FRESH – FRI-based single-image super-resolution algorithm

Xiaoyao Wei and Pier Luigi Dragotti

***Abstract*—In this paper we consider the problem of single image super-resolution and propose a novel algorithm that outperforms state-of-the-art methods without the need of learning patches pairs from external datasets. We achieve this by modeling images and, more precisely, lines of images as piecewise smooth functions and propose a resolution enhancement method for this type of functions. The method makes use of the theory of sampling signals with finite rate of innovation (FRI) and combines it with traditional linear reconstruction methods. We combine the two reconstructions by leveraging from the multiresolution analysis in wavelet theory and show how an FRI reconstruction and a linear reconstruction can be fused using filter-banks. We then apply this method along vertical, horizontal and diagonal directions in an image to obtain a single-image super-resolution algorithm. We also propose a further improvement of the method based on learning from the errors of our super-resolution result at lower resolution levels. Simulation results show that our method outperforms state-of-the-art algorithms under different blurring kernels.**

***Index Terms*—super resolution, resolution enhancement, wavelet theory, sampling, finite rate of innovation (FRI).**

I. Introduction

SINGLE-IMAGE super-resolution refers to the problem of obtaining a high-resolution (HR) version of a single low-resolution (LR) image. This differs from the more traditional multi-frame super-resolution problem where one has access to multiple shifted versions of the LR image and tries to estimate a single HR image from these multiple images (see [1] for a nice overview). The single image super-resolution problem is highly ill-posed since it is possible to find many high-resolution images that can lead to the same low-resolution one. Thus prior knowledge of the properties of natural images has to be used to regularize the problem.

Strategies to solve the resolution enhancement problem are typically categorized into three broad methods: interpolation based methods, constrained reconstruction based methods, and learning based methods. Interpolation based techniques [2], [3], e.g. bilinear, bicubic interpolation algorithms have their roots in sampling theory and their essence is to recover the continuous-time signal from the given discrete pixels. They are computationally simple, however they typically are based on a slow-varying image model (sum of weighted and shifted versions of a basis function) so often produce images without high frequency details. Reconstruction based approaches (e.g., [4]–[10]) define constraints for the target high-resolution image. Commonly used priors includes statistical prior of natural images [4], [5], total-variation prior [6], gradient-profile prior [8]. There are also learning based algorithms which infer missing high frequency information based on a dictionary containing pairs of low-resolution and high-resolution patches. The dictionary is either trained externally using a database of low-resolution (LR) and high-resolution (HR) image pairs [11]–[16] or internally using self-similarities of the image at different scales [17]–[20].

Xiaoyao and Pier Luigi are with the Communications and Signal Processing Group, Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK. e-mail: ivy.wei@imperial.ac.uk; p.dragotti@imperial.ac.uk.

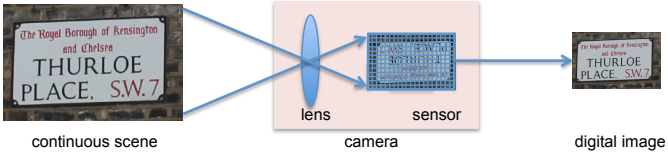


Fig. 1. Image formation process. The incoming irradiance light field is blurred by the lens and sampled by the image sensor.

In this paper we connect the single image super-resolution problem to the one of sampling and reconstructing piecewise regular functions. We note that the high-resolution, high-frequency information of an image is lost during the acquisition process (see Fig. 1) because of lens blur (usually modeled by the point spread function) and limited density of imaging sensors. This process is very similar to the way acquisition is modeled in traditional sampling theory where the analogue signal is low-pass filtered (equivalent to the blurring due to lenses) and then sampled (equivalent to the sensor grid in a digital camera). Linear interpolation methods have the merit of making this connection explicit. However, they are not effective in practice because they can only recover globally smooth functions, whereas images and scan-lines of images are piecewise regular (see Fig. 2).

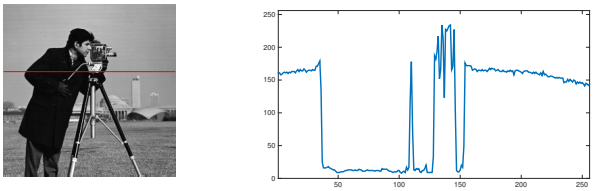


Fig. 2. Natural images and scan-lines of natural images are approximately 2-D and 1-D piecewise smooth functions respectively.

The recently developed theory of sampling signals with finite rate of innovation (FRI) [21]–[28] overcomes in some

cases the limitation of the linear interpolation methods by exploiting the fact that many signals, like piecewise polynomial functions, are fully specified by a finite number of parameters. FRI theory has shown that these signals, although non-bandlimited, can be uniquely reconstructed by only a small number of samples taken with specific acquisition devices. In other words, there is a unique mapping between a specific low-resolution version to the infinite-resolution version for these signals and there is a constructive way for recovery. The FRI theory is later extended to the approximate FRI framework that works with any sampling kernel [28].

This insight inspires a novel method for sampling continuous-time image scan-lines or 1-D piecewise smooth functions: a piecewise smooth signal can be modelled as the sum of a piecewise polynomial and a globally smooth part and we propose a hybrid reconstruction method based on classical linear recovery of the smooth part and non-linear recovery of the piecewise polynomial part using FRI on the same set of samples. We then leverage from wavelet theory and the corresponding multi-resolution analysis [29] to adapt this hybrid reconstruction method to the resolution enhancement problem. In particular, enhancing the resolution of a signal is equivalent to finding the detail wavelet coefficients at finer scales. We do this using FRI and we combine the details with the coarse linear approximation. Because of the connection with wavelet theory this can be achieved using filter banks. This leads to a fast and extremely effective algorithm to enhance the resolution of 1-D piecewise smooth functions.

We extend this approach to images by approximating the point-spread-function with a scaling function in the wavelet theory (typically a spline of a certain order) and apply the 1-D method along vertical, horizontal and diagonal directions. These reconstruction are then combined with the low-resolution version of the image using a 2-D filter-bank.

Finally, inspired by the works in single image super-resolution based on self-learning, we propose correcting the error in our FRI upsampling result by learning from the pair of input LR image and the corresponding FRI image of same size recovered from an even lower scale. Contrary to [17], [19], we use self-learning algorithms only to refine our FRI based method. As a result, we do not require learning at every small increment of scales and our upsampled images have less artefacts. The end result is an FRi-based singlE-image Super-resolution algoritHm - FRESH, which outperforms state-of-the-art methods in most situations. Fig. 3 shows a comparison on an example.

The paper is organized as follows. In Section II we review the classical sampling theory whose reconstruction process is linear, and the recently developed FRI sampling theory. In Section III, we first show our interpretation of sampling from multi-resolution property of wavelet transform and propose how to sample piecewise smooth signals, then by relating the resolution enhancement problem to sampling problem we propose a method for enhancing the resolution of piecewise smooth signals using filter-banks. In Section IV, we show how the 1-D upsampling method is extended to enhance the resolution of 2-D images. We then propose an improved upsampling method which corrects errors in the FRI upsampled image by exploiting the similarities between images at different resolutions. We show simulation results in Section V and conclude in Section VI.

(a) the original image (b) linear reconstr. PSNR=25.91dB (c) A+ [16] PSNR=27.34dB (d) our algorithm PSNR=27.72dB

Fig. 3. Upsampling results (factor 4) of *woman* by different methods. The downsampling kernel is bior4.4.

II. Overview of Recent Developments in Sampling Theory

A. Sampling Problem and Classical Linear Reconstruction

Fig. 4 depicts the typical sampling setup, where the original continuous-time signal $x(t)$ is filtered with a linear-time invariant filter with impulse response $h(t)$ and then is sampled with sampling period T. Under this model the samples y_n are given by

$$y_n = \langle x(t), \tilde{\varphi}(t/T - n) \rangle, \tag{1}$$

where $\tilde{\varphi}(t)$ is the sampling kernel and is the scaled and time-reversed version of $h(t)$.

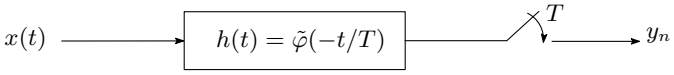


Fig. 4. Sampling set-up. Here $x(t)$ is the input signal, $h(t)$ is the impulse response of the acquisition device and T is the sampling period. The samples are given by $y_n = \langle x(t), \tilde{\varphi}(t/T - n) \rangle$.

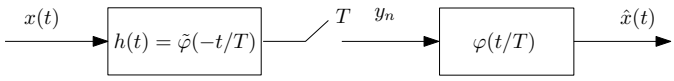


Fig. 5. Linear reconstruction set-up. The reconstruction of the input signal is given by $\hat{x}(t) = \sum_n y_n \varphi(t/T - n)$.

