Signal Processing: *Image Communication* 36 (2015) 83–94

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Image interpolation using shearlet based iterative refinement

H. Lakshman [a,*], W.-Q Lim [b], H. Schwarz [a], D. Marpe [a], G. Kutyniok [b], T. Wiegand [a,c]

[a] *Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Berlin, Germany*
[b] *Department of Mathematics, Technical University of Berlin, Germany*
[c] *Department of Electrical Engineering, Technical University of Berlin, Germany*

ARTICLE INFO

Article history:
Received 9 November 2014
Received in revised form
8 June 2015
Accepted 8 June 2015
Available online 16 June 2015

Keywords:
Interpolation
Sparse representation
Shearlets

ABSTRACT

This paper proposes an image interpolation algorithm exploiting sparse representation for natural images. It involves three main steps: (a) obtaining an initial estimate of the high resolution image using FIR filtering, (b) promoting sparsity in a selected dictionary through hard thresholding to obtain an approximation, and (c) extracting high frequency information from the approximation to refine the initial estimate. For the sparse modeling, a shearlet dictionary is chosen to yield a multiscale directional representation. The proposed algorithm is compared to several state-of-the-art methods to assess its objective and subjective performance. Compared to the cubic spline interpolation method, an average PSNR gain of around 0.8 dB is observed over a dataset of 200 images.

Image interpolation refers to generating a high resolution (HR) image from an input low resolution (LR) image. The resolution of an image can be defined in various ways, e.g., based on:

- the number of pixels in the image,
- the characteristics of the physical sensing device in the camera,
- the effective sharpness as perceived by a human observer.

To quantify the resolution based on the first method is simple, but the latter two are considerably more complex.

Interpolation tasks have regained attention because images/videos are being viewed on displays of different sizes, like mobile phones, tablets, laptops and PCs. More recently, 4K displays are becoming popular, however, content to be displayed might be available in a lower resolution. Interpolation also finds many applications in computer vision, graphics, compression, editing, surveillance and texture mapping. Details synthesis in image interpolation can also be used as a tool for spatial scalability in video coding.

Many established methods are available for interpolation, e.g., FIR filtering and spline based schemes. These techniques may be sufficient for certain applications, but can cause blurring, ringing or other visual artifacts. The main aim of this paper is to overcome these shortcomings using the assumption that the desired HR image can be represented as a sparse linear combination of few basic elements. Images show geometric structures like edges, and conventional Fourier or DCT domains are not well suited for accurate modeling or extraction of such geometric structures. In our proposed method, we use a Shearlet dictionary for modeling HR images, since it provides optimally sparse representations for a large class of multidimensional data [25]. We show that enforcing sparsity on the coefficients of the Shearlet representation helps to improve the regularity along edges in the resulting HR images.

* Corresponding author.
E-mail address: haricharan.lakshman@hhi.fraunhofer.de (H. Lakshman).

http://dx.doi.org/10.1016/j.image.2015.06.004
0923-5965/

1. State-of-the-art

1.1. Linear methods

Signal processing theory for band limited signals advocates sampling higher than the Nyquist rate and recovery using a sinc interpolation [38,46]. The assumption of band limitedness does not hold for most images due to the existence of sharp edges. However, conventional schemes adhere to this philosophy and approximate the ideal low pass filter to produce acceptable results for many practical applications. Techniques like bilinear or bicubic interpolation are some popular examples that have very low computational complexity. Extending the sampling theory to shift-invariant spaces without band limiting constraints has led to a generalized interpolation framework, e.g., B-spline [45] and MOMS interpolation [5] that provide improvements in image quality for a given support of basis functions. However, these linear models cannot capture the fast evolving statistics around edges. Increasing the degree of the basis functions in these linear models helps to capture higher order statistics but results in longer effective support in the spatial domain and hence produces artifacts like ringing around edges.

1.2. Directional methods

To improve the linear models, directional interpolation schemes have been proposed. These perform interpolation along the edge directions and try to avoid filtering across the edges. Some schemes in this class use edge detectors [2,40]. The method in *New edge directed interpolation* (NEDI) [28] computes local covariances in the input image and uses them to adapt the interpolation at the higher resolution, so that the support of the interpolator is along the edges. However, the resulting images still show some artifacts (cf. Section 5). The iterative back projection [23] technique improves image interpolation when the downsampling process is known. Its basic idea is that the reconstructed HR image from the LR image should produce the same observed LR image when it is passed through the same blurring and downsampling process. However, the downsampling filter may not be known in many cases, or the input image may be camera captured, where the optical anti-alias filter used within the sampling system is not known during the subsequent image processing stages. Therefore, it is desirable to design a method that does not rely directly on the knowledge of the downsampling process.

1.3. Sparsity based methods

Image interpolation can be seen as an estimation problem where the input data are inadequate. Naturally, the solution to this problem is not unique due to the lack of information in the HR grid. A popular idea used in such underdetermined problems is to exploit the structure of the desired solution. For images, sparsity in transform domains has proven itself to be a very useful prior [14,35,36]. Sparse approximation can be viewed as approximating a signal with only a few expansion coefficients [37]. Sparsity priors have also been proposed for image interpolation, e.g., in [33,47,32]. The method in [33] uses a contourlet transform for sparse approximation and is designed for an observation model that assumes that the LR image is the low pass subband of a wavelet transform. It uses the same transform in a recovery framework, so it relies directly on knowledge of the downsampling process. We follow a similar recovery principle, but design a system so that it works for typical anti-aliased LR images instead of requiring a specific wavelet transform. The method in [47] involves jointly training two dictionaries for the low- and high-resolution image patches. It then performs a sparsity based recovery, but involves high search complexity to determine a sparse approximation in the trained dictionary (observed to be more than 100x slower than [33]). The method in [32] considers the case when the LR image produced by sub-sampling a HR image is aliased. The method in [9] learns a series of compact sub-dictionaries and assigns adaptively a sub-dictionary to each local patch as the sparse domain. The K-SVD algorithm proposed in [1] and its extensions are commonly used for learning an overcomplete dictionary. These methods depend on the similarity of training and test patches, and the number of the selected examples, which are typical issues in learning-based algorithms. Furthermore, analytically determined transforms have structures that can be exploited to produce a fast implementation, which might be hard to impose during dictionary learning.

1.4. Discussion of the proposed method

We recognize the fact that linear models such as interpolation based on FIR filters are faithful in interpolating the low frequency components but distort the high frequency components in the upsampled image. An iterative framework, based on [20,33], is proposed that combines the output from an initial interpolator and detail components from a denoised approximation. The method used here for denoising is the so-called shrinkage or thresholding approach, i.e., by transforming the signal to a specific domain, setting the transform coefficients below a certain (absolute) value to zero and inverse transforming the coefficients to get back an approximation. The domain used for transforming is chosen so that the coefficients with large absolute values capture most of the geometric features and the coefficients with low absolute values constitute noise or finer details. To this end, multi-resolution transforms or multi-resolution directional transforms are preferred. The concepts of multi-resolution and directionality in transforms are reviewed in Section 3. Using this, a framework for details synthesis in interpolation is proposed in Section 4. In fact, wavelet domain thresholding has been successfully applied to many denoising problems [11,12]. Due to the subsampling in orthogonal wavelet transforms, they are not translation invariant. But, unlike a typical compression scenario, the number of transform coefficients generated during modeling or denoising need not be the same as the number of input samples. This is exploited by removing the subsampling in the wavelet transform and is shown to yield better denoising results [13,15]. Super-resolution methods