Super-resolution Image Reconstuction Performance

Sina Jahanbin, Richard Naething

May 16, 2005

Abstract

As applications involving the capture of digital images become more ubiquitous – and at the same time more ambitious – there is a driving need for digital images of higher resolutions and quality. However, there is a limit to the spatial resolution that can be recorded by any digital device. Super-resolution (SR) image reconstruction is the process of combining several low resolution images into a single higher resolution image. This allows the use of lower resolution (and thus lower cost) imaging systems than could otherwise be used for a given application. Due to these obvious benefits many SR reconstruction methods have been developed. We present an overview of existing SR methods and address the current need for an objective method to compare these techniques based on computational complexity and output quality.

1 Introduction

Any given set of source low resolution (LR) images only captures a finite amount of information from a scene; the goal of SR is to extract the independent information from each image in that set and combine the information into a single high resolution (HR) image. The only requirement is that each LR image must contain some information that is unique to that image. This means that when these LR images are mapped onto a common reference plane their samples must be subpixel shifted from samples of other images – otherwise the images would contain only redundant information and SR reconstruction would not be possible.

Most methods in SR are strictly reconstruction based; that is, they are based primarily on uniform and non-uniform sampling theorems and do not attempt to create any information not found in the LR images. There are also learning SR methods that create new information based on generative models [1][2].

SR techniques can prove useful in many different applications, and these applications can have different requirements in terms of both quality and computational complexity. The quality may also vary for different methods based on characteristics of the input image. The implementation complexity may be affected by implementation specifics, such as the availability of specific optimized libraries. Finally the artifacts caused by poor SR performance can be more visually distracting than blurring from interpolation [3]. For these and other reasons choosing between SR methods is a complex task.

2 Existing Work

Most papers on SR implementations provide subjective results by comparing the SR image to a bilinear interpolated image or the source HR image from which the LR images were created. This provides neither a clear method of comparing different SR methods nor a way of demonstrating a particular SR methods suitability for a desired application.

Comparisons between various SR techniques have been primarily concerned with what assumptions are made in modeling the SR problem. Some of these assumptions include assuming the blurring process to be known [4] or that regions of interest among multiple frames are related through global parametric transformations [1]. Other models take into account arbitrary sampling lattices, a digital sensor elements physical dimensions, a non-zero aperture time, focus blurring, and more advanced additive noise models [5].

Many times these assumptions are chosen to simplify a model and are usually biased toward a particular method. Methods employing models with fewer restrictions are assumed to have higher performance. However, methods that do not make these assumptions have not demonstrated objectively that removing these assumptions yields better SR reconstruction performance. Signal-to-noise ratio, peak signal-to noise ratio (PSNR), root mean squared error, mean absolute error, and mean square error (MSE) of super-resolved images versus interpolated images have all been used as objective measures of SR accuracy; however, the prominent method of presenting results in literature has clearly been subjective visual quality.

3 Proposed Method

We present a comparison of four SR implementations applied to a standard set of source images. These source images have been chosen to demonstrate the effect of frequency content on various SR methods performance. One image (Lena) has information contained primarily in low-frequencies, while the other (Mandrill) has more visually distinct higher frequency content.

We start with a high resolution source image that represents continuous space for our model. Sub-pixel shifted LR images are created from this source image by down-sampled with a factor of four in each spatial dimension. In addition a target HR image is created with half the spatial resolution of the source image. The LR images are corrupted with zero-mean Gaussian noise with a variance of 0.001.

SR performance is then presented in terms of MSE, PSNR, and Structural SIMilarity (SSIM) [6] index as measured between the target HR and reconstructed SR images. SSIM is a method that provides a quality measurement of images based on structural content. Previous perceptual models have been based on MSE with error weighted by different visibility models of the human visual system (HVS). As Wang *et al*. has shown this has several drawbacks: error visibility may not be strongly correlated with a loss perceptual quality, it is difficult to quantifying loss of quality, and MSE based methods are very sensitive to multiplicative noise [7]. As SSIM is based on structural content rather than simply a HVS adjusted noise measurement it does not suffer from these issues, yet is strongly correlated with perceptual image quality.

Additionally the relative computational complexity of each method is shown. When applicable various parameters influence on output quality and run time is also presented.

