

# Review of Image Interpolation and Super-resolution

Wan-Chi Siu and Kwok-Wai Hung

The Hong Kong Polytechnic University, Hong Kong, China

E-mail: enwcsiu@polyu.edu.hk

E-mail: kwokwai.hung@connect.polyu.hk

**Abstract**— Image/video interpolations and super-resolution are topics of great interest. Their applications include HDTV, image coding, image resizing, image manipulation, face recognition and surveillance. The objective is to increase the resolution of an image/video through upsampling, deblurring, denoising, etc. This paper reviews the development of various approaches on image interpolation and super-resolution theory for image/video enlargement in multimedia applications. Some basic formulations will be derived such that readers can make use of them to design their own, practical and efficient interpolation algorithms. New results, such as hole filling using nonlocal means for 3D video synthesis and fast interpolation using a simplified image model will be introduced. New directions and trends will also be discussed at the end of the paper.

## I. INTRODUCTION

Image interpolation and super-resolution are topics of great interest. The aim is to improve the resolution of a low-resolution (LR) image/video to obtain a high-resolution (HR) one which is able to preserve the characteristics of natural images/videos. The major difference between interpolation and super-resolution is that interpolation only involves upsampling the low-resolution image, which is often assumed to be aliased due to direct down-sampling. The interpolation algorithms often exploit this aliasing property and perform de-aliasing of the LR image during the upsampling process. As a result, the high-frequency components of the upsampled HR image can be better recovered for preserving the characteristics of natural images [1]. However, natural images do not usually observe severe aliases, such that the upsampled HR image does not usually recover sufficiently the high-frequency components which lead to a blurry image.

Besides the blurry effect, noises due to CCD's limitations and transmission errors, etc, have to be handled. Super-resolution aims to address all these undesirable effects, including the resolution degradation, blur and sometimes noise effects. Hence, super-resolution usually involves three major processes which are upsampling (interpolation), deblurring and denoising.

The applications of image interpolation and super-resolution are very wide, including HDTV, image/video coding, image/video resizing, image manipulation, face recognition, view synthesis and surveillance. Due to many reasons, such as the cost of camera, insufficient bandwidth, storage limitation and limited computational power, the resolution of an image/video is always constrained. One intuitive example is the transmission of low-resolution content over the internet due to limited network bandwidth.

Image interpolation and super-resolution algorithms play the role to produce a high quality and high resolution image/video from the observed low quality and low resolution image/video.

The rest of the organization of this paper is as follows. Section II describes the formulation and theory of major classes of the image interpolation algorithms. Section III classifies super-resolution algorithms and briefly explains one major class using the FIR Wiener filter. In both sections, some experimental results, including the execution times, are given. Furthermore, some future trends are included in each of the image interpolation or super-resolution section, and a conclusion is given at the end of the paper.

## II. IMAGE INTERPOLATION

### A. Polynomial-based interpolation

Image interpolation aims to produce a high-resolution image by upsampling the low-resolution image. As explained in the introduction section, interpolation algorithms often assume that the observed LR image is a direct downsampled version of the HR image. Hence, the de-aliasing ability during the upsampling process is important, i.e. the recovery of the high-frequency signal from the aliased low-frequency signal [1].

For real-time applications, conventional polynomial-based interpolation methods such as bilinear and bicubic interpolation are often used due to their computational simplicity [2-5]. For example, the bicubic convolution interpolator requires only several arithmetic operations per pixel, such that real-time processing can easily be achieved [4]. The basic idea is to model the image signal by a low-order polynomial function (using some observed signals). However, polynomial functions are not good at modeling the signal's discontinuities (e.g. edges). Hence, the conventional polynomial-based interpolation methods often produce annoying artifacts such as aliasing, blur, halo, etc. around the edges. To resolve this problem, some adaptive polynomial-based and step function-based interpolation methods were proposed [6-9].

### B. Edge-directed interpolation

#### B1. Explicit interpolation

Since edges are visually attractive to the human perceptual system, some edge-directed interpolation methods have been developed to address the edge reconstructions [10-33]. In fact, the adaptive polynomial-based methods can be regarded as edge-directed methods as well. The basic idea of edge-

directed methods is to preserve the edge sharpness during the upsampling process. The intuitive way is to explicitly estimate the edge orientation and then interpolate along the edge orientation [10-13]. For achieving low computation, some methods further quantize the edge orientations [12-13]. However, the interpolation quality of this intuitive way is constrained by the estimation accuracy of the edge orientation. Since, the edges of natural images are often blurred, blocky, aliased and noisy, the estimation accuracy of edge orientations is usually unstable. The interpolation quality of these methods can be improved by weighting the edge orientations, as described in the next section.

### B2. Fusion of edge orientations

One major improvement to the explicit methods [10-13] is to adaptively fuse the results of several estimates of different edge orientations [14-18, 32-33]. The fusion process is usually computationally inexpensive. In [15-16, 32-33], two directionally interpolated results are fused using the linear minimum mean squares error estimation (LMMSE). The hidden markov random field (HMRF) was used to fuse two diagonal interpolated results and the bicubic interpolated results together [18]. Compared with the LMMSE methods [15-16, 32-33], the HMRF method also considers the consistency of fusion results by making use of the transition of states. Due to the advances in graphic processing unit (GPU), the GPU is often able to assist the directional image interpolation to achieve real-time upsampling [32-33].

### B3. New edge-directed interpolation (NEDI)

New edge-directed interpolation (NEDI) uses the FIR Wiener filter, equivalently, the linear minimum mean squares error estimator [19] for linear prediction. Let us briefly describe the formulation of NEDI as an introduction. The 4-order linear estimation model is given by

$$Y_i = \sum_{k=1}^4 A_k \cdot Y_{i\Delta k} + \varepsilon_i, \quad \text{for } i=1,2,\dots,P \quad (2.1)$$

where  $\varepsilon_i$  is the estimation error,  $P$  is the number of data point samples,  $A_k$  is the model parameters and each available data point sample ( $Y_i$ ) has four neighboring data points,  $Y_{i\Delta k}$ . Figure 1a shows the spatial positions of  $Y_{i\Delta k}$  and  $Y_i$ . The model parameters  $A_k$  can be found by using the least squares estimation:

$$\{\hat{A}_k\} = \arg \min_{\{\hat{A}_k\}} \sum_{i=1}^P \left[ Y_i - \sum_{k=1}^4 A_k \cdot Y_{i\Delta k} \right]^2 \quad (2.2)$$

where the matrix form of (2.2) is given by

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A}} \|\mathbf{Y} - \mathbf{Y}_A \mathbf{A}\|_2^2 \quad (2.3)$$

and the matrices are defined as

$$\mathbf{Y} = \{Y_i\}^T, \mathbf{Y}_A = \{Y_{i\Delta k}\}^T, \mathbf{A} = \{A_k\}^T \quad (2.4)$$

The sizes of matrices  $\mathbf{Y}$ ,  $\mathbf{Y}_A$  and  $\mathbf{A}$  are  $P \times 1$ ,  $P \times 4$  and  $4 \times 1$  respectively. The close form solution of (2.3) is given as

$$\hat{\mathbf{A}} = (\mathbf{Y}_A^T \mathbf{Y}_A)^{-1} \mathbf{Y}_A^T \mathbf{Y} \quad (2.5)$$

where  $\hat{\mathbf{A}}$  is called the ordinary least squares (OLS) estimator. Due to ‘‘geometry duality’’ [20], the missing data point  $X$  can be interpolated by its four neighboring data points,  $X_k$ , as follows:

$$X = \sum_{k=1}^4 \hat{A}_k \cdot X_k \quad (2.6)$$

For the missing data point between two available LR data points (e.g. the missing data point between  $X_1$  and  $X_2$  in Figure 1b), its value is obtained by rotating the spatial positions of the neighbors and missing data points by 45 degree with a scaling factor of  $1/\sqrt{2}$ . More details can be found in [19]. Note that the linear interpolation method described in this paper can be applied to high activity areas only, which can be identified by using a local variance larger than 8, as used in NEDI. The number of samples is  $P=64$ , i.e. using an  $8 \times 8$  window.

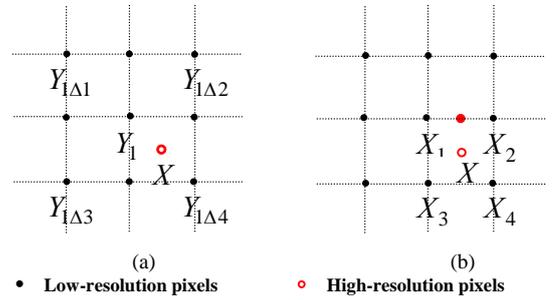


Fig. 1 Graphical illustration of spatial positions of LR and HR pixels.

