
P B Chopade
PhD Research Scholar
Sinhgad College of Engineering
Pune-411041, India

P M Patil
Joint Director
KJ's Educational Institute
Pune-411048, India

ABSTRACT
Super-resolution is the process of combining one or more low–resolution images to obtain a high–resolution image. It has been a very interesting topic for the research over the last few years. It is used for practical applications in more realistic problems faced in different areas, which range over satellite and aerial imaging of biomedical image processing as well as in daily routine life like different biometrics in offices and industries. In this survey article, more focus is given on basic algorithms and their classification based methodology used to implement it. Due to its vast scope of applications researchers are developing a novel superresolution algorithm for a specific intention based on single and multiframe image resolution. The proposed comprehensive survey gives an overview of most of published works based on its performance analysis. In this survey, the basic concepts of the algorithms are explained and then their performance analyses through which each of these methods have developed are mentioned in detail. Furthermore, different issues in supersolution algorithms for single and multiframe such as models and registration algorithms, optimization of the deployment of methods, improvement in quality of image.

Keywords
Single and multi frame superresolution, Image registration, wavelet transform, Assessment of SR algorithms.

1. INTRODUCTION
Super-resolution (SR) is a process of combining one or more low resolution images to obtain high resolution image. It has been used in for many various applications such as Satellite and Aerial imaging, Medical image processing, automated mosaicking, sign and number plate recognition and ultrasound imaging.

Super-resolution is an algorithm that provides details finer than a sampling grid of a given imaging device by increasing the number of pixels per unit area in an image. We need to understand about the hardware limitation of increasing the number of pixels per unit area. Resolution can be increased in hardware approach by i) decreasing the pixel size and ii) increasing the sensor size. This hardware approach has a limitation like decreasing pixel size, will increase shot noise and increasing the sensor size will increase the capacitance of the system. The hardware approach is highly costly for large scale imaging. Therefore, algorithm-based approaches are preferred to the hardware based approach.

Classification of super-resolution is based on different parameters. These parameters include domain deployed, the number of low resolution images involved and different reconstruction method. Super-resolution algorithms are classified based on single and multi frames used for as an input. The single frame based algorithms mostly applied some learning methods and hallucinate the missing data of the high-resolution images establishing relationship between low resolution and high resolution images from a standard database. In multi frame based super-resolution algorithms, we establish relationship between targeted high resolution image and the low resolution images having some relative geometric or photometric motions from the targeted high resolution image. These algorithms used the differences between the low resolution images and hence are referred to as reconstruction based super-resolution algorithms. Before going into the details of the SR algorithms, the most common imaging modes are described in the next section.

2. IMAGING MODELS
The imaging model of reconstruction based super-resolution algorithms gives a procedure by which the observed images have been obtained. In easiest way, this process can be modelled as

\[ g(m,n)=\frac{1}{q^2} \sum_{x=qn}^{(n+1)q-1} \sum_{y=qn}^{(m+1)q-1} f(x,y) \]  

Where \( g \) is a low resolution image, \( f \) is a high resolution image, \( q \) is decimation factor. The imaging model in Eq. (1) states that the low resolution image has been obtained by averaging the high resolution pixel values over neighborhood of \( q^2 \) pixels. This model is very basic and is very easy to understand, but realistic model can be obtained by adding different parameters like blurring, warping and noise. Adding these parameters in Eq.(1), results as

\[ g(m,n)=d(h(w(f(x,y))))+\eta(m,n) \]  

Where \( w \) is a warping function, \( h \) is a blurring function, \( d \) is decimation operator and \( \eta \) is an additive noise.

If the number of low resolution images is more than one, the imaging model of Eq.(2) becomes:

\[ g_k(m,n)=d(h_k(w_k(f(x,y))))+\eta(m,n) \]  

Imaging model consist of:
Wrapping: It stands for any transformations between observed low resolution image and original high resolution image.

Blurring: The blurring function models any blurring effect that is imposed on the LR observed image, for example by the optical system or by atmospheric effects.

Sampling: It provides down or up sampling of image.

3. FREQUENCY DOMAIN
Super-resolution algorithms of this domain, transform input low resolution images to the frequency domain and then formulate the high resolution image in the same domain.
Algorithms from this domain can be further classified as into two groups: Fourier-Transform based and Wavelet-Transform based methods, which are briefed in the following subsections.

