsymbol{Y}}_k$ is the simulated k-th LR frame from the current HR estimation. In [41], the authors applied this idea to real applications by incorporating a multiple motion tracking algorithm to deal with partially occluded objects, transparent objects and some objects of interest. The back-projection algorithm is simple and flexible in handing many observations with different degradation procedures. However, the solution of back-projection is not unique, depending on the initialization and the choice of back-projection kernel. As shown in [26] and [10], the back-projection algorithm is none other than an ML estimator. The choice of BPF implies some underlying assumption about the noise covariance of the observed low-resolution pixels [10]. Treating the motion estimates $M(\nu)$ as unknown, Tom *et. al* [98] proposed an ML SR image estimation algorithm to estimate the subpixel shifts, noise of the image, and the HR image simultaneously. The proposed ML estimation is treated by the Expectation-Maximization (EM) algorithm.

As in the image denoising and single image expansion case, direct ML estimator without regularization in SR where number of observations is limited can be severely ill-posed, especially when the zoom factor is large (e.g. greater than 2). The ML estimator is usually very sensitive to noise, registration estimation errors and PSF estimation errors [10], and therefore proper regularization on the feasible solution space is always desirable. This leads to the mainstream SR reconstruction approaches based on MAP.

1.3.4.2 Maximum a Posteriori

Many works [46, 84, 15] in SR reconstruction have followed the MAP approach in Eqn. 1.11, where the techniques vary in the observation model assumptions and the prior term $Pr(\boldsymbol{X})$ for the desired solution. Different kinds of priors for natural images have been proposed in the literature, but none of them stands out as the lead. In the following, we list three commonly used image priors for the SR reconstruction techniques.

1. **Gaussian MRF**. The Gaussian Markov Random Field (GMRF) [37, 33] takes the form
$$A(\boldsymbol{X}) = X^T Q X, \tag{1.15}$$
where Q is a symmetric positive matrix, capturing spatial relations between adjacent pixels in the image by its off-diagonal elements. Q is often defined as $\Gamma^T\Gamma$, where Γ acts as some first or second derivative operator on the image $\boldsymbol{X}$. In such a case, the log likelihood of the prior is
$$\log p(\boldsymbol{X}) \propto \|\Gamma \boldsymbol{X}\|^2, \tag{1.16}$$
which is well known as the *Tikhonov regularization* [95, 26, 63], the most commonly used method for regularization of ill-posed problems. Γ is usually referred as Tikhonov matrix. Hardie *et al.* [35] proposed a joint MAP framework for simultaneous estimation of the high resolution

image and motion parameters with Gaussian MRF prior for the HR image. Bishop *et. al* [96] proposed a simple Gaussian process prior where the covariance matrix Q is constructed by spatial correlations of the image pixels. The nice analytical property of Gaussian process prior allows a Bayesian treatment of the SR reconstruction problem, where the unknown high resolution image is integrated out for robust estimation of the observation model parameters (unknown PSFs and registration parameters). Although the GMRF prior has many analytical advantages, a common criticism for it associated with super-resolution reconstruction is that the results tend to be overly smooth, penalizing sharp edges that we desire to recover.

2. **Huber MRF**.The problem with GMRF can be ameliorated by modeling the image gradients with a distribution with heavier tails than Gaussian, leading to the popular Huber MRF (HMRF) where the Gibbs potentials are determined by the Huber function,

$$\rho(a) = \begin{cases} a^2 & |a| \leq \alpha \\ 2\alpha|a| - \alpha^2 & \textit{otherwise}, \end{cases} \tag{1.17}$$

where a is the first derivative of the image. Such a prior encourages piecewise smoothness, and can preserve edges well. Schultz and Stevenson [83] applied this Huber MRF to single image expansion problem, and later to the SR reconstruction problem in [84]. Many later works on super-resolution employed the Huber MRF as the regularization prior, such as [11, 12, 15, 13, 73, 74] and [3].

3. **Total Variation**. The Total variation (TV) norm as a gradient penalty function is very popular in the image denoising and deblurring literature [81, 54, 16]. The TV criterion penalizes the total amount of change in the image as measured by the ℓ_1 norm of the magnitude of the gradient

$$A(\boldsymbol{X}) = \|\nabla \boldsymbol{X}\|_1 \tag{1.18}$$

where ∇ is a gradient operator that can be approximated by Laplacian operators [81]. The ℓ_1 norm in the TV criterion favors sparse gradients, preserving steep local gradients while encouraging local smoothness[13]. Farsiu *et al.* [30] generalized the notation of TV and proposed the so called bilateral TV (BTV) for robust regularization.

For more comparisons of these generic image priors on effecting the solution of super-resolution, one can further refer to [10] and [25].

1.3.4.3 Joint MAP restoration

Multiple frame SR reconstruction can be divided into two subproblems: LR registration and HR estimation. Many previous algorithms treat these two processes as two distinct processes: first do registration and then estimation by

MAP, which is suboptimal as registration and estimation are interdependent. Motion estimation and HR estimation can benefit each other if interactions between them are allowed. In joint MAP restoration, Eqn. 1.11 is extended to include the motion and PSF estimates as unknowns for inference:

$$\begin{aligned}\{\boldsymbol{X}, \nu, h\} &= \arg\max_{\boldsymbol{X},\nu,h} Pr(\underline{\boldsymbol{Y}}|\boldsymbol{X}, M(\nu, h))Pr(\boldsymbol{X})Pr(M(\nu, h)) \\ &= \arg\min_{\boldsymbol{X},\nu,h} -\log\left[Pr(\underline{\boldsymbol{Y}}|\boldsymbol{X}, M(\nu, h))\right] - \log\left[Pr(\boldsymbol{X})\right] \\ &\quad - \log\left[Pr(M(\nu, h))\right].\end{aligned} \tag{1.19}$$

Tom *et al.* [98] divided the SR problem into three subproblems, namely registration, restoration and interpolation. Instead of solving them independently they simultaneously estimated registration and restoration by maximizing likelihood using Expectation-Maximization (EM). Later they included interpolation into the framework and estimated all of the unknowns using EM in [97]. [35] applied the MAP framework for simultaneous estimation of the high-resolution image and translation motion parameters (PSF is taken as a known prior). The high resolution image and motion parameters are estimated using a cyclic coordinate-descent optimization procedure. The algorithm converges slowly but improves the estimation a lot. Segall *et al.* [86, 85] presented an approach of joint estimation of dense motion vectors and HR images applied to compressed video. Woods *et al.*[105] treated the noise variance, regularization and registration parameters all as unknowns and estimated them jointly in a Bayesian framework based on the available observations. Chung *et al.* [19] proposed a joint optimization framework and showed superior performance to the coordinate descent approach [46]. The motion model they handled is affine transformations. To handle more complex multiple moving objects problem in SR setting, Shen *et al.* [87] addressed the problem by MAP formulation combining motion estimation, segmentation and SR reconstruction together. The optimization is done in a cyclic coordinate descent process similar to [46].