NEDI uses the LR image to estimate the HR covariances for the FIR Wiener filter, as shown in (2.5) and (2.6). There are many algorithms which are based on the idea of NEDI [21-30]. In [21], a six-order filter design was proposed to avoid error accumulation in the original two steps design. In [22], an eight-order filter was also proposed to include the vertical-horizontal correlation. The filter weights of the FIR Wiener filter were regularized for stability [23]. In [31], instead of using the LR image to estimate the HR covariances; the relevant HR image pairs were searched from an offline dictionary to estimate the HR covariances. Let us clarify this filtering process with the following over-simplified example.

**Example:** A data sequence is given by  $\{Y(0), Y(1), Y(2), Y(3), Y(4), Y(5), Y(6), Y(7), Y(8), Y(9)\}$ . We have to find  $Y(10)$  using the FIR Wiener filter. Let also the length of the filter be 2, i.e.  $N=2$ . Hence the observed sample data point  $\mathbf{Y}$  and their corresponding interpolating points,  $\mathbf{Y}_A$ , are as shown below

$$\mathbf{Y} = [Y(2) \quad Y(3) \quad Y(4) \quad Y(5) \quad Y(6) \quad Y(7) \quad Y(8) \quad Y(9)]^T$$

$$\mathbf{Y}_A = \begin{bmatrix} Y(0) & Y(1) & Y(2) & Y(3) & Y(4) & Y(5) & Y(6) & Y(7) \\ Y(1) & Y(2) & Y(3) & Y(4) & Y(5) & Y(6) & Y(7) & Y(8) \end{bmatrix}^T$$

This means that  $\{Y(9), Y(8), \dots, Y(2)\}$  are used as sample points to look for the required statistics. The interpolation points for  $Y(9)$  are  $Y(7)$  and  $Y(8)$ ; for  $Y(8)$  are  $Y(6)$  and  $Y(7)$ ;  $\dots$ ; for  $Y(2)$  are  $Y(0)$  and  $Y(1)$ . Let us assume further that the cross-correlation matrix  $\mathbf{Y}_A^T \mathbf{Y} = [18 \ 11]^T$  and the

auto-correlation matrix  $\mathbf{Y}_A^T \mathbf{Y}_A = \begin{bmatrix} 13 & 8 \\ 8 & 5 \end{bmatrix}$ . The filter weights can be estimated by equation (2.5), as follows:

$$\mathbf{A} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = (\mathbf{Y}_A^T \mathbf{Y}_A)^{-1} \mathbf{Y}_A^T \mathbf{Y} = \begin{bmatrix} 13 & 8 \\ 8 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 18 \\ 11 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}. \text{ Hence,}$$

the data point  $Y(10)$  can be estimated by (eqn.2.6), i.e.

$$Y(10) = A_1 Y(8) + A_2 Y(9) = 2 Y(8) - Y(9).$$

#### B4. Soft-decision interpolation (SAI)

The soft-decision adaptive interpolation (SAI) [24] was proposed to interpolate a block of pixels at one time using the idea of NEDI. It has the benefit of using the block-based estimation, by constraining the statistical consistency within the block region, which comprises of both observed LR and unobserved HR pixels. Within a local window as shown in Figure 2, the observed LR pixels  $\mathbf{Y}$  in the LR image are used to estimate the HR pixels  $\mathbf{X}$  in the HR image. The sizes of  $\mathbf{X}$  and  $\mathbf{Y}$  in the SAI are  $12 \times 1$  and  $21 \times 1$  respectively. Let us formulate the SAI as a maximum a posteriori (MAP) estimation problem, as follows

$$\begin{aligned} \arg \max_{\mathbf{X}} p(\mathbf{X} / \mathbf{Y}) &= \arg \max_{\mathbf{X}} \log[p(\mathbf{Y} / \mathbf{X})p(\mathbf{X})] \\ &= \arg \min_{\mathbf{X}} \left( \|\mathbf{X} - \mathbf{Y}_X \mathbf{A}\|_2^2 + \|\tilde{\mathbf{Y}} - \mathbf{X}_Y \mathbf{A}\|_2^2 + \lambda \|\tilde{\mathbf{X}} - \mathbf{X}_X \mathbf{B}\|_2^2 \right) \end{aligned} \quad (2.7)$$

where the likelihood and prior are assumed to follow the Gaussian distributions, elements of  $\mathbf{A}$  and  $\mathbf{B}$  are the model parameters (which are estimated using the NEDI method),  $\mathbf{Y}_X$  are diagonal neighbors of  $\mathbf{X}$ ,  $\tilde{\mathbf{Y}}$  are the centermost five elements of  $\mathbf{Y}$  (bounded by dotted cross in Figure 2),  $\mathbf{X}_Y$  are diagonal neighbors of  $\tilde{\mathbf{Y}}$ ,  $\tilde{\mathbf{X}}$  is a  $4 \times 1$  vector representing the centermost four elements of  $\mathbf{X}$  (bounded by dotted square in Figure 2), and  $\mathbf{X}_X$  are horizontal-vertical neighbors of  $\tilde{\mathbf{X}}$ . Hence, all the elements of  $\mathbf{X}$  and  $\mathbf{Y}$  are defined in the posterior.  $\lambda$  is the regularization factor which is recommended to be 0.5 [24]. In (2.7), the SAI uses three auto-regressive models to model the image signals. Such a formulation constrains the statistical consistency within a local region, which is usually locally stationary. Hence, this method is called the soft-decision adaptive interpolation. To solve (2.7), we can use a more compact form of the argument, as follows

$$\arg \min_{\mathbf{X}} (\mathbf{C}\mathbf{X} - \mathbf{D}\mathbf{Y})^T (\mathbf{C}\mathbf{X} - \mathbf{D}\mathbf{Y}) \quad (2.8)$$

which remaps the elements into a very compact form. Matrices  $\mathbf{C}$  and  $\mathbf{D}$  are defined as,  $\mathbf{C} = [\mathbf{I}_{12} \ \mathbf{C}_1 \ \lambda \ \mathbf{C}_2]^T$  and  $\mathbf{D} = [\mathbf{D}_1 \ \mathbf{D}_2 \ \mathbf{0}_{4 \times 21}]^T$  (see reference [24] for details). The argument in (2.8) can be solved by differentiating the cost function with respect to  $\mathbf{X}$  to obtain the following closed-form solution

$$\mathbf{X} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{D}\mathbf{Y} \quad (2.9)$$

#### B5. Robust soft-decision interpolation using weighted least squares (RSAI)

It is well known that the least squares estimation is not robust to outliers, hence the weighted least squares was

proposed to improve the accuracy and robustness of the SAI [25-26]. The robust SAI (RSAI) incorporates the weights to all residuals in the cost function of the SAI, as follows

$$\arg \min_{\mathbf{X}} \left( \left\| \mathbf{W}_1 (\mathbf{X} - \mathbf{Y}_X \mathbf{A}) \right\|_2^2 + \left\| \mathbf{W}_2 (\tilde{\mathbf{Y}} - \mathbf{X}_Y \mathbf{A}) \right\|_2^2 + \lambda \left\| \mathbf{W}_3 (\tilde{\mathbf{X}} - \mathbf{X}_X \mathbf{B}) \right\|_2^2 \right) \quad (2.10)$$

where  $\mathbf{W}_i$  contains the weighting parameters for residuals in  $(\mathbf{X} - \mathbf{Y}_X \mathbf{A})$ ,  $(\tilde{\mathbf{Y}} - \mathbf{X}_Y \mathbf{A})$  and  $(\tilde{\mathbf{X}} - \mathbf{X}_X \mathbf{B})$  respectively. To solve (2.10), equations similar to (2.8)-(2.9) can be used to derive the cost function and the resultant closed-form solution. As verified in [25], there exists severe outliers due to the mismatch of "geometric duality", which is a fundamental assumption of estimating the HR parameters using the LR image. Hence, using the weighted least squares to adaptively weight the residuals for both HR parameters and HR pixels estimation results in a significant improvement of PSNR and SSIM, as shown in Table 1 and Table 2. As a matter of fact that the RSAI is currently one of the best performers for image interpolation.

