SIViP (2011) 5:329–342
DOI 10.1007/s11760-010-0204-6

ORIGINAL PAPER

A survey on super-resolution imaging

Jing Tian · Kai-Kuang Ma

Received: 13 July 2010 / Revised: 14 December 2010 / Accepted: 15 December 2010 / Published online: 3 February 2011

Abstract The key objective of *super-resolution* (SR) imaging is to reconstruct a higher-resolution image based on a set of images, acquired from the same scene and denoted as 'low-resolution' images, to overcome the limitation and/or ill-posed conditions of the image acquisition process for facilitating better content visualization and scene recognition. In this paper, we provide a comprehensive review of SR image and video reconstruction methods developed in the literature and highlight the future research challenges. The SR image approaches reconstruct a single higher-resolution image from a set of given lower-resolution images, and the SR video approaches reconstruct an image sequence with a higher-resolution from a group of adjacent lower-resolution image frames. Furthermore, several SR applications are discussed to contribute some insightful comments on future SR research directions. Specifically, the SR computations for multi-view images and the SR video computation in the temporal domain are discussed.

Keywords Super-resolution imaging · Regularization · Resolution enhancement

1 Introduction

Super-resolution (SR) imaging [1–20] aims to overcome or compensate the limitation or shortcomings of the image acquisition device/system and/or possibly ill-posed acquisition conditions to produce a higher-resolution image based on a set of images that were acquired from the same scene (Fig. 1). With rapid development and deployment of image processing for visual communications and scene understanding, there is a strong demand for providing the viewer with high-resolution imaging not only for providing better visualization (fidelity issue) but also for extracting additional information details (recognition issue). For examples, a high-resolution image is beneficial to achieve a better classification of regions in a multi-spectral remote sensing image or to assist radiologist for making diagnosis based on a medical imagery. In video surveillance systems, higher-resolution video frames are always welcomed for more accurately identifying the objects and persons of interest.

The most direct approach on obtaining higher-resolution images is to improve the image acquisition device (e.g., digital camera) by reducing the pixel size on the sensor (e.g., charge-coupled device). However, there is a limitation in reducing the sensor's pixel size in the sensor technology. When the sensor's pixel size becomes too small, the captured image quality will be inevitably degraded [5]. This is due to the fact that the noise power remains roughly the same, while the signal power decreases proportional to the sensor's pixel size reduction. Furthermore, higher cost is required to increase the chip size. Owing to the above-mentioned, the SR image processing becomes a promising alternative. The SR imaging research has grown very rapidly, after it was first addressed by Tsai and Huang [21] in 1984. In view of this, the purpose of this paper is to provide a comprehensive and updated survey for the SR research literature and contribute several inspirations for future SR research.

To understand the SR imaging, several fundamental concepts are required to be clarified. First, it is important to note that an image's *resolution* is fundamentally different from its *physical size*. In our context, the objective of SR imaging is to produce an image with a clearer content from

J. Tian · K.-K. Ma (✉)
School of Electrical and Electronic Engineering,
Nanyang Technological University, Singapore 639798, Singapore
e-mail: ekkma@ntu.edu.sg

J. Tian
e-mail: pg05061499@ntu.edu.sg

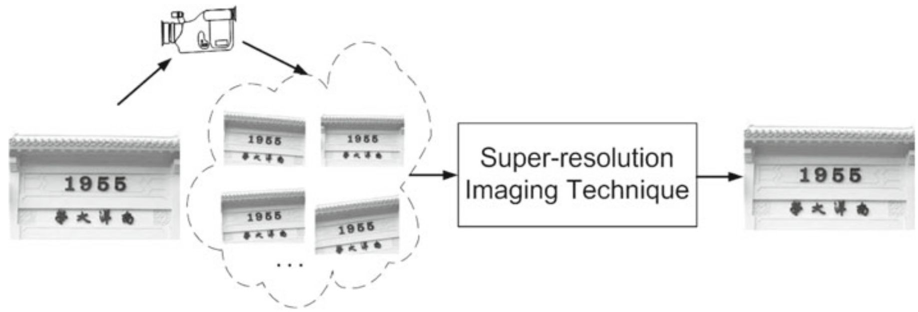

Fig. 1 A framework of super-resolution imaging

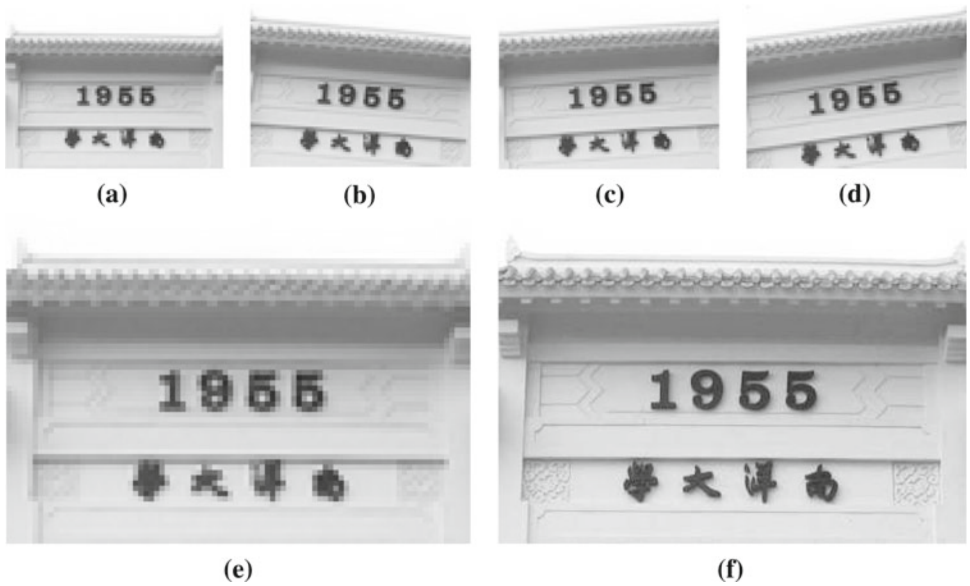

Fig. 2 **a**–**d** Four test images *NTU* (70 × 90, each); **e** An enlarged image (280 × 360) by applying pixel duplication method on the image (**a**); **f** An enlarged image (280 × 360) by applying a SR algorithm [50] based on the images (**a**)–(**d**)

its low-resolution counterpart (e.g., producing Fig. 2f from Fig. 2a–d), rather than simply achieving a larger size of image (e.g., Fig. 2e is produced by applying pixel duplication on Fig. 2a). In other words, the main goal and the first priority of super-resolution imaging is to 'fuse' the contents of multiple input images in order to produce one output image containing with more clear and detailed contents. The physical size of the output image (in terms of total number of pixels) could be the same as any one of the input images or subject to further enlargement using an image interpolation method. Second, in our context, the term *resolution* of super-resolution is referred to the *spatial* resolution of the image, not the *temporal* resolution of the image sequence. The latter is commonly expressed in terms of the number of frames captured per second (i.e., *frame rate*). Third, it is worthwhile to note that the term *super-resolution* has been used in other research areas as well. For example, in the field of optics, super-resolution refers to a set of restoration procedures that seek to recover the information beyond the diffraction limit [22]. In another example on the scanning antenna research, the super-resolution technique is exploited to resolve two closely spaced targets when a one-dimensional stepped scanning antenna is used [23].

The limitation of SR computation mainly comes from the following factors [24–29]: interpolation error, quantization error, motion estimation error and optical-blur. Baker and Kanade [30] showed that, for a sufficiently large resolution enhancement factor, any smoothness prior image model will result in reconstructions with very little high-frequency content. Lin et al. studied a numerical perturbation model of reconstruction-based SR algorithms for the case of translational motion [31] and the learning-based SR algorithms [32]. Robinson and Milanfar [33] analyzed this issue using *Cramer-Rao* bounds. A thorough study of SR performance bounds would understand the fundamental limits of the SR imaging to find the balance between expensive optical imaging system hardwares and image reconstruction algorithms.

The paper is organized as follows. Section 2 presents a description of the general model of imaging systems (observation model) that provides the SR image and video computation formulations, respectively. Section 3 presents the SR image reconstruction approaches that reconstruct a single high-resolution image from a set of given low-resolution images acquired from the same scene. Section 4 presents the SR video reconstruction approaches that produce an image sequence with a higher-resolution from a group of adjacent low-resolution uncompressed or compressed image frames. Each section begins with some introductory remarks, followed by an extensive survey of existing approaches.

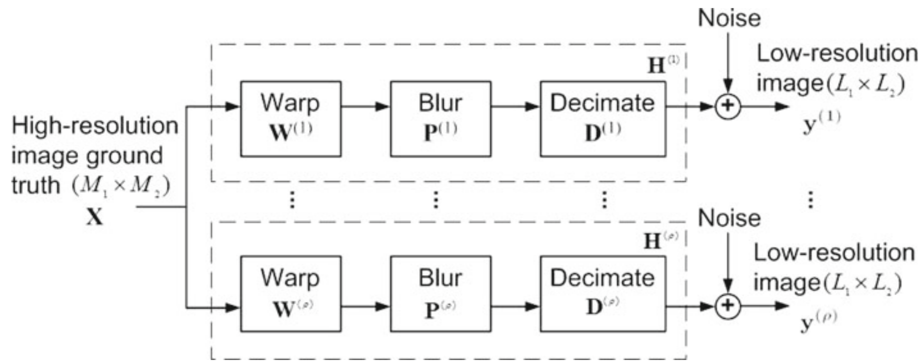

Fig. 3 The observation model, establishing the relationship between the original high-resolution image and the observed low-resolution images. The observed low-resolution images are the warped, blurred, down-sampled and noisy version of the original high-resolution image

Section 5 discusses several research challenges that remain open in this area for future investigation. More specifically, we discuss the SR computations for multi-view images and the SR video computation in the temporal domain. Finally, Sect. 6 concludes this paper.

2 Observation model

In this section, two observation models of the imaging system are presented to formulate the SR image reconstruction problem and the SR video reconstruction problem, respectively.

2.1 Observation model for super-resolution image

As depicted in Fig. 3, the image acquisition process is modeled by the following four operations: (i) geometric transformation, (ii) blurring, (iii) down-sampling by a factor of $q_1 \times q_2$, and (iv) adding with white Gaussian. Note that the geometric transformation includes translation, rotation, and scaling. Various blurs (such as motion blur and out-of-focus blur) are usually modeled by convolving the image with a low-pass filter, which is modeled by a *point spread function* (PSF). The given image (say, with a size of $M_1 \times M_2$) is considered as the high-resolution ground truth, which is to be compared with the high-resolution image reconstructed from a set of low-resolution images (say, with a size of $L_1 \times L_2$ each; that is, $L_1 = M_1/q_1$ and $L_2 = M_2/q_2$) for conducting performance evaluation. To summarize mathematically,

$$\mathbf{y}^{(k)} = \mathbf{D}^{(k)}\mathbf{P}^{(k)}\mathbf{W}^{(k)}\mathbf{X} + \mathbf{V}^{(k)}, \quad (1)$$

$$= \mathbf{H}^{(k)}\mathbf{X} + \mathbf{V}^{(k)}, \quad (2)$$

where $\mathbf{y}^{(k)}$ and $\mathbf{X}$ denote the kth $L_1 \times L_2$ low-resolution image and the original $M_1 \times M_2$ high-resolution image, respectively, and $k = 1, 2, \ldots, \rho$. Furthermore, both $\mathbf{y}^{(k)}$ and $\mathbf{X}$ are represented in the lexicographic-ordered vector form, with a size of $L_1L_2 \times 1$ and $M_1M_2 \times 1$, respectively, and each $L_1 \times L_2$ image can be transformed (i.e., lexicographic ordered) into a $L_1L_2 \times 1$ column vector, obtained by ordering the image row by row. $\mathbf{D}^{(k)}$ is the decimation matrix with a size of $L_1L_2 \times M_1M_2$, $\mathbf{P}^{(k)}$ is the blurring matrix of size $M_1M_2 \times M_1M_2$, and $\mathbf{W}^{(k)}$ is the warping matrix of size $M_1M_2 \times M_1M_2$. Consequently, three operations can be combined into one transform matrix $\mathbf{H}^{(k)} = \mathbf{D}^{(k)}\mathbf{P}^{(k)}\mathbf{W}^{(k)}$ with a size of $L_1L_2 \times M_1M_2$. Lastly, $\mathbf{V}^{(k)}$ is a $L_1L_2 \times 1$ vector, representing the white Gaussian noise encountered during the image acquisition process. Note that $\mathbf{V}^{(k)}$ is assumed to be independent with $\mathbf{X}$. Over a period of time, one can capture a set of (say, ρ) observations $\{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \ldots, \mathbf{y}^{(\rho)}\}$. With such establishment, the goal of the SR image reconstruction is to produce *one* high-resolution image $\mathbf{X}$ based on $\{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \ldots, \mathbf{y}^{(\rho)}\}$.

It is important to note that there is another observation model commonly used in the literature (e.g., [34–37]). The only difference is that the order of warping and blurring operations is reversed; that is, $\mathbf{y}^{(k)} = \mathbf{D}^{(k)}\mathbf{W}^{(k)}\mathbf{P}^{(k)}\mathbf{X} + \mathbf{V}^{(k)}$. When the imaging blur is spatio-temporally invariant and only global translational motion is involved among multiple observed low-resolution images, the blur matrix $\mathbf{P}^{(k)}$ and the motion matrix $\mathbf{W}^{(k)}$ are commutable. Consequently, these two models coincide. However, when the imaging blur is spatio-temporally variant, it is more appropriate to use the second model. The determination of the mathematical model for formulating the SR computation should coincide with the imaging physics (i.e., the physical process to capture low-resolution images from the original high-resolution ones) [2].

2.2 Observation model for super-resolution video

The observation model for SR video computation can be formulated by applying the observation model of SR image at each temporal instant. More specifically, at each temporal instant n, the relationship between the low-resolution image and the high-resolution image can be mathematically formulated as

$$\mathbf{y}_n = \mathbf{H}_n\mathbf{X}_n + \mathbf{V}_n, \quad (3)$$

where $\mathbf{H}_n$ is defined as in (2), $\mathbf{y}_n$ and $\mathbf{X}_n$ represent an $L_1 \times L_2$ low-resolution image and an $M_1 \times M_2$ high-resolution image in the temporal instant n, respectively. Both of these two images are represented in the lexicographic-ordered vector form, with a size of $L_1L_2 \times 1$ and $M_1M_2 \times 1$, respectively. Additionally, $\mathbf{V}_n$ is modeled as a zero-mean white Gaussian noise vector (with a size of $L_1L_2 \times 1$), encountered during the image acquisition process. Furthermore, $\mathbf{V}_n$ is assumed to be independent with $\mathbf{X}_n$.

In summary, the SR video problem is an *inverse problem*, and its goal is to make use of a set of low-resolution images $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n\}$ with a dimension of $L_1 \times L_2$ each to compute the high-resolution ones $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ with a dimension of $M_1 \times M_2$ each (i.e., the enlarged resolution ratio is $q_1 \times q_2$).

3 Super-resolution image reconstruction

The objective of SR image reconstruction is to produce an image with a higher resolution based on one or a set of images captured from the same scene. In general, the SR image techniques can be classified into four classes: (i) *frequency-domain*-based approach [21,38–43], (ii) *interpolation*-based approach [44–47], (iii) *regularization*-based approach [48–66], and (iv) *learning*-based approach [67–74]. The first three categories get a higher-resolution image from a set of lower-resolution input images, while the last one achieves the same objective by exploiting the information provided by an image database.

3.1 Frequency-domain-based SR image approach

The first frequency-domain SR method can be credited to Tsai and Huang [21], where they considered the SR computation for the noise-free low-resolution images. They proposed to first transform the low-resolution image data into the *discrete Fourier transform* (DFT) domain and combined them according to the relationship between the aliased DFT coefficients of the observed low-resolution images and that of the unknown high-resolution image. The combined data are then transformed back to the spatial domain where the new image could have a higher resolution than that of the input images. Rhee and Kang [75] exploited the *Discrete cosine transform* (DCT) to perform fast image deconvolution for SR image computation. Woods et al. [76] presented an iterative *expectation maximization* (EM) algorithm [77] for simultaneously performing the registration, blind deconvolution, and interpolation operations.

The frequency-domain-based SR approaches have a number of advantages. First, it is an intuitive way to enhance the details (usually the high-frequency information) of the images by extrapolating the high-frequency information presented in the low-resolution images. Secondly, these frequency-domain-based SR approaches have low computational complexity. However, the frequency-domain-based SR methods are insufficient to handle the real-world applications, since they require that there only exists a global displacement between the observed images and the linear space-invariant blur during the image acquisition process.

Recently, many researchers have begun to investigate the use of the wavelet transform for addressing the SR problem to recover the detailed information (usually the high-frequency information) that is lost or degraded during the image acquisition process. This is motivated by that the wavelet transform provides a powerful and efficient multi-scale representation of the image for recovering the high-frequency information [38]. These approaches typically treat the observed low-resolution images as the low-pass filtered subbands of the unknown wavelet-transformed high-resolution image, as shown in Fig. 4. The aim is to estimate the finer scale subband coefficients, followed by applying the inverse wavelet transform, to produce the high-resolution image. To be more specific, take the 2×2 SR computation as an example. The low-resolution images are viewed as the representation of wavelet coefficients after several levels (say, N levels) of decomposition. Then, the high-resolution image can be produced by estimating the $(N + 1)$th scale wavelet coefficients, followed by applying the inverse wavelet decomposition.

In [39,40], Ei-Khamy et al. proposed to first register multiple low-resolution images in the wavelet domain, then fuse the registered low-resolution wavelet coefficients to obtain a single image, followed by performing interpolation to get a higher-resolution image. Ji and Fermuller [41,42] proposed a robust wavelet SR approach to handle the error incurred in both the registration computation and the blur identification computation. Chappalli and Bose incorporated a denoising stage into the conventional wavelet-domain SR framework to develop a simultaneous denoising and SR reconstruction approach in [43].