1.3.4.4 Bayesian Treatments

Due to limited low-resolution observations, the SR reconstruction problem is ill-posed in nature. Joint MAP estimation of motion parameters, PSF, and HR image may face the problem of overfitting [96]. While motion and blurring is difficult to model in general, simple models spanned by few parameters are sufficient for SR applications in many scenarios. Given the low resolution observations, however, estimating these parameters by integrating over the unknown high resolution image is a useful approach. Bishop [96] proposed such a Bayesian approach for SR where the unknown high-resolution image is integrated out and the marginal is used to estimate the PSF and motion parameters. To make the problem analytically tractable, a Gaussian process prior (GMRF) is used to model the high resolution image. Even though an unfavorable GMRF is used for the high resolution image, the PSF and motion parameters can still be estimated quite accurately. Then the estimated parameters are fixed, and a MAP estimation of the HR image is performed. An in-depth analysis similar to this for blind deconvolution is discussed in [53]. Such a Bayesian approach outperforms the joint MAP approaches in Subsection 1.3.4.3, which will easily get overfitting with the PSF parameters. However, the integration over the high resolution image is computationally heavy and the Gaussian prior over the image leads the final reconstruction toward excess smoothness. Instead of marginalization over the unknown high resolution image, Pickup *et al.* proposed in their recent works [73, 74, 72] to integrate over the unknown PSF and motion parameters as in Eqn. 1.8, which is motivated to overcome the uncertainty of the registration parameters [79]. The registration parameters are estimated beforehand and then treated as Gaussian variables with the pre-estimated values as the means to model their uncertainty. The HR image estimation can be combined with any favorable image prior for MAP estimation after integrating the observation model parameters. Sharper results can be obtained with such an approach compared with [96] as reported in [73, 74, 72].

Such Bayesian treatments by marginalizing the unknowns demonstrate promising power for SR recovery. However, in order for integration to be tractable, image priors or registration parameters have to take simple parametric forms, limiting these models in dealing with more complex cases that may happen in real videos. Computation could also be a concern for such algorithms in realistic applications.

1.3.5 Example-based approach

Previous super-resolution approaches rely on aggregating multiple frames that contain complementary spatial information. Generic image priors are usually deployed to regularize the solution properly. The regularization becomes especially crucial when insufficient number of measurements is supplied, as in the extreme case, only one single low-resolution frame is observed. In such cases, generic image priors do not suffice as an effective regularization for SR [2]. A recently emerging methodology for regularizing the ill-posed super-resolution reconstruction is to use examples, in order to break the super-resolution limit caused by inadequate measurements. Different from previous approaches where the prior is in a parametric form regularizing on the whole image, the example-based methods develop the prior by sampling from other images, similar to [24],[38] in a local way.

One family of example-based approaches is to use the examples directly, with the representative work proposed by Freeman *et al.* [31]. Such approaches usually work by maintaining two sets of training patches, $\{\boldsymbol{x}_i\}_{i=1}^n$ sampled from the high resolution images, and $\{\boldsymbol{y}_i\}_{i=1}^n$ sampled from the low resolution images correspondingly. Each patch pair $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ is connected by the observation model $\boldsymbol{y}_i = DH\boldsymbol{x}_i + \boldsymbol{v}$. This high- and low-resolution co-occurrence model is then applied to the target image for predicting the high resolution image in a patch-based fashion, with a Markov Random Field (MRF) model as shown in Fig. 1.5. The observation model parameters have to be known as a prior, and the training sets are tightly coupled with the image targeted. Patch size should also be chosen properly. If the patch size is very small, the co-occurrence prior is too weak to make the prediction meaningful. On the other hand, if the patch size is too large, one may need a huge training set to find proximity patches for the current observations.

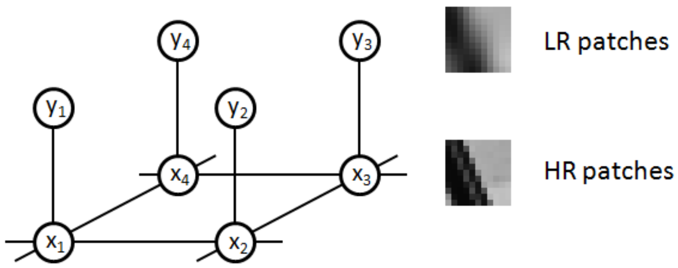

FIGURE 1.5
The MRF model for single frame super-resolution.

A naive way to do super-resolution with such a coupled training sets is, for each low resolution patch in the low-resolution image, find its nearest neighbor $\tilde{\boldsymbol{y}}$ in $\{\boldsymbol{y}_i\}_{i=1}^n$, and then put the corresponding $\tilde{\boldsymbol{x}}$ from $\{\boldsymbol{x}_i\}_{i=1}^n$ to the high resolution image grid. Unfortunately, this simple approach will produce disturbing artifacts due to noise and the ill-posed nature of super-resolution [25]. Relaxing the nearest neighbor search to k-nearest neighbors can ensure that the proximity patch we desire will be included. Freeman *et al.* [31] proposed a belief propagation [108] algorithm based on the above MRF model to select the best high resolution patch found by k-nearest neighbors that has best compatibility with adjacent patches. Sun *et al.*[90] extended this idea using the sketch prior to enhance only the edges in the image, aiming to speed up the algorithm. The IBP [39] algorithm is then applied as a post processing step to ensure data consistence on the whole image. Wang *et al.* [103] further followed this line of work and proposed a statistical model that can handle unknown PSF.

The above methods are based on image patches directly, requiring large training sets to include any patterns possibly encountered in testing. Chang *et al.* [17] proposed another simple but effective method based on neighbor embedding [93], with the assumption of correspondence between the two manifolds formed by the low- and high-resolution image patches. For each low resolution image patch $\boldsymbol{y}_k^t$ from the test image (superscript "t" distinguishes the test patch from the training patches), the algorithm finds its k nearest neighbors $\mathcal{N}_t$ from $\{\boldsymbol{y}_i\}_{i=1}^n$, and computes the reconstruction weights by neighbor embedding

$$\begin{aligned} \hat{w}_s = & \arg\min_{w_s} \|\boldsymbol{y}_k^t - \sum_{\boldsymbol{y}_s \in \mathcal{N}_t} w_s \boldsymbol{y}_s\|^2, \\ s.t. & \sum_{\boldsymbol{y}_s \in \mathcal{N}_t} w_s = 1. \end{aligned} \tag{1.20}$$

The reconstruction weights are then applied to generate the corresponding high resolution patch $\hat{\boldsymbol{x}}_k^t = \sum_{\boldsymbol{y}_s \in \mathcal{N}_t} \hat{w}_s \boldsymbol{x}_s$. To handle the compatibility problem between adjacent patches, simple averaging in the overlapping regions is performed. The algorithm works nicely even with smaller patch databases than [108].However, fixing k for each low resolution patch may result in overfitting or underfitting. Yang *et al.* [107] proposed another patch-based single frame super-resolution method. The method is derived from the compressive sensing theory, which ensures that linear relationships among high-resolution signals can be precisely recovered from their low-dimensional projections [9],[22]. The algorithm models the training sets as two dictionaries: $D_h = [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n]$ and $D_l = [\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_n]$. Given a test low resolution image patch $\boldsymbol{y}_k^t$, the approach basically seeks the supports by an ℓ_1 minimization [23]

$$\begin{aligned} \hat{w} = & \arg\min_{w} \|w\|_1 \\ s.t. & \quad \|\boldsymbol{y}_t - D_l w\|^2 \leq \sigma^2, \end{aligned} \tag{1.21}$$

which can be rewritten with Lagrange multiplier as an unconstrained optimization problem known as Lasso in the statistics literature [94]. The corresponding high resolution patch is recovered by $\boldsymbol{x}_k^t = D_h \hat{w}$. Compared to the neighbor embedding method with fixed k neighbors, Yang's method adaptively chooses the fewest necessary supports for reconstruction, avoiding overfitting. Moreover, the ℓ_1 minimization formulation is more robust to noise than the previous mentioned patch-based methods. In a later version [42], this approach is further extended by learning a coupled dictionary instead of using the raw patches, allowing the algorithm to be much more efficient.

One criticism with the aforementioned methods with direct examples is that operating on local patches cannot guarantee global optimality of the estimation. Another kind of example-based approach seeks to perform MAP estimation with local priors on the image space sampled from examples. The pioneering work by Baker and Kanade [2] formulated an explicit regularization which demands proximity between the spatial derivatives of the unknown image to those of the found examples. The examples are formed by a pyramid derivative set of features, instead of raw data directly. Similar method is applied to text super-resolution in [75]. Datsenko and Elad [21] presented a global MAP estimation where the example-based regularization is given by a binary weighted average instead of the nearest neighbor, bypassing outliers due to noise. This work is further extended and elaborated in [21], where the binary weighting scheme is relaxed. Another noteworthy approach for example-based approach is by Protter *et al.* [78], generalized from the nonlocal means denoising algorithm [8]. Instead of sampling examples from other training images, the algorithm explores self-similarities within the image (or sequence) and extract the example patches from the target image (or sequence) itself. A recent work by Glasner *et al.* further explored self-similarities in images for SR by combining the classical algorithm based on subpixel displacements and the example-based method based on patch pairs extracted from the target image.

The use of examples can be much more effective when dealing with narrow families of images, such as text and face images. A group of algorithms have emerged targeting face super-resolution in recent years due to its importance in surveillance scenarios. Face super-resolution is usually referred as face hallucination, following the early work by Baker and Kanade [2]. Capel *et al.* [14] proposed an algorithm where PCA [45] subspace models are used to learn parts of the faces. Liu *et al.* [58], [57] proposed a two-step approach toward super-resolution of faces, where the first step uses the eigenface [100] to generate a medium resolution face, followed by the nonparametric patch-based approach [31] in the second step. Such an Eigenface-based approach has been explored in several later works [32],[104] too. Yang *et al.* [106] proposed a similar two-step approach. Instead of using the holistic PCA subspace, [106] uses local Nonnegative Matrix Factorization (NMF)[51] to model faces and the patch-based model in the second step is adopted from [107]. Jia and Gong [43], [44] proposed the tensor approach to deal with more face variations, such as illuminations and expressions. Although these face hallucination algorithms work surprisingly well, they only apply to frontal faces, and only few works have been devoted on evaluating face hallucination for recognition [32],[36].

Example-based regularization is effective in our SR problem when insufficient observations are available. There are still a number of questions we need to answer regarding this kind of approaches. First, how to choose the optimal patch size given the target image. Perhaps a multi-resolution treatment is needed. Second, how to choose the database. Different images have different statistics, and thereby need different databases. An efficient method for dictionary adaptation to the current target image may suggest a way out. Third, how to use the example-based prior more efficiently. The computation issue could be a difficulty for practical applications. Readers are suggested to refer to [25] for more detailed analysis on example-based regularization for inverse problems.