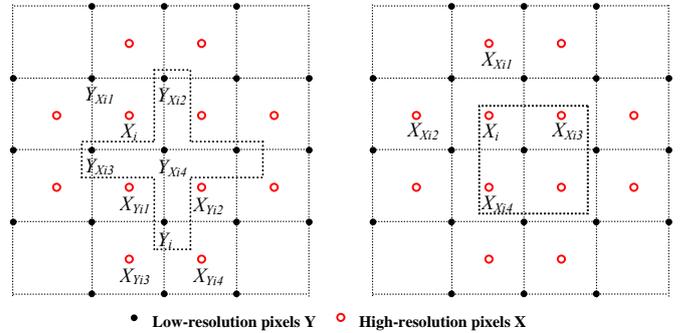


Fig. 2 Graphical illustration of 21 LR and 12 HR pixels within the local window used by SAI. Left figure shows one element of  $X_i \in \mathbf{X}$  and  $Y_i \in \mathbf{Y}$ , and their diagonal neighbors

#### B6. Bilateral soft-decision interpolation for real-time applications (BSAI)

The SAI and RSAI are able to produce the best quality, but the computational cost is high. To largely reduce the computation, the least squares estimation should be avoided or reduced to single parameter estimation. Hence, the bilateral soft-decision interpolation (BSAI) was proposed to use the bilateral filter for replacing the least squares estimation for the model parameters and to reduce the weighted least squares soft-decision estimation to a single parameter estimation [17]. The BSAI uses bilateral filter weights  $A_k$  to replace the least squares parameters and adopts the following cost function (modified from SAI and RSAI) for estimating a HR pixel at one time (instead of a block of HR pixels at one time in the SAI and RSAI)

$$\arg \min_X \left\{ \left( X - \sum_{k=0}^3 A_k X_k \right)^2 + \sum_{k=0}^3 U_k \left( X_k - \sum_{i=0}^3 A_i X_{ki} \right)^2 \right\} \quad (2.11)$$

where  $X_k$ 's are the neighbors of  $X$  (which is the HR pixel to be interpolated),  $X_{ki}$ 's are the neighbors of  $X_k$ . Figure 3 shows the spatial configurations of  $X_k$  and  $X$ , and  $X_{0i}$  and  $X_0$  as examples.

In (2.11), the first term constrains the result to approach the interpolated result using the bilateral filter, and the second term constrains the interpolated result to be continuous with its neighbors (soft-decision estimation). The weight  $U_k$  is defined by the bilateral filter weights (since the continuity property of edge depends on the edge orientation, which is exploited by the bilateral filter weight) added with a constant for stabilization. To solve (2.11), we differentiate the cost function with respect to the variable  $X$  and obtain the following closed-form solution,

$$X = \frac{\sum_{k=0}^3 A_k X_k + \sum_{k=0; l=3-k}^3 U_k A_l \left( X_k - \sum_{i=0; i \neq 3-k}^3 A_i X_{ki} \right)}{1 + U_0 A_3^2 + U_1 A_2^2 + U_2 A_1^2 + U_3 A_0^2} \quad (2.12)$$

which is a very compact and efficient equation. In real-time implementation using C++ codes, BSAI requires 0.062 second to interpolate a 384x256 image to double its size. The major advantage of BSAI is that it is the near-best performer for the image interpolation (as shown in Table 1 to Table 3) but its computational cost is sufficiently low for real-time applications.

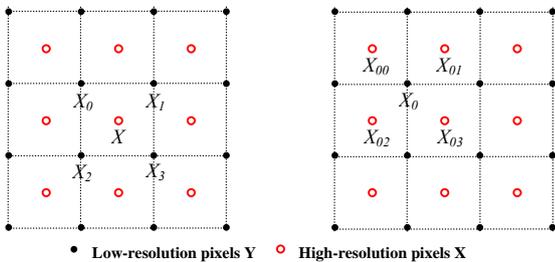


Fig. 3 Spatial configurations of  $X_k X$  (left), and  $X_{0i} X_0$  (right).

### B7. Recent trends of image interpolation applications

Recently, image interpolation does benefit from the development of sparse representation. A family of linear estimators corresponding to different priors are mixed using the sparse representation to give the final estimates for image interpolation [34]. Its performance is on a par with the SAI [24]. Obviously, the sparse representation, which has shown to be successful in compressive sensing, denoising, restoration and super-resolution, etc, is a fruitful research direction of image interpolation.

Another trend of the image interpolation application is the real-time upsampling of a LR video sequence for future very high-definition TVs, such as videos with 4K resolution. The HR video sequence can be obtained by fusing several low-resolution (LR) frames into one high-resolution (HR) frame. For such real-time applications, nonlocal means (weighted sum filter) can be directly applied due to its simplicity in computing the filter weights and its high performance. The nonlocal means can be used with a linear motion model to better estimate the filter weights, which is optimized for the upsampling [35]. After fast upsampling, some simple restoration techniques can be applied to deblur the video. In the future, there should be more of these fast algorithms

which can perform upsampling in real-time, of which the real-time requirement is the major difficulty for their development.

Recently view synthesis attracts much attention in the image processing community. In 3D videos and multi-view synthesis, the major issue is to fill the newly exposed area (hole region). The hole filling in view synthesis resembles the video (multi-frame) interpolation scenario. Hence, the multi-frame interpolation techniques such as nonlocal means can be adopted to fill the hole. Specifically, the nonlocal means can be modified to cope with irregular hole sizes and the extra depth information [36]. Similar to eqn.(2.6), the nonlocal means for hole filling can be defined as

$$x = \sum_{i \in N} w_i x_i \quad (2.13)$$

where  $x$  is a pixel's intensity inside the hole region,  $x_i$  are neighbors of  $x$ ,  $N$  is the set of neighbors and  $w_i$  is the weight of the neighbor  $x_i$ . A new closeness term (which comprises of the intensity and depth difference) was proposed to determine the weights of nonlocal means [36], as follows

$$c(x, x_i) = \sum_{j \in W_x; k \in W_{x(i)}} [x(j) - x_i(k)]^2 [d_x(j) - d_{x(i)}(k)]^2 \quad (2.14)$$

where  $x(j)$  and  $d_x(j)$  represent the nearby intensity and depth values of  $x$  within the local window  $W_x$ , while  $x_i(k)$  and  $d_{x(i)}(k)$  represent the nearby intensity and depth values of  $x_i$  within the local window  $W_{x(i)}$ . After some further development, the weight of the modified nonlocal means for hole filling is defined as

$$w_i = \frac{\exp(-c(x, x_i) / \sigma_c) \cdot \exp(-d(x, x_i) / \sigma_d)}{\sum_{i \in N} \exp(-c(x, x_i) / \sigma_c) \cdot \exp(-d(x, x_i) / \sigma_d)} \quad (2.15)$$

where  $d(.,.)$  measures the geometric distance between two pixels and  $\sigma$  stands for the variance. Experimental results show that this new approach outperforms the conventional spatial and temporal approaches for hole filling during the view synthesis. However, we believe that more interpolation techniques can be novelty designed suitable for view synthesis, which is a good future direction for image interpolation.

### C. PSNR and SSIM comparisons of several image interpolation algorithms

Let us compare the subjective and objective quality of some image interpolation algorithms as described in this paper. Table I and Table II show the PSNR and SSIM values of various image interpolation algorithms [4, 15, 17, 19, 24, 25] using 24 natural images from Kodak. We have also considered the execution time of these image interpolation algorithms using C++ codes as shown in Table III (see also figure 5). It is shown that the RSAI [25] achieves the highest average PSNR and SSIM but its execution time is longer than the second and the third best performers, i.e. SAI [24] and BSAI [17]. BSAI requires much less computation compared to that of the SAI but its quality is closed to the SAI. For real-time applications, BSAI is the best choice; and for offline applications, RSAI is the best choice.

### D. Classification of image interpolation algorithms

In this section, we are going to classify broadly image interpolation algorithms, as shown in figure 4. They are the polynomial-based approaches which are fast and can be adaptive to local statistics, and the edge-directed approaches which directly address the edge reconstruction criterion. Edge-directed methods can intuitively interpolate along one edge orientation, or fuse several estimates of different edge orientations, or minimize the linear mean squares error, for example, using the FIR Wiener filter, whereas the soft-decision interpolation can also be applied.