3.2 Interpolation-based SR image approach

The *interpolation*-based SR approach constructs a high-resolution image by projecting all the acquired low-resolution images to the reference image, then fuse together all the information available from each image, due to the fact that each low-resolution image provides an amount of additional information about the scene, and finally deblurs the image. Note that the single image interpolation algorithm cannot handle the SR problem well, since it cannot produce those high-frequency components that were lost during the image acquisition process. The quality of the interpolated image generated by applying any single input image interpolation algorithm is inherently limited by the amount of data available in the image.

The interpolation-based SR approach usually consists of the following three stages, as depicted in Fig. 5: (i) the registration stage for aligning the low-resolution input images, (ii) the interpolation stage for producing a higher-resolution

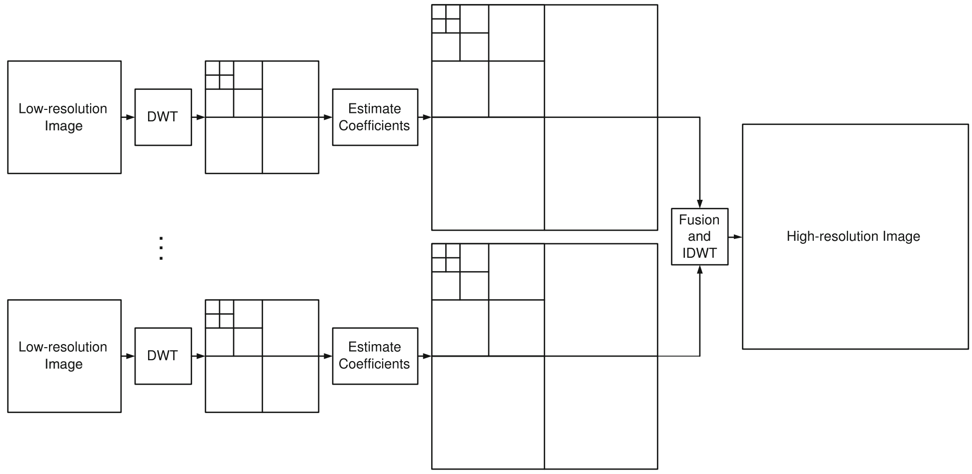

Fig. 4 A conceptual framework of wavelet-domain-based SR image approach

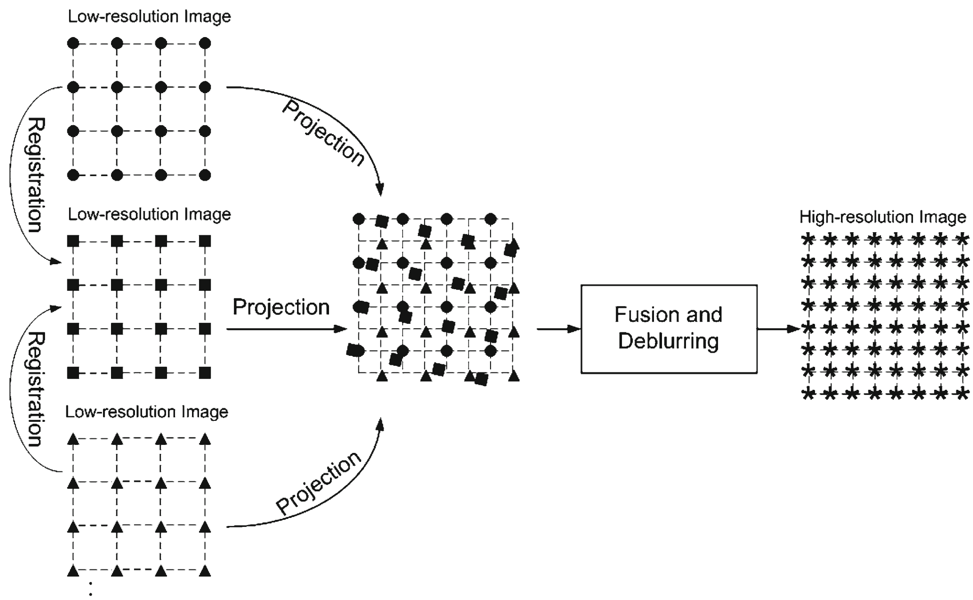

Fig. 5 The interpolation-based approach for SR image reconstruction

image, and (iii) the deblurring stage for enhancing the reconstructed high-resolution image produced in the step ii). The interpolation stage plays a key role in this framework. There are various ways to perform interpolation. The simplest interpolation algorithm is the *nearest neighbor* algorithm, where each unknown pixel is assigned with an intensity value that is same as its neighboring pixels. But this method tends to produce images with a blocky appearance. Ur and Gross [44] performed a nonuniform interpolation of a set of spatially shifted low-resolution images by utilizing the generalized multichannel sampling theorem. The advantage of this approach is that it has low computational load, which is thus quite suitable for real-time applications. However, the optimality of the entire reconstruction process is not guaranteed, since the interpolation errors are not taken into account. Bose and Ahuja [45] used the *moving least square* (MLS) method to estimate the intensity value at each pixel position of the high-resolution image via a polynomial approximation using the pixels in a defined neighborhood of the pixel position under consideration. Furthermore, the coefficients and the order of the polynomial approximation are adaptively adjusted for each pixel position.

Three steps that are depicted in Fig. 5 can be conducted iteratively. Irani and Peleg [46] proposed an *iterative back-projection* (IBP) algorithm, where the high-resolution image is estimated by iteratively projecting the difference between the observed low-resolution images and the simulated low-resolution images. However, this method might not yield unique solution due to the ill-posed nature of the SR problem. A *projection onto convex sets* (POCS) was proposed by Patti and Tekalp [47] to develop a set-theoretic algorithm to produce the high-resolution image that is consistent with the information arising from the observed low-resolution images and the prior image model. These information are associated with the constraint sets in the solution space; the intersection of these sets represents the space of permissible solutions.

By projecting an initial estimate of the unknown high-resolution image onto these constraint sets iteratively, a fairly good solution can be obtained. This kind of method is easy to be implemented; however, it does not guarantee uniqueness of the solution. Furthermore, the computational cost of this algorithm is very high.

3.3 Regularization-based SR image approach

Motivated by the fact that the SR computation is, in essence, an ill-posed inverse problem [78], numerous regularization-based SR algorithms have been developed for addressing this issue [49,50,52,54–56]. The basic idea of these regularization-based SR approaches is to use the regularization strategy to incorporate the prior knowledge of the unknown high-resolution image. From the Bayesian point of view, the information that can be extracted from the observations (i.e., the low-resolution images) about the unknown signal (i.e., the high-resolution image) is contained in the probability distribution of the unknown. Then, the unknown high-resolution image can be estimated via some statistics of a probability distribution of the unknown high-resolution image, which is established by applying Bayesian inference to exploit the information provided by both the observed low-resolution images and the prior knowledge of the unknown high-resolution image.

Two most popular Bayesian-based SR approaches are *maximum likelihood* (ML) estimation approach [48] and *maximum a posterior* (MAP) estimation approach [49,50, 52,54–56].

3.3.1 ML estimation approach

The first ML estimation-based SR approach is proposed by Tom and Katsaggelos in [48], where the aim is to find the ML estimation of the high-resolution image (denoted as $\hat{\mathbf{X}}^{\mathrm{ML}}$) by

$$\hat{\mathbf{X}}^{\mathrm{ML}} = \underset{\mathbf{X}}{\operatorname{argmax}}\, p\left(\mathbf{Y}|\mathbf{X}\right). \tag{4}$$

The key of the above ML estimation approach is to establish the conditional pdf $p\left(\mathbf{Y}|\mathbf{X}\right)$ in (4). First, since the low-resolution images are obtained independently from the original (high-resolution) image, the conditional pdf $p\left(\mathbf{Y}|\mathbf{X}\right)$ can be expressed as

$$p\left(\mathbf{Y}|\mathbf{X}\right) = \prod_{k=1}^{\rho} p\left(\mathbf{y}^{(k)}|\mathbf{X}\right). \tag{5}$$

Furthermore, according to (2), the conditional pdf $p\left(\mathbf{y}^{(k)}|\mathbf{X}\right)$ can be expressed as

$$p\left(\mathbf{y}^{(k)}|\mathbf{X}\right) \propto \exp\left(-\frac{1}{2\sigma_k^2}\left\|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\right\|^2\right). \tag{6}$$

Substituting (6) into (5), the conditional pdf $p(\mathbf{Y}|\mathbf{X})$ can be obtained as

$$\begin{aligned} p(\mathbf{Y}|\mathbf{X}) &\propto \prod_{k=1}^{\rho} \exp\left(-\frac{1}{2\sigma_k^2}\left\|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\right\|^2\right) \\ &= \exp\left(-\sum_{k=1}^{\rho}\frac{1}{2\sigma_k^2}\left\|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\right\|^2\right). \end{aligned} \tag{7}$$

Substituting (7) into (4), we can obtain the estimation of the high-resolution image as

$$\hat{\mathbf{X}}^{\mathrm{ML}} = \underset{\mathbf{X}}{\operatorname{argmax}} \exp\left(-\sum_{k=1}^{\rho}\frac{1}{2\sigma_k^2}\left\|\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{X}\right\|^2\right). \tag{8}$$