1.3.6 Set theoretic restoration

Besides the optimization approaches derived from stochastic view as discussed above, another stream of methods is through the well known Projection onto Convex Sets (POCS) [110]. The POCS methods approach the SR problem by formulating multiple constraining convex sets containing the desired image as a point within the sets. Defining such convex sets is flexible and can incorporate different kinds of constraints or priors, even nonlinear and nonparametric constraints. As an example, we introduce several commonly used convex sets in the POCS methods. The data consistency or reconstruction constraints can be modeled as K convex sets:

$$\mathcal{C}_k = \left\{ \boldsymbol{X} \middle| \| D_k H_k F_k \boldsymbol{X} - \boldsymbol{Y}_k \|^2 \leq \sigma^2, 1 \leq k \leq K \right\}. \tag{1.22}$$

Smoothness constraints can be defined as

$$\mathcal{C}_\Gamma = \left\{ \boldsymbol{X} \middle| \| \Gamma \boldsymbol{X} \|_p < \sigma_2 \right\}. \tag{1.23}$$

where $p = 1, 2, \infty$ denotes different norms. Amplitude constraints can also be modeled:

$$\mathcal{C}_A = \left\{ \boldsymbol{X} \middle| A_1 \leq \boldsymbol{X}[m, n] \leq A_2 \right\}. \tag{1.24}$$

With a group of $\mathcal{M}$ convex sets, the desired solution lies in the intersection of these sets $\boldsymbol{X} \in \mathcal{C}_s = \bigcap_{i=1}^{\mathcal{M}} \mathcal{C}_i$. The POCS technique suggests the following recursive algorithm for finding a point within the intersection set given an initial guess:

$$\boldsymbol{X}_{k+1} = P_{\mathcal{M}} P_{\mathcal{M}-1} \cdots P_2 P_1 \boldsymbol{X}_k, \tag{1.25}$$

where $\boldsymbol{X}_0$ is an initial guess, and P_i is the projection operator which projects a point onto a closed, convex sets $\mathcal{C}_i$.

Early POCS techniques for SR reconstruction were proposed by Stark *et al.* [88]. Extensions were proposed to handle space-varying PSF, motion blur, sensor blur and aliasing sampling effects in [68], [67], [69]. Many super-resolution works only consider non-zero aperture size (the lens blur, PSF), but not finite aperture time (motion blur) which is quite common in real low resolution videos. [69] is the early work to take into account the motion blur in SR reconstruction of videos based on POCS technique. As the motion blurring caused by a finite aperture time will in general be space- and perhaps time-varying, it cannot not be factored out of the SR restoration problem and performed as a separate post processing step. The POCS technique can conveniently handle such problems. Extending this method, Eren *et al.* [28] proposed a POCS-based approach for robust, object-based SR reconstruction. The proposed method employs a validity map to disable projections based on observations with inaccurate motion estimation, and a segmentation map for object-based processing. Elad *et al.* [26] analyzed and compared the ML, MAP and POCS methods for super-resolution, and proposed a hybrid approach. Patti [66] extended their earlier work in the image observation model to allow high order interpolation and modified constraint sets to reduce the edge ringing artifacts.

The advantage of the POCS technique lies in its simplicity to incorporate any kinds of constraints and priors that may present as impossible for those stochastic approaches. However, POCS is notorious for its heavy computation and slow convergence. The solution is not unique, depending on the initial guess. The POCS methods also assume prioris on the motion parameters and system blurs. They cannot estimate those registration parameters and high resolution image as in the stochastic approaches simultaneously. The hybrid approach combining stochastic view and POCS philosophy suggests a promising way to pursue.

1.4 Challenge issues for super-resolution

In the previous sections, we have discussed several basic techniques for SR reconstruction. Although many different approaches have been proposed since the SR concept was introduced, most approaches work well on toy data rather than in real problems. In building a practical SR system, many challenge issues still lay ahead preventing the SR techniques from wide applications. In the following, we list several challenge issues that we think are important for the future development and application of SR techniques.

1.4.1 Image Registration

Image registration is critical for the success of multiframe SR reconstruction, where complementary spatial samplings of the HR image are fused. The image registration is a fundamental image processing problem that is well known as ill-posed. The problem is even more difficult in the SR setting, where the observations are low-resolution images with heavy aliasing artifacts. The performance of the standard image registration algorithms decreases as the resolution of the observations goes down, resulting in more registration errors. Artifacts caused by these registration errors are visually more annoying than the blurring effect resulting from interpolation of a single image. Traditional SR reconstruction usually treat image registration as a distinct process from the HR image estimation. Therefore, the recovered HR image quality depends largely on the image registration accuracy from the previous step. Many image registration techniques derived from different principles have been proposed in the literature [7, 114]. However, Robinson *et al.* [79] showed that the registration performance is bounded even for the simplest case of global translation.

LR image registration and the HR image estimation are actually interdependent [80]. On one hand, accurate subpixel motion estimation benefits HR image estimation. On the other hand, high quality HR image can facilitate accurate motion estimation. Therefore, tailored to the SR reconstruction problem, the LR image registration can be addressed together with the HR image reconstruction, leading to joint ML [97] or MAP [35, 87, 76] framework for simultaneous estimation. These joint estimation algorithms capture the dependence between LR image registration and HR image estimation, and performance improvements are observed. However, with limited observations, the joint estimation for registration parameters and HR image may result in overfitting. To overcome this overfitting problem, Tipping and Bishop [96] employed a Bayesian approach for estimating both registration and blur parameters by marginalizing the unknown high resolution image. The algorithms shows noteworthy estimation accuracy both for registration and blur parameters, however the computation cost is very high. Pickup *et al.* [73, 74, 72] instead cast the Bayesian approach in another way by marginalizing the unknown registration parameters, to address the uncertainty inherent with the image registration [79].

The stochastic approaches associating the HR image estimation toward image registration do demonstrate promising results, however such parametric methods are limited in the motion models they can effectively handle. Usually, some simple global motion models are assumed. Real videos are complicated comprising arbitrary local motions, where parametrization of the motion models maybe intractable. Optical flow motion estimation can be applied to such scenarios. However, the insufficient measurements for local motion estimations make these algorithm vulnerable to errors, which may cause disasters for SR reconstruction [112]. Another promising approach toward SR reconstruction is the nonparametric methods, which try to bypass the explicit motion estimation. Protter *et al* [78] extended the non-local means denoising algorithm to SR reconstruction, where fuzzy motion estimation based on block matching is used. Later they proposed a probabilistic motion model in [77], which is a nonparametric model analogy to [72]. Both [78] and [77] can handle complex motion patterns in real videos. Compared to the classical SR methods based on optical flow motion estimation, Protter's methods reduce the errors caused by misalignment by a weighting strategy over multiple possible candidates. Takeda *et al.* [92] on the other hand applied an 3-D steering kernel proposed in their early work [91] to video, which also avoids explicit motion estimation, for denoising and SR reconstruction. The 3-D steering kernel captures both spatial and temporal structure, encoding implicit motion information, and thus can be applied to both spatial and temporal SR for video with complex motion activities. While methods without explicit motion estimation indeed produce promising results toward the practical applicability of SR techniques, further improvements may include computation efficiency, combining adaptive interpolation or regression together with deblurring, and generalizing observation models to 3-D motions in video, e.g. out-of-plane rotation.

1.4.2 Computation Efficiency

Another difficulty limiting practical application of SR reconstruction is its intensive computation due to large number of unknowns, which require expensive matrix manipulations. Real applications always demand efficiency of the SR reconstruction to be of practical utility, e.g., in the surveillance video scenarios, it is desired for the SR reconstruction to be real time. Efficiency is also desirable for SR systems with users in the loop for tuning parameters. Many SR algorithms targeting efficiency fall into the previously discussed interpolation-restoration approach, such as [27], [1], [61] and [34]. In [34], Hardie showed the computation superiority of his algorithm over previous efficient algorithms proposed in [1] and [61], and claimed that the algorithm can be applied in real time with global translation model. However, the computation goes up significant when non-translation model occurs, which can be ameliorated by massive parallel computing. Others tried to examine particular modeling scenarios to speed up the optimization problem. Zomet [115] and and Farsiu [30] studied the application of D_k, H_k and F_k directly as the corresponding image operations of downsampling, blurring and shifting, bypassing the need to explicitly construct the matrices, bringing significant speed ups. [6] combined a slightly modified version of [27] and [30] and implemented a real-time SR system using FPGA, a nice attempt to practical use of SR.

However, such algorithms require precise image registration, which is computation intensive in the first place. Moreover, these algorithm can only handle simple motion models efficiently up to now, far from application in real complex video scenarios. For videos with arbitrary motions, [92] suggests promising directions for seeking efficient algorithms. It is also interesting to see how parallel computing, e.g., GPU, and hardware implementations affect the future applications of SR techniques.

1.4.3 Robustness Aspects

Traditional SR techniques are vulnerable to the presence of outliers due to motion errors, inaccurate blur models, noise, moving objects, motion blur etc. These inaccurate model errors cannot be treated as Gaussian noise as the usual assumption with ℓ_2 reconstruction residue. Robustness of SR is of interest because the image degradation model parameters cannot be estimated perfectly, and sensitivity to outliers may result in visually disturbing artifacts, which are intolerable in many applications, e.g., video standard conversion. However, not enough work has been devoted to such an important aspect. Chiang and Boulte [18] used median estimation to combine the upsampled images to cope with outliers from non-stationary noise. Zomet *et al.* [116] cast the problem in a different way, where a robust median-based gradient is used for the optimization to bypass the influence of outliers. Farsiu *et al.* [30],[82] changed the commonly used ℓ_2 norm into ℓ_1 norm for robust estimation similar to [18] and robust regularization. [113] introduced a simultaneous super-resolution with Huber norm as the prior for robust regularization. Pham *et al.* [71] proposed a robust certainty to each neighboring sample for interpolating unknown data, with the same photometric-based weighting scheme used in bilateral filtering. Similar uncertainty scheme is also used in the probabilistic motion model [77] for taking care of optical flow motion estimation errors based on block matching. Many of these algorithms showed improvements for outliers assumed on the toy data, where more experimental evaluations are needed to see how much the robustness efforts can benefit real SR performance.

1.4.4 Performance Limits

The SR reconstruction has become a hot research topic since it was introduced, and thousands of SR papers have bloomed into publications. However, not much work has been devoted to fundamental understanding of the performance limits of these SR reconstruction algorithms. Such a performance limit understanding is important. For example, it will shed light on SR camera design, helping to analyze factors such as model errors, zooming factors and number of frames etc. In general, an ambitious analysis of the performance limits for all SR techniques could be intractable. First, SR reconstruction is a complex task which consists of many interdependent components. Second, it is still unknown what is the most informative prior given the SR task, especially for the example-based approaches. Last, a good measure instead of simple MSE is still needed for performance evaluation. It has been recognized that an estimation with higher MSE does not have to be visually more appealing. For example, bicubic interpolation usually achieves smaller MSE compared with those recovered by some example-based approaches [107].

Several works attempting at the performance understanding have been proposed over the last several years. [2] analyzed the numerical conditions of the SR linear systems, and concluded that as the zoom factor increases the general image prior is of less and less help for SR. [56] derived the performance limits based on matrix perturbation, but with assumption that image registration is known as a prior. With simple translation model, Robinson and Milanfar in [79], use the Cramér-Rao (CR) bounds to analyze the registration performance limit. They extend this work in [80] to give a thorough analysis of SR performance with factors such as motion estimation, decimation factor, number of frames and prior information. The analysis is based on the MSE criterion and the motion model is again assumed to be simple global translational. Eekeren *et al.* [102] evaluated several SR algorithms on real-world data exploring several influential factors empirically. Even though these efforts at understanding performance bounds are far from enough about SR, they indeed suggest ways for people to follow.

While it is hard to draw consistent conclusions for different SR techniques, in terms of performance evaluation, some benchmark and realistic datasets are needed for fair comparison and algorithm understanding. Future research should pursue more theoretical analysis and performance evaluation for directing SR technique developments.

Bibliography

[1] M. S. Alam, J. G. Bognar, R. C. Hardie, and B. J. Yasuda. Infrared image registration and high-resolution reconstruction using multiple trans-

lationally shifted aliased video frames. *IEEE Transactions on Instrumentation and Measurement*, 49(5):915–923, 2000.

[2] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1167–1183, 2002.

[3] S. Borman and R. L. Stevenson. Simultaneous multi-frame MAP super-resolution video enhancement using spatio-temporal priors. In *Proceedings of IEEE International Conference on Image Processing*, volume 3, pages 469–473, 1999.

[4] Sean Borman and Robert L. Stevenson. Super-resolution from image sequences - A review. In *Proceedings of the 1998 Midwest Symposium on Circuits and Systems*, pages 374–378, 1998.

[5] N. K. Bose, H. C. Kim, and H. M. Valenzuela. Recursive implementation of total least squares algorithm for image reconstruction from noisy, undersampled multiframes. In *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*, volume 5, pages 269–272, 1993.

[6] O. Bowen and C. S. Bouganis. Real-time image super resolution using an FPGA. In *International Conference on Field Programmable Logic and Applications*, pages 89–94, 2008.

[7] L. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.

