Imaging-Consistent Super-Resolution*

Ming-Chao Chiang

Columbia University
Department of Computer Science
New York, NY 10027
chiang@cs.columbia.edu

Terrance E. Boult

Lehigh University
Department of EECS
Bethlehem, PA 18015
tboult@eecs.lehigh.edu

Abstract

This paper introduces two algorithms for enhancing image resolution from an image sequence. The "image-based" approach presumes that the images were taken under the same illumination conditions and uses the intensity information provided by the image sequence to construct the high-resolution image. This ideal, however, is almost always not true when the illumination varies. The "edge-based" approach, based on edge models and a local blur estimate, circumvents these difficulties. The paper presents the theory and the experimental results using these two algorithms.

1 Introduction

The idea of super-resolution, combining images by combining peices from an image sequence into a single image with higher resolution than any of the individual images, has been around for years. Previous research on super-resolution, [Huang and Tsai-1984, Gross-1986, Peleg *et al.*-1987, Keren *et al.*-1988, Irani and Peleg-1991, Irani and Peleg-1993, Bascle *et al.*-1996], ignore the impact of image warping techniques. They also presume that the images were taken under the same illumination conditions. The objective of this paper is to address techniques to improve the quality of super-resolution imaging and to deal with lighting variations. We show that image warping techniques may have a strong impact on the quality of image resolution enhancement.

Image warping requires the underlying image to be "resampled" at non-integer locations; it requires spatially varying image reconstruction. When the goal of warping is to produce output for human viewing, only mildly accurate image intensities are needed. In these cases, techniques using bi-linear interpolation have been found sufficient. However, as a step for applications such as super-resolution, the precision of the warped intensity values is often important. For these problems, bi-linear image reconstruction may not be sufficient; the spatially varying nature of the reconstruction limits the "efficient" alternative reconstruction methods. This paper shows how ideas of imaging-consistent reconstruction/restoration algorithms [Boult and Wolberg-1993] and the integrating resampler [Chiang and Boult-1996b], can be used for warping while maintaining superior image quality.

*This work is supported in part by NSF PYI IRI-90-57951, NSF grant CDA-9413782, and ONR/DARPA via DOD MURI program ONR N00014-95-1-0601. Several other agencies and companies have also supported parts of this research.

2 Image-Based Super-Resolution

The idea of super-resolution is based on the fact that each image in the sequence provides small amount of additional information. There are, of course, some fundamental limits on what this combination can do. If the images were noise-free, focused and Nyquist sampled, then multiple images would add nothing. However images are blurred and with the noise and aliasing present in images, deblurring is unstable. If time is not a concern, then standard DSP techniques can address these problems, formulating fusion as millions of coupled equations. The goal is then to come up with an efficient approximation.

Our approach treats recognizes four separate components, the matching (to determine alignment), the warping (to align the data and increasing sampling rate), the fusion (to produce a less noisy image), and an optional debluring stage to remove lens blur. For now we are using tradational matching on image fields (normalized SSD or correlation) and tradational debluring. We are concentrating our efforts on warping and fusion. Warping is considered in the next section.

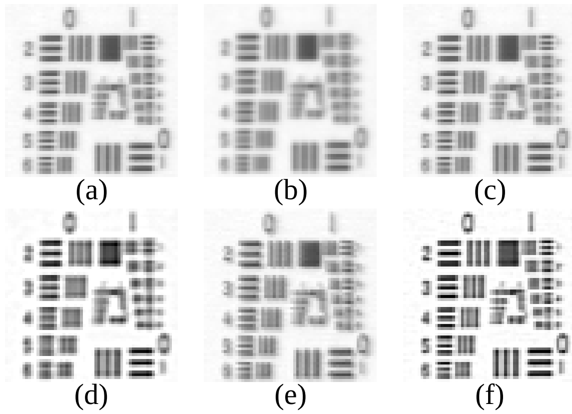


Figure 1: Orginal images and down-sampled version of super-resolution results. (a) shows one of the eight original images. (b) shows the downsazmpled super-resolution using bi-linear resampling. (c) down-sampled super-resolution using QRS (d) shows a deblured orginal (i.e. deblurred (a). (e) shows down-sampled super-resolution by back-projection (f) shows supre-resolution with QRS followed by deblurring followed by down-sampling.