Traditionally, see Fig. 5, $x(t)$ is reconstructed using a linear filter with response $\varphi(t)$ where the pair $\{\tilde{\varphi}(t), \varphi(t)\}$ is chosen so that $\langle \varphi(t-n), \tilde{\varphi}(t-k) \rangle = \delta_{n-k}$. Under this model the sampling and reconstruction process can be interpreted as computing the orthogonal projection of $x(t)$ onto the shift-invariant subspace $\mathbf{V}$ spanned by $\varphi(t)$ and its shifted versions: $\mathbf{V} = \text{span}\{\varphi(t/T - n)\}_{n \in \mathbb{Z}}$. Therefore perfect reconstruction of $x(t)$ is achieved when $x(t) \in \mathbf{V}$. Interestingly, this framework includes the classical Shannon sampling theorem, where, in that case, $\varphi(t)$ is the sinc function, $\tilde{\varphi}(t) = \varphi(t)$ and $\mathbf{V}$ is the shift-invariant space of functions bandlimited to $f_s/2 = 1/2T$. However, $\varphi(t)$ does not have to be limited to the sinc function and alternative choices are possible. For example, $\varphi(t)$ could be a polynomial B-spline of a certain order. We also note that slightly more sophisticated forms of

linear reconstruction are also possible where, for example, one tries to impose *consistency*. Specifically, the goal is to reconstruct a signal $\hat{x}(t)$ that would lead to the same samples y_n if $\hat{x}(t)$ were to be sampled again. This new constraint leads to a different synthesis filter and to a reconstruction that gives an oblique projection rather than orthogonal projection of $x(t)$ onto $\mathbf{V}$. For more details on the topic, we refer to the insightful review [30].

B. Non-linear Reconstruction Method based on FRI

Signals that are neither bandlimited nor belong to a fixed subspace cannot be reconstructed perfectly using the classical linear reconstruction methods. However recently it was shown that it is possible to develop sampling schemes for classes of signals having parametric representations with finite number of degrees of freedom, called signals with finite rate of innovation (FRI) [21]. Examples of FRI signals include streams of Diracs and piecewise polynomial signals. The reconstruction process for this scheme is non-linear and is based on the use of the annihilating filter method. In what follows we show how to recover a piecewise polynomial signal $p(t)$ from the samples $y_n = \langle p(t), \tilde{\varphi}(t/T - n)\rangle$ taken with sampling kernel $\tilde{\varphi}(t)$ of Fig. 4. Emphasis is given to the key aspects of this method which will be useful to understand the signal upsampling approach which will be introduced in the following sections.

We consider a piecewise polynomial function $p(t)$ with pieces of maximum degree $R-1$. We denote the R-th derivative of $p(t)$ by $p^{(R)}(t)$, and note that $p^{(R)}(t)$ is a stream of (differentiated) Diracs. By using the link between discrete differentiation and derivation in continuous domain (see Appendix A for more details) we note that the R-th finite differences

$$z_n^{(R)} = \sum_{k=0}^{R} (-1)^{R-k} \binom{R}{k} y_{n+k} \text{ with } y_n = \langle p(t), \tilde{\varphi}(t/T - n)\rangle \tag{2}$$

corresponds to the samples $\langle p^{(R)}(t), \tilde{\varphi}_{\text{eq}}(t/T - n)\rangle$ obtained by acquiring $p^{(R)}(t)$ with the new kernel $\tilde{\varphi}_{\text{eq}}(t) = \tilde{\varphi}(t) * \beta_{R-1}(t)$, where $\beta_{R-1}(t)$ is the polynomial B-spline of degree $R-1$. Because of this connection in what follows we focus on the reconstruction of streams of differentiated Diracs to solve the problem of sampling piecewise polynomials.

We consider a sampling kernel satisfying the generalised Strang-Fix conditions [28]:

$$\hat{\tilde{\varphi}}(j\omega_m) \neq 0 \text{ and } \hat{\tilde{\varphi}}(j\omega_m + j2\pi l) = 0 \qquad l \in \mathbb{Z}\backslash\{0\} \tag{3}$$

where $\hat{\tilde{\varphi}}(j\omega)$ is the Fourier transform of $\tilde{\varphi}(t)$, and $j\omega_m = m'\lambda$ where $m' = m - (M+1)/2$ and $m = 0, \ldots, M$. This kernel can reproduce exponentials at frequencies $j\omega_m$. This means that there exists coefficients $b_{m,n}$ such that $\sum_{n\in\mathbb{Z}} b_{m,n}\tilde{\varphi}(t/T - n) = \mathrm{e}^{j\omega_m t/T}$. The new kernel $\tilde{\varphi}_{\text{eq}}(t) = \tilde{\varphi}(t) * \beta_{R-1}(t)$ also satisfies the generalised Strang-Fix conditions and can reproduce exponentials:

$$\sum_{n\in\mathbb{Z}} c_{m,n}\tilde{\varphi}_{\text{eq}}(t/T - n) = \mathrm{e}^{j\omega_m t/T} \quad \text{with } m = 0, 1, \ldots, M \tag{4}$$

for a proper choice of coefficients $c_{m,n}$.

The stream of differentiated Diracs $p^{(R)}(t)$ can be expressed as follows:

$$p^{(R)}(t) = \sum_{k=0}^{K-1} \sum_{r=0}^{R-1} a_{k,r}\delta^{(r)}(t - t_k), \tag{5}$$

and the samples $z_n^{(R)}$ can be written as:

$$z_n^{(R)} = \langle p^{(R)}(t), \tilde{\varphi}_{\text{eq}}(t/T - n)\rangle. \tag{6}$$

Consider now the weighted sum of the samples $z_n^{(R)}$: $\tau_m = \sum_n c_{m,n} z_n^{(R)}$, where the weights $c_{m,n}$ are those in (4) that reproduce $\mathrm{e}^{j\omega_m t/T}$. We can prove that:

$$\tau_m = \sum_{k=0}^{K-1} \sum_{r=0}^{R-1} \hat{a}_{k,r}(m')^r u_k^{m'}, \qquad \begin{matrix} m = 0, 1, \ldots, M \text{ and} \\ m' = m - (M+1)/2, \end{matrix} \tag{7}$$

where $\hat{a}_{k,r} = (-\lambda/T)^r a_{k,r}$ and $u_k = \mathrm{e}^{\lambda t_k/T}$. See Appendix B for more details.

Given τ_m we can then recover the locations t_k's of the Diracs as well as the amplitudes a_k's using the Prony's method. This is explained in more detail in Appendix B.

When $\tilde{\varphi}_{\text{eq}}(t)$ is an exponential reproducing kernel we can retrieve the t_k's and $a_{k,r}$'s exactly and can achieve exact reconstruction of $p^{(R)}(t)$. For any other kernel $\tilde{\varphi}_{\text{eq}}(t)$ which does not satisfy the generalized Strang-Fix conditions we want to find a linear combination of $\tilde{\varphi}_{\text{eq}}(t)$ with its shifted versions that provides the best approximation to a specific exponential [28]. More precisely, we want to find coefficients $c_{m,n}$ such that:

$$\sum_{n\in\mathbb{Z}} c_{m,n}\tilde{\varphi}_{\text{eq}}(t - n) \approx \mathrm{e}^{j\omega_m t}. \tag{8}$$

For the sake of clarity, we use $c_{m,n} = c_0\,\mathrm{e}^{j\omega_m n}$ and then we can show that the error in approximating the exponential is [28]:

$$\epsilon_{\text{approx},m}(t) = \mathrm{e}^{j\omega_m t}[1 - c_0 \sum_{l\in\mathbb{Z}} \hat{\tilde{\varphi}}_{\text{eq}}(j(\omega_m + 2\pi l))\,\mathrm{e}^{j2\pi l t}]. \tag{9}$$

Note that if the Fourier transform of $\tilde{\varphi}_{\text{eq}}(t)$ decays sufficiently quickly, which is true for any low-pass filter, we can assume the terms $\hat{\tilde{\varphi}}_{\text{eq}}(j(\omega_m + 2\pi l))$ are close to zero for $l \in \mathbb{Z}\backslash\{0\}$. In this case, the approximation error is small and is minimised when $c_{m,n} = \hat{\tilde{\varphi}}_{\text{eq}}(j\omega_m)^{-1}\,\mathrm{e}^{j\omega_m n}$. This approximate Strang-Fix theory enables accurate reconstruction of the input FRI signal with arbitrary sampling kernels. In particular, it shows that any acquisition device that behaves approximately like a low-pass filter can be used in the FRI framework and this fact will be particularly useful for the rest of the paper since point spread functions are approximately low-pass filters. We conclude by summarizing FRI method discussed so far in Algorithm 1.