[8] A. Buades, B. Coll, and J. M. More. A non-local algorithm for image denoising. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 60–65, 2005.

[9] E. Candes. Compressive sensing. In *Proceedings of International Congress of Mathematicians*, volume 3, pages 1433–1452, 2006.

[10] D. Capel. *Image Mosaicing and Super-resolution.* Springer, 2004.

[11] D. Capel and A. Zisserman. Automated mosaicing with super-resolution zoom. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 885–891, 1998.

[12] D. Capel and A. Zisserman. Super-resolution enhancement of text image sequences. In *Proceedings of the International Conference on Pattern Recognition*, volume 1, pages 1600–1605, 2000.

[13] D. Capel and A. Zisserman. Super-resolution enhancement of text image sequences. In *Proceedings of 15th International Conference on Pattern Recognition*, pages 600–605, 2000.

[14] D. Capel and A. Zisserman. Super-resolution from multiple views using learnt image models. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 627–634, 2001.

[15] D. Capel and A. Zisserman. Computer vision applied to super-resolution. *IEEE Signal Processing Magazine*, 20(3):75–86, 2003.

[16] T. F. Chan, S. Osher, and J. Shen. The digital TV filter and nonlinear denosing. *IEEE Transaction on Image Processing*, 10(2):231–241, 2001.

[17] H. Chang, D. Y. Yeung, and Y. Xiong. Super-resolution through neighbor embedding. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 275–282, 2004.

[18] M. C. Chiang and T. E. Boulte. Efficient super-resolution via image warping. *Image and Vision Computing*, 18(10):761–771, 2000.

[19] J. Chung, E. Haber, and J. Nagy. Numerical methods for coupled super-resolution. *Inverse Problems*, 22(4):1261–1272, 2006.

[20] Marco Crisani, Dong Seon Cheng, Vittorio Murino, and Donato Pannullo. Distilling information with super-resolution for video surveillance. In *Proceedings of the ACM 2nd International Workshop on Video Surveillance and Sensor Networks*, pages 2–11, 2004.

[21] D. Datsenko and M. Elad. Example-based single document image super-resolution: a global MAP approach with outlier rejection. *Multidimensional System and Signal Processing*, 18(2-3):103–121, 2007.

[22] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[23] D. L. Donoho. For most large underdetermined systems of linear equations, the minimal ℓ^1-norm near-solution approximates the sparsest near-solution. *Communications on Pure and Applied Mathematics*, 59(7):907–934, 2006.

[24] A. A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. In *Proceedings of IEEE International Conference on Computer Vision*, pages 1033–1038, 1999.

[25] M. Elad and D. Datsenko. Example-based regularization deployed to super-resolution reconstruction of a single image. *The Computer Journal*, 52(1):15–30, 2007.

[26] M. Elad and A. Feuer. Restoration of single super-resolution image from several blurred, noisy and down-sampled measured images. *IEEE Transaction on Image Processing*, 6(12):1646–1658, 1997.

[27] M. Elad and Y. Hel-Or. A fast super-resolution reconstruction algorithm for pure translational motion and common space invariant blur. *IEEE Transactions on Image Processing*, 10(8):1187–1193, 2001.

[28] P. E. Eren, M. I. Sezan, and A. M. Tekalp. Robust, object-based high resolution image reconstruction from low-resolution video. *IEEE Transactions on Image Processing*, 6(10):1446–1451, 1997.

[29] X. Jia F. Li and D. Fraser. Universal HMT based super resolution for remote sensing images. In *IEEE International Conference on Image Processing*, pages 333–336, 2008.

[30] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar. Fast and robust multi-frame super-resolution. *IEEE Transaction on Image Processing*, 13(10):1327–1344, 2004.

[31] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *IEEE Computer Graphics and Applications*, 22(2):56–65, 2002.

[32] B. K. Gunturk, A. U. Batur, Y. Altunbasak, M. H. Hayes, and R. M. Mersereau. Eigenface-domain super-resolution for face recognition. *IEEE Transactions on Image Processing*, 12(5):597–606, 2003.

[33] K. M. Hanson and G. W. Wecksung. Bayesian approach to limited-angle reconstruction in computed tomography. *Journal of Optical Society of America*, 73(11):1501–1509, 1983.

[34] R. Hardie. A fast image super-resolution algorithm using an adaptive Wiener filter. *IEEE Transactions on Image Processing*, 16(12):2953–2964, 2007.

[35] R. C. Hardie, K. J. Barnard, and E. E. Armstrong. Join MAP registration and high resolution image estimation using a sequence of undersampled images. *IEEE Transactions on Image Processing*, 6(12):1621–1633, 1997.

[36] P. H. Hennings-Yeomans, S. Baker, and B. V. K. V. Kumar. Simultaneous super-resolution and feature extraction for recognition of low-resolution faces. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[37] G. T. Herman, H. Hurwitz, A. Lent, and H-P. Lung. On the Bayesian approach to image reconstruction. *Information and Control*, 42(1):60–71, 1979.

[38] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin. Image analogies. In *Proceedings of the 28th annual conference Computer Graphics and interactive techniques*, pages 327–340, 2001.

[39] M. Irani and S. Peleg. Super resolution from image sequences. In *Proceedings of 10th International Conference on Pattern Recognition*, volume 2, pages 115–120, 1990.

[40] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP: Graphical Models and Imaging Processing*, 53(3):231–239, 1991.

[41] M. Irani and S. Peleg. Motion analysis for image enhancement: resolution, occlusion and tranparency. *Journal of Visual Communications and Image Representation*, 4(4):324–335, 1993.

[42] T. S. Huang J. Yang, J. Wright and Y. Ma. Super-resolution via sparse representation. *Submitted to IEEE Transactions on Image Processing*, 2009.

[43] K. Jia and S. Gong. Multi-model tensor face for simultaneous super-resolution and recognition. In *Proceedings of IEEE International Conference on Computer Vision*, volume 2, pages 1683–1690, 2005.

[44] K. Jia and S. Gong. Generalized face super-resolution. *IEEE Transactions on Image Processing*, 17(6):873–886, 2008.

[45] I. T. Jolliffe. *Principal Component Analysis.* Series: Springer Series in Statistics, 2002.