For fusion we have experimented with simple averaging, which is good if there are no outliers, averages with trimed tails and median. These produce decreasingly accurate estimates with increasing robustness to outliers. As the matching is sometimes inaccurate and because of aliasing artifacts, a few outliers are common, thus the trimed tails is probably the best overall technique.

In [Chiang and Boult-1996a] we presented inital results and compared our technique the leading existing work of [Irani and Peleg-1993] (which is refered as back-projection in the following. Figs., 2, 3 and 1 show some example results. In all cases the resulting super-resolution images are a scale-up by a factor of 4. We note that previous work on this topic reported results only scaling by a factor of 2.

If we down-sample our the super-resolution estimation, we should get an increase in image quality. A few examples of this are shown in Fig. 1. It can be easily seen from Fig. 1 that image warping techniques indeed have a strong impact on the quality enhancement, even with the image resolution is not increased. In particular, Fig. 1f is significantly clearer than the original (Fig. 1a) or a deblurred version thereof (Fig. 1g). Thus, super-resolution provides added benefits even if the final sampling rate is exactly the same as the original.

Fig. 2 shows the final results of our first experiment. Fig. 2a shows Fig. 1a blown up by a factor

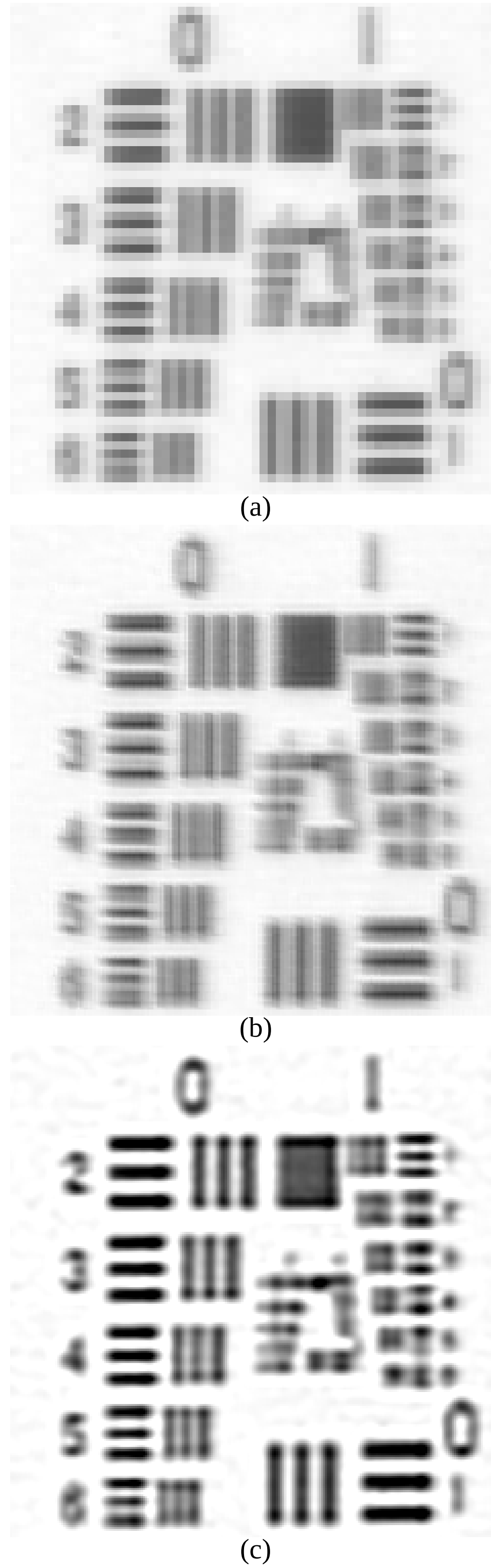
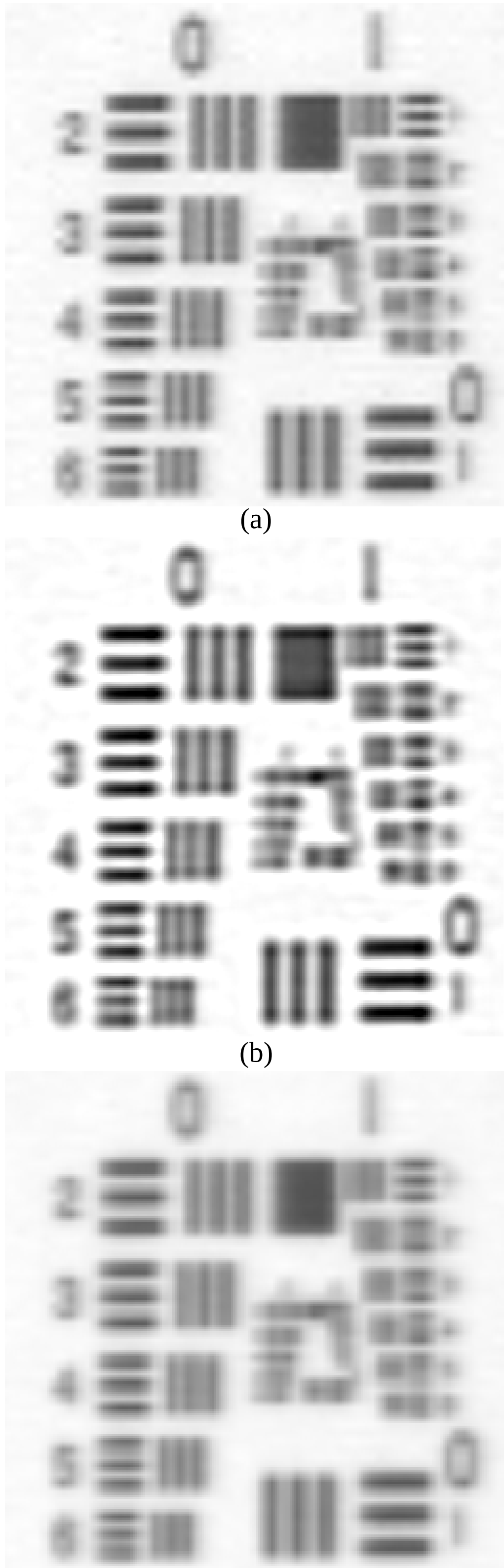