4 Methods of Super-resolution

There are many existing SR methods including non-uniform interpolation, frequency domain, deterministic and stochastic regularization, projection onto convex sets (POCS), hybrid techniques, optical flow, and other approaches [4][1][8]. Additionally several methods provide parameters that can effect tradeoffs between such factors as fidelity and smoothness or quality and computation time.

4.1 Non-uniform Interpolation

The basis of non-uniform interpolation SR techniques is the non-uniform sampling theory which allows for the reconstruction of functions from samples taken at non-uniformly distributed locations. This was developed by Clark *et al*. [9] and later extended to two-dimensional signals by Kim and Bose [10]. SR image enhancement is a logical application of this new theory, but one that requires very accurate registration between images. Non-uniform interpolation is a basic and intuitive method of super-resolution and has relatively low computational complexity, but it assumes that the blur and noise characteristics are identical across all LR images [4].

4.2 Frequency Domain

Tsai and Huang [11] proved that in the absence of noise or blurring it is possible to reconstruct a HR image from multiple LR images based on the aliasing present in the LR images. This was accomplished by relating the aliased discrete fourier transform coefficients of the LR images to a sampled continuous fourier transform of an unknown HR image. Kim and Bose extended this to blurred and noisy LR images, provided the noise has zero mean and the blur and noise are identical across all LR images, using a recursive implementation based on the weighted least square theory [10].

4.3 Regularization

SR image reconstruction is generally an ill posed problem. However, it can be stabilized with a regularization procedure. Without loss of generality, we can define a model to relate LR images with the original HR image and additive noise as:

$$Y_k = W_k X + n_k \; for \; k = 1, \ldots, p$$

By assuming that registration parameters are estimated, the inverse problem can be solved by deterministic regularization by taking proper prior information about the solution. For example, a constrained least square (CLS) method can be used to find x such that

$$\left[\sum_{k=1}^{p} \|y_k - W_k x\|^2 + \alpha \|Cx\|^2 \right]$$

becomes minimum. In this method a smoothness constraint is used as priori knowledge for reconstruction. Parameter α, which is known as the regularization parameter, controls the trade off between fidelity and smoothness in the solution. Current research is focused on simultaneous blur identification and robust super-resolution.

4.4 Projection Onto Convex Sets

Low resolution images usually suffer from blurring caused by a sensor's point spread function (PSF) and additionally from aliasing caused by under-sampling. Stark and Oskoui [12] have proposed a POCS technique that accounts for both the blurring introduced by the sensors as well as the effects of under-sampling. In their model a low resolution image sequence is denoted by $g(m_1, m_2, k)$. It is assumed that an estimate of the high resolution image at time $k = t_r$ is desired. A family of closed, convex constraint sets can be defined, one for each pixel within the low-resolution image sequence

$$C_{t_r}(m_1, m_2, k) = \left\{ y(n_1, n_2, t_r) : |r^{(y)}(m_1, m_2, k)| \leq \delta_0 \right\}$$

where

$$r^{(y)}(m_1, m_2, k) \doteq g(m_1, m_2, k) - \sum_{(n_1, n_2)} y(n_1, n_2, t_r) h_{t_r}(n_1, n_2; m_1, m_2, k)$$

is the residual associated with an arbitrary member, y, of the constraint set. h_{t_r} combines the effect of the blur PSF and relative motion of object and sensor. The quantity δ_0 is an a priori bound reflecting the statistical confidence with which the actual image, y, is a member of the set $C_{t_r}(m_1, m_2, k)$ This family of constraints is referred to as data consistency constraints. An estimate of the high-resolution version of the reference image is determined iteratively starting from some arbitrary initialization. Successive iterations are obtained by projecting the previous estimate onto the consistency set with an amplitude constraint set that restricts the gray levels of the estimate to the range [0, 255].

4.5 Optical Flow

Some applications can benefit from the generalization of SR techniques to support the imaging of objects that are non-planar, non-rigid, or which are subject to self-occlusion when rotated. One such application is SR reconstruction of facial images. Baker and Kande present optical flow as a solution to this problem [13]. Zhao and Sawhney present a comparison of three different flow methods: least-squares based flow, consistent flow (CONS), and bundled flow with CONS flow as initial input. They demonstrated that it worked well when small amount of noise were present, but that it was very sensitive to flow accuracy [1].

