Super-resolution Image Reconstuction Performance

Sina Jahanbin, Richard Naething

March 30, 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. This may be due to optical distortions, motion blur (caused by motion of the scene outpacing the shutter speed), under-sampling, or noise [1]. 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. A comparison of the performance differences between these methods would only be valid given suitable objective measurements, such as the application of SR images in a face recognition system.

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 LR images can come from a variety of sources: they can be taken from different frames of a video sequence, different still images taken from a single camera that has undergone translation or rotation, or multiple cameras capturing a single scene. 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 [2][3].

2 Steps of Super-resolution

There are several major steps in super-resolution reconstruction: registration, warping, blurring, motion cancellation, and merging the converted LR frames into a final HR image [4]. Registration is the process of determining where sampled values of the LR image should lie on the HR image, thus creating a point-to-point mapping to be used in warping. Warping is the process of converting the samples of LR images, by using the relationship determined in registration, to the HR image. This typically involves a projection from the LR image plane to the HR plane, and interpolation of the LR sample values to the HR resolution. The images are processed to remove blur and noise. Finally the warped, processed LR images are merged to form the HR image.

3 Methods of Super-resolution

There are a number of different algorithms developed to perform SR reconstruction. These include non-uniform interpolation, frequency domain, deterministic and stochastic regularization, projection onto convex sets (POCS), hybrid techniques, optical flow, and other approaches [1][2][5].

3.1 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.* [6] and later extended to two-dimensional signals by Kim and Bose [7]. SR image enhancement is a logical application of this new theory, but one that requires very accurate registration between images. Early SR applications used detailed camera placement to allow for accurate interpolation. Komatsu *et al.* developed a method to overcome the limitations of insufficient registration accuracy by employing multiple digital sensors with different pixel sizes. This ensures that pixels of multiple images will not coincide regardless of camera placement [8]. 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 [1].

3.2 Frequency Domain

Tsai and Huang [9] 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. This was accomplished using a recursive implementation based on the weighted least square theory [7].

3.3 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) methods 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. In [10] an iterative regularized approach is introduced to increase the resolution of a video sequence. A multiple input smoothing convex functional is defined and used to obtain a globally optimal high resolution video sequence. In [11] the work by Hong *et al.* is extended to calculate an optimal regularization parameter systematically using the L-curve method. Current research is focused on simultaneous blur identification and robust super-resolution.

3.4 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

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

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].

3.5 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 [2].

3.6 Generative Methods

One example of the recent work in generative methods, that is, methods that use additional information not contained in the LR image set to restore a HR image, is the "recogstruction" research conducted by Baker and Kanade [3]. Where most earlier papers used smooth a priori assumptions, this technique relies on strong class based priors to provide far more information than simple smooth priors used in existing SR algorithms. They claim significantly better results both in subjective and root-mean-square (RMS) pixel error. However, the use of strong class based priors means that the method will find what it is looking for even if it does not exist in the image set. For instance, applying this method with a face priors, but to a LR scene of a grove of trees, will yield a face like image. An open research area is how these priors will effect applications such as face recognition that depend mainly on differences in a set of images that all fit the prior.

4 Comparison of SR Techniques

Comparisons of SR techniques have been primarily concerned with what assumptions are made in the modeling of the SR problem. Some of these assumptions include assuming the blurring process to be known [1] or that regions of interest among multiple frames are related through global parametric transformations [2]. 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 [21]. Many times these assumptions are chosen to simplify a model and are usually biased toward a particular method. In addition, methods that do not make these assumptions have not demonstrated objectively that removing these assumptions yields better SR reconstruction performance. Signal-to-noise ratio (SNR), peak signal-to noise ratio (PSNR), RMS, mean absolute error (MAE), and mean square error (MSE) have all been used as objective measures of SR accuracy; however, the prominent method of presenting results is clearly subjective visual quality.