[46] E. Kaltenbacher and R. C. Hardie. High-resolution infrared image reconstruction using multiple low resolution aliased frames. In *Proceedings of the IEEE National Aerospace Electronics Conference*, volume 2, pages 702–709, 1996.

[47] J. A. Kennedy, O. Israel, A. Frenkel, R. Bar-Shalom, and A. Haim. Super-resolution in PET imaging. *IEEE Transactions on Medical Imaging*, 25(2):137–147, 2006.

[48] D. Keren, S. Peleg, and R. Brada. Image sequence enhancement using subpixel displacements. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 742–746, 1988.

[49] S. P. Kim, N. K. Bose, and H. M. Valenzuela. Recursive reconstruction of high resolution image from noisy undersampled multiframes. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(6):1013–1027, 1990.

[50] H. Knutsson and C.-F. Westin. Normalized and differential convolution. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 515–523, 1993.

[51] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorizaiton. *Nature*, 401(6755):788–791, 1999.

[52] S. Lerttrattanapanich and N. K. Bost. High resolution image formation from low resolution frames using delaunay triangulation. *IEEE Transaction on Image Processing*, 11(12):1427–1441, 2002.

[53] A. Levin, Y. Weiss, F. Durand, and W. Freeman. Understanding and evaluating blind deconvolution algorithms. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1964–1971, 2009.

[54] Y. Li and F. Santosa. A computational algorithm for minimizing total variation in image restoration. *IEEE Transactions on Image Processing*, 5(6):987–995, 1996.

[55] Frank Lin, Clinton B. Fookes, Vinod Chandran, and Sridha Sridharan. Investigation into optical flow super-resolution for surveillance applications. In *The Austrilian Pattern Recognition Society Worshop on Digital Image Computing*, 2005.

[56] Z. Lin and H.-Y. Shum. Fundamental limits on reconstruction-based superresolution algorithms under local translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):83–97, 2004.

[57] C. Liu, H. Y. Shum, and W. T. Freeman. Face hallucination: theory and practice. *International Journal of Computer Vision*, 75(1):115–134, 2007.

[58] C. Liu, H. Y. Shum, and C. S. Zhang. Two-step approach to hallucinating faces: global parametric model and local nonparametric model. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 192–198, 2001.

[59] J. Maintz and M. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[60] K. Malczewski and R. Stasinski. Toeplitz-based iterative image fusion scheme for MRI. In *IEEE International Conference on Image Processing*, pages 341–344, 2008.

[61] B. Narayanan, R. C. Hardie, K. E. Barner, and M. Shao. A computationally efficient super-resolution algorithm for video processing using partition filters. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(5):621–634, 2007.

[62] N. Nguyen and P. Milanfar. An efficient wavelet-based algorithm for image super-resolution. In *Proceedings of International Conference on Image Processing*, volume 2, pages 351–354, 2000.

[63] N. Nguyen, P. Milanfar, and G. H. Golub. A computationally efficient image superresolution algorithm. *IEEE Transactions on Image Processing*, 10(5):573–583, 2001.

[64] A. Papulis. Generalized sampling expansion. *IEEE Transactions on Circuits and Systems*, 24(11):652–654, 1977.

[65] Sung C. Park, Min K. Park, and Moon G. Kang. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20(3):21–36, 2003.

[66] A. J. Patti and Y. Altunbasak. Aritifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. *IEEE Transaction on Image Processing*, 10(1):179–186, 2001.

[67] A. J. Patti, M. Sezan, and A. M. Tekalp. Robust methods for high quality stills from interlaced video in the presence of dominant motion. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(2):328–342, 1997.

[68] A. J. Patti, M. I. Sezan, and A. M. Tekalp. High-resolution image reconstruction from a low-resolution image sequence in the presence of time-varing motion blur. In *Proceedings of the IEEE International Conference on Image Processing*, volume 1, pages 343–347, 1994.

[69] A. J. Patti, M. I. Sezan, and A. M. Tekalp. Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. *IEEE Transactions on Image Processing*, 6(8):1064–1076, 1997.

[70] S. Peleg and Y. Yeshurun. Superresolution in MRI: application to human white matter fiber tract visualization by diffusion tensor imaging. *Magnazine Resonance in Medicine*, 45(1):29–35, 2001.

[71] T. Q. Pham, L. J. Vliet, and K. Schutte. Robust fusion of irregularly sampled data using adaptive normalized convolution. *EURASIP Journal on Applied Signal Processing*, 2006.

[72] L. C. Pickup, D. P. Capel, S. J. Robert, and A. Zisserman. Overcoming registration uncertainty in image super-resolution: maximize or marginalize? *EURASIP Journal on Advances in Signal Processing*, 2007.

[73] L. C. Pickup, D. P. Capel, S. J. Roberts, and A. Zisserman. Bayesian image super-resolution, continued. In *Proceedings of Advances in Neural Information and Proceedings Systems*, pages 1089–1096, 2006.

[74] L. C. Pickup, D. P. Capel, S. J. Roberts, and A. Zisserman. Bayesian methods for image super-resolution. *The Computer Journal*, 52(1):101–113, 2009.

[75] L. C. Pickup, S. J. Robert, and A. Zisserman. A sampled texture prior for image super-resolution. In *Proceedings of Advances in Neural Information and Processing System*, pages 1587–1594, 2003.

[76] L. C. Pickup, S. J. Roberts, and A. Zisserman. Optimizing and learning for super-resolution. In *British Machine Vision Conference*, volume 2, pages 439–448, 2006.

[77] M. Protter and M. Elad. Super resolution with probabilistic motion estimation. *IEEE Transactions on Image Processing*, 18(8):1899–1904, 2009.

[78] M. Protter, M. Elad, H. Takeda, and P. Milanfar. Generalizing the nonlocal-means to super-resolution reconstruction. *IEEE Transactions on Image Processing*, 18(1):36–51, 2009.

[79] D. Robinson and P. Milanfar. Fundamental performance limits in image registration. *IEEE Transactions on Image Processing*, 13(9):1185–1199, 2004.

[80] D. Robinson and P. Milanfar. Statistical performance analysis of super-resolution. *IEEE Transactions on Image Processing*, 15(6):1413–1428, 2006.

[81] L. Rudin, S. Osher, and E. Fatemi. Nonlinaer total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, 1992.