5 Results

For the Lena Image both MSE and PSNR show the Vandewalle *et al.* method is significantly improved over the other SR images. The MSSIM measure in this case is the only quality measurement that favors the Irani-Peleg SR image. Our informal subjective assessment of the images also favors the Irani-Peleg image.

The Mandrill image has more pronounced high frequency detail. The LCAV method does a partic-

Method	Complexity		Source Image	Quality Measurements		
	Itterative	Relative Complexity		MSE	PSNR	MSSIM
Kim *et al*. [10] [14]	$\surd$	Medium	Lena	0.0087725	20.569	0.66687
			Mandrill	0.014926	18.261	0.52432
Irani-Peleg [15] [16]	$\surd$	Medium	Lena	0.0044577	23.509	**0.81647**
			Mandrill	0.011321	19.461	0.54386
Wavelet [17] [18]	$\surd$	High	Lena	0.0063555	21.969	0.68759
			Mandrill	0.012211	19.133	0.52639
Vandewalle *et al*. [3][19]	$\surd$	Low	Lena	**0.0021182**	**26.74**	0.81441
			Mandrill	**0.008321**	**20.798**	**0.68132**
Bilinear Interpolation		Negligable	Lena	0.0061053	22.143	0.67013
			Mandrill	0.012956	18.875	0.4256

Table 1: Result for SR techniques

ularly good job recovering details such as the whiskers of the mandrill. Not surprisingly Vandewalle *et al*. was favored by all three objective measurements. However, the Kim *et al*. method again performed worse in both MSE and PSNR measures than the bilinear interpolated image, while visually (and in MSSIM) it is clear that Kim *et al*. provides a much better image.

6 Conclusion

From these results, particularly the Mandril SR images, it is clear that MSE and PSNR are not particularly good indicators of SR quality. MSSIM appears to more adequately assess the quality of SR images. Perceptive image quality analysis is an emerging field, and obviously MSSIM is not perfect, but this technique is very promising.

7 Future Work

This method relies upon being able to model a real-life continuous scene with a single HR image. This is not sufficient to test methods that are designed to perform SR of non-planar or non-rigid scenes such as the optical flow method proposed by Baker and Kande [13]. Additionally our method only models translation and rotation that is uniform across a frame.

Figure 1: HR Lena

Figure 2: Bilinear Upsampled

Figure 3: Kim *et al*.

Figure 4: Irani-Peleg

Figure 5: Harr Wavelet

Figure 6: Vandewalle *et al*.

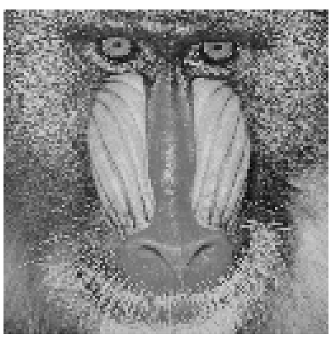

Figure 7: HR Mandrill

Figure 8: Bilinear Upsampled

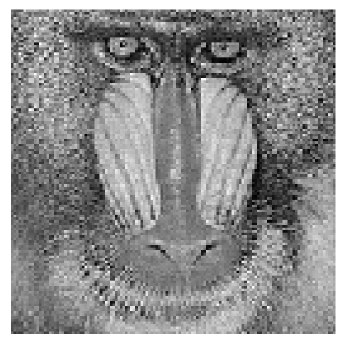

Figure 9: Kim *et al*.

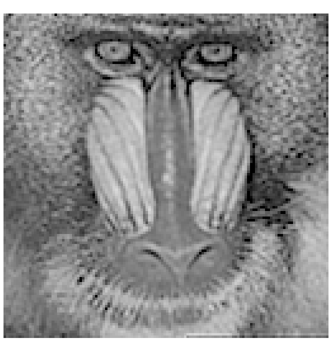

Figure 10: Irani-Peleg

Figure 11: Harr Wavelet

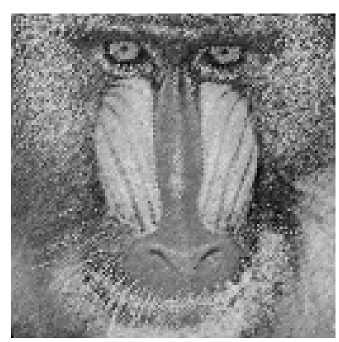

Figure 12: Vandewalle *et al*.