III. Sampling and Resolution Enhancement of 1-D Piecewise Smooth Signals

We now go back to the original problem of enhancing the resolution of images and make the following observations. First of all we note that the image formation process in a digital camera (refer to Fig. 1) can be seen as a 2-D version

Algorithm 1: FRI method for reconstructing piecewise polynomial signals

input : samples y_n
output: reconstruction of $p(t)$

1 Calculate R-th finite difference $z_n^{(R)}$ of the samples y_n (refer to (2)).
2 Compute $\tau_m = \sum_n c_{m,n} z_n^{(R)}$, with $c_{m,n} = \hat{\tilde{\varphi}}_{\text{eq}}(j\omega_m)^{-1}\, \mathrm{e}^{j\omega_m n}$.
3 Apply Prony's method to the sequence τ_m to solve for locations of discontinuity and then for the amplitudes.

of the sampling set-up of Fig. 4 where the sampling kernel now is the point spread function of the camera. Moreover, images are piecewise regular functions (see Fig. 2), therefore, enhancing images to infinite resolution can be interpreted as the problem of sampling and reconstructing 2-D piecewise smooth functions.

In this section, we consider a 1-D version of this problem and discuss the 2-D case in Sec. IV. We consider the sampling and reconstruction of 1-D continuous-time functions first and then the resolution enhancement of 1-D discrete-time signals.

A. Sampling of Piecewise Smooth Signals

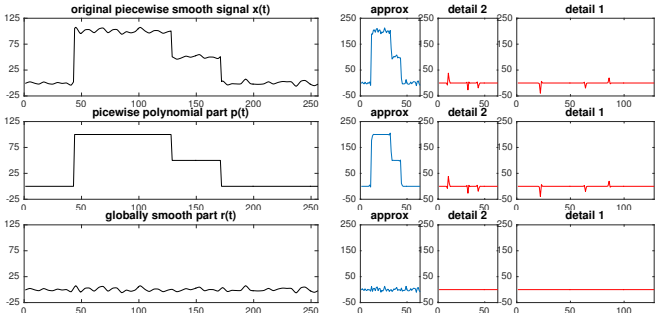
We consider the case where the sampling kernel (or point spread function in 2-D case) is the scaling function of a wavelet transform because the wavelet framework naturally relates the linear and FRI non-linear reconstruction methods to the notion of resolution enhancement and therefore provides a proper way to combine them. Here we emphasize that the FRI sampling method itself is universal since it works with any kernel.

Now we provide an interpretation of the problem of sampling an input signal $x(t)$ from the multi-resolution representation of $x(t)$. Denote with $\varphi(t)$ and $\psi(t)$ the scaling and wavelet functions respectively, and with $\varphi_{J,n}(t) = 2^{-J/2}\varphi(2^{-J}t - n)$ and $\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n), \quad J, m, n \in \mathbb{Z}$ the set of dilated and shifted versions of the scaling and wavelet function. Consider the following multi-resolution representation of a signal $x(t)$ in terms of the scaling and wavelet functions:

$$x(t) = \underbrace{\sum_{n=-\infty}^{\infty} y_{J,n}\varphi_{J,n}(t)}_{x_J(t)} + \sum_{m=-\infty}^{J}\sum_{n=-\infty}^{\infty} d_{m,n}\psi_{m,n}(t), \tag{10}$$

where $y_{J,n} = \langle x(t), \tilde{\varphi}_{J,n}\rangle$ and $d_{m,n} = \langle x(t), \tilde{\psi}_{m,n}\rangle$. Here $\tilde{\varphi}_{J,n}$, $\tilde{\psi}_{m,n}$ are the dual bases of $\varphi_{J,n}$ and $\psi_{m,n}$ respectively. We also note that $x_J(t)$ in (10) represents an approximation of $x(t)$ at resolution 2^J. Adding more and more levels of details $\sum_{n=-\infty}^{\infty} d_{m,n}\psi_{m,n}(t)$ to the coarse version $x_J(t)$ gives finer and finer resolution approximations and eventually the original signal $x(t)$.

The inner products $y_{J,n} = \langle x(t), \tilde{\varphi}_{J,n}\rangle$ are equivalent to the samples obtained by sampling $x(t)$ with sampling kernel $\tilde{\varphi}_{J,n}$ and sampling period $T = 2^J$ (see Fig. 4). Moreover, the coarse approximation $x_J(t) = \sum_{n=-\infty}^{\infty} y_{J,n}\varphi_{J,n}(t)$ in (10) corresponds to the linear reconstruction process in classical sampling theory discussed in Sec. II-A (see also Fig. 5), which finds projection of $x(t)$ onto the shift-invariant subspace spanned by $\{\varphi_{J,n}(t)\}_{n\in\mathbb{Z}}$ [30]. However, for the purpose of resolution enhancement, we are after an algorithm which is able to recover details from $y_{J,n}$ which are beyond the coarse approximation.



(a) (b)

Fig. 7. The details we need for resolution enhancement are due only to the piecewise polynomial part. (a) The piecewise smooth signal $x(t) = p(t) + r(t)$, the piecewise polynomial part $p(t)$ and the globally smooth part $r(t)$. (b) Wavelet decomposition of $x(t)$, $p(t)$ and $r(t)$ respectively.

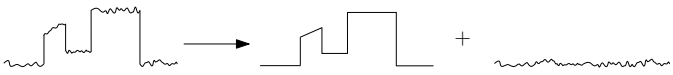


Fig. 6. We model piecewise smooth signals by the sum of a piecewise polynomial signal and a globally smooth signal.

We assume $x(t)$ is piecewise smooth and model piecewise smooth functions as the combination of a piecewise polynomial signal $p(t)$ and a globally smooth function $r(t)$ (see Fig. 6). We assume the smooth part $r(t)$ lives in the shift-invariant subspace generated by integer shifts of $\varphi(2^{-J}t)$. The piecewise smooth function can then be expressed as:

$$\begin{aligned} x(t) &= p(t) + r(t) \\ &= \underbrace{\sum_{n=-\infty}^{\infty} y^p_{J,n}\varphi_{J,n}(t) + \sum_{m=-\infty}^{J}\sum_{n=-\infty}^{\infty} d^p_{m,n}\psi_{m,n}(t)}_{p(t)} \\ &\quad + \underbrace{\sum_{n=-\infty}^{\infty} y^r_{J,n}\varphi_{J,n}(t)}_{r(t)} \\ &= \sum_{n=-\infty}^{\infty} \underbrace{(y^p_{J,n} + y^r_{J,n})}_{y_{J,n}}\varphi_{J,n}(t) + \sum_{m=-\infty}^{J}\sum_{n=-\infty}^{\infty} d^p_{m,n}\psi_{m,n}(t). \end{aligned} \tag{11}$$

By comparing (11) with (10), we notice that the details $d^p_{m,n}$ we need for resolution enhancement are due only to the piecewise polynomial part $p(t)$ (see also Fig. 7), and the remaining part can simply be obtained through linear recovery using coefficients $y_{J,n}$.

This observation together with the fact that the approximate reproduction formula of (8) enables reconstructing approximately piecewise polynomial signals using any kernel $\tilde{\varphi}(t)$,

leads to the proposed hybrid reconstruction strategy highlighted in Fig. 8, which recovers $p(t)$ using FRI method and $r(t)$ using the traditional linear reconstruction approach.

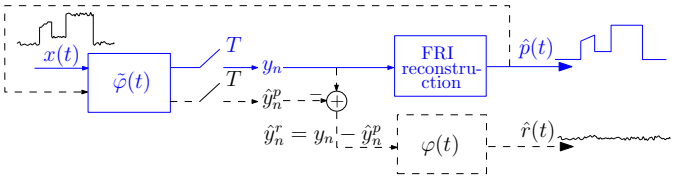


Fig. 8. The Schematic diagram of our proposed sampling and reconstruction strategy for piecewise smooth signals. The blue solid part: FRI reconstruction of the piecewise polynomial function $p(t)$. The black dashed part: linear reconstruction of the smooth residual $r(t)$.

Specifically, our proposed scheme first reconstructs the piecewise polynomial part $p(t)$ using the approximate Strang-Fix theory of Algorithm 1 from the samples $y_{J,n}$ by treating the globally smooth residual as noise (the blue solid part in Fig. 8). Given the estimated picewise polynomial $\hat{p}(t)$, we compute $\hat{y}^p_{J,n} = \langle \hat{p}(t), \tilde{\varphi}_{J,n}(t) \rangle$ which can then be removed from the samples $y_{J,n}$ to obtain the contribution $\hat{y}^r_{J,n} = y_{J,n} - \hat{y}^p_{J,n}$ due to the smooth residual $r(t)$. Then $r(t)$ can be reconstructed by classical linear method using the dual of the sampling kernel, i.e. $r(t) = \sum_{n=-\infty}^{\infty} \hat{y}^r_{J,n} \varphi_{J,n}(t)$ (the black dashed part in Fig. 8). The estimation of $x(t)$ is then the summation of estimated piecewise polynomial and estimated smooth part.

B. Resolution Enhancement of 1-D Piecewise Smooth Signal

Given the discrete-time sequence $y_{J,n} = \langle x(t), \tilde{\varphi}_{J,n} \rangle$, rather than trying to reconstruct the original continuous-time signal $x(t)$, one could be more interested in just trying to enhance the resolution of $y_{J,n}$ and the multi-resolution decomposition of (10) provides the right framework to achieve this goal. Assume that $\tilde{\varphi}(t)$ is a valid scaling function satisfying the two-scale relation:

$$\tilde{\varphi}(t) = \sqrt{2} \sum h_0[n] \tilde{\varphi}(2t - n), \tag{12}$$

and that we aim to enhance the resolution of $y_{J,n}$ by a factor 2^K for some positive integer K. It is then natural to seek for the signal $y_{J-K,n}$ which corresponds to the sequence obtained by sampling $x(t)$ with scaling function $\tilde{\varphi}_{J-K}(t)$ at finer scale 2^{J-K}.