3.3.2 MAP estimation approach

In contrast to the ML SR approach that only considers the relationship among the observed low-resolution images and the original high-resolution image, the MAP SR approach further incorporates the prior image model to reflect the expectation of the unknown high-resolution image. These approaches provide a flexible framework for the inclusion of prior knowledge of the unknown high-resolution image, as well as for modeling the dependence between the observed low-resolution images and the unknown high-resolution image. The aim of the MAP SR approach is to find the MAP estimation of the high-resolution image (denoted as $\hat{\mathbf{X}}^{\mathrm{MAP}}$) via

$$\hat{\mathbf{X}}^{\mathrm{MAP}} = \underset{\mathbf{X}}{\operatorname{argmax}}\, p\left(\mathbf{X}|\mathbf{Y}\right). \tag{9}$$

To evaluate the above pdf, the Bayes rule is applied to rewrite $p\left(\mathbf{X}|\mathbf{Y}\right)$ as

$$p\left(\mathbf{X}|\mathbf{Y}\right) \propto p\left(\mathbf{Y}|\mathbf{X}\right) p\left(\mathbf{X}\right). \tag{10}$$

There are three fundamental issues required to be solved for the above-mentioned Bayesian-based SR approaches: (i) determination of the prior image model (i.e., $p(\mathbf{X})$ in (10)); (ii) determination of optimal prior image model's hyperparameter value; and (iii) computation of the high-resolution image by solving the optimization problem defined in (9). All of them will be discussed in detail as follows.

First, the prior image model plays a key role for determining the reconstructed high-resolution image. Since the image is viewed as a locally smooth data field, the *Markov random field* (MRF) [79] is considered as a useful prior image model, which bears the following form [79]

$$p(\mathbf{X}) = \frac{1}{Z(\lambda)}\exp\left\{-\lambda U(\mathbf{X})\right\}, \tag{11}$$

where λ is the so-called *hyperparameter*, and $Z(\lambda)$ is a normalization factor, which is called the *partition function* and defined as $Z(\lambda) = \int_{\chi}\exp\left\{-\lambda U(\mathbf{X})\right\} d\mathbf{X}$, where χ represents all possible configurations of the unknown high-resolution image $\mathbf{X}$, and the *energy function* $U(\mathbf{X})$. The *Gaussian Markov random field model* [80] can be selected for formulating the prior image model, which has been proved to be effective for super-resolution problem [49,50]; hence, the $U(\mathbf{X})$ in (11) can be rewritten as [80]

$$U(\mathbf{X}) = \mathbf{X}^T \mathbf{C}\mathbf{X}, \tag{12}$$

where $\mathbf{C}$ is an $M_1 M_2 \times M_1 M_2$ matrix, whose entries are given by $C(i, j) = \gamma$, if $i = j$; $C(i, j) = -1$, if both i and j fall in the γ-neighborhood; otherwise, $C(i, j) = 0$.

Besides the aforementioned MRF model (i.e., (11)), an edge-enhanced prior image model is proposed in [51] by incorporating an edge-enhancing term into the MRF image model (12). Moreover, several other models have been exploited, such as, the *Huber-Markov random field* model (HMRF) [52], the *conditional random field* (CRF) [53], the *discontinuity adaptive MRF* (DAMRF) [54], the two-level Gaussian nonstationary model [55], the prior image model with the incorporation of the segmentation information [56–58], and the prior image model particularly for text images [59] or face images [60,61].

The second issue is how to determine the optimal hyperparameter value of the chosen prior image model. It is important to note that the hyperparameter controls the degree of regularization and intimately affects the quality of the reconstructed high-resolution image. To tackle this problem, some works have been developed to automatically determine the prior image model's hyperparameter value during the process of estimating the unknown high-resolution image [62–66], rather than applying an experimentally determined fixed value [81]. Tipping and Bishop [62] exploited the *expectation maximization* (EM) algorithm for estimating the hyperparameter value by maximizing its marginal likelihood function. However, the EM algorithm used in this method yields tremendous computational cost, and more importantly, the EM algorithm does not always converge to the global optimum [82]. Pickup et al. [63] improved the method [62] by reducing its computational load and further modified the prior image model by considering illumination changes among the captured low-resolution images. He and Kondi [64] proposed a method to jointly estimate both the hyperparameter value and the high-resolution image by optimizing a single cost function. However, this method assumes that the displacements among the low-resolution images are restricted to integer-valued shifts on the high-resolution grid; obviously, this assumption often mismatches the real-world image acquisition. In [76], Woods et al. proposed to apply the *expectation maximization* (EM) method to compute the maximum likelihood estimation of the hyperparameter. This method requires that the geometric displacements between the low-resolution images have the globally uniform translations only.

The third major challenge of Bayesian-based SR approaches is that the above-mentioned distribution of the unknown high-resolution image is usually very complicated; thus, its statistics are difficult to be directly computed. For that, the Monte Carlo method [83] has been considered as a promising approach, since it provides an extremely powerful and efficient way to compute the statistics of any complicate distribution. The Monte Carlo method [83] aims to provide an accurate estimation of the unknown target (i.e., the high-resolution image) through stochastic simulations. For that, a fairly large number of samples are generated according to the target probability distribution, followed by discarding those initially generated unreliable samples such that more accurate statistical measurement of the target distribution can be obtained. By further incorporating the first-order Markov chain property (that is, the generation of the current sample only depends on the previous one) into the Monte Carlo method [83] for generating sufficiently large number of reliable samples as mentioned above, the established *Markov chain Monte Carlo* (MCMC) process is exploited to develop a novel stochastic SR approach in [50]. They exploit the MCMC method for producing the high-resolution image, based on the reliable image samples distributed according to the joint posterior distribution of the unknown high-resolution image and the prior image model's parameter given the observed low-resolution images.

3.4 Learning-based SR image approach

Recently, learning-based techniques are proposed to tackle the SR problem [67–70]. In these approaches, the high-frequency information of the given single low-resolution image is enhanced, by retrieving the most likely high-frequency information from the given training image samples based on the local features of the input low-resolution image. Hertzmann [67] proposed an image analogy method to *create* the high-frequency details for the observed low-resolution image from a training image database, as illustrated in Fig. 6. It contains two stages: an off-line training stage and a SR reconstruction stage. In the off-line training stage, the image patches serve as ground truth and are used to generate low-resolution patches through the simulating the image acquisition model (2). Pairs of low-resolution patches and the corresponding (ground truth) high-frequency patches are collected. In the SR reconstruction stage, the patches extracted from the input low-resolution images are compared with those stored in the database. Then, the best matching patches are selected according to a certain similarity measurement criterion (e.g., the *nearest distance*) as the corresponding high-frequency patches used for producing the high-resolution image. Chang et al. [68] proposed that the generation of the high-resolution image patch depends on multiple nearest neighbors in the training set in a way similar to the concept of manifold learning methods, particularly the *locally linear embedding* (LLE) method [68]. In contrast to the generation

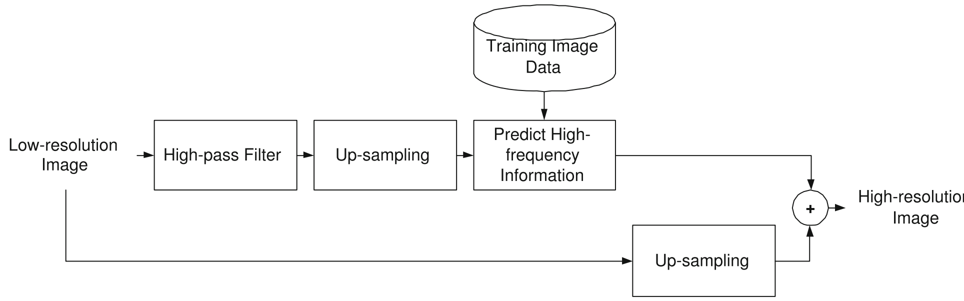

Fig. 6 A conceptual framework of learning-based SR image approach [69]

of a high-resolution image patch, which depends on only one of the nearest neighbors in the training set as used in the aforementioned SR approaches [67,69,70], this method requires fewer training samples.

Another way is to jointly exploit the information learnt from a given high-resolution training data set, as well as that provided by multiple low-resolution observations. Datsenko and Elad [71] first assigned several candidate high-quality patches at each pixel position in the observed low-resolution image; these are found as the nearest-neighbors in an image database that contains pairs of corresponding low-resolution and high-resolution image patches. These found patches are used as the prior image model and then merged into an MAP cost function to arrive at the closed-form solution of the desired high-resolution image.

Recent research on the studies of image statistics suggests that image patches can be represented as a sparse linear combination of elements from an over-complete image patch dictionary [72–74]. The idea is to seek a sparse representation for each patch of the low-resolution input, followed by exploiting this representation to generate the high-resolution output. By jointly training two dictionaries for the low-resolution and high-resolution image patches, the sparse representation of a low resolution image patch can be applied with the high-resolution image patch dictionary to generate a high-resolution image patch.