References

[1] W. Zhao and H. S. Sawhney, “Is super-resolution with optical flow feasible?,” in *ECCV ’02: Proceedings of the 7th European Conference on Computer Vision-Part I*, pp. 599–613, Springer-Verlag, 2002.

[2] S. Baker and T. Kanade, “Limits on super-resolution and how to break them,” in *Proceedings of the 2000 IEEE Conference on Computer Vision and Pattern Recognition*, June 2000.

[3] P. Vandewalle, S. Süsstrunk, and M. Vetterli, “A Frequency Domain Approach to Registration of Aliased Images with Application to Super-Resolution,” *EURASIP Journal on Applied Signal Processing (special issue on Super-resolution)*, 2005.

[4] S. C. Park, M. K. Park, and M. G. Kang, “Super-resolution image reconstruction: A technical review,” *IEEE Signal Processing Mag.*, vol. 20, pp. 21–36, May 2003.

[5] A. J. Patti, M. I. Sezan, and A. M. Tekalp, “Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time,” *IEEE Trans. Image Processing*, vol. 6, pp. 1064–1076, Aug. 1997.

[6] Z. Wang, “Ssim index, version 1.0,” 2003. http://www.cns.nyu.edu/~zwang/files/research/ssim/ssim_index.m.

[7] Z. Wang, A. C. Bovik, and L. G. Lu, “Video quality assessment based on structural distortion measurement,” *Signal Processing: Image Communication (special issue on Objective Video Quality Metrics)*, vol. 19, pp. 121–132, Feb. 2004.

[8] M. Elad and A. Feuer, “Restoration of a single superresolution image from several blurred, noisy, and under-sampled measured images,” *IEEE Trans. Image Processing*, vol. 6, pp. 1646–1658, Dec. 1997.

[9] J. J. Clark, M. R. Palmer, and P. D. Lawrence, “A transformation method for the reconstruction of functions from non-uniformly spaced samples,” *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 33, no. 4, pp. 1151–1165, 1985.

[10] S. P. Kim, N. K. Bose, and H. M. Valenzuela, “Recursive reconstruction of high resolution image from noisy undersampled multiframes,” *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 38, pp. 1013–1027, June 1990.

[11] R. Y. Tsai and T. S. Huang, “Multiframe image restoration and registration,” tech. rep., Greenwhich, CT, 1984.

[12] H. Stark and P. Oskoui, “High resolution image reconstruction from lower-resolution image sequences and space varying image restoration,” *J. Opt. Soc. Am. A*, vol. 6, pp. 1715–1726, 1989.

[13] S. Baker and T. Kanade, “Super resolution optical flow,” Tech. Rep. CMU-RI-TR-99-36, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, October 1999.

[14] S. Lertrattanapanich, “Recursive reconstruction of high resolution image from noisy undersampled multi-frames.” http://www.dsplab.ee.psu.edu/Research/Surapong/RLS/RLS_matlab.html.

[15] M. Irani and S. Peleg, “Improving resolution by image registration,” *CVGIP: Graph. Models Image Process.*, vol. 53, no. 3, pp. 231–239, 1991.

[16] D. Hilvert, “Anti-lamenessing engine, version 0.8.1,” 2005. http://auricle.dyndns.org/ALE/.

[17] N. Nguyen and P. Milanfar, “A wavelet-based interpolation-restoration method for superresolution (wavelet superresolution),” *Circuits Systems Signal Processing*, vol. 19, no. 4, pp. 321–338, 2000.

[18] S. Lertrattanapanich, “Wavelet-based interpolation-restoration method superresolution matlab code.” http://www.dsplab.ee.psu.edu/Research/Surapong/ Superresolution/Superresolution.html.

[19] P. Vandewalle, S. Süsstrunk, and M. Vetterli, “Lcav super-resolution source code and images.” http://lcavwww.epfl.ch/reproducible_research/VandewalleSV05.