Because of the two-scale equation (12) we can relate $y_{J,n}$ to $y_{J-K,n}$ using the K-level filter bank of Fig. 9. More precisely,

$$y_J = (y_{J-K} * h_0^{(K)}) \downarrow^{2^K}, \tag{13}$$

where $h_0^{(K)}$ in z-domain is $H_0^{(K)}(z) = H_0(z)H_0(z^2)\dots H_0(z^{2^{K-1}})$ and is derived based on the fact that the cascade of K analysis filters $H_0(z)$ each followed by subsampling by 2 is equivalent to one filter $H_0^{(K)}(z) = H_0(z)H_0(z^2)\dots H_0(z^{2^{K-1}})$ followed by subsampling by 2^K. Here $H_0(z)$ is the z-transform of h_0 in (12).

Therefore, one reasonable linear upsampling of $y_{J,n}$ is the version obtained by simply feeding $y_{J,n}$ to the wavelet reconstruction stage, and can be expressed as:

$$\hat{y}_{J-K} = y_J \uparrow^{2^K} * g_0^{(K)}, \tag{14}$$

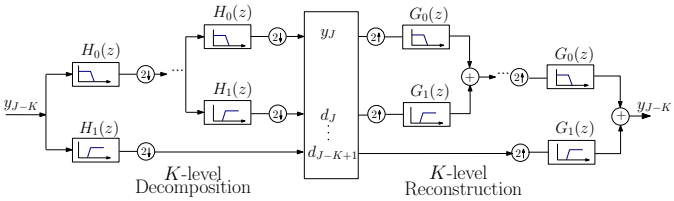


Fig. 9. K-level biorthogonal filter bank. Given the approximation coefficients $y_J = \langle x(t), \tilde{\varphi}_{J,n} \rangle$ we are looking for a higher resolution version $y_{J-K} = \langle x(t), \tilde{\varphi}_{J-K,n} \rangle$.

where $g_0^{(K)}$ in z-domain is $G_0^{(K)}(z) = G_0(z)G_0(z^2)\dots G_0(z^{2^{K-1}})$ and $\hat{y}_{J-K}$ is the projection of y_{J-K} to the subspace spanned by $\{g_0^{(K)}[n - 2^K k]\}_{k \in \mathbb{Z}}$. Here $G_0(z)$ is the synthesis low-pass filter.

However, as Fig. 9 also indicates, this linear reconstruction does not allow us to retrieve the missing detail coefficients $d_J \dots d_{J-K+1}$, so in order to get a better estimate of y_{J-K} we estimate the detail coefficients using FRI.

More specifically, knowing the low-pass filter $\tilde{\varphi}_J(t)$ we first apply FRI method of Algorithm 1 to estimate the piecewise polynomial part $p(t)$ from the approximation coefficients $y_{J,n}$, with the assumption that the contribution to $y_{J,n}$ only comes from $p(t)$ and that the smooth part $r(t)$ is noise. We call this estimated piecewise polynomial signal $\hat{p}^{\text{FRI}}(t)$. We then put $\hat{p}^{\text{FRI}}(t)$ onto a grid of resolution 2^{J-K-L} with $L > 0$, which is a grid finer than the resolution 2^{J-K} we are looking for, and we denote this discretized polynomial with $\hat{p}^{\text{FRI}}_{J-K-L}[n]$. Recall that in our piecewise smooth model the detail coefficients are due only to the piecewise polynomial, we therefore obtain the detail coefficients at resolution 2^J to 2^{J-K+1} from $\hat{p}^{\text{FRI}}_{J-K-L}[n]$ through $(K+L)$-level filter-bank decomposition. Finally the estimation of $y_{J-K,n}$ is obtained by computing K-level wavelet reconstruction using the approximation coefficients $y_{J,n}$ and the estimated detail coefficients $d_J[n] \dots d_{J-K+1}[n]$. We summarize this resolution enhancement method in Fig. 10.

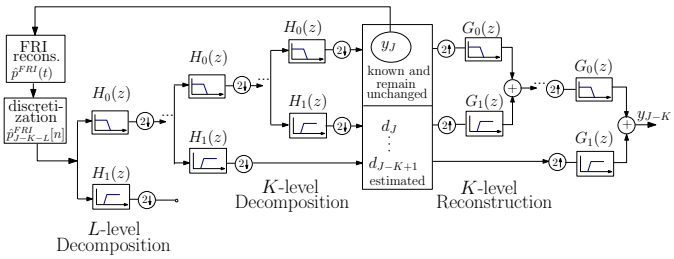
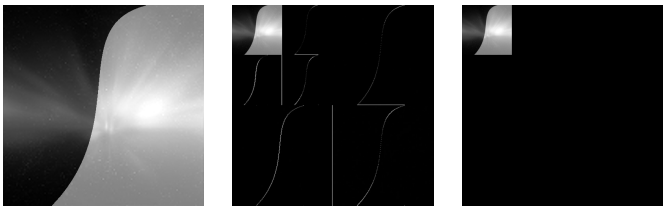


Fig. 10. The schematic diagram of resolution enhancement of a piecewise smooth signal by factor of 2^K using a biorthogonal filter bank. From the given approximation coefficients $y_J = \langle x(t), \tilde{\varphi}_{J,n} \rangle$ and the detail coefficients $d_J, \dots, d_{J-K+1}$ estimated using FRI, we are able to recover a higher resolution version $y_{J-K} = \langle x(t), \tilde{\varphi}_{J-K,n} \rangle$.

IV. Image Up-sampling

Equipped with the resolution enhancement method of the previous section, we now approach image upsampling by modeling lines (along different directions) of images as 1-D piecewise smooth functions and extend the method of 1-D

case to 2-D images. For clarity and simplicity, we denote the image at original low-resolution with y_0 and its upsampled version by factor 2^K with y_{-K}. The low-resolution image y_0 of size $N \times N$ is the low-pass version of a K-level 2D wavelet transform applied to the high-resolution image y_{-K} of size $2^K N \times 2^K N$ with all the high-pass coefficients discarded (see Fig. 11).



(a) The high-resolution piecewise smooth image

(b) Low-pass and high-pass subbands of a 2-level 2D wavelet transform of (a)

(c) We only have access to the low-pass subband in (b)

Fig. 11. Assume we only have access to the low-pass subband y_0 of a 2D wavelet transform applied to the high-resolution image y_{-K}. We want to estimate the high-pass coefficients using FRI in order to recover the high-resolution image.

The 2D wavelet decomposition leads to a set of approximation coefficients and three types of high-pass coefficients, representing horizontal, vertical and diagonal high-frequency details respectively. Similar to the 1D case, a simple linear reconstruction of y_{-K} from approximation coefficients y_0 merely increases the size of y_0 without increasing the resolution. The goal is to estimate the high-pass coefficients for the purpose of resolution enhancement.

In part A, we introduce the basic FRI image upsampling algorithm, and we explain in part B an improved method which uses the basic algorithm and also exploits patch repetitions across scales to correct errors of the basic upsampling method.

A. Basic Image Up-sampling Algorithm

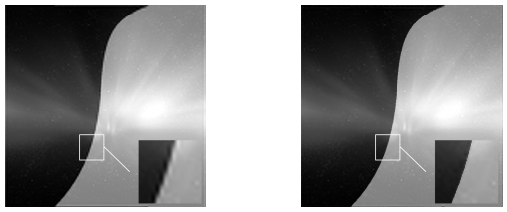
The basic idea is that high-pass coefficients can be estimated from the set of FRI recovered images (stacks of high-resolution piecewise polynomials) along horizontal, vertical and diagonal directions. The proposed image upsampling method is summarized in the block diagram of Fig. 12 and described in further detail in this section.

We first linearly interpolate column by column the low-resolution image y_0 to size $2^K N \times N$ using the synthesis filter $g_0^{(K)}$ (refer to (14)). Then from the N samples of each horizontal line we reconstruct a piecewise polynomial of length $2^K N$ using the resolution enhancement method described in Sec. III-B. In this way we get an image $\hat{y}^{\text{FRI}}_{-K,h}$ of size $2^K N \times 2^K N$ which contains estimates of the vertical edges of y_{-K}. Similarly, we operate on the other coordinate of y_0 to recover an image $\hat{y}^{\text{FRI}}_{-K,v}$ that contains horizontal edges.

Then we obtain a first reconstruction $\hat{y}^{\text{FRI}}_{-K}$ by inputting $\hat{y}^{\text{FRI}}_{-K,h}$ to the decomposition channels that capture vertical edges, $\hat{y}^{\text{FRI}}_{-K,v}$ to the channels that capture horizontal edges and either one of them to the channels that capture diagonal edges.

We then improve this reconstruction by operating diagonally. We do this to remove jaggies in $\hat{y}^{\text{FRI}}_{-K}$. We first downsample each 45 degree diagonal line of $\hat{y}^{\text{FRI}}_{-K}$ by factor 2, and then recover it with FRI and we call the recovered image $\hat{y}^{\text{FRI}}_{-K,d1}$. We then do the same on -45 degree diagonal lines of $\hat{y}^{\text{FRI}}_{-K}$ and obtain the image $\hat{y}^{\text{FRI}}_{-K,d2}$. Then we reconstruct a final image by selecting patch by patch (size 4×4 with 1-pixel overlap) from $\hat{y}^{\text{FRI}}_{-K,d1}$ and $\hat{y}^{\text{FRI}}_{-K,d2}$. For patch with dominant gradient direction closer to 45 degrees we use the patch from $\hat{y}^{\text{FRI}}_{-K,d1}$ and otherwise from $\hat{y}^{\text{FRI}}_{-K,d2}$. Then we ensure the consistency between our reconstruction and the input data y_0 by replacing the approximation coefficients of our reconstruction with y_0. The final unsampled image with upsampling factor 2^K is denoted with y^{FRI}_{-K}.

Fig. 13 demonstrates that our proposed method is able to upsample a piecewise smooth image with sharp edges. It is evident that we gain by adding our estimated high-frequency information.



(a) linear reconstruction, PSNR=36.49dB

(b) our reconstruction, PSNR=42.81dB

Fig. 13. Upsampling results of a piecewise smooth image by linear reconstruction and by proposed method.

B. Exploiting Cross Scale Similarities

Inevitably, there are errors in our FRI reconstruction y^{FRI}_{-K}, some are due to imperfection of the piecewise smooth model—for example not all edges behave like an immediate transition between two nearby pixels. Inspired by the idea of deriving a HR patch from an input LR patch with a linear transformation learnt from internal LR and HR dictionary patches [19], we propose estimating and correcting the error in upsampled FRI image by learning the relationship between the ground truth input LR image and our FRI reconstruction recovered from an even lower resolution version of the input LR image. More precisely, a specific patch in y^{FRI}_{-1} is expected to be corrected by a linear transformation $\mathbf{M}$ which transforms its similar patches in y^{FRI}_{-1+m} to corresponding patches in y_{-1+m}, where y_{-1+m} is an intermediate scale 1.25 times smaller than y_{-1} (1.6 times larger than y_0). This is possible because there are patch repetitions across small-scale factors (typically 1.25). This is also the prior typically used in most single-image super-resolution techniques without external dictionary [17]–[19].

Now we explain in detail our proposed algorithm for upsampling by 2. Its main idea is depicted in Fig. 14. For an upsampling factor $2^K (K > 1)$, we iterate the algorithm K times.

First of all, we use the basic FRI upsampling method explained in Section IV-A to upsample input low-resolution image y_0 to y^{FRI}_{-1} (see Fig. 14a).

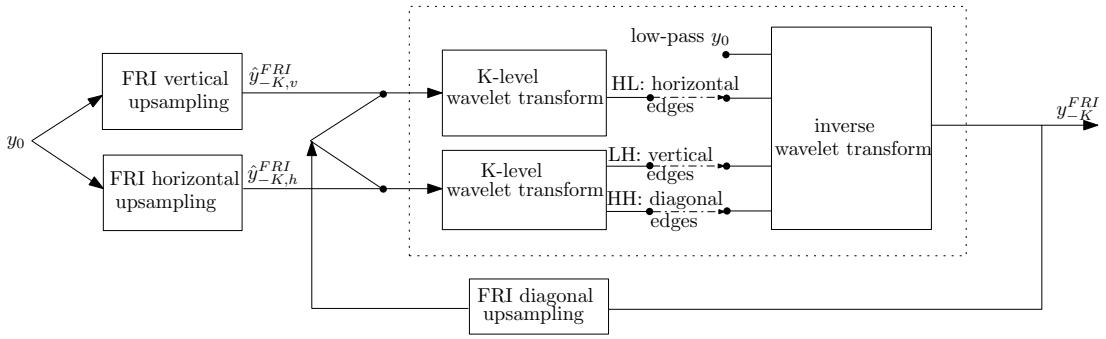


Fig. 12. The block diagram of our proposed upsampling scheme.

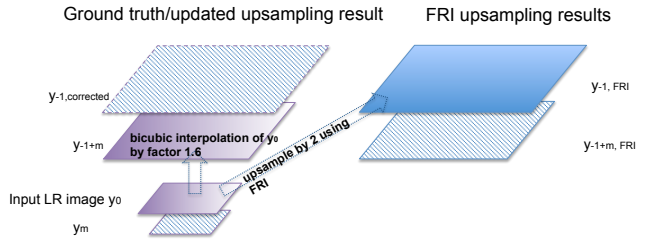
Then for updating the y^{FRI}_{-1} we try to create the FRI image y^{FRI}_{-1+m} and ground truth image y_{-1+m} pair, whose resolution is slightly lower than y_{-1}. A temporary version of y_{-1+m} is obtained by bicubic interpolation of y_0 by factor 1.6 (see Fig. 14a), and its corresponding FRI image y^{FRI}_{-1+m} is obtained by downsampling y_{-1+m} by factor 2 followed by upsampling by 2 using basic FRI image upsampling method of Section IV-A (see Fig. 14b).

Having obtained the first version of y_{-1+m} and y^{FRI}_{-1+m} pair, we use them to update y^{FRI}_{-1} as follows (see Fig. 14c): first, for each patch $b^{\text{FRI}}_{-1,i}$ in y^{FRI}_{-1} (size 5×5 with 1-pixel shift each time), we search for P (e.g. 4) similar patches $\{b^{\text{FRI}}_{-1+m,i,j}\}_{j=1}^{P}$ in y^{FRI}_{-1+m}. We do the search locally within a small window of 25 by 25 pixels centered around the relative center-coordinates of $b^{\text{FRI}}_{-1,i}$. Now we compute the transformation $\mathbf{M}_i \in \mathbb{R}^{25\times 25}$ that maps the P vectorized patches in y^{FRI}_{-1+m} to the corresponding P vectorized patches in y_{-1+m}. This tranformation $\mathbf{M}_i$ is expected to correct $b^{\text{FRI}}_{-1,i}$ and the way to compute it is explained in detail later. We apply the $\mathbf{M}_i$ learnt to correct $b^{\text{FRI}}_{-1,i}$:

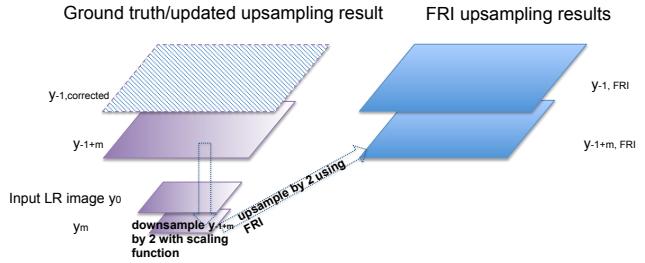
$$b^{\text{corrected}}_{-1,i}(:) = \mathbf{M}_i b^{\text{FRI}}_{-1,i}(:), \tag{15}$$

where $b(:)$ denote the vectorized version of patch b. All the corrected patches $b^{\text{corrected}}_{-1,i}$ are then combined to obtain a corrected high-resolution image $y^{\text{corrected}}_{-1}$ by averaging contributing patch values at each pixel. Then we ensure data fidelity by replacing the low-pass coefficients of $y^{\text{corrected}}_{-1}$ with the ground-truth y_0.

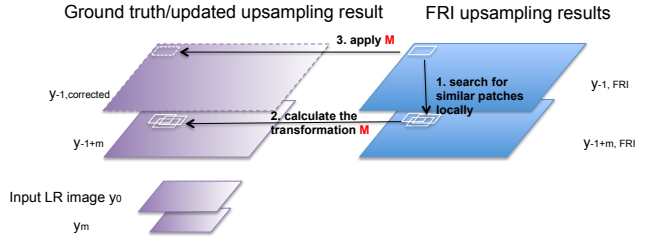
Lastly, we want to update y_{-1+m}, y^{FRI}_{-1+m} and subsequently $y^{\text{corrected}}_{-1}$ because the current y_{-1+m} we learnt from is simply a bicubic interpolation of y_0 which is blurred and is not well served as the ground truth image at resolution 2^{-1+m}. So we update it (see Fig. 14d) by downsampling $y^{\text{corrected}}_{-1}$ by factor 1.25 using bicubic interpolation. Its corresponding FRI version y^{FRI}_{-1+m} is updated by downsampling y_{-1+m} by 2 followed by upsampling by 2 using the basic FRI image upsampling method. Then given the new y_{-1+m} and y^{FRI}_{-1+m} pair we re-calculate the transformation $\mathbf{M}_i$ for each patch $b^{\text{FRI}}_{-1,i}$ in y^{FRI}_{-1} and apply the new $\mathbf{M}_i$ to $b^{\text{FRI}}_{-1,i}$ as in (15). Again, we combine all the patches and ensure the low-pass version of $y^{\text{corrected}}_{-1}$ is y_0. We note that we could further repeat this updating step. However, we have numerical evidence that further iterations would not improve the result significantly, so we stick with one updating iteration.