[82] M. Elad S. Farsiu, D. Robinson and P. Milanfar. Advances and challenges in super-resolution. *International Journal of Imaing Systems and Technology*, 14(2):47–57, 2004.

[83] R. R. Schultz and R. L. Stevenson. A Bayesian approach to image expansion for improved definition. *IEEE Transactions on Image Processing*, 3(3):233–242, 1994.

[84] R. R. Schultz and R. L. Stevenson. Extraction of high-resolution frames from video sequences. *IEEE Transactions on Image Processing*, 5(6):996–1011, 1996.

[85] C. A. Segall, A. K. Katsaggelos, R. Molina, and J. Mateos. Bayesian resolution enhancement of compressed video. *IEEE Transactions on Image Processing*, 13(7):898–910, 2004.

[86] C. A. Segall, R. Molina, and A. K. Katsaggelos. High resolution images from low-resolution compressed video. *IEEE Signal Processing Magazine*, 20(3):37–38, 2003.

[87] H. Shen, L. Zhang, B. Huang, and P. Li. A MAP approach for joint motion estimation, segmentation and super-resolution. *IEEE Transactions on Image Processing*, 16(2):479–490, 2007.

[88] H. Stark and P. Oskoui. High-resolution image recovery from image-plane arrays, using convex projections. *Journal of Optical Society of America A*, 6(11):1715–1726, 1989.

[89] W. Su and S. P. Kim. High-resolution restoration of dynamic image sequences. *International Journal of Imaging Systems and Technology*, 5(4):330–339, 1994.

[90] J. Sun, N. N. Zheng, H. Tao, and H. Shum. Image hallucination with primal sketch priors. In *Proceedings of IEEE Computer Society Conferece on Computer Vision and Pattern Recognition*, volume 2, pages 729–736, 2003.

[91] H. Takeda, S. Farsiu, and P. Milanfar. Kernel regression for image processing and reconstruction. *IEEE Transactions on Image Processing*, 16(2):349–366, 2007.

[92] H. Takeda, P. Milanfar, M. Protter, and M. Elad. Super-resolution without explicit subpixel motion estimation. *IEEE Transaction on Image Processing*, 18(9):1958–1975, 2009.

[93] J. B. Tenenbaum, V. Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[94] R. Tibshirani. Regression shrinkge and selection via the Lasso. *Journal of Royal Statistical Society: Series B (Statistical Methodology)*, 59(1):267–288, 1996.

[95] A. N. Tikhonov and V. A. Arsenin. *Solution of ill-posed problems*. Winston & Sons, Washington, 1997.

[96] Michael E. Tipping and Christopher M. Bishop. Bayesian image super-resolution. In *Proceedings of Advances in Neural Information Proceeding Systems*, pages 1279–1286, 2003.

[97] B. C. Tom and A. K. Katsaggelos. Reconstuction of a high-resolution image by simultaneous registration, restoration and interpolation of low-resolution images. In *Proceedings of the IEEE International Conference on Image Processing*, volume 2, page 2539, 1995.

[98] B. C. Tom, A. K. Katsaggelos, and N. P. Galatsanos. Reconstruction of a high resolution image from registration and restoration of low resolution images. In *Proceedings of IEEE International Conference on Image Processing*, pages 553–557, 1994.

[99] R. Y. Tsai and T. S. Huang. Multipleframe image restoration and registration. In *Advances in Computer Vision and Image Processing*, pages 317–339. Greenwich, CT: JAI Press Inc., 1984.

[100] M. Turk and A. Pentland. Face recognition using eigenfaces. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 586–591, 1991.

[101] H. Ur and D. Gross. Improved resolution from subpixel shifted pictures. *CVGIP: Graphical Models and Image Processing*, 54(2):181–186, 1992.

[102] A. W. M. van Eekeren, K. Schutte, O. R. Oudegeest, and L. J. van Vilet. Perforamnce evaluation of super-resolution reconstruction methods on real-world data. *EURASIP Journal on Advances in Signal Processing*, 2007.

[103] Q. Wang, X. Tang, and H. Shum. Patch based blind image super-resolution. In *Proceedings of IEEE International Conferece on Computer Vision*, volume 1, pages 709–716, 2005.

[104] X. Wang and X. Tang. Hallucinating face by eigentransformation. *IEEE Transactions on Systems, Man, and Cybernetics*, 35(3):425–434, 2003.

[105] N. A. Woods, N. P. Galatsanos, and A. K. Katsaggelos. Stochastic methods for joint registration, restoration and interpolation of multiple undersampled images. *IEEE Transactions on Image Processing*, 15(1):210–213, 2006.

[106] Jianchao Yang, Hao Tang, Yi Ma, and Thomas Huang. Face hallucination via sparse coding. In *Proceedings of IEEE International Conference on Image Processing*, pages 1264–1267, 2008.

[107] Jianchao Yang, John Wright, Thomas Huang, and Yi Ma. Image super-resolution as sparse representation of raw image patches. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[108] J. S. Yedidia, W. T. Freeman, and Y. Weiss. Generalized belief propagation. In *Proceedings of Advances in Neural Information Processing Systems*, pages 689–695, 2001.

[109] L. J. Yen. On non-uniform sampling of bandwidth limited signals. *IRE Transactions on Circuits Theory*, 3(4):251–257, 1956.

[110] D. C. Youla and H. Webb. Image registration by the method of convex projections: Part 1-thoery. *IEEE Transactions on Medical Imaging*, 1(2):81–94, 1982.

[111] D. M. Young. *Iterative solution of large linear systems.* New York: Academic, 1971.

[112] W. Zhao and H. S. Sawhney. Is super-resolution with optical flow feasible? In *Proceedings of the 7th European Conference on Computer Vision*, pages 599–613, 2002.

[113] M. V. W. Zibetti and J. Mayer. Outlier robust and edge-preserving simultaneous super-resolution. In *Proceedings of IEEE International Conference on Image Processing*, pages 1741–1744, 2006.

[114] B. Zitová and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.

[115] A. Zomet and S. Peleg. Efficient super-resolution and applications to mosaics. In *Proceedings of International Conference on Pattern Recognition*, pages 579–583, 2000.

[116] A. Zomet, A. Rav-acha, and S. Peleg. Robust super-resolution. In *Proceedings of the IEEE Workshop on Applications of Computer Vision*, pages 645–650, 2001.