Figure 2: Fig. 1a pixel-replicated by a factor of 4; (b) super-resolution by back-projection; (c) super-resolution using QRS with deblurring.



(a)

(b)

(c)

Figure 3: Effects of debluring and reconstruction kernel. (a) shows Fig. 2(c) without deblurring. (b) shows super-resolution using bi-linear resampling with deblurring; and (c) shows (b) before deblurring.

of 4 using pixel replication. Fig. 2b shows super-resolution by our implementation of Irani's back-projection method using bi-linear resampling to simulate the image formation process and Fig. 3a as the initial guess. Fig. 2d, super-resolution using QRS followed by deblurring. Fig. 3 shows the effects of debluring and of using bi-linear reconstruction for warpping.

Fig. 4 shows our second example, shows a more complex gray level image. The tread-wheels of the toy tank are visible in the super-resolution image but not in the orginals, and the "tank-number" is (just) readable in the supre-resulution image while not in the orginals.

Results from our first two experiments show that the image-based method we propose herein is not only computationally cheaper, but it also gives results comparable to or better than those using back-projection. In general, our method is often more than two or three times faster. See [Chiang and Boult-1996a, Chiang-1996] for details. Moreover, it is easily seen from Figs. 1, 2, and 3 that integrating resampler outperforms traditional bi-linear resampling.

3 Imaging-Consistent Warping

We consider the imaging model in figure 5. Due to the limit of space, we only briefly review the integrating resampler, more details (including the image formation process and the sensor model) can be found in [Chiang and Boult-1996b, Chiang-1996].

An algorithm is called *imaging-consistent* if it is the exact solution for some input function, which, according to the sensor model, would have generated the measured input. For image reconstruction, we achieve this by computing a functional restortation (i.e., f_2), then blurring it again by the pixel's PSF. This actually defines a whole class of image restoration/reconsturction techniques, depending on the model for f_2. Probably the simplest method to consider is based on a piecewise quadratic model for the image. If we assume a Rect PSF filter for the photosite, the imaging consistent algorithm is easy to derive, see [Chiang and Boult-1996b]. To ensure that the method is local and function is continuous, and that the method is local, we define the value of the reconstruction at the pixel boundaries k_i and k_{i+1}, to be equal to E_i and E_{i+1}, where we compute E_i with some technique, e.g. cubic convolution. Given the values E_i at the pixel edges, combined with the imaging-consistent constaint (the integral across the pixel must equal the measured intensity) results in exactly three constraints. From this, one can determine