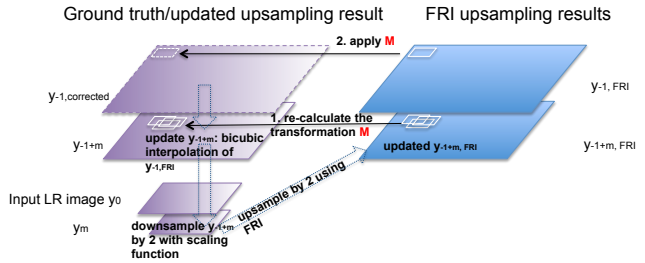
(a) Creating y^{FRI}_{-1} and the first version of the intermediate level y_{-1+m}.



(b) Creating the FRI version of the intermediate level, y^{FRI}_{-1+m}.



(c) Correcting y^{FRI}_{-1} using the intermediate image pair y_{-1+m} and y^{FRI}_{-1+m} and obtaining $y^{\text{corrected}}_{-1}$.



(d) Updating the intermediate image pair y_{-1+m} and y^{FRI}_{-1+m} followed by updating $y^{\text{corrected}}_{-1}$.

Fig. 14. The schematic diagram of how to correct our FRI upsampled image y^{FRI}_{-1}.

The linear transformation $\mathbf{M}_i \in \mathbb{R}^{D^2\times D^2}$ that maps the P FRI patches $\{b^{\text{FRI}}_{-1,i,j} \in \mathbb{R}^{D\times D}\}_{j=1}^{P}$ to ground truth patches $\{b_{-1,i,j} \in \mathbb{R}^{D\times D}\}_{j=1}^{P}$ can be found by minimizing the empirical fitting error between all P pairs of examples. However, this problem is underdetermined and Tikhonov regularization is added to solve it stably:

$$\begin{aligned}\mathbf{M}_i &= \underset{\mathbf{M}_i\in\mathbb{R}^{D^2\times D^2}}{\operatorname{argmin}} \sum_{j=1}^{P}\|b_{i,j}(:) - \mathbf{M}b^{\text{FRI}}_{i,j}(:)\|_2^2 + \lambda\|\mathbf{M}\|_F^2 \\ &= \underset{\mathbf{M}_i\in\mathbb{R}^{D^2\times D^2}}{\operatorname{argmin}} \|\mathbf{B}_i - \mathbf{M}\mathbf{B}^{\text{FRI}}_i\|_2^2 + \lambda\|\mathbf{M}\|_F^2,\end{aligned} \tag{16}$$

where λ is a regularization parameter, $\mathbf{B}_i$ and $\mathbf{B}^{\text{FRI}}_i$ are matrices with $\{b_{i,j}(:)\}_{j=1}^{P}$ and $\{b^{\text{FRI}}_{i,j}(:)\}_{j=1}^{P}$ as their columns respectively. The solution of (16) can be written in a close-form as follows:

$$\mathbf{M}_i = \mathbf{B}_i\mathbf{B}^{\text{FRI}\top}_i\left(\mathbf{B}^{\text{FRI}}_i\mathbf{B}^{\text{FRI}\top}_i + \lambda\mathbf{I}\right)^{-1}, \tag{17}$$

where $\mathbf{I}$ is the identity matrix.

To conclude, we summarize the complete upsampling method in Algorithm 2.

Algorithm 2: FRI-based single image super-resolution algorithm

input : y_0
output: y_{-1}: upsampled version of y_0

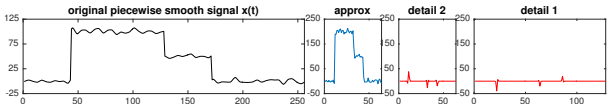
1 Upsample y_0 to y^{FRI}_{-1} by the basic FRI algorithm in part A.
2 Create the intermediate level y_{-1+m} by upsampling y_0 by a factor 1.6 using bicubic interpolation
3 Create the corresponding intermediate FRI level y^{FRI}_{-1+m} by first downsampling y_{-1+m} by 2 followed by upsampling it using the basic FRI algorithm.
4 **for** *patch* $b^{FRI}_{-1,i}$ *in* y^{FRI}_{-1} **do**
5 Search for P similar patches $\{b^{\text{FRI}}_{-1+m,i,j}\}_{j=1}^{P}$ in y^{FRI}_{-1+m}, locally within the small window centered around the relative center-coordinates of $b^{\text{FRI}}_{-1,i}$.
6 Calculate the linear transformation $\mathbf{M}_i$ that transforms the patches $\{b^{\text{FRI}}_{-1+m,i,j}\}_{j=1}^{P}$ to the corresponding ground truth patches $\{b_{-1+m,i,j}\}_{j=1}^{P}$ in y_{-1+m} (refer to (17)).
7 Correct $b^{\text{FRI}}_{-1,i}$ by applying the linear transformation $\mathbf{M}_i$ to it (refer to (15)).
8 **end**
9 Combine all the corrected patches and replace the low-pass coefficients of the corrected y^{FRI}_{-1} with the ground-truth y_0, and call the image $y^{\text{corrected}}_{-1}$.
10 Update the intermediate level y_{-1+m} by downsampling y^{FRI}_{-1} with bicubic interpolation.
11 Update y^{FRI}_{-1+m} by downsampling the updated y_{-1+m} by 2 followed by upsampling by 2 with the basic FRI algorithm.
12 Repeat step 4 to 9 with the updated pair of intermediate images except step 5 need not to be recalculated.

V. Simulation Results

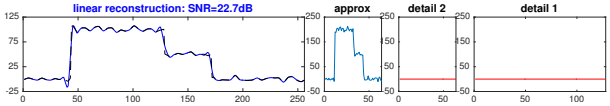
A. 1-D Piecewise Smooth Signal Upsampling

In the following simulations, we show the resolution enhancement results using our novel hybrid reconstruction method introduced in Sec. III-B. In this section we assume our samples y_0 (refer to Fig. 9) are the low-pass coefficients of 2-level wavelet decomposition applied to a high-resolution piecewise smooth signal and we want to recover it to its original resolution.

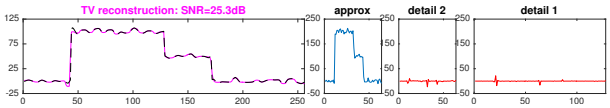
First, we demonstrate in Fig. 15 that when the signal is exactly the discrete-time version of model (11), our method, compared to the linear reconstruction and the total variation method, is able to achieve nearly perfect reconstruction.



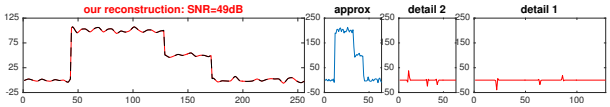
(a) the original high-resolution piecewise smooth signal and its wavelet decomposition



(b) the linear reconstruction (22.7dB) and its wavelet decomposition



(c) TV reconstruction (25.3dB) and its wavelet decomposition



(d) our reconstruction (49.0dB) and its wavelet decomposition

Fig. 15. Our method is able to accurately recover a piecewise smooth signal from its approximation coefficients.

Then we also test our algorithm in the case where the high-resolution signal is not exactly in our model but is a scan-line of an image (see Fig. 2). The result of comparisons between our reconstruction, the linear reconstruction, and the total variation method in Table I shows that our model is an accurate representation of the scan-lines such that our proposed method is still efficient in this case. The result also proves the universality of our method in the sense that it works robustly with different downsampling kernels.

TABLE I
Recovering the high-resolution image scan-line from its approximation coefficients of different wavelet decomposition with different methods.

scaling function	linear reconstr.	TV reconstr.	our reconstr.
linear spline	21.7 dB	23.1 dB	24.2 dB
cubic spline	22.2 dB	23.1 dB	24.4 dB
bior4.4	23.9 dB	23.0 dB	25.0 dB

B. Image Upsampling

1) Upsampling of artificially downsampled images: In this part, we test upsampling of low-resolution images obtained by downsampling the original ones by a factor 4 using the 2-level 2-D wavelet decomposition with biorthogonal 4.4 filter, and those obtained by downsampling with a linear spline. We compare our basic upsampling method in Sec. IV-A and our improved method in Sec. IV-B with the linear reconstruction method and some of the state-of-the-art algorithms, we show the upsampling results of two different downsampling kernels in terms of PSNR and SSIM (structural similarity index [31]) in Table II and Table III respectively. Visual comparisons on 'Zebra' (downsampling kernel of bior4.4) and 'Comic' (downsampling kernel of linear spline) are shown in Fig. 16 and Fig. 17 respectively. Note that for self-learning method [17], we use a third-party implementation [32] and we cannot guarantee the implementation duplicates the original results. For other methods we ensure the comparison is fair because we modified the blurring kernel in the source codes to the specific kernel we use (bior4.4/linear spline), and for dictionary-based methods of [15], [16], the dictionary was re-trained with the same kernel used in upsampling process.