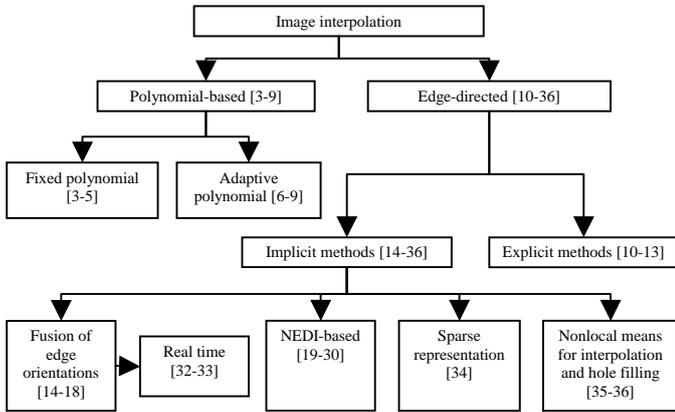


Fig. 4 Classifications of the image interpolation algorithms

TABLE I  
PSNR (dB) of the different image interpolation algorithms

Images	Bicubic [4]	LMSE [15]	NEDI [19]	SAI [24]	BSAI [17]	RSAI [25]
Kodim01	25.208	25.069	24.827	25.223	25.339	25.422
Kodim02	33.069	33.113	33.042	33.327	33.482	33.533
Kodim03	33.913	34.173	34.157	34.711	34.663	34.839
Kodim04	34.019	33.868	33.754	34.276	34.421	34.419
Kodim05	26.06	26.16	26.045	27.109	26.98	27.062
Kodim06	26.593	26.588	26.597	26.788	26.778	26.966
Kodim07	33.347	33.417	33.017	34.377	34.28	34.326
Kodim08	22.403	22.45	22.074	22.51	22.517	22.692
Kodim09	31.652	31.676	31.417	32.225	32.171	32.177
Kodim10	31.696	31.72	32.208	32.768	32.415	32.863
Kodim11	28.218	28.183	28.194	28.465	28.547	28.617
Kodim12	32.499	32.654	32.371	32.868	33.057	33.01
Kodim13	22.931	22.824	22.822	23.03	22.975	23.17
Kodim14	28.596	28.584	28.445	29.023	28.982	29.073
Kodim15	32.654	32.719	32.751	33.116	33.207	33.293
Kodim16	30.073	30.106	30.055	30.166	30.286	30.355
Kodim17	31.59	31.681	31.716	32.181	32.108	32.289
Kodim18	27.311	27.2	26.976	27.491	27.416	27.528
Kodim19	27.001	27.024	26.088	26.727	27.287	27.364
Kodim20	30.834	31.25	31.264	31.729	31.697	31.809
Kodim21	27.594	27.483	27.306	27.756	27.729	27.851
Kodim22	29.74	29.604	29.564	29.92	29.908	30.047
Kodim23	35.041	34.892	35.283	36.197	35.91	36.236
Kodim24	26.06	25.918	25.831	26.133	26.06	26.277
Average	29.504	29.514	29.408	29.921	29.925	<b>30.051</b>

### III. SUPER-RESOLUTION

#### A. Single-image and multi-frame super-resolution

Super-resolution (SR) aims to produce a high-resolution (HR) image using one or several observed low-resolution (LR) images by upsampling, deblurring and denoising. For multi-frame SR, there can involve registration and fusion processes. It is interesting to point out that some SR algorithms do not involve the denoising process, or some

interpolation algorithms are also referred to as super-resolution algorithms. Generally speaking, the super-resolution methods can be classified into single-image super-resolution (only one LR image is observed) [37-54] and multi-frame super-resolution [55-71]. The multi-frame super-resolution for a video sequence can moreover use a recursive estimation approach of video frames, which is called the dynamic super-resolution [55-58]. Dynamic super-resolution makes use of the previously reconstructed HR frame to estimate the current HR frame.

TABLE II  
SSIM [79] of the different image interpolation algorithms

Images	Bicubic [4]	LMSE [15]	NEDI [19]	SAI [24]	BSAI [17]	RSAI [25]
Kodim01	0.7659	0.7509	0.7393	0.7647	0.7673	0.769
Kodim02	0.8842	0.8746	0.8802	0.8867	0.889	0.8907
Kodim03	0.9223	0.9173	0.923	0.9279	0.9279	0.93
Kodim04	0.9035	0.8894	0.8926	0.904	0.9052	0.9051
Kodim05	0.8616	0.8549	0.8475	0.8782	0.8758	0.8754
Kodim06	0.8041	0.7965	0.799	0.8085	0.8087	0.8139
Kodim07	0.949	0.9423	0.9447	0.9567	0.9559	0.9563
Kodim08	0.7771	0.7751	0.7572	0.7804	0.7838	0.7841
Kodim09	0.9045	0.8979	0.9021	0.9093	0.9137	0.9119
Kodim10	0.9046	0.8982	0.9065	0.9139	0.9145	0.9159
Kodim11	0.8344	0.8261	0.8276	0.8379	0.8408	0.8422
Kodim12	0.8856	0.8787	0.8812	0.8876	0.8911	0.8911
Kodim13	0.729	0.7126	0.7073	0.7273	0.7265	0.7305
Kodim14	0.8495	0.8386	0.8354	0.8537	0.8536	0.8547
Kodim15	0.9088	0.9028	0.9078	0.9117	0.9137	0.9147
Kodim16	0.8426	0.8343	0.8394	0.8441	0.8456	0.8482
Kodim17	0.9162	0.9096	0.9155	0.9249	0.9245	0.9259
Kodim18	0.8502	0.837	0.8321	0.8522	0.8522	0.8523
Kodim19	0.8454	0.8351	0.8378	0.8465	0.8509	0.8518
Kodim20	0.9121	0.904	0.9118	0.9175	0.9198	0.9196
Kodim21	0.8685	0.8569	0.8595	0.8707	0.8729	0.873
Kodim22	0.8598	0.8432	0.8453	0.8596	0.8598	0.8612
Kodim23	0.9507	0.9407	0.9528	0.9556	0.9554	0.9569
Kodim24	0.8528	0.8419	0.8384	0.8565	0.855	0.8572
Average	0.8659	0.8566	0.8577	0.8698	0.8710	<b>0.8722</b>

#### B. Reconstruction-based and learning-based super-resolution

There are two major categories of SR algorithms: they are the reconstruction-based and learning-based algorithms. Reconstruction-based algorithms [37-39, 55-65, 67-71] use the data fidelity function, usually with the prior knowledge to regularize the ill-posed solution. Gradients (edges) are the main features to be used as the prior knowledge. Some more general prior knowledge using nonlocal self-similarity was also proposed [62]. Learning-based algorithms [40-54, 66] moreover utilize the information from training images. Generally, when a LR image is observed, a patch within that LR image is extracted and searched within the trained dictionary for estimating a suitable HR patch that reconstructs the HR image. Some recent investigations [41-43] make use of the sparse representation for estimating the HR pair and training the dictionary. By combining the use of prior knowledge and dictionary training, some unified frameworks [42-44] for single-image SR were proposed to further regularize the estimated solution from the training sets.

#### C. Generic and face super-resolution

For the learning-based algorithms, they are most suitable for specific applications, such as face super-resolution. Since

the face SR is crucial in applications like face recognition, surveillance, etc, a number of super-resolution methods were proposed for hallucinating the faces [48-54]. These are training-based methods which make use of the common characteristics of human faces (e.g. eigenface) to design the formulation for dictionary training and HR image reconstruction.

TABLE III  
Execution time (second) of image interpolation algorithms using C++ codes (\* is the estimated time)

Bicubic [4]	LMMSE [15]	NEDI [19]	SAI [24]	BSAI [17]	RSAI [25]
0.006	*0.5	*5	3.732	0.062	12.78

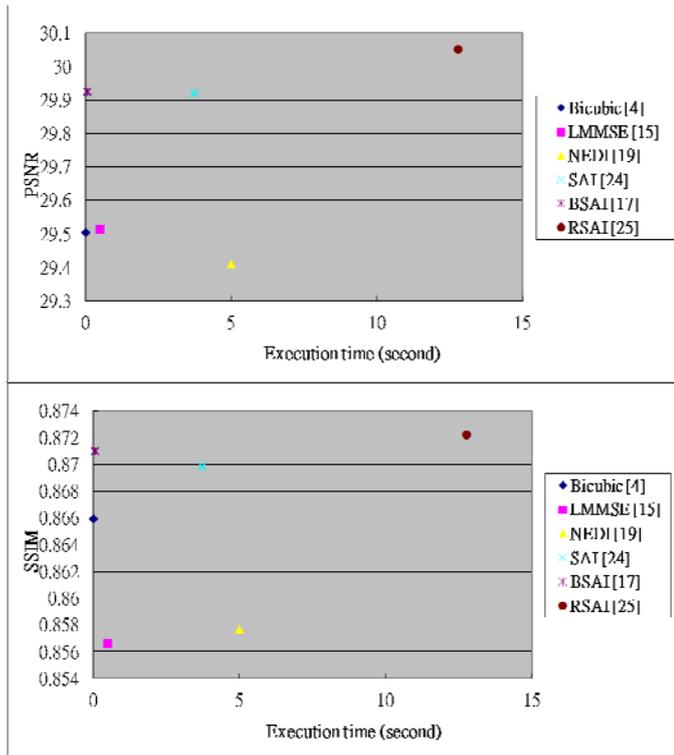


Fig. 5 Plot of PSNR and SSIM [79] against the execution time of image interpolation algorithms

#### D. Blind and non-blind super-resolution

The super-resolution methods can also be classified as blind and non-blind methods. The blind methods treat the point spread function (PSF) that represents the blur, and the registration parameters as variables to be estimated simultaneously with the HR image [67-71]. The Bayesian or MAP framework is widely used to jointly estimate the variables. Non-blind methods are still widely used because the PSF can be approximated, can be set according to the user preference, or is known due to the knowledge of the camera. The registration parameters can also be separately estimated. Usually, the non-blind methods apply the regularizations to make them robust to inaccurate estimations of PSF, registration parameters, etc. The non-blind methods are widely used due to its simplicity in formulations, which are also easier to be parallelized for the practical applications.

#### E. FIR Wiener filter for super-resolution

The finite-impulse response (FIR) Wiener filter, or equivalently, the linear minimum mean squares error (LMMSE) estimator, can be applied to perform super-resolution reconstruction. The block-based FIR Wiener filters were proposed for multi-frame SR [65-66]. In [65], a wide-sense stationary correlation function based on the geometric distance between pixels was proposed to estimate the covariances for the FIR Wiener filter, and elegant results were reported. The partition filters partitions an image into blocks for applying the FIR Wiener filter, where the filter weights are learned from an offline dictionary, which is retrieved during the online estimation [66].

Similarly, the Gaussian process regression (GPR) can be applied for super-resolution reconstruction. GPR provides a sophisticated Bayesian framework to estimate the HR image and the hyper-parameters of the process (e.g. noise variance, and variance of the correlation function) using a pilot HR image, which can be obtained by the bicubic interpolation [37]. The nonlocal means has been proposed as the correlation function in GPR [37]. Using the nonlocal means as the correlation function, the iterative scheme of FIR Wiener filter can alternatively update the estimated covariances and HR image, which can address the disadvantage of inaccurate pilot HR image [47].

Let us briefly introduce the iterative Wiener filter (IWF) algorithm [47] which currently produces the best PSNR and SSIM results among some state-of-the-art algorithms, as shown in Table IV and Table V. The formulation of the FIR Wiener filter which minimizes the linear mean squares error is given by

$$\mathbf{W}_i = \mathbf{R}_i^{-1} \mathbf{P}_i \quad (3.1)$$

where the filter weight  $\mathbf{W}_i$  is related with the autocorrelation matrix  $\mathbf{R}_i$  for the observation vector and cross-correlation matrix  $\mathbf{P}_i$  for the unobserved vector and observation vectors. The unobserved vector,  $\mathbf{y}_i \in \mathbf{y}$ , is defined as pixels inside a block within the unobserved HR image  $\mathbf{y}$  and observation vector,  $\mathbf{x}_i \in \mathbf{x}$ , is defined as pixels geometrically closest to the unobserved vector within the observed LR images  $\mathbf{x}$ . Figure 6 shows a graphical illustration of these definitions. The FIR Wiener filter estimates the unobserved vector by

$$\hat{\mathbf{y}}_i = \mathbf{W}_i^T \mathbf{x}_i \quad (3.2)$$

and the unobserved HR image becomes  $\hat{\mathbf{y}} = \{\hat{\mathbf{y}}_i\}$ . The weights of the FIR Wiener filter can be defined as

$$\begin{aligned} \mathbf{W}_i &= \left[ c_{\max} E\{\mathbf{x}_i \mathbf{x}_i^T / c_{\max}\} \right]^{-1} c_{\max} E\{\mathbf{x}_i \mathbf{y}_i^T / c_{\max}\} \\ &= \left[ E\{\mathbf{x}_i \mathbf{x}_i^T / c_{\max}\} \right]^{-1} E\{\mathbf{x}_i \mathbf{y}_i^T / c_{\max}\} \end{aligned} \quad (3.3)$$

where  $c_{\max}$  is a constant for normalization. Hence, elements of the scaled correlation matrices are always smaller than 1. Let us consider the nonlocal means filter to approximate the elements, as follows

$$E\{p_j p_k / c_{\max}\} \approx \exp(-E(p_j - p_k)^2 / \sigma) \quad (3.4)$$

where  $p_j, p_k \in \{\mathbf{x}_i, \mathbf{y}_i\}$  and the hyper-parameter  $\sigma$  controls the decay speed of the correlation function. Note that the correlation function depends on the unobserved HR image  $\mathbf{y}$

and its blurred version  $\mathbf{H}\mathbf{y}$ , where  $\mathbf{H}$  is the PSF. An iterative scheme can be applied iteratively to update the estimated correlation matrices and unobserved HR image to give the final estimate, as follows.

*Input: LR image  $\mathbf{x}$*

*Output: HR image  $\hat{\mathbf{y}}$*

1. Initialization
  - (a) Create an initial estimate of the unobserved HR image  $\hat{\mathbf{y}}^{(0)}$ .
  - (b) Create the initial estimate of the blurred HR image  $\mathbf{H}\hat{\mathbf{y}}^{(0)}$ .
2. Iterate on  $n$  until  $n$  reaches the desired number of iterations  $N$ .
  - (a) Compute the filter weight in (3.3) by using the normalized correlation function in (3.4).
  - (b) Use the computed filter weight and (3.2) to update the estimated HR image to  $\hat{\mathbf{y}}^{(n+1)}$ .
  - (c) Use IBP [59] to refine the estimated HR image  $\hat{\mathbf{y}}^{(n+1)}$  to fit the image model.
  - (d) Update the blurred HR image using  $\mathbf{H}\hat{\mathbf{y}}^{(n+1)}$ .
  - (e) Repeat 2(a) to 2(d) until the termination.

Algorithm 1 The iterative scheme of the FIR Wiener filter for SRR

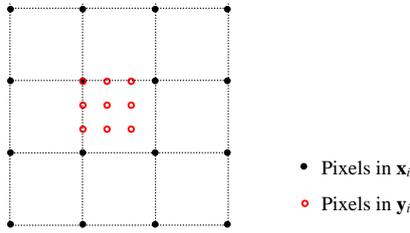


Fig. 6 An example illustration of the observation vector  $\mathbf{x}_i$  and the unobserved vector  $\mathbf{y}_i$  when the magnification factor  $q=3$  and number of elements in  $\mathbf{x}_i$  is 16.

**Example:** A data sequence is given by  $\{x(0), x(1), x(3), x(4)\}$ , where  $x(2)$  is a missing data point between  $x(1)$  and  $x(3)$ . We have to find  $x(2)$  using the Iterative Wiener filter. Let us define the observation vector  $\mathbf{x}_i$  and unobserved vector  $\mathbf{y}_i$  as below

$$\mathbf{x}_i = [x(1) \quad x(3)]^T \quad \text{and} \quad \mathbf{y}_i = x(2)$$

Let us initialize the unobserved vector as the average of the nearest two data points as follows

$$\mathbf{y}_i^{(0)} = (x(1) + x(3)) / 2$$

Let us compute the auto-correlation matrix and cross-correlation matrix using the correlation function in (3.4) with the nearest three data points for the expectation  $E(\cdot)$ :

First iteration:

$$\begin{aligned} \mathbf{R}_i^{(0)} &= \begin{bmatrix} \exp(-E(x(1) - x(1))^2 / \sigma) & \exp(-E(x(1) - x(3))^2 / \sigma) \\ \exp(-E(x(3) - x(1))^2 / \sigma) & \exp(-E(x(3) - x(3))^2 / \sigma) \end{bmatrix} \\ &= \begin{bmatrix} \exp(-0) / \sigma & \exp(-E(x(1) - x(3))^2 / \sigma) \\ \exp(-E(x(1) - x(3))^2 / \sigma) & \exp(-0) / \sigma \end{bmatrix} \end{aligned}$$

where the definition of the non-diagonal element is

$$\exp(-E(x(1) - x(3))^2 / \sigma) = \exp(-E(x(3) - x(1))^2 / \sigma)$$

$$= \exp(-[(x(0) - \mathbf{y}_i^{(0)})^2 + (x(1) - x(3))^2 + (\mathbf{y}_i^{(0)} - x(4))^2] / \sigma)$$

which is assumed to be 0.5. The auto correlation matrix is given by

$$\mathbf{R}_i^{(0)} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

The cross correlation matrix is given by

$$\begin{aligned} \mathbf{P}_i^{(0)} &= \begin{bmatrix} \exp(-E(x(1) - x(2))^2 / \sigma) \\ \exp(-E(x(3) - x(2))^2 / \sigma) \end{bmatrix} \\ &= \begin{bmatrix} \exp(-[(x(0) - x(1))^2 + (x(1) - \mathbf{y}_i^{(0)})^2 + (\mathbf{y}_i^{(0)} - x(3))^2] / \sigma) \\ \exp(-[(\mathbf{y}_i^{(0)} - x(1))^2 + (x(3) - \mathbf{y}_i^{(0)})^2 + (x(4) - x(3))^2] / \sigma) \end{bmatrix} \end{aligned}$$

which is assumed to be  $[0.4 \ 0.6]^T$ . Then the filter weight is

$$\mathbf{W}_i^{(0)} = (\mathbf{R}_i^{(0)})^{-1} \mathbf{P}_i^{(0)} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} = \begin{bmatrix} 0.1333 \\ 0.5333 \end{bmatrix}$$

which is used to update the output vector

$$\mathbf{y}_i^{(1)} = \mathbf{W}_i^T \mathbf{x}_i = \begin{bmatrix} 0.1333 & 0.5333 \end{bmatrix} \begin{bmatrix} x(1) \\ x(3) \end{bmatrix}$$

For the next iterations, the estimated output vector is substituted in computing the auto-correlation matrix and cross correlation matrices until convergence or the termination criterion is met.