	Performance Measurements					Source Type	
	Subj.	SNR	PSNR	MSE	Other	Still	Video
Nonuniform Sampling Based							
Komatsu [8]	√	√				√	
Frequency Domain							
Kim and Bose [7]	√					√	
Kim and Su [14]	√	√				√	
Rhee and Kang [15]	√		√			√	
Deterministic Regularization							
Hong, Kang, and Katsaggelos [10]	√		√				√
Alam *et al.* [16]	√						√
Bose and Koo [17]	√		√			√	
Stochastic Regularization							
Tom and Katsaggelos	√					√	
Schultz and Stevenson [18]	√	√					
Hardie, Barnard, and Armstrong [11]					MAE	√	
Cheeseman *et al.* [19]	√					√	
Projection onto Convex Sets							
Tekalp, Ozkan, and Sezan [20]	√					√	
Patti, Sezan, and Tekalp [21]	√						√
Eren, Sezan, Tekalp [22]	√						√
Patti, and Altunbasak [23]	√			√		√	
POCS & Maximum a posteriori hybrid							
Elad and Feuer [5]	√			√		√	
Optical Flow							
Baker and Kanade [13]	√					√	
Zhao and Sawhney [2]	√					√	
Other Methods							
Irani and Peleg [24]	√					√	
Baker and Kanade [3]					RMS	√	

Table 1: Result for SR techniques

5 Conclusion

As different methods of SR have been developed using models with unequal assumptions of the underlying problem, and because the results provided have been primarily based on subjective measurements, it is difficult to find an unbiased comparison on what SR methods are more appropriate for a given task. 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 does not provide a clear method of comparing different implementations suitability for a desired application.

6 Future Work

We propose developing an objective measurement for comparison of SR methods. One possible objective measurement is a universal image quality measures for human vision systems and computer vision systems. An alternative would be to use the HR images as the input to some other image processing system, such as a face-recognition algorithm, and examine how different SR techniques affect the recognition accuracy.

References

[1] 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.

[2] 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.

[3] 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.

[4] T. Boult and M. Chang, "Efficient image warping and super-resolution," *Proceedings of the Third Workshop on Applications of Computer Vision*, pp. 56–61, Dec. 1996.

[5] M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE Trans. Image Processing*, vol. 6, pp. 1646–1658, Dec. 1997.

[6] 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.

[7] 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.

[8] T. Komatsu, K. Aizawa, T. Igarashi, and T. Saito, "Signal-processing based method for acquiring very high resolution image with multiple cameras and its theoretical analysis," *Proc. Inst. Elec. Eng.*, vol. 140, pp. 19–25, Feb. 1993.

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

[10] M. C. Hong, M. G. Kang, and A. Katsaggelos, "An iterative weighted regularized algorithm for improving the resolution of video sequences," *Proc. Int. Conf. Image Processing*, vol. 2, pp. 474–477, 1997.

[11] R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong, and E. A. Watson, "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Opt. Eng.*, vol. 37, pp. 247–260, Jan. 1998.

[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. P. Kim and W. Y. Su, "Recursive high-resolution reconstruction of blurred multiframe images," *IEEE Trans. Image Processing*, vol. 2, pp. 534–539, Oct. 1993.

[15] S. H. Rhee and M. G. Kang, "Discrete cosine transform based regularized high-resolution image reconstruction algorithm," *Opt. Eng.*, vol. 38, pp. 1348–1356, Aug. 1999.

[16] M. S. Alam, J. G. Bognar, R. C. Hardie, and B. J. Yasuda, "Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames," *IEEE Trans. Instrum. Meas.*, vol. 49, pp. 915–923, Oct. 2000.

[17] N. K. Bose, S. Lertrattanapanich, and J. Koo, "Advances in superresolution using l-curve," *Proc. Int. Symp. Circuits and Systems*, vol. 2, pp. 433–436, 2001.

[18] R. R. Schulz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. Image Processing*, vol. 5, pp. '996–1011, June 1996.

[19] P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "Super-resolved surface reconstruction from multiple images," Tech. Rep. Tech. Rep. FIA-94-12, NASA Ames Research Center, Moffett Field, CA, Dec. 1994.

[20] A. M. Tekalp, M. K. Ozkan, and M. I. Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space varying image restoration," *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, pp. 169–172, Mar. 1992.

[21] 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.

[22] P. E. Eren, M. I. Sezan, and A. M. Tekalp, "Robust, object-based high-resolution image reconstruction from low-resolution video," *IEEE Trans. Image Processing*, vol. 6, pp. 1446–1451, Oct. 1997.

[23] A. J. Patti and Y. Altunbasak, "Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants," *IEEE Trans. Image Processing*, vol. 10, pp. 179–186, Jan. 2001.

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