3.1 Fourier Transform
Gerchberg [1] (1974) and then Santis and Gori [2] introduced the first SR algorithms. They had used iterative methods in the frequency domain, based on the Fourier transform [3], which could extend the spectrum of a given signal beyond its diffraction limit and therefore increase its resolution. Tsai and Huang’s system [4] (1984) was the first multiple-image SR algorithm in the frequency domain. In these algorithms, different translated (shifted) images, $g_k$ of same area are used of continuous scene. Here we can defined $g_k(m,n)=f(x,y), \text{where } x=m+\Delta m_k \text{and } y=n+\Delta n_k$. These shifts or translations between the low resolution images are considered by shifting property of the Fourier transform:

$$F_{g_k}(m,n) = e^{i2\pi(\Delta m_k m+\Delta n_k n)}F_f(m,n) \quad \cdots \quad 5$$

Where $F_{g_k}$ and $F_f$ are the Fourier transforms of the $k$th low resolution image and the high resolution scene respectively.

Low resolution images are the discrete version of continuous scene. Therefore

$$g_k(m,n) = f(mTm + mk, nTn + nK) \quad \cdots \quad 6$$

Where $Tm$ and $Tn$ are the sampling periods of low resolution image along columns and rows. If we Fourier transform of the above equation, we will get a discrete Fourier Transform $G_k$, and its continuous Fourier transform, $F_g$ are related such as

$$G_k(m,n) = \frac{1}{TmTn} \sum_{p=1}^{\infty} \sum_{n=0}^{\infty} F_{g_k}\left(\frac{m}{MTm} + \frac{p}{Tm}, \frac{n}{NTn} + \frac{n}{Tn}\right) \quad \cdots \quad 7$$

Where $M$ and $N$ are the maximum values of the dimensions of low resolution images $m,n$, respectively. Combining eq(5) into eq(7) and writing the results in matrix forms, gives us

$$G = \Phi F_f \quad \cdots \quad 8$$

in which $\Phi$ relates the discrete Fourier Transform of the LR images $G$ to the continuous Fourier Transform of the HR scene$F_f$. SR here is therefore reduced to finding $F_f$ in Eq. (8) which is usually solved by a Least Squares (LS) algorithm.

3.2 Wavelet Transform
The wavelet transform as an alternative for the Fourier transform has been widely used in frequency-domain based superresolution algorithms. Wavelet Transform is used to decompose the input image into structurally correlated sub-images. This gives us the self-similarities between local neighboring regions [5], [6]. For example, in [6] the input image is first decomposed into subbands. Then, the input image and the high-frequency subbands are both interpolated. The results of a Stationary Wavelet Transform of the high-frequency subbands are used to improve the interpolated subbands. Then, the super-resolved image is generated by combining all of these subbands using an inverse Discrete Wavelet Transform (DWT). Similar methods based on the DWT have been developed for SR in [7], [8], [9], [10], [11]. In [12], [13], [14], [15] the results obtained by DWT are used as a regularization term in Maximum a Posteriori (MAP) formulation of the problem. In [16], [17] they have been used with Compressive Sensing (CS) methods and in [18] within a PCA-based face hallucination algorithm. Wavelet based methods may have difficulties in efficient implementation of degraded convolution filters, while they can be done efficiently using the Fourier transform. Therefore, these two transforms have sometimes been combined together into the Fourier-Wavelet Regularized Deconvolution [19]. In addition to the above mentioned methods in the frequency domain, some other SR algorithms of this domain have borrowed the methods that have been usually used in the spatial domain; among them are: [20], [21], [22] which have used a Maximum Likelihood (ML) method , [23], [24] which have used a regularized ML method, [25], [26] which have used a MAP method, and [27], [28] which have implemented a Projection Onto Convex Set (POCS) method.

Chopade and Patil[29] have proposed a wavelet based super resolution technique which is discussed here in brief. Existing algorithms of super resolution is depends on the interpolation techniques. The interpolation technique has its own limitation that it generate image contained blurred effect in the edges. To preserve edges of image multi resolution analysis (MRA) is useful technique.In discrete wavelet transform (DWT), MRA technique used properly. The power of DWT is to give good time resolution for high frequencies and good frequency resolution for low frequencies. But in DWT, down sampling is present due to which the information loss in the respective sub-bands is bound to happen. Hence, in order to mitigate this effect, the three detail sub-bands without down-sampling factor (called as stationary wavelet transform) has been introduced in the interpolated detail sub-bands to obtain super-resolved image [30]. However, such type of wavelet based super-resolution algorithm introduced the spatial domain noise. This type of noise can be reduced with the help of wavelet-coefficient based thresholding technique. Thus, Chappedi and Bose [31] investigated the effect of soft-thresholding level on the reconstructed image quality by the use of existing wavelet based on the two-step lifting step. However, these algorithms required the irrational coefficients based wavelet filters that lead to increase the hardware computational complexity. In order to reduce this issue, a wavelet filter based on dyadic-integer coefficient is proposed to reduce the hardware complexity in the field of super-resolution.