4 Super-resolution video reconstruction

The SR video approaches reconstruct an image sequence with a higher resolution from a group of adjacent lower-resolution uncompressed image frames or compressed image frames. The major challenges in SR video problem lie in several parts. The first challenge is how to extract additional information from each low-resolution image to improve the quality of the final high-resolution images. The second challenge is how to consider the uncertainty (i.e., estimation error) incurred in motion estimation of the SR computation. For that, a reliable motion estimation is highly essential for achieving high-quality super-resolution reconstruction. The SR algorithms are usually required to be robust with respect to these errors [84]. For that, various motion estimation methods have been developed, including parametric modeling [85], block matching [86,87], and optical flow estimation [88–92].

The existing SR video approaches can be classified into the following four categories: (i) the sliding-window-based SR video approach [93–96], (ii) the simultaneous SR video approach [97–99], (iii) the sequential SR video approach [100–104], and (iv) the learning-based SR video approach [105–107]. In the following, a review of each category is presented.

4.1 Sliding-window-based SR video approach

The *sliding-window*-based approach [93–96], as illustrated in Fig. 7, is the most commonly-used and direct approach to conduct SR video. The sliding window selects a set of consecutive low-resolution frames for producing one high-resolution image frame; that is, the window is moved across the input frames to produce successive high-resolution frames sequentially. The major drawback of this approach is that the temporal correlations among the consecutively reconstructed high-resolution images are not considered.

4.2 Simultaneous SR video approach

Many works ([97–99], for examples) have been done to simultaneously reconstruct multiple high-resolution images at one time, as illustrated in Fig. 8. Borman and Stevenson [97] proposed an algorithm to reconstruct multiple high-resolution images simultaneously by imposing temporal smoothness constraint on the prior image model, while Zibetti and Mayer [98] reduced its computational complexity by only considering the temporally smooth low-resolution frames in the observation model. Alvarez et al. [99] proposed a multichannel SR approach for the compressed image sequence, which applies the sliding window to select a set of low-resolution frames to reconstruct another set (same number of frames) of high-resolution frames simultaneously. However, all these three algorithms require large memory storage, since these multiple high-resolution and low-resolution images are required available at the same time during the reconstruction process. Moreover, how many high-resolution

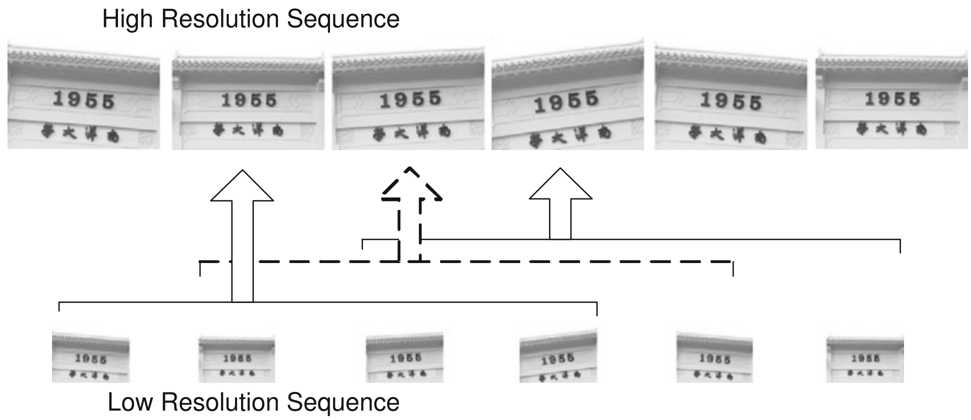

Fig. 7 The sliding-window-based super-resolution video scheme

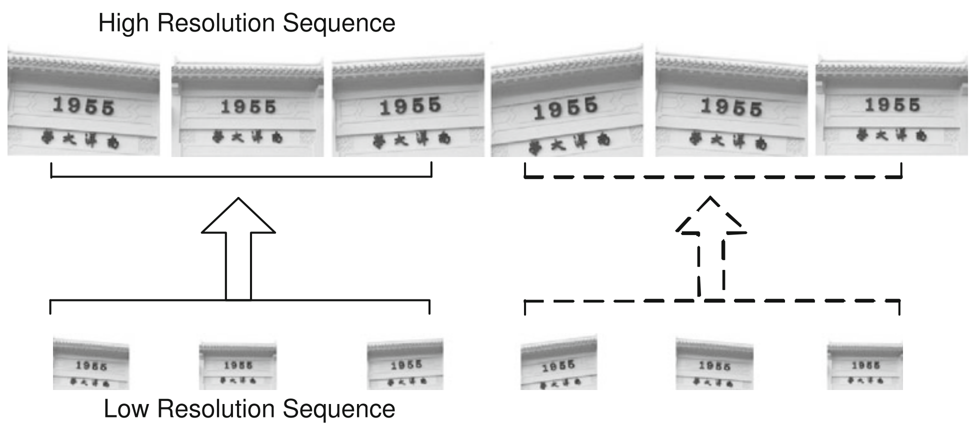

Fig. 8 The simultaneous multiple high-resolution frames reconstruction scheme

frames should be exploited for simultaneous reconstructions is another important issue, which has not been addressed.

4.3 Sequential SR video approach

The major challenge in the SR video problem is how to exploit the temporally correlated information provided by the established high-resolution images and available temporally-correlated low-resolution images respectively to improve the quality of the desired high-resolution images. Elad et al. [100–102] proposed an SR image sequence algorithm based on adaptive filtering theory, which exploits the correlation information among the high-resolution images. However, the information provided by the previously observed low-resolution images is neglected; that is, only a single low-resolution image is used to compute the least-squares estimation for producing one high-resolution image. On the other hand, the computational complexity and the convergence issues of the SR algorithm developed in [100] are analyzed by Costa and Bermudez [103].

A three-equation-based state-space filtering is proposed by Tian and Ma [104], by incorporating an extra observation equation into the framework of the conventional two-equation-based Kalman filtering to build up a three-equation-based state-space model. In [104], a full mathematical derivation for arriving at a closed-form solution is provided, which exploits the information from the previously reconstructed high-resolution frame, the currently observed low-resolution frame as well as the previously observed low-resolution frame for producing the next high-resolution frame. The above-mentioned steps will be sequentially processed across the image frames. This is in contrast to the conventional two-equation-based Kalman filtering approach, in which only the previously reconstructed high-resolution frame and the currently observed low-resolution frame are exploited.

4.4 Learning-based SR video approach

Bishop et al. [105] proposed a learning-based SR video method, based on the principle of learning-based SR still image approach developed in [69]. Their method uses a learnt data set of image patches capturing the relationship between the low-frequency and the high-frequency bands of natural images and uses a prior image model over such patches. Furthermore, their method uses the previously enhanced frame to provide part of the training set for producing the current high-resolution frame. Dedeoglu et al. [106] and Kong et al.

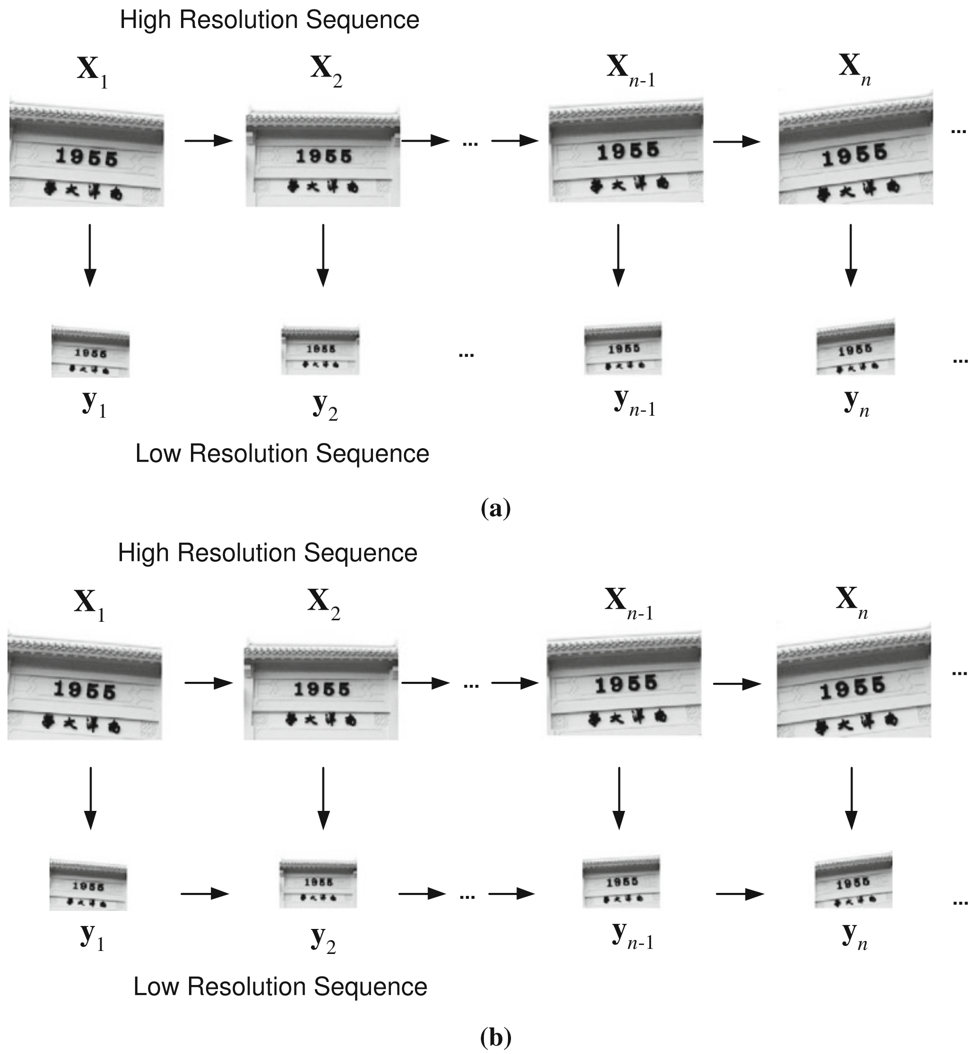

Fig. 9 The relationship among the low-resolution frames $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ and the high-resolution frames $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ using: **a** the conventional two-equation-based Kalman filtering [100]; **b** the three-equation-based state-space filtering proposed in [104]

[107] exploit the same idea, while enforcing both the spatial and the temporal smoothness constraints among the reconstructed high-resolution frames.