Our proposed method is universal in the sense that it works with different blurring kernels. The results demonstrate that our proposed basic (fast) method, with no learning involved, outperforms other reconstruction-based algorithms, e.g. total variation [6], contourlet [7] and even one of the dictionary learning methods [15]. The improved method – FRESH is robust and outperforms state-of-the-art methods in different categories. Moreover, our method leads to visually pleasant edges.

TABLE II
COMPARISONS OF UPSAMPLING RESULTS (FACTOR 4) GIVEN BY DIFFERENT METHODS IN TERMS OF PSNR. SAMPLING KERNEL: BIOR4.4.

PSNR (dB) & SSIM [31]	linear	TV [6]	contourlet[a] [7]	sparse coding [15]	A+ [16]	self-learn [17]	our basic method	FRESH
Peppers	29.91	30.98	30.19	31.03	31.73	31.31	31.13	**31.85**
	0.823	0.835	0.824	0.837	**0.846**	0.837	0.836	**0.845**
Lena	29.49	29.97	29.82	30.17	**30.64**	29.99	30.16	30.55
	0.835	0.837	0.839	0.844	**0.853**	0.843	0.841	0.849
Cameraman	28.35	28.85	28.59	29.21	29.63	29.04	29.16	**29.82**
	0.872	0.883	0.872	0.885	0.892	0.886	0.886	**0.894**
Butterfly	21.50	23.76	21.66	22.37	23.19	**24.41**	23.40	24.24
	0.744	0.853	0.736	0.796	0.836	0.859	0.833	**0.864**
Bird	29.37	29.99	29.75	30.23	30.94	30.04	30.29	**31.04**
	0.877	0.887	0.881	0.893	**0.905**	0.888	0.891	**0.904**
Comic	20.83	21.10	–	21.07	21.29	21.00	21.13	**21.32**
	0.621	0.647	–	0.636	0.654	0.654	0.642	**0.659**
Zebra	23.61	24.34	–	24.38	24.71	24.44	24.50	**25.11**
	0.711	0.721	–	0.726	0.731	0.732	0.726	**0.737**
Woman	25.91	26.71	–	26.66	27.34	27.36	26.91	**27.72**
	0.842	0.857	–	0.861	0.876	0.872	0.859	**0.879**
Average	26.12	26.96	–	26.89	27.43	27.20	27.09	**27.71**
	0.791	0.815	–	0.810	0.824	0.821	0.814	**0.829**

[a] contourlet interpolation software does not support upsampling of non-square images.

TABLE III
COMPARISONS OF UPSAMPLING RESULTS (FACTOR 4) GIVEN BY DIFFERENT METHODS IN TERMS OF PSNR. SAMPLING KERNEL: LINEAR SPLINE.

PSNR (dB) & SSIM [31]	linear	TV [6]	contourlet[a] [7]	sparse coding [15]	A+ [16]	self-learn [17]	our basic method	FRESH
Peppers	29.95	31.09	30.25	30.77	31.63	31.49	31.28	**31.95**
	0.821	0.835	0.824	0.831	**0.844**	0.836	0.836	**0.845**
Lena	29.53	30.02	29.90	29.84	30.58	30.13	30.28	**30.66**
	0.834	0.838	0.84	0.835	**0.850**	0.837	0.842	**0.850**
Cameraman	28.40	28.89	28.66	28.88	29.54	29.27	29.34	**30.05**
	0.870	0.884	0.871	0.877	0.890	0.881	0.888	**0.896**
Butterfly	21.55	23.91	21.72	22.08	23.03	24.22	23.52	**24.61**
	0.737	0.856	0.736	0.782	0.830	0.860	0.835	**0.870**
Bird	29.42	30.05	29.82	29.86	30.73	30.23	30.43	**31.26**
	0.876	0.887	0.881	0.883	0.901	0.888	0.892	**0.906**
Comic	20.85	21.15	–	20.92	21.27	21.09	21.22	**21.41**
	0.619	0.648	–	0.616	0.647	0.632	0.644	**0.662**
Zebra	23.67	24.41	–	24.13	24.61	24.65	24.61	**25.30**
	0.711	0.722	–	0.709	0.728	0.728	0.728	**0.741**
Woman	25.95	26.78	–	26.36	27.22	27.23	27.04	**27.86**
	0.838	0.858	–	0.851	0.872	0.867	0.861	**0.881**
Average	26.17	27.04	–	26.60	27.33	27.29	27.22	**27.89**
	0.788	0.816	–	0.798	0.820	0.816	0.816	**0.831**

[a] contourlet interpolation software does not support upsampling of non-square images.

2) Upsampling of images taken with a camera: Finally, we show that the proposed algorithm is also able to upsample the images taken with a real camera, where the blurring due to lens is not exactly a scaling function as assumed previously but can still be modelled as a spline. We demonstrate in Fig. 18 that the algorithm achieves visually good performance for upsampling factor of 4. In the following result, the original photographs are taken with Canon 400D, and its point spread function is modelled by the fifth order spline. The upsampling is performed only on the luminance component of the input image and the chrominance component are simply upscaled by bicubic interpolation.

C. Computation complexity and discussions

Upsampling an image of size $N \times N$ to $2^K N \times 2^K N$ with the basic algorithm proposed in Sec. IV-A requires number of line upsampling operations in the order of $2^K N$ and also $2^K N \times 2^K N$ block selecting operations.

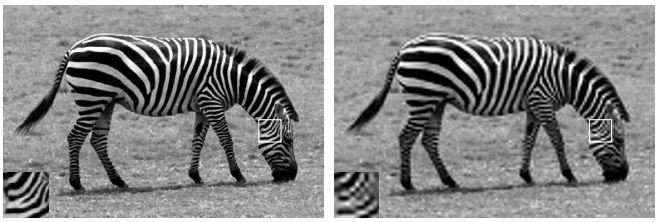
Unlike the basic algorithm, the improved algorithm proposed in Sec. IV-B requires upsampling one level per time (factor 2 each time) and involves searching for similar blocks, which is computational more expensive.

Table IV shows the execution time (averaged over 5 test images) of the C++ implementations of the basic method and FRESH on a Mac mini with 2.6GHz Intel Core i7 CPU and 16GB RAM.

TABLE IV
COMPUTATION COST FOR UPSAMPLING AN IMAGE USING THE BASIC METHOD AND FRESH ON A MAC MINI WITH 2.6GHZ INTEL CORE I7 CPU AND 16GB RAM.

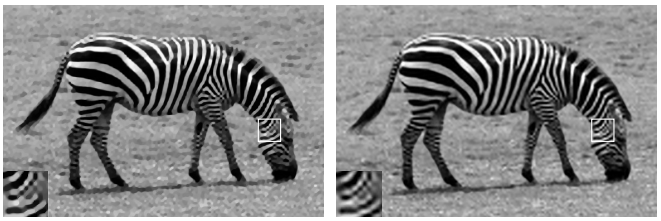
time (seconds)	upsample 64×64 images by 2	upsample 64×64 images by 4
basic algorithm	0.45	2.68
FRESH	1.65	8.95

Since we are using a wavelet scheme to handle the upsampling, the scale factor can only be powers of 2. Moreover, the point spread function of the camera needs to be fairly close to the scaling function of a wavelet transform. The wavelet scheme may seem to limit the classes of acquisition filters we are able to handle, nevertheless we want to emphasize that the point spread function of cameras in general can be accurately modelled by splines which are valid scaling functions and numerical results on images taken with real cameras confirm the universality of our approach.



(a) the original image

(b) linear reconstruction
PSNR=23.61dB, SSIM=0.711



(c) total variation [6]
PSNR=24.34dB, SSIM=0.721

(d) sparse coding [15]
PSNR=24.38dB, SSIM=0.726

(e) A+ [16]
PSNR=24.71dB, SSIM=0.731

(f) self-learning [17]
PSNR=24.44dB, SSIM=0.732

(g) our basic algorithm
PSNR=24.50dB, SSIM=0.726

(h) our improved algorithm
PSNR=25.11dB, SSIM=0.737

Fig. 16. Upsampling results (factor 4) of *Zebra* by different methods. The downsampling kernel is bior4.4.

VI. Conclusions

In this paper, we have proposed a scheme for upsampling piecewise smooth signals and its extension to images by modelling images as lines of piecewise smooth signals. We show that the method proposed improves classical linear reconstruction results by making use of an additional non-linear

(a) the original image

(b) linear reconstruction
PSNR=20.85dB, SSIM=0.619

(c) total variation [6]
PSNR=21.15dB, SSIM=0.648

(d) sparse coding [15]
PSNR=20.92dB, SSIM=0.616

(e) A+ [16]
PSNR=21.27dB, SSIM=0.647

(f) self-learning [17]
PSNR=21.09dB, SSIM=0.632

(g) our basic algorithm
PSNR=21.22dB, SSIM=0.644

(h) our improved algorithm
PSNR=21.41dB, SSIM=0.662

Fig. 17. Upsampling results (factor 4) of *Comic* by different methods. The downsampling kernel is linear spline.

(a) (b) (c)

(d) (e) (f)

Fig. 18. Upsampling results (factor 4) of images taken with Canon 400D. Our upsampling results are sharper than the bicubic interpolation results. (a)(d) original images. (b)(e) bicubic interpolation. (c)(f) our upsampling results.

reconstruction method based on FRI theory. The method is further improved by using a self-learning approach which also makes use of FRI. The resulting algorithm outperforms state-of-the-art methods and does not require the use of external datasets.

Appendix A
The Relationship Between Sampling Piecewise Polynomial and streams of Diracs

Consider a function $\varphi(t/T)$ with Fourier transform $T\hat{\varphi}(\omega T)$ and the difference $\varphi(t/T)-\varphi(t/T-1)$. The Fourier transform of $\varphi(t/T)-\varphi(t/T-1)$ is

$$\begin{aligned}\varphi(t/T)-\varphi(t/T-1) \Longleftrightarrow & T\hat{\varphi}(\omega T)(1-\mathrm{e}^{-j\omega T}) \\ =& j\omega T\hat{\varphi}(\omega T)\cdot T\frac{1-\mathrm{e}^{-j\omega T}}{j\omega T} \\ =& j\omega T\hat{\varphi}(\omega T)\cdot T\hat{\beta}_0(\omega T).\end{aligned} \tag{18}$$

Therefore

$$\varphi(t/T)-\varphi(t/T-1)=\frac{d}{dt}[\varphi(t/T)*\beta_0(t/T)], \tag{19}$$

where β_0 is called B-spline of order 0 and its Fourier transform is $\frac{1-\mathrm{e}^{-j\omega}}{j\omega}$. Let z_n denote the finite difference $y_{n+1}-y_n$. It follows that

$$\begin{aligned}z_n=y_{n+1}-y_n &= \langle x(t),\varphi(t/T-n-1)-\varphi(t/T-n)\rangle \\ &= \langle x(t),-\frac{d}{dt}[\varphi(t/T-n)*\beta_0(t/T-n)]\rangle \\ &= \langle \frac{dx(t)}{dt},[\varphi(t/T-n)*\beta_0(t/T-n)]\rangle.\end{aligned} \tag{20}$$

Using a similar derivation it is also possible to prove that the R-th finite difference of the samples taken with $\varphi(t)$ is equivalent to the set of samples obtained by sampling $p^{(R)}(t)$ with the new kernel $\varphi(t)*\beta_{R-1}(t)$, where $\beta_{R-1}(t)$ is the B-spline of degree $R-1$ and this new kernel still meets the Strang-Fix conditions. Therefore we are able to use the result of sampling stream of differentiated Diracs for sampling piecewise polynomials.

APPENDIX B
PRONY'S METHOD FOR RECOVERING A STREAM OF DIFFERENTIATED DIRACS

The weighted sum of the samples $z_n^{(R)}$: $\tau_m = \sum_n c_{m,n} z_n^{(R)}$ can be written in a power sum form:

$$\begin{aligned}\tau_m &= \sum_n c_{m,n} z_n^{(R)} = \langle p^R(t), \sum_n c_{m,n} \tilde{\varphi}_{\text{eq}}(t/T - n)\rangle \\ &= \int_{-\infty}^{\infty} p^{(R)}(t)\, \mathrm{e}^{j\omega_m t/T} dt \\ &= \int_{-\infty}^{\infty} \sum_{k=0}^{K-1} \sum_{r=0}^{R-1} a_{k,r} \delta^{(r)}(t - t_k)\, \mathrm{e}^{j\omega_m t/T}\, dt \\ &= \sum_{k=0}^{K-1} \sum_{r=0}^{R-1} \hat{a}_{k,r} (m')^r u_k^{m'}, \qquad \begin{array}{l} m = 0, 1, \ldots, M \text{ and} \\ m' = m - (M+1)/2, \end{array}\end{aligned} \tag{21}$$

where $\hat{a}_{k,r} = (-\lambda/T)^r a_{k,r}$ and $u_k = \mathrm{e}^{\lambda t_k/T}$.

The locations t_k's of the differentiated Diracs can be retrieved from τ_m using Prony's method (annihilating filter method). The key is to note that given a filter $\{h_m\}_{m=0}^{KR}$ whose z-transform is:

$$H(z) = \sum_{m=0}^{KR} h_m z^{-m} = \prod_{k=0}^{K-1} (1 - u_k z^{-1})^R, \tag{22}$$

then this filter can annihilate the sequence τ_m. That is, $h_m * \tau_m = 0$. The KR unknown coefficients of h_m can be found by writing $h_m * \tau_m = 0$ in matrix/vector form using at least $2KR$ consecutive τ_m. From the roots of the annihilating filter we obtain the locations t_k exactly. Then the exact amplitudes can be found by solving, for example, the first KR equations in (21).

ACKNOWLEDGMENT

We wish to thank Matteo Maggioni for the C++ implementation of the algorithm. Moreover, we would like to thank the anonymous reviewers for their comments and suggestions. This work is supported by the European Research Council (ERC) starting investigator award Nr. 277800 (RecoSamp). The material in this paper was presented in part at the International Conference of Image Processing, Quebec City, in September 2015 [33].

REFERENCES

[1] Sina Farsiu, Dirk Robinson, Michael Elad, and Peyman Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2, pp. 47–57, 2004.

[2] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, 1981.

[3] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited," *IEEE Transactions on Medical Imaging*, vol. 19, no. 7, pp. 739–758, 2000.

[4] M. F. Tappen, B. C. Russell, and W. T. Freeman, "Exploiting the sparse derivative prior for super-resolution and image demosaicing," in *In IEEE Workshop on Statistical and Computational Theories of Vision*. Citeseer, 2003.

[5] R. Fattal, "Image upsampling via imposed edge statistics," in *ACM Transactions on Graphics (TOG)*. ACM, 2007, vol. 26, p. 95.

[6] J. M. Bioucas-Dias and M. A. Figueiredo, "A New TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2992–3004, 2007.

[7] N. Mueller, Y. Lu, and M. N. Do, "Image interpolation using multiscale geometric representations," in *Proc. SPIE Conf. on Electronic Imaging*, San Jose, USA, 2007.

[8] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum, "Image super-resolution using gradient profile prior," in *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008.* IEEE, 2008, pp. 1–8.

[9] Q. Shan, Z. Li, J. Jia, and C.-K. Tang, "Fast image/video upsampling," in *ACM Transactions on Graphics (TOG)*. ACM, 2008, vol. 27, p. 153.

[10] P. Sen and S. Darabi, "Compressive image super-resolution," in *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*. IEEE, 2009, pp. 1235–1242.

[11] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *Computer Graphics and Applications, IEEE*, vol. 22, no. 2, pp. 56–65, 2002.

[12] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* IEEE, 2004, vol. 1, pp. I–I.

[13] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 6, pp. 1127–1133, 2010.

[14] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.

[15] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Curves and Surfaces*, pp. 711–730. Springer, 2012.

[16] R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *Computer Vision–ACCV 2014*, pp. 111–126. Springer, 2014.

[17] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *2009 IEEE 12th International Conference on Computer Vision*. IEEE, 2009, pp. 349–356.

[18] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 2, pp. 12, 2011.

[19] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. Alberi Morel, "Single-image super-resolution via linear mapping of interpolated self-examples," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5334–5347, 2014.

[20] J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 5197–5206.

[21] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, 2002.

[22] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, Part 1, pp. 1741–1757, 2007.

[23] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 31–40, 2008.

[24] L. Baboulaz and P. L. Dragotti, "Exact feature extraction using finite rate of innovation principles with an application to image super-resolution," *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 281–298, 2009.

[25] R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1827–1842, 2011.

[26] J. A. Urigüen, Y. C. Eldar, P. L. Dragotti, and Z. Ben-Haim, "Sampling at the rate of innovation: Theory and applications," *Compressed Sensing: Theory and Applications*, p. 148, 2012.

[27] X. Wei, T. Blu, and P. L. Dragotti, "Finite rate of innovation with non-uniform samples," in *IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC)*, August 2012, pp. 369 –372.

[28] J. A. Urigüen, T. Blu, and P. L. Dragotti, "FRI sampling with arbitrary kernels," *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5310–5323, 2013.

[29] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[30] M. Unser, "Sampling-50 years after shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, 2000.

[31] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[32] C.-Y. Yang, C. Ma, and M.-H. Yang, "Single-image super-resolution: A benchmark," in *Computer Vision–ECCV 2014*, pp. 372–386. Springer, 2014.

[33] X. Wei and P. L. Dragotti, "Sampling piecewise smooth signals and its application to image up-sampling," in *2015 IEEE International Conference on Image Processing*, 2015.