First, the two low-resolution images $I_1$ and $I_2$ are obtained from the original image based on the half-pixel shift in both the directions (row and column-wise). Both of these frames are rotated by an angle of 45° using quincunx sampling so as to determine the measure of similarity of the pixel values between two frames. These rotated images are interpolated to denote the missing pixels in super-resolution image. The proposed specific class of DICWF is applied to these two rotated high-resolution images and the interpolation is performed at one scale. Next, the reconstruction DICWFs are applied to these interpolated rotated images and combined into single image. This single image is the post-rotated back to its original orientation so as to obtain the super-resolved image. The steps of the super-resolution algorithm based on the proposed class DICWFs are mentioned in their article.
4. SPATIAL DOMAIN
Depending on the number of available low resolution images, superresolution algorithms can be classified into two groups: single image based and multiple image based algorithms. The algorithms included in these groups are explained in the following subsections.

4.1 Multiple Image based SR Algorithms
Multiple image (or classical) SR algorithms are mostly reconstruction-based algorithms, i.e., they try to figure out the aliasing artifacts that are present in the observed low resolution images due to under-sampling process. A few algorithms are studied in the following subsections.

Following are the few algorithms based on multiple image based super resolution algorithms.

- Iterative Back Projection
- Iterative Adaptive Filtering
- Directs Methods
- Projection onto Convex Sets
- Maximum Likelihood
- Maximum A Posteriori

4.2 Iterative Back Projection
Iterative Back Projection (IBP) methods were among the first methods developed for spatial-based super resolution. We defined the imaging model given in Eq. (3). In iterative back projection, we try to minimize degradation factor d. To do so, usually an initial guess for the HR targeted image is generated and then it is refined. Such a guess can be obtained by registering the LR images over an HR grid and then averaging them [32], [33], [34], [35]. To refine this initial guess t, the imaging model given in Eq. (3) is used to simulate the set of the available LR observations, gk where k = 1:k. The error between the simulated LR images and the observed ones, which is computed by

\[
\sqrt{\frac{1}{k} \sum_{k=1}^{h} (g_k - g_k^{(t)})^2}
\]

(Here t is the number of iterations) and then these error is back-projected to the coordinates of the HR image to improve the initial guess [36]. This process is either repeated for a specific number of iterations or until no further improvement can be achieved in high resolution image. To do so, usually the following Richardson iteration is used in this group of algorithms:

\[
f^{(t+1)} = f^{(t)}(x, y) + 1/k \sum_{k=1}^{w-1} ((g_k - g_k^{(t)}) d) * h
\]

In which t is an iteration parameter, w−1 is the inverse of the warping kernel, d is decimation operator, h is a debluring kernel. The block diagram of example algorithm based on iterative back projection is given in fig. 2.

4.3 Maximum Likelihood
If we considered the noise term in the imaging model given in Eq. (5) is a Gaussian noise with zero mean and variance σ as well as if estimation of the super-resolved image is given as \(\hat{f}\), the total probability of an low resolution image gk is [37], [38], [39], [40], [41]:

\[
P(gk : f) = \prod_{m,n} 1/2\sqrt{\pi} \exp(-\frac{(g_k - g_k)^2}{2\sigma^2})
\]

Maximum likelihood concepts used to get high resolution image from low resolution images, But it is sensitive to small disturbances, such as noise or errors in the estimation of the imaging parameters. It is also gives distorted result if the number of LR images is less than the square of the improvement factor. This means that there might not be a unique solution. To deal with these problems, there is need of some additional information to constrain the solution. Such information can be collected from a priori knowledge of the desired image. The apriori term can prefer a specific solution over other solutions when the solution is not unique. The involvement of a priori knowledge can convert the ML problem to a MAP problem. If the number of LR images over-determines the super-resolved images, the results of ML and MAP are the same and since the computation of ML is easier, it is preferred over MAP. However, if the number of the LR images is insufficient for the determination of the super-resolved image, the involvement of a priori knowledge plays an important role and MAP outperforms ML [42]

5. SINGLE IMAGE BASED SR ALGORITHMS
In the process of sub-sampling or decimation of an image, the high frequency information may get lost. In ML and MAP algorithms the generic smoothness priors can regularize the solution but cannot help recover the lost frequencies, especially for high improvement factors [43]. In single image based super resolution algorithms, these generic priors are replaced by more meaningful and class wise priors like. This is because images from the same class have similar statistics and the accuracy of multiple-image based SR algorithms is highly dependent on the estimation accuracy of the motions between the LR observations, which gets more unstable in real world applications where different objects in the same scene can have different and complex motions. In situations like these, single image based SR algorithms may work better [44]. There algorithms are either reconstruction based (similar to multiple image based algorithms) or learning based. These are described in the following two subsections.

- Learning based Single Image super resolution Algorithms
- Reconstruction based Single Image SR algorithms

5.1 Learning based Single Image SR Algorithms
These algorithms are learning based or Hallucination algorithms which were first introduced in [45] 1985 in which a neural network was used to improve the resolution of fingerprint images. These algorithms contain a training step in which the relationship between some HR examples and their LR counterparts are learned. This learned knowledge is then incorporated into the a priori term of the reconstruction. The training database of learning based SR algorithms needs to have a proper generalization capability [41]. To measure this, the two factors of sufficiency and predictability have been introduced in [41]. Using a larger database does not necessarily generate better results, on the contrary, a larger number of irrelevant examples not only increases the computational time of searching for the best matches, but it also disturbs this search [46]. To deal with this, in [46] it is suggested to use a content-based classification of image patches during the training.
6. ASSESSMENT OF SR ALGORITHMS

For the assessment of the results of super resolution algorithms both subjective and objective methods have been used. In the subjective method, usually human observers assess the quality of the produced image. In few articles the results of the SR algorithms are presented in the form of Mean Opinion Scores (MOP) and Variance Opinion Score (VOP) [47]. In objective methods, the results of SR algorithms are usually compared against the ground truth using measures like MSE and (Peak) Signal to Noise Ratio (PSNR).

MSE is define as:

\[
MSE = \frac{\sum_{k=0}^{N-1} \sum_{M-1}^{M-1} (\hat{f}(x, y) - f(x, y))^2}{\sum_{k=0}^{N-1} \sum_{M-1}^{M-1} (f(x, y))^2}
\]

The value of MSE is smaller, the closer the result is to the ground truth.

PSNR is defined:

\[
PSNR = 20 \log_{10} \frac{255}{\sqrt{MSE}}
\]

where MSE is the square root of the MSE. Though these two measures have been very often used by SR researchers, they do not represent the Human Visual System (HVS) very well [48]. Therefore, other measures, such as the Correlation Coefficient (CC) and Structural Similarity measure (SSIM) have also been involved in SR algorithms [49], [50]. CC is defined by:

\[
CC = \frac{\sum_{k=0}^{N-1} \sum_{M-1}^{M-1} d_1 d_2}{\sum_{k=0}^{N-1} \sum_{M-1}^{M-1} d_1^2 \sum_{k=0}^{N-1} \sum_{M-1}^{M-1} d_2^2}
\]

Where,

\[
d_1 = f(x, y) - \mu_f \& d_2 = \hat{f}(x, y) - \mu_f
\]

\[
\mu_f \text{ and } \mu_f \text{ are the mean of } f \text{ and } \hat{f} \text{, respectively. The maximum value of CC is one which means a perfect match between the reconstructed image and the ground truth. The other measure, SSIM, is defined by:}
\]

\[
SSIM = \frac{(2\mu_f \mu_\hat{f} + C_1)(2\sigma_{f,\hat{f}} + C_2)}{(\mu_f^2 + \mu_\hat{f}^2 + C_1)(\sigma_f^2 + \sigma_\hat{f}^2 + C_2)}
\]

where C1 and C2 are constants, \( \sigma_f \) and \( \sigma_\hat{f} \) are the standard deviations of the associated images, and \( \sigma_{f,\hat{f}} \) is defined by:

\[
\sigma_{f,\hat{f}} = \frac{1}{M \times N-1} \sum_{k=0}^{N-1} \sum_{m=0}^{M-1} d_1 d_2
\]

To use SSIM, the image is usually first divided into sub-images, and then Eq. (67) is applied to every subimage and the mean of all the values is used as the actual measure between the super-resolved image and the ground truth. A mean value of SSIM close to unity means a perfect match between the two images. It is discussed in that SSIM favors blur over texture misalignment, and therefore may favor algorithms which do not provide enough texture details. Beside these assessment measures, some authors have shown that feeding some real world applications by images that are produced by some SR algorithms improves the performance of those applications to some degree.

7. CONCLUSION

This survey paper reviews most of the papers published on the topic of super-resolution and proposes a broad classification for these works. Besides giving the details of most of the methods, we mention the basic ideas of the methods when they’ve been available in the reviewed papers. Furthermore, it highlights the most common ways for dealing with problems like the number of LR images, the assessment of the algorithms developed. Comparing the frequency domain methods against the spatial domain methods, the former are more interesting from the theoretical point of view but have many problems when applied to real world scenarios, e.g., they mostly have problems in the proper modeling of the motion in real world applications. Spatial domain methods have better evolved for coping with such problems.

Among these methods, the single-image based methods are more application-dependent while the multiple image based methods have been applied to more general applications. The multiple-image based methods are generally composed of two different steps: motion estimation, then fusion. These had and still have limited success because of their lack of robustness to motion error. This survey papers definitely will helpful to beginners who are interested to research in super resolution domain.

8. ACKNOWLEDGMENTS

The author P B Chopade is working as an Assistant Professor at MES’s College of Engineering, Pune, India. Also we would like to thank the institute.

---

Fig. 1 The imaging model used in most super-resolution algorithms [51]
9. REFERENCES