5 Future challenges

In this section, two research challenges are discussed for future SR research: 1) multi-view SR imaging [108–110], and 2) *temporal* SR video (by increasing the frame rate) [111–114].

5.1 Multi-view SR imaging

The objective of SR stereo imaging is to reconstruct a pair of images from their low-resolution counterparts, in order to enhance visualization and improve content recognition accuracy. This is different from the conventional SR image reconstruction problem that was discussed in Sects. 3 and 4, where a set of similar images are acquired from the same scene using a single camera with multiple shots to produce a single higher-resolution image only.

Regarding the SR image reconstruction for stereo images, there are only three references, [108–110], can be found in the literature. Motivated by the fact that the SR computation is, in essence, an ill-posed inverse problem [62,63], a regularization strategy is usually exploited for numerically solving the SR computation by incorporating a prior knowledge of higher-resolution image under reconstruction. Kimura et al. [108] proposed a *maximum a posterior* (MAP) approach to use multiple pairs of stereo images to estimate both the higher-resolution disparity field and a single higher-resolution image. Furthermore, Kimura et al.'s method imposed the same prior model for both the higher-resolution image and the higher-resolution disparity field. However, the prior model of the disparity field should be different from that of the higher-resolution image [115]. The formulation of the prior model of the disparity field needs further investigation. Bhavsar and Rajagopalan [109,110] proposed to exploit an *iterated conditional modes* (ICM) algorithm to reconstruct the higher-resolution images via estimating their MAP estimators. However, the ICM algorithm depends very much on the initial estimator, and it does not always converge to the global minimum [116]. Furthermore, the prior image model used in their approach neglects the constraints between the reconstructed pair of high-resolution images.

5.2 Temporal SR video

Recently, *temporal* resolution enhancement of digital video has received growing attention [111–114]. Temporal resolution refers to the number of frames captured per second and is also commonly known as the *frame rate*. It is related to the amount of motion perceived across the frames. A higher frame rate could result in less, or even completely avoid, smearing artifacts due to the movements of moving objects encountered in the scene. A typical frame rate suitable for a pleasing view is about 25 frames per second or above. Conventional ways utilize single video sequence as input and exploit a frame-rate up-conversion technique (e.g., [117–119]) to increase temporal resolution. Robertson and Stevenson [111] considered to increase the frame rate of compressed video by inserting a number of frames between received adjacent frames of the sequence. They exploited a Bayesian framework to prevent the spatial compression artifacts in the received frame from propagating to the adjacent reconstructed high-resolution frames. Shechtman et al. [112] proposed a joint space-time SR framework by using multiple image sequence which has different spatial resolutions and different frame rates. Watanabe et al. [113] developed a method for obtaining a high spatio-temporal resolution video from two video sequences. The fist sequence has a high spatial resolution but a low frame rate, while the second sequence has a low spatial resolution but a high frame rate. Both of these two sequences are captured by a dual sensor camera that can capture these sequences with the same field of view simultaneously. Maor et al. [114] proposed an algorithm to generate high-definition video sequences using commercial digital camcorders. Under the constraint that the data read-out rate of a camcorder is limited, a digital video sequence is first nonuniformly sampled, and then its resolution is further increased by using a SR technique to generate a high-resolution video clip.

6 Conclusions

The SR imaging has been one of the fundamental image processing research areas. It can overcome or compensate the inherent hardware limitations of the imaging system to provide a more clear image with a richer and informative content. It can also be served as an appreciable front-end pre-processing stage to facilitate various image processing applications to improve their targeted terminal performance. In this survey paper, our goal is to offer new perspectives and outlooks of SR imaging research, besides giving an updated overview of existing SR algorithms. It is our hope that this work could inspire more image processing researchers endeavoring on this fascinating topic and developing more novel SR techniques along the way.

References

1. Chaudhuri, S.: Super-Resolution Imaging. Kluwer Academic Publishers, Boston (2001)
2. Katsaggelos, A.K., Molina, R., Mateos, J.: Super Resolution of Images and Video. Morgan & Claypool, San Rafael (2007)
3. Bannore, V.: Iterative-Interpolation Super-Resolution Image Reconstruction: A Computationally Efficient Technique. Springer, Berlin (2009)
4. Milanfar, P.: Super-Resolution Imaging. CRC Press, Boca Raton (2010)
5. Kang, M.G., Chaudhuri, S.: Super-resolution image reconstruction. IEEE Signal Process. Mag. **20**, 19–20 (2003)
6. Bose, N.K., Chan, R.H., Ng, M.K.: Special issue on high resolution image reconstruction. Int. J. Imaging Syst. Technol. **14**(2–3), 35–145 (2004)
7. Ng, M.K., Chan, T., Kang, M.G., Milanfar, P.: Special issue on super-resolution imaging: analysis, algorithms, and applications. EURASIP J. Appl. Signal Process. **2006**, 1–2 (2006)
8. Ng, M.K., Lam, E., Tong, C.S.: Special issue on superresolution imaging: theory, algorithms and applications. Multidimens. Syst. Signal Process. **18**, 55–188 (2007)
9. Hardie, R.C., Schultz, R.R., Barner, K.E.: Special issue on super-resolution enhancement of digital video. EURASIP J. Appl. Signal Process. **2007**, 1–3 (2007)
10. Katsaggelos, A.K., Molina, R.: Special issue on super resolution. Comput. J. **52**, 1–167 (2008)
11. Hunt, B.R.: Superresolution of images: algorithms, principles, performance. Int. J. Imaging Syst. Technol. **6**(4), 297–304 (1995)
12. Borman, S., Stevenson R.L.: Spatial resolution enhancement of low-resolution image sequences: a comprehensive review with directions for future research. University of Notre Dame, Technical Report (1998)
13. Borman, S., Stevenson, R.L.: Super-resolution from image sequences—a review. In: Proceedings of the IEEE Midwest Symposium on Circuits and Systems, pp. 374–378. Notre Dame, IN (1998)
14. Park, S.C., Park, M.K., Kang, M.G.: Super-resolution image reconstruction: a technical overview. IEEE Signal Process. Mag. **20**, 21–36 (2003)
15. Ng, M.K., Bose, N.K.: Mathematical analysis of super-resolution methodology. IEEE Signal Process. Mag. **20**, 62–74 (2003)
16. Capel, D., Zisserman, A.: Computer vision applied to super resolution. IEEE Signal Process. Mag. **20**, 75–86 (2003)
17. Choi, E., Choi, J., Kang, M.G.: Super-resolution approach to overcome physical limitations of imaging sensors: an overview. Int. J. Imaging Syst. Technol. **14**, 36–46 (2004)
18. Farsiu, S., Robinson, D., Elad, M., Milanfar, P.: Advances and challenges in super-resolution. Int. J. Imaging Syst. Technol. **14**, 47–57 (2004)

19. Ouwerkerk, J.D.: Image super-resolution survey. Image Vis. Comput. **24**, 1039–1052 (2006)
20. Cristobal, G., Gila, E., Sroubek, F., Flusser, J., Miravet, C., Rodriguez, F.B.: Superresolution imaging: a survey of current techniques. In: Proceedings of the SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations, pp. 70740C.1–70740C.18 (2008)
21. Tsai, R.Y., Huang, T.S.: Multiframe image restoration and registration. In: Huang, T.S. (ed.) Advances in Computer Vision and Image Processing. JAI Press Inc., London (1984)
22. Max, B., Wolf, E.: Principles of Optics. Cambridge University Press, Cambridge (1997)
23. Dropkin, H., Ly, C.: Superresolution for scanning antenna. In: Proceedings of the IEEE International Radar Conference, pp. 306–30. (1997)
24. Caner, G., Tekalp, A.M., Heinzelman, W.: Performance evaluation of super-resolution reconstruction from video. In: Proceedings of the IS&T/SPIE's Annual Symposium on Electronic Imaging, pp. 177–185. Santa Clara, CA (2003)
25. Tanaka, M., Okutomi, M.: Theoretical analysis on reconstruction-based super-resolution for an arbitrary PSF. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, pp. 947–954. San Diego, CA (2005)
26. Rajagopalan, A.N., Kiran, V.P.: Motion-free superresolution and the role of relative blur. Opt. Soc. Am. J. A Opt. Image Sci. **20**, 2022–2032 (2003)
27. Pham, T.Q., van Vliet, L.J., Schutte, K.: Influence of signal-to-noise ratio and point spread function on limits of superresolution. In: Proceedings of the SPIE Conference on Image Processing: Algorithms and Systems IV, pp. 169–180. San Jose, CA (2005)
28. Wang, Z.-Z., Qi, F.-H.: Analysis of multiframe super-resolution reconstruction for image anti-aliasing and deblurring. Image Vis. Comput. **23**, 393–404 (2005)
29. Eekeren, A.W.M., Schutte, K., Oudegeest, O.R., Vliet, L.J.: Performance evaluation of super-resolution reconstruction methods on real-world data. EURASIP J. Appl. Signal Process. **2007**(43953), 1–11 (2007)
30. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 1167–1183 (2002)
31. Lin, Z., Shum, H.-Y.: Fundamental limits of reconstruction-based superresolution algorithms under local translation. IEEE Trans. Pattern Anal. Mach. Intell. **26**, 83–97 (2004)
32. Lin, Z.-C., He, J.-F., Tang, X.-O., Tang, C.-K.: Limits of learning-based superresolution algorithms. In: Proceedings of the IEEE International Conference on Computer Vision. Rio de Janeiro, Brazil (2007)
33. Robinson, D., Milanfar, P.: Statistical performance analysis of super-resolution. IEEE Trans. Image Process. **15**, 1413–1428 (2006)
34. Chiang, M.-C., Boult, T.E.: Efficient super-resolution via image warping. Image Vis. Comput. **28**, 761–771 (2000)
35. Patti, A., Altunbasak, Y.: Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. IEEE Trans. Image Process. **10**, 179–186 (2001)
36. Lertrattanapanich, S., Bose, N.K.: High resolution image formation from low resolution frames using Delaunay triangulation. IEEE Trans. Image Process. **11**, 1427–1441 (2002)
37. Wang, Z.-Z., Qi, F.-H.: On ambiguities in super-resolution modeling. IEEE Signal Process. Lett. **11**, 678–681 (2004)
38. Nguyen, N., Milanfar, P.: A wavelet-based interpolation-restoration method for superresolution. Circuits Syst. and Signal Process. **19**, 321–338 (2000)
39. Ei-Khamy, S.E., Hadhoud, M.M., Dessouky, M.I., Salam, B.M., Ei-Samie, F.E.: Regularized super-resolution reconstruction of images using wavelet fusion. Opt. Eng. **44**, 097001.1–097001.10 (2005)
40. Ei-Khamy, S.E., Hadhoud, M.M., Dessouky, M.I., Salam, B.M., Ei-Samie, F.E.: Wavelet fusion: a tool to break the limits on LMMSE image super-resolution. Int. J. Wavel. Multiresolut. Inf. Process. **4**, 105–118 (2006)
41. Ji, H., Fermuller, C.: Wavelet-based super-resolution reconstruction: theory and algorithm. In: Proceedings of the European Conference on Computer Vision, pp. 295–307. Graz, Austria (2006)
42. Ji, H., Fermuller, C.: Robust wavelet-based super-resolution reconstruction: theory and algorithm. IEEE Trans. Pattern Anal. Mach. Intell. **31**, 649–660 (2009)
43. Chappalli, M.B., Bose, N.K.: Simultaneous noise filtering and super-resolution with second-generation wavelets. IEEE Signal Process. Lett. **12**, 772–775 (2005)
44. Ur, H., Gross, D.: Improved resolution from subpixel shifted pictures. CVGIP Graph. Models Image Process. **54**, 181–186 (1992)
45. Bose, N.K., Ahuja, N.A.: Superresolution and noise filtering using moving least squares. IEEE Trans. Image Process. **15**, 2239–2248 (2006)
46. Irani, M., Peleg, S.: Improving resolution by image registration. CVGIP Graph. Models Image Process. **53**, 231–239 (1991)
47. Patti, A.J., Tekalp, A.M.: Super resolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. IEEE Trans. Image Process. **6**, 1446–1451 (1997)
48. Tom, B.C., Katsaggelos, A.K.: Reconstruction 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, pp. 539–542. Washington, DC (1995)
49. Hardie, R.C., Barnard, K.J., Armstrong, E.E.: Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. IEEE Trans. Image Process. **6**, 1621–1633 (1997)
50. Tian, J., Ma, K.-K.: Stochastic super-resolution image reconstruction. J. Vis. Commun. Image Represent. **21**, 232–244 (2010)
51. Tian, J., Ma, K.-K.: Edge-adaptive super-resolution image reconstruction using a Markov chain Monte Carlo approach. In: Proceedings of the International Conference on Information, Communications and Signal Processing. Singapore (2007)
52. Schultz, R.R., Stevenson, R.L.: Extraction of high-resolution frames from video sequences. IEEE Trans. Image Process. **5**, 996–1011 (1996)
53. Kong, D., Han, M., Xu, W., Tao, H., Gong Y.: A conditional random field model for video super-resolution. In: Proceedings of the IEEE International Conference on Pattern Recognition, pp. 619–622. Hong Kong (2006)
54. Suresh, K.V., Rajagopalan, A.N.: Robust and computationally efficient superresolution algorithm. Opt. Soc. Am. J. A Opt. Image Sci. **24**, 984–992 (2007)
55. Belekos, S., Galatsanos, N.P., Katsaggelos, A.K.: Maximum a posteriori video super-resolution using a new multichannel image prior. IEEE Trans. Image Process. **19**, 1451–1464 (2010)
56. Shen, H., Zhang, L., Huang, B., Li, P.: A MAP approach for joint motion estimation, segmentation, and super resolution. IEEE Trans. Image Process. **16**, 479–490 (2007)
57. Humblot, F., Mohammad-Djafari, A.: Super-resolution using hidden Markov model and Bayesian detection estimation framework. EURASIP J. Appl. Signal Process. **2006**(36971), 1–16 (2006)
58. Mohammad-Djafari, A.: Super-resolution: A short review, a new method based on hidden Markov modeling of hr image and future challenges. Comput. J. **52**, 126–141 (2008)