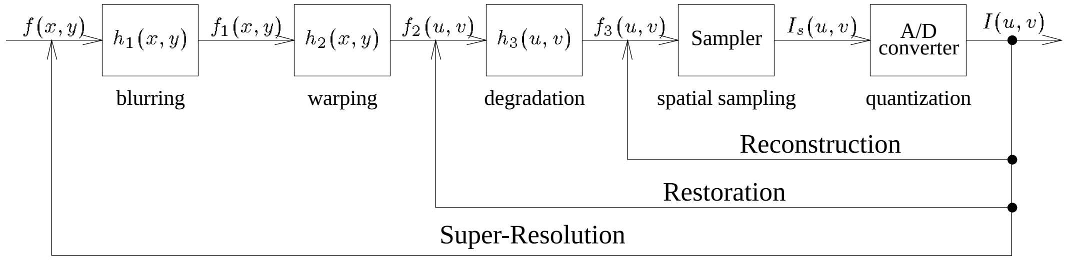


Figure 5: The image formation process and the relationship between restoration, reconstruction, and super-resolution.

the quadratic polynomial for f_2. This gives the intra-pixel restoration. For super-resolution, we consider only this intra-pixel restoration (abbreviated QRS in the following discussion). Reconstruction can be derived by simply blurring the resulting restoration by a PSF of the same scale as input.

To define the integrating resamplers, we generalize the idea of the imaging-consistent algorithms described above. Whereas imaging-consistent algorithms simply assume the degradation models are identical for both input and output; the integrating resamplers go one step further, allowing (1) both input and output to have their own degradation model, and (2) the degradation model to vary its size for each output pixel.

When we are resampling the image and warping its geometry in a nonlinear manner, this new approach allows us to efficiently do both pre-filtering and post-filtering. Because we have already determined a functional form for the input, no spatially-varying filtering is needed, as would be the case if direct inverse mapping were done. The integrating resampler [Chiang and Boult-1996b] also handles antialiasing of partial pixels in a straightforward manner.

4 Edge-Based Super-Resolution

For almost all applications involving an image sequence, the problem of lighting variation arises, even when they are taken consecutively in a well controlled environment. If the images are not from a short time span, variations are often significant. The idea we propose herein is a simple solution, we fuse edge and blur information and use that to combine it with one of the orginal intesnity image to reconstruct the super-resolution image. This does not necessarily solve the lighting variation problem. But, this effectively avoids the problem of lighting variation since we are now dealing with a single image and the edge positions that are less sensitive to the change of lighting. This means that we can, at least, get rid of most of the undesirable effect of lighting variations.

However, to fuse all the edges together, it requires that the edges be first detected and then warped. It also requires a image reconstruction technique that directly incorporates both the edge and intensityies. This will allow the reference image to be reestimated and scaled up based on the edge models and local blur estimation. We have generalized the idea of the imaging-consistent reconstruction algorithms to deal direclty discontinuities in an image [Boult and Wolberg-1993, Chiang and Boult-1996a, Chiang-1996]. More details can be found in [Chiang and Boult-1996a, Chiang-1996].

Given the image sequence, our edge-based super-resolution algorithm is shown, as follows:

1. Estimate the edges and blur models using the procedure described in Section 4.1.
2. Estimate the motions involved in the image sequence.
3. Choose one of the images as the reference image (the one with lighting right). Scale the reference image up and deblur at the same time it is being scaled up.
4. Warp all the edges/blur models to the reference image and fuse them.
5. Use the fused edge/blur models and the reference image to compute the super-resolution intensity image.
6. Optional debluring stage.

4.1 Edge Localization& Local Blur Estimation

Typically, edge detection involves the estimation of first and second derivatives of the luminance function, followed by selection of zero crossing and extrema. While the world does not need yet another edge detector, we define using our imaging-consistent model because allows us to work in a consistent framework and to tie the edge-model into the image reconstruction algorithm. The edge localization/detection is obtained by differentating the

(a)

(b)

Figure 4: Results from a toy-tank sequence. (a) one of the original images blown up by a factor of 4 and deblurred; (b) super-resolution using QRS followed by deblurring.

functional form of the image reconstruction model and considering only significant maxima of the first derivative.[1] The edge model is tied into the image reconstruction in that each pixel is now modeled as potentially having a discontinuity, while still satsifying the imaging-consistent constraint. The model we use is still peicewise quadratic and the integral accross the pixel, including any step discontinuity, must still equal the measured data. If a discontinuity is included within a pixel, the approximations used for the pixel boundaries are recomputed using data from only one side of the discontinuity.