#### F. PSNR and SSIM comparison of several super-resolution algorithms

Table IV and Table V show the PSNR and SSIM values of 8 natural images (512×512) using several (single-image) super-resolution algorithms [37, 43, 47, 59, 65] to upsample 3 times and deblur using a 3×3 box filter. Let us consider further the execution time of these super-resolution algorithms using MATLAB codes in Table VI, as plotted in figure 7. Table IV and Table V show that the iterative FIR Wiener filter (IWF) [47] achieves the highest average PSNR and SSIM values but its execution time is far less than that required for the second best performers, namely ASDS [43]. AWF [65] requires much less computation than the IWF but its quality is worse than IWF. As a result, we conclude that AWF would be a good choice among the tested algorithms for real-time applications, while IWF is likely to be the best choice for offline applications. Figures 10 and 11 show the subjective comparisons of IWF and the bilinear interpolation.

#### G. Summary of super-resolution algorithms

Let us also give a relational diagram for some major super-resolution algorithms. Figure 8 gives a brief classification of the super-resolution algorithms. The super-resolution algorithms can be classified as single-image and multi-frame approaches. For single-image algorithms, they are mostly learning-based approaches which aim at reconstructing the generic images and face images. The reconstruction-based algorithms often make use of gradient or patch redundancy as the prior knowledge to regularize the solution. The multi-frame algorithms can be divided into static and dynamic algorithms of which the recursive structure is applied. Moreover, the multi-frame algorithms can also be classified as non-blind and blind approaches, which simultaneously

estimate the registration parameters and the PSF together with the HR image.

#### IV. CONCLUSION

For image/video interpolation and super-resolution, more intensive research is expected due to the popularity of ultra high-definition TV and free viewpoint TV. Specifically, the multi-frame super-resolution and hole filling interpolation for view synthesis are going to be popular directions in 2D and 3D video applications. For a practical use of the super-resolution, fast algorithms are demanding and the blind algorithms will receive more attention in the future development. Furthermore, recent works show that face recognition can be benefited from a customized face super-resolution which maximizes the differences between two face manifolds. However, the theoretical study of such justification requires more investigations. Among the techniques for interpolation and super-resolution, the sparse representations should be a promising direction, and significant results have already been available in image processing applications.

To complete this review, let us also include a short highlight of review works in the literature. An early review of super-resolution algorithms is given in [72] and a statistical performance analysis of super-resolution is shown in [76]. The limitation and challenges of super-resolution was reviewed in [73-75] which show that a major limitation of multi-frame super-resolution is the registration accuracy. However, this may be resolved by making use of the optical flow techniques [77], and the influence of the inaccurate registration can also be alleviated by using an appropriate regularization [78].

TABLE IV

PSNR (dB) of the estimated images using different algorithms

Images	IBP [59]	AWF [65]	GPR [37]	ASDS [43]	IWF [47]
Bike	22.808	22.876	22.312	23.6	<b>23.784</b>
House	20.353	20.328	20.019	20.48	<b>20.682</b>
Game	25.546	25.574	25.076	<b>25.975</b>	25.966
Statue	28.614	28.678	28.151	29.269	<b>29.27</b>
Woman	25.717	25.765	25.223	26.083	<b>26.238</b>
Lighthouse	23.76	23.717	23.442	23.885	<b>23.948</b>
Parrot	29.439	29.639	28.69	30.387	<b>30.467</b>
Lena	31.334	31.5	30.774	32.696	<b>32.721</b>
Average	25.946	26.01	25.461	26.547	<b>26.635</b>

TABLE V

SSIM [79] of the estimated images using different algorithms

Images	IBP [59]	AWF [65]	GPR [37]	ASDS [43]	IWF [47]
Bike	0.7161	0.7199	0.6686	0.7582	<b>0.7597</b>
House	0.6446	0.6409	0.6096	0.6572	<b>0.6668</b>
Game	0.7402	0.7421	0.7039	<b>0.7569</b>	0.7548
Statue	0.8379	0.8407	0.817	0.8551	<b>0.857</b>
Woman	0.7629	0.7669	0.7172	0.7762	<b>0.7765</b>
Lighthouse	0.7559	0.753	0.732	0.7665	<b>0.7671</b>
Parrot	0.9084	0.9089	0.8963	0.917	<b>0.9182</b>
Lena	0.8792	0.8807	0.8636	<b>0.8923</b>	0.892
Average	0.7807	0.7816	0.751	0.7974	<b>0.799</b>

TABLE VI

Execution time (minute) of super-resolution algorithms using MATLAB codes

IBP [59]	AWF [65]	GPR [37]	ASDS [43]	IWF [47]
0.02	0.04	123	32	5.6

#### ACKNOWLEDGMENT

This work is supported by the Center for Signal Processing, the Hong Kong Polytechnic University (U-G863/P10-024), RGC-PolyU 5278/08E, and SmartEN Marie Curie ITN-Network.

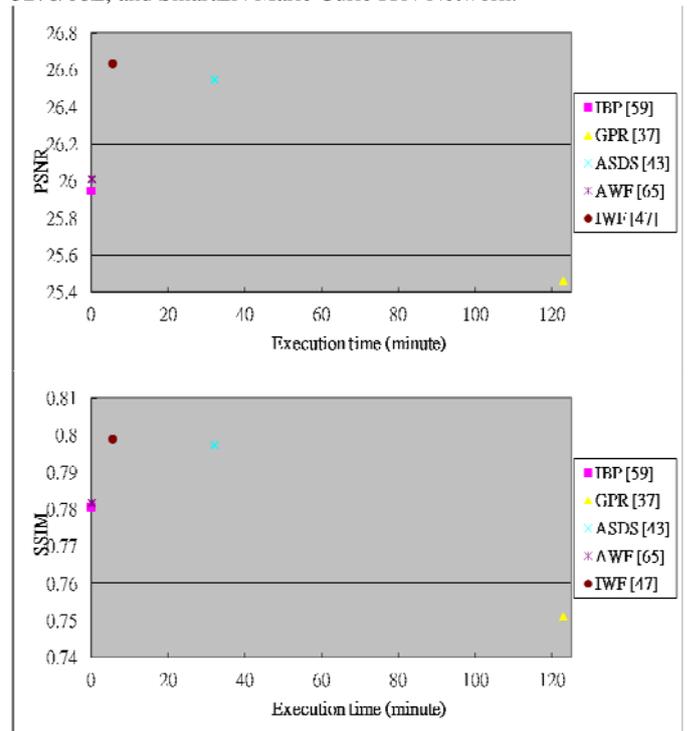


Fig. 7 Plot of PSNR (dB) and SSIM [79] against the execution time of super-resolution algorithms

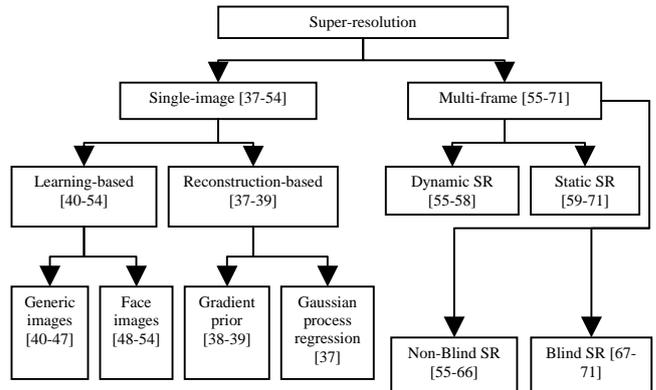


Fig. 8 Classifications of the super-resolution algorithms

#### REFERENCES

- [1] Aly, H.A and Dubois, E., "Image up-sampling using total-variation regularization with a new observation model," IEEE Trans. on Image Processing, vol.14, no.10, pp. 1647-1659, Oct 2005
- [2] T. Blu, P. Thévenaz, and M. Unser, "Linear interpolation revitalized," IEEE Trans. on Image Processing, vol. 13, no. 5, pp. 710-719, May 2004.
- [3] H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," IEEE Trans. on Acoust., Speech, Signal Process, vol. ASSP-26, no. 6, pp. 508-517, Dec. 1978.
- [4] R. G. Keys, "Cubic convolution interpolation for digital image processing," IEEE Trans. on Acoust., Speech, Signal Process, vol. ASSP-29, no. 6, pp. 1153-1160, Dec. 1981.
- [5] T. Lehmann, C. Gönnner, and K. Spitzer, "Addendum: B-spline interpolation in medical image processing," IEEE Trans. on Med. Imag., vol. 20, no. 7, pp. 660-665, Jul. 2001.

- [6] S. W. Lee and J. K. Paik, "Image interpolation using adaptive fast B-spline filtering," Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP 1993), vol. 5, pp. 177-180, 1993.
- [7] F.G.B. De Natale, G.S. Desoli and D.D. Giusto, "Adaptive least-squares bilinear interpolation (ALSBI): a new approach to image-data compression," Electronics Letters, vol.29, no.18, pp.1638-1640, 2 Sept. 1993
- [8] Jong-Ki Han and Seung-Ung Baek, "Parametric cubic convolution scaler for enlargement and reduction of image," IEEE Trans. on Consumer Electronics, vol.46, no.2, pp.247-256, May 2000
- [9] K. Jensen and D. Anastassiou, "Subpixel edge localization and the interpolation of still images," IEEE Trans. on Image Processing, vol. 4, no. 3, pp. 285-295, Mar. 1995.
- [10] Qing Wang and Ward, R.K., "A New Orientation-Adaptive Interpolation Method," IEEE Trans. on Image Processing, vol.16, no.4, pp.889-900, April 2007
- [11] Siyoung Yang, Yongha Kim and Jechang Jeong, "Fine edge-preserving technique for display devices," IEEE Trans. on Consumer Electronics, vol.54, no.4, pp.1761-1769, November 2008
- [12] Kwan Pyo Hong, Joon Ki Paik, Hyo Ju Kim and Chul Ho Lee, "An edge-preserving image interpolation system for a digital camcorder," IEEE Trans. on Consumer Electronics, vol.42, no.3, pp.279-284, Aug 1996
- [13] M. J. Chena, C. H. Huang and W. L. Lee, "A fast edge-oriented algorithm for image interpolation," Image and Vision Computing, Vol. 23, pp. 791-798, 2005.
- [14] Min Li and Nguyen, T.Q., "Markov Random Field Model-Based Edge-Directed Image Interpolation," IEEE Trans. on Image Processing, vol.17, no.7, pp.1121-1128, July 2008
- [15] Lei Zhang and Xiaolin Wu, "An edge-guided image interpolation algorithm via directional filtering and data fusion," IEEE Trans. on Image Processing, vol.15, no.8, pp.2226-2238, Aug. 2006
- [16] G. Shi, W. Dong, X. Wu, and L. Zhang, "Context-based adaptive image resolution upconversion," Journal of Electronic Imaging, Vol 19(1), 013008, Jan-Mar 2010
- [17] Kwok-Wai Hung and Wan-Chi Siu, "Fast image interpolation using bilateral filter", pp. 1-14, *IET Image Processing*, doi:10.1049/iet-*ipr*.2011.005, The Institution of Engineering and Technology, 2012.
- [18] Amin Behnad and Xiaolin Wu, "Image interpolation with hidden Markov model", Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, pp. 874-877, 15-19 March 2010
- [19] X. Li and M. T. Orchard, "New edge-directed interpolation," IEEE Transactions on Image Processing, vol. 10, pp. 1521-1527, Oct. 2001.
- [20] S. G. Mallat, *A Wavelet Tour of Signal Processing*. New York: Academic, 1998.
- [21] Wing-Shan Tam, Chi-Wah Kok and Wan-Chi Siu, "A Modified Edge Directed Interpolation for Images," Journal of Electronic Imaging, Vol. 19(1), pp.13011\_1-20, Jan-March 2010.
- [22] Dung T. Vo, Joel Sole, Peng Yin, Cristina Gomila and Truong Q. Nguyen, "Selective Data Pruning-Based Compression using High Order Edge-Directed Interpolation", IEEE Trans. on Image Processing, vol 19, No. 2, pp 399-409, February 2010.
- [23] Xianming Liu, Debin Zhao, Ruiqin Xiong, Siwei Ma, Wen Gao and Huifang Sun, "Image Interpolation Via Regularized Local Linear Regression," IEEE Trans. on Image Processing, vol.20, no.12, pp.3455-3469, Dec. 2011
- [24] Xiangjun Zhang and Xiaolin Wu, "Image Interpolation by Adaptive 2-D Autoregressive Modeling and Soft-Decision Estimation," IEEE Trans. on Image Processing, vol.17, no.6, pp.887-896, June 2008
- [25] Kwok-Wai Hung and Wan-Chi Siu, "Robust Soft-Decision Interpolation Using Weighted Least Squares," IEEE Trans. on Image Processing, vol.21, no.3, pp.1061-1069, March 2012
- [26] Kwok-Wai Hung and Wan-Chi Siu, "Improved Image Interpolation using Bilateral Filter for Weighted Least Square Estimation," Proc. IEEE Int. Conf. Image Processing (ICIP 2010), pp. 3297-3300, 26-29 September, 2010, Hong Kong
- [27] Chi-Shing Wong and Wan-Chi Siu, "Further Improved Edge-directed Interpolation and Fast EDI for SDTV to HDTV Conversion," Proc. 18th European Signal Processing Conference (EUSIPCO 2010), pp. 309-313, 23-27 August, 2010, Aalborg Denmark.
- [28] Chi-Shing Wong and Wan-Chi Siu, "Adaptive Directional Window Selection For Edge-Directed Interpolation," Proc. ICCCN, 2010
- Workshop on Multimedia Computing and Communications, MMC1, No.4, pp. 1-6., 2-5 August, 2010, Zurich, Switzerland.
- [29] Ketan Tang, Oscar C. Au, Lu Fang, Zhiding Yu and Yuanfang Guo, "Image Interpolation Using Autoregressive Model and Gauss-Seidel Optimization," Proc. Sixth Int. Conf. Image and Graphics (ICIG 2011), pp. 66-69, 12-15 August, 2011, Hefei, Anhui, China
- [30] Jie Ren, Jiaying Liu, Wei Bai and Zongming Guo, "Similarity modulated block estimation for image interpolation," Proc. IEEE Int. Conf. Image Processing (ICIP 2011), pp.1177-1180, 11-14 Sept. 2011, Brussels, Belgium
- [31] K. Ni and T. Q. Nguyen, "An Adaptive k-Nearest Neighbor Algorithm for MMSE Image Interpolation," IEEE Trans. on Image Processing, Vol. 18, No. 9, pp 1976-1987. September 2009.
- [32] Jie Cao, Ming-Chao Che, Xiaolin Wu and Jie Liang, "GPU-aided directional image/video interpolation for real time resolution upconversion," IEEE International Workshop on Multimedia Signal Processing (MMSP), pp. 1-6, 5-7 Oct. 2009
- [33] Wei Lei, Ruiqin Xiong, Siwei Ma and Luhong Liang, "GPU based fast algorithm for tanner graph based image interpolation," IEEE International Workshop on Multimedia Signal Processing (MMSP), pp.1-4, 17-19 Oct. 2011
- [34] S. Mallat and G. Yu, "Super-Resolution with Sparse Mixing Estimators", IEEE Trans. on Image Processing, vol.19, issue 11, pp. 2889 - 2900, 2010.
- [35] Kwok-Wai Hung and Wan-Chi Siu, "Fast Video Interpolation/Upsampling using Linear Motion Model," Proc. IEEE Int. Conf. Image Processing (ICIP 2011), pp. 1341-1344, 11-14 September, 2011, Brussels, Belgium.
- [36] Kwok-Wai Hung and Wan-Chi Siu, "Depth-assisted Nonlocal Means Hole Filling for Novel View Synthesis," Proc. pp.2737-40, IEEE Int. Conf. Image Process. (ICIP 2012), 30 Sept.-3 Oct., 2012, Orlando, USA.
- [37] He He and Wan-Chi Siu, "Single image super-resolution using Gaussian process regression," Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR2011), pp.449-456, 20-25 June 2011, Colorado, USA.
- [38] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong and A. K. Katsaggelos, "SoftCuts: A soft edge smoothness prior for color image super-resolution," IEEE Trans. on Image Process., vol.18, no.5, pp.969-981, May 2009.
- [39] J. Sun, J. Sun, Z. Xu and H. Y. Shum, "Gradient Profile Prior and Its Applications in Image Super-Resolution and Enhancement," IEEE Trans. on Image Processing, vol.20, no.6, pp. 1529-1542, June 2011
- [40] Xinbo Gao, Kaibing Zhang, Dacheng Tao and Xuelong Li, "Joint Learning for Single-Image Super-Resolution via a Coupled Constraint," IEEE Trans. on Image Processing, vol.21, no.2, pp.469-480, Feb. 2012
- [41] J. Yang, J. Wright, T.S. Huang and Y. Ma, "Image Super-Resolution Via Sparse Representation," IEEE Trans. on Image Processing, vol.19, no.11, pp.2861-2873, Nov. 2010
- [42] K. I. Kim and Y. Kwon, "Single-Image Super-Resolution Using Sparse Regression and Natural Image Prior," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.32, no.6, pp.1127-1133, June 2010
- [43] W. Dong, L. Zhang, G. Shi and X. Wu, "Image Deblurring and Super-Resolution by Adaptive Sparse Domain Selection and Adaptive Regularization," IEEE Trans. on Image Processing, vol.20, no.7, pp.1838-1857, July 2011
- [44] D. Glasner, S. Bagon and M. Irani, "Super-Resolution from a Single Image", Proc. Int. Conf. Computer Vision (ICCV 2009), 2009, Kyoto, Japan.
- [45] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based superresolution", IEEE Computer Graphics and Applications, vol.22(2), pp. 56-65, March/April 2002.
- [46] R. He and Z. Zhang, "Locally affine patch mapping and global refinement for image super-resolution" Pattern Recognition, Volume 44, pp. 2210-2219, Issue 9, September 2011.
- [47] Kwok-Wai Hung and Wan-Chi Siu, "Single-Image Super-Resolution Using Iterative Wiener Filter," Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP 2012), pp. 1269-1272, 25-30, March, 2012, Kyoto, Japan
- [48] B.K. Gunturk, A.U. Batur, Y. Altunbasak, M.H. Hayes III and R.M. Mersereau, "Eigenface-domain super-resolution for face recognition," IEEE Trans. on Image Processing, vol.12, no.5, pp. 597-606, May 2003.
- [49] H. Huang, H. He, X. Fan and J. Zhang, "Super-resolution of human face image using canonical correlation analysis", Pattern Recognition, Volume 43, Issue 7, pp. 2532-2543, July 2010.

[50] Y. Zhuang, J. Zhang and F. Wu, "Hallucinating faces: LPH super-resolution and neighbor reconstruction for residue compensation", *Pattern Recognition*, Volume 40, Issue 11, pp. 3178-3194, November 2007.

[51] Wei Zhang and Wai-Kuen Cham, "Hallucinating Face in the DCT Domain," *IEEE Trans. on Image Processing*, vol.20, no.10, pp.2769-2779, Oct. 2011

[52] Kai Guo, Xiaokang Yang, Rui Zhang, Guangtao Zhai, Songyu Yu: Face super-resolution using 8-connected Markov Random Fields with embedded prior. *ICPR 2008*: 1-4

[53] Hui Zhuo and Kin-Man Lam, "Eigentransformation-based face super-resolution in the wavelet domain," *Pattern Recognition Letters*, vol 33(6), pp. 718-727, 2012

[54] Yu Hu, Kin-Man Lam, Guoping Qiu and Tingzhi Shen, "From Local Pixel Structure to Global Image Super-Resolution: A New Face Hallucination Framework," *IEEE Trans. on Image Processing*, vol. 20(2), pp. 433-445, 2011

[55] S. Farsiu, M. Elad, and P. Milanfar, "Video-to-Video Dynamic Super-Resolution for Grayscale and Color Sequences", *EURASIP Journal on Applied Signal Processing*, No. Article ID 61859, pp.1-15, 2006.

[56] C. Bishop, A. Blake, and B. Marthi, "Super-resolution enhancement of video", *Proc. of the Ninth International Workshop on Artificial Intelligence and Statistics*, January 2003.

[57] K. Simonyan, S. Grishin, D. Vatolin and D. Popov, "Fast video superresolution via classification," *Proc. IEEE Int. Conf. Image Processing (ICIP 2008)* pp.349-352, 12-15 Oct. 2008, San Diego, U.S.A.

[58] Kwok-Wai Hung and Wan-Chi Siu, "New Motion Compensation Model via Frequency Classification for Fast Video Super-Resolution," *Proc. IEEE Int. Conf. Image Processing (ICIP 2009)*, pp. 1193-1196, 7-11 November, 2009, Cairo, Egypt.

[59] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graph. Models and Image Proc.*, vol. 53, pp.231-9, May 1991.

[60] A. Marquina and S. J. Osher, "Image super-resolution by TV-regularization and Bregman iteration," *J. Sci. Comput.*, vol. 37, pp. 367-382, 2008.

[61] M. Protter and M. Elad, "Super Resolution With Probabilistic Motion Estimation," *IEEE Trans. on Image Processing*, vol.18, no.8, pp.1899-1904, Aug. 2009

[62] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal- means to super-resolution reconstruction," *IEEE Trans. on Image Processing*, vol. 18, no. 1, pp. 36-51, Jan. 2009.

[63] H. Takeda, P. Milanfar, M. Protter, and M. Elad, "Super-resolution without explicit subpixel motion estimation," *IEEE Trans. on Image Processing*, vol.18, no.9, pp.1958-1975, Sept. 2009.

[64] Yue Zhuo, Jiaying Liu, Jie Ren and Zongming Guo. "Nonlocal Based Super Resolution with Rotation Invariance and Search Window Relocation", *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP 2012)*, pp. 853-856, 25-30, March, 2012, Kyoto, Japan

[65] R.C. Hardie, "A Fast Image Super-Resolution Algorithm Using an Adaptive Wiener Filter," *IEEE Trans. on Image Processing*, vol.16, no.12, pp.2953-2964, Dec. 2007

[66] B. Narayanan, R.C. Hardie, K.E. Barner and M. Shao, "A Computationally Efficient Super-Resolution Algorithm for Video Processing Using Partition Filters," *IEEE Trans. on Circuits and Systems for Video Technology*, vol.17, no.5, pp.621-634, May 2007

[67] R.C. Hardie, K.J. Barnard, E.E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Trans. on Image Processing*, vol.6, no.12, pp.1621-1633, Dec 1997

[68] Y. He, K.H. Yap, L. Chen, and L.P. Chau, "A nonlinear least square technique for simultaneous image registration and super-resolution," *IEEE Trans. on Image Processing*, vol. 16, no. 11, pp. 2830-2841, Nov. 2007.

[69] N. A. Woods, N. P. Galatsanos, and A. K. Katsaggelos, "Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images," *IEEE Trans. on Image Processing*, vol. 15, no. 1, pp. 201-213, Jan. 2006.

[70] N. Nguyen, P. Milanfar, and G. Golub, "Efficient generalized crossvalidation with applications to parametric image restoration and resolution enhancement," *IEEE Trans. on Image Processing*, vol. 10, no. 9, pp. 1299-1308, Sep. 2001.

[71] Xuesong Zhang, Jing Jiang and Silong Peng, "Commutability of Blur and Affine Warping in Super-Resolution With Application to Joint

Estimation of Triple-Coupled Variables," *IEEE Trans. on Image Processing*, vol.21, no.4, pp.1796-1808, April 2012.

[72] S.C. Park, M.K. Park, and M.G. Kang, "Super-Resolution Image Reconstruction: A Technical Overview", *IEEE Signal Processing Magazine*, Vol. 20, pp. 21-36, May 2003.

[73] S. Baker and T.Kanade, "Limits on Super-Resolution and How to Break Them," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.24, Issue 9, pp. 1167-1183, Sept. 2002.

[74] Z. Lin, J. He, X. Tang, and C. K. Tang, "Limits of Learning-Based Superresolution Algorithms," *International Journal of Computer Vision*, vol. 80, no. 3, Aug. 2008.

[75] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *Int. J. Imag. Syst. Technol.*, vol. 14, no. 2, pp. 47-57, Oct. 2004.

[76] D. Robinson and P. Milanfar, "Statistical performance analysis of super-resolution," *IEEE Trans. on Image Processing*, vol.15, no.6, pp.1413-1428, June 2006

[77] W. Zhao and H. S. Sawhney, "Is super-resolution with optical flow feasible?," *Proc. European Conference on Computer Vision (ECCV 2002)*. pp. 599-613, 2002, Copenhagen, Denmark.

[78] Eun Sil Lee and Moon Gi Kang, "Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration," *IEEE Trans. on Image Processing*, vol.12, no.7, pp. 826- 837, July 2003

[79] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, "Image quality assessment: from error measurement to structural similarity," *IEEE Trans. on Image Process.*, vol. 3, no. 4, pp. 600-612, Apr. 2004.



Fig. 10 Super-resolution results: From top to bottom, they are the LR input image and the estimated HR images using bilinear interpolation and IWF [47].



Fig. 11 Super-resolution results: From top to bottom, they are the LR input image and the estimated HR images using bilinear interpolation and IWF [47].