59. Capel, D., Zisserman, A.: Super-resolution enhancement of text image sequences. In: Proceedings of the IEEE International Conference on Pattern Recognition, pp. 600–605. Barcelona, Spain (2000)
60. Chakrabarti, A., Rajagopalan, A.N., Chellappa, R.: Super-resolution of face images using kernel PCA-based prior. IEEE Trans. Multimed. **9**, 888–892 (2007)
61. Kim, K.I., Franz, M., Scholkopf, B.: Kernel hebbian algorithm for single-frame super-resolution. In: Proceedings of the ECCV Workshop on Statistical Learning in Computer Vision, pp. 135–149. Prague, Czech Republic (2004)
62. Tipping, M.E., Bishop, C.M. : Bayesian image super-resolution. In: Becker, S., Thrun, S., Obermeyer, K. (eds.) Advances in Neural Information Processing Systems, MIT Press, Cambridge (2002)
63. Pickup, L.C., Capel, D., Roberts, S.J., Zisserman, A. : Bayesian image super-resolution, continued. In: Scholkopf, B., Platt, J., Hoffman, T. (eds.) Advances in Neural Information Processing Systems, pp. 1089–1096. MIT Press, Cambridge (2006)
64. He, H., Kondi, L.P.: Resolution enhancement of video sequences with simultaneous estimation of the regularization parameter. J. Electron. Imaging **13**, 586–596 (2004)
65. Zibetti, M., Bazan, F., Mayer, J.: Determining the regularization parameters for super-resolution problems. Signal Process. **88**, 2890–2901 (2008)
66. Zibettia, M.V.W., Baznb, F.S.V., Mayer, J.: Estimation of the parameters in regularized simultaneous super-resolution. Pattern Recognit. Lett. **32**, 69–78 (2011)
67. Hertzmann, A., Jacobs, C.E., Oliver, N., Curless, B., Salesin, D.H.: Image analogies. In: Proceedings of the SIGGRAPH, pp. 327–340. Los Angeles (2001)
68. Chang, H., Yeung, D.Y., Xiong, Y.: Super-resolution through neighbor embedding. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, pp. 275–282. Washington, D.C. (2004)
69. Freeman, W.T., Pasztor, E.C., Carmichael, O.T.: Learning low-level vision. Int. J. Comput. Vis. **40**, 25–47 (2000)
70. Pickup, L.C., Roberts, S.J., Zisserman, A. : A sampled texture prior for image super-resolution. In: Thrun, S., Saul, L., Scholkopf, B. (eds.) Advances in Neural Information Processing Systems, MIT Press, Cambridge (2003)
71. Datsenko, D., Elad, M.: Example-based single document image super-resolution: a global MAP approach with outlier rejection. Multidimens. Syst. Signal Process. **18**, 103–121 (2007)
72. Wang, J., Zhu, S., Gong, Y.: Resolution enhancement based on learning the sparse association of image patches. Pattern Recognit. Lett. **31**, 1–10 (2010)
73. Kim, K.I., Kwon, Y.: Single-image super-resolution using sparse regression and natural image prior. IEEE Trans. Pattern Anal. Mach. Intell. **32**, 1127–1133 (2010)
74. Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution via sparse representation. IEEE Trans. Image Process. **19**, 2861–2873 (2010)
75. Rhee, S., Kang, M.G.: Discrete cosine transform based regularized high-resolution image reconstruction algorithm. Opt. Eng. **38**, 1348–1356 (1999)
76. Woods, N.A., Galatsanos, N.P., Katsaggelos, A.K.: Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images. IEEE Trans. Image Process. **15**, 201–213 (2006)
77. Dempster, P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. B **39**(1), 1–38 (1977)
78. Bertero, M., Boccacci, P.: Introduction to Inverse Problems in Imaging. IOP Publishing Ltd, Philadelphia (1998)
79. Li, S.Z.: Markov Random Field Modeling in Computer Vision. Springer, New York (1995)
80. Rue, H.: Gaussian Markov Random Fields: Theory and Applications. Chapman & Hall, Boca Raton (2005)
81. Bose, N.K., Lertrattanapanich, S., Koo, J.: Advances in super-resolution using L-curve. In: Proceedings of the IEEE International Symposium on Circuits and Systems, pp. 433–436. Sydney, Australia (2001)
82. Wu, C.F.J.: On the convergence properties of the EM algorithm. Ann. Stat. **11**, 95–103 (1983)
83. Metropolis, N., Rosenbluth, M.N., Rosenbluth, A.W., Teller, A.H., Teller, E.: Equations of state calculations by fast computing machines. J. Chem. Phys. **21**, 1087–1092 (1953)
84. Schultz, R.R., Li, M., Stevenson, R.L.: Subpixel motion estimation for super-resolution image sequence enhancement. J. Vis. Commun. Image Represent. **9**, 38–50 (1998)
85. Protter, M., Elad, M.: Super resolution with probabilistic motion estimation. IEEE Trans. Image Process. **18**, 1899–1904 (2009)
86. Barreto, D., Alvarez, L.D., Abad, J. : Motion estimation techniques in super-resolution image reconstruction: a performance evaluation. In: Tsvetkov, M., Golev, V., Murtagh, F., Molina, R. (eds.) Virtual Observatory. Plate Content Digitalization, Archive Mining and Image Sequence Processing, pp. 254–268. Sofia, Bulgary (2005)
87. Callico, G., Lopez, S., Sosa, O., Lopez, J.F., Sarmiento, R.: Analysis of fast block matching motion estimation algorithms for video super-resolution systems. IEEE Trans. Consum. Electron. **54**, 1430–1438 (2008)
88. Baker, S., Kanade, T.: Super resolution optical flow. The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Technical Report, Oct (1999)
89. Zhao, W., Sawhney, H.: Is super-resolution with optical flow feasible? In: Proceedings of the European Conference on Computer Vision, pp. 599–613. Copenhagen, Denamrk (2002)
90. Lin, F., Fookes, C., Chandran, V., Sridharan, S.: Investigation into optical flow super-resolution for surveillance applications. In: Proceedings of the APRS Workshop on Digital Image Computing, pp. 73–78 (2005)
91. Le, H.V., Seetharaman, G.: A super-resolution imaging method based on dense subpixel-accurate motion fields. J. VLSI Signal Process. **42**, 79–89 (2006)
92. Fransens, R., Strecha, C., Gool, L.V.: Optical flow based super-resolution: a probabilistic approach. Comput. Vis. Image Underst. **106**, 106–115 (2007)
93. Suresh, K.V., Kumar, G.M., Rajagopalan, A.N.: Superresolution of license plates in real traffic videos. IEEE Trans. Intell. Trans. Syst. **8**, 321–331 (2007)
94. Narayanan, B., Hardie, R.C., Barner, K.E., Shao, M.: A computationally efficient super-resolution algorithm for video processing using partition filters. IEEE Trans. Circuits Syst. Video Technol. **17**, 621–634 (2007)
95. Ng, M.K., Shen, H.-F., Zhang, L.-P., Lam, E.: A total variation regularization based super-resolution reconstruction algorithm for digital video. EURASIP J. Appl. Signal Process. **2007**(74585), 1–16 (2007)
96. Patanavijit, V., Jitapunkul, S.: A Lorentzian stochastic estimation for a robust iterative multiframe super-resolution reconstruction with Lorentzian-Tikhonov regularization. EURASIP J. Appl. Signal Process. **2007**(34821), 1–21 (2007)
97. Borman, S., Stevenson, R.L.: Simultaneous multi-frame MAP super-resolution video enhancement using spatio-temporal priors. In: Proceedings of the IEEE International Conference on Image Processing, pp. 469–473. Kobe, Japan (1999)
98. Zibetti, M., Mayer, J.: A robust and computationally efficient simultaneous super-resolution scheme for image

sequences. IEEE Trans. Circuits Syst. Video Technol. **17**, 1288–1300 (2007)
99. Alvarez, L.D., Molina, R., Katsaggelos, A.K.: Multi-channel reconstruction of video sequences from low-resolution and compressed observations. In: The 8th Iberoamerican Congress on Pattern Recognition, pp. 46–53. Havana, Cuba (2003)
100. Elad, M., Feuer, A.: Superresolution restoration of an image sequence: adaptive filtering approach. IEEE Trans. Image Process. **8**, 387–395 (1999)
101. Elad, M., Feuer, A.: Super-resolution reconstruction of image sequences. IEEE Trans. Pattern Anal. Mach. Intell. **21**, 817–834 (1999)
102. Farsiu, S., Elad, M., Milanfar, P.: Video-to-video dynamic super-resolution for grayscale and color sequences. EURASIP J. Appl. Signal Process. **2006**(61859), 1–15 (2006)
103. Costa, G.H., Bermudez, J.C.M.: Statistical analysis of the LMS algorithm applied to super-resolution image reconstruction. IEEE Trans. Signal Process. **55**, 2084–2095 (2007)
104. Tian, J., Ma, K.-K.: A state-space super-resolution approach for video reconstruction. Signal Image Video Process. **3**, 217–240 (2009)
105. Bishop, C.M., Blake, A., Marthi, B.: Super-resolution enhancement of video. In: Proceedings of the International Conference on Artificial Intelligence and Statistics. Key West, FL (2003)
106. Dedeoglu, G., Kanade, T., August, J.: High-zoom video hallucination by exploiting spatio-temporal regularities. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, pp. 151–158. Washington, DC (2004)
107. Kong, D., Han, M., Xu, W., Tao, H., Gong, Y.: Video superresolution with scene-specific priors. In: Proceedings of the British Machine Vision Association. Edinburgh (2006)
108. Kimura, K., Nagai, T., Nagayoshi, H., Sako, H.: Simultaneous estimation of super-resolved image and 3D information using multiple stereo-pair images. In: Proceedings of the IEEE International Conference on Image Processing, pp. 417–420. (2007)
109. Bhavsar, A.V., Rajagopalan, A.N.: Resolution enhancement for binocular stereo. In: Proceedings of the International Conference on Pattern Recognition, pp. 1–4. (2008)
110. Bhavsar, A.V., Rajagopalan, A.N.: Resolution enhancement in multi-image stereo. IEEE Trans. Pattern Anal. Mach. Intell. **99**, 1721–1728 (2010)
111. Robertson, M.A., Stevenson, R.L.: Temporal resolution enhancement in compressed video sequences. EURASIP J. Appl. Signal Process. **2001**, 230–238 (2001)
112. Shechtman, E., Caspi, Y., Irani, M.: Space-time super-resolution. IEEE Trans. Pattern Anal. Mach. Intell. **27**, 531–545 (2005)
113. Watanabe, K., Iwai, Y., Nagahara, H., Yachnida, M., Suzuki, T.: Video synthesis with high spatio-temporal resolution using motion compensation and image fusion in wavelet domain. In: Proceedings of the Asian Conference on Computer Vision, pp. 480–489. Hyderabad, India (2006)
114. Maor, N., Feuer, A., Goodwin, G.C.: Compression at the source for digital camcorders. EURASIP J. Image Video Process. **2007**(41516), 1–11 (2007)
115. Ishikawa, H., Geiger, D.: Rethinking the prior model for stereo. In: Proceedings of the European Conference on Computer Vision, pp. 526–537. (2006)
116. Szeliski, R., Veksler, O., Agarwala, A., Rother, C.: A comparative study of energy minimization methods for Markov random fields. In: Proceedings of the European Conference on Computer Vision, pp. 16–29. (2006)
117. Krishnamurthy, R., Woods, J., Moulin, P.: Frame interpolation and bidirectional prediction of video using compactly encoded optical flow fields and label fields. IEEE Trans. Circuits Syst. Video Technol. **9**, 713–726 (1999)
118. Zhang, Y., Zhao, D., Ji, X., Wang, R., Gao, W.: A spatio-temporal auto regressive model for frame rate upconversion. IEEE Trans. Circuits Syst. Video Technol. **19**, 1289–1301 (2009)
119. Zhang, Y., Zhao, D., Ma, S., Wang, R., Gao, W.: A motion-aligned auto-regressive model for frame rate up conversion. IEEE Trans. Image Process. **19**, 1248–1258 (2010)