[1] Yea there is a threshold hiding there. Future work will address how to better determine significant vs insignificnat edges, and, more importantly, if this should be done before or after the fusion of the "edge-models".

In [Chiang and Boult-1996a], we showed that deblurring after image fusion is most effective for super-resolution imaging. However, that work presumes that the blur is not dominated by depth-of-field effects. This allows us to replace a spatially-varying point spread function with a cascade of two simpler components: a spatially-invariant blur and a geometric warp. Unfortunately, this assumption is almost always not true in practice.

In [Chiang and Boult-1997]. we propose a new algorithm for local blur estimation. The idea of this algorithm is to model a blurred edge with two components: a step edge and a blur kernel. We assumes the step edge is from v to $v + \delta u(x)$ where v is the unknown intensity value and δ is the unknown amplitude of the edge. The blur of this edge is modeled by a "truncated" Gaussian blur kernel

$$G(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

where σ is the unknown standard deviation, Given the functional form of our reconstruction, we can solve directly for the three parameters of the blured edge model.

For the examples we again presume that "motion" is computed, which for general lighting changes is much more difficult. For the examples here we use a normalized SSD computation.

5 Image-Based vs. Edge-Based

Due to the limit of space, we only briefly compare the two algorithms proposed herein. Both algorithms take time roughly proportional to the number of images in the image sequence, with the image-base fusion being the faster of the two, producing a 500x500 superresolution image in a few seconds on a Ultrasparc.

If the variation of lighting is small, such as in an controled indoor environment, the image-based approach is more appropriate because it uses the intensity information provided by the whole image sequence to construct the super-resolution image and thus is better at removing the noise and undesirable artifacts. On the other hand, the edge-based algorithm is more appropriate if the variation of illumination is large.

If the images the variation of lighting is intermediate, a possible solution is probably a hybrid of the

(a) (b) (c) (d)

Figure 6: Edge-based example using 32 81x81 images: (a) and (b) are two of the original images blown up by a factor of 4 with pixel replication using pixel replication; (c) is super-resolution using the image-based algorithm without deblurring at the end; and (d) shows results with debluring.

two algorithms we propose herein. The idea is that instead of choosing a single reference image of the edge-based super-resolution algorithm, use the averaging or median of a sub-sequence out of the image sequence as the reference image, presuming that the variation of lighting is not so significant within the sub-sequence.

6 Future Work

Further work is needed before the super-resolution algorithm is robust enough for general use in VSAM applications, inparticular we need to incorporate a more robust sub-pixel matching algorithms and include better debluring algorithms. Quantitatively analysis of both approaches is now under way using recognition rates as the benchmark.

7 Conclusion

This paper introduces two algorithms for enhancing image resolution from an image sequence. The image-based approach presumes that the images were taken under the same illumination conditions and uses the intensity information provided by the image sequence to construct the super-resolution image. The edge-based approach, based on edge models and a local blur estimate, circumvents the difficulties caused by lighting variations. We show that image warping techniques may have a strong impact on the quality of image resolution enhancement.

References

[Bascle *et al.*, 1996] B. Bascle, A. Blake and A. Zisserman. Motion deblurring and super-resolution from an image sequence. *Computer Vision—ECCV*, pages 573–581, Apr 1996.

[Boult and Wolberg, 1993] T. E. Boult and George Wolberg. Local image reconstruction and subpixel restoration algorithms. *CVGIP: Graphical Models and Image Processing*, 55(1):63–77, Jan 1993.

[Chiang and Boult, 1996a] Ming-Chao Chiang and T.E. Boult. Efficient image warping and super-resolution. *Proc. of IEEE Workshop on Applications of Computer Vision*, pages 56–61, Dec 1996.

[Chiang and Boult, 1996b] Ming-Chao Chiang and T.E. Boult. The integrating resampler and efficient image warping. *Proceedings of the ARPA Image Understanding Workshop*, pages 843–849, Feb 1996.

[Chiang and Boult, 1997] Ming-Chao Chiang and T.E. Boult. Local blur estimation and super-resolution. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, June 1997. To appear.

[Chiang, 1996] Ming-Chao Chiang. Imaging-consistent warping and super-resolution. Thesis Proposal, Columbia Univ., Dept of CS, Dec 1996.

[Gross, 1986] D. Gross. Super-resolution from subpixel shifted pictures. Master's thesis, Telaviv University, Oct 1986.

[Huang and Tsai, 1984] T. S. Huang and R. Y. Tsai. Multi-frame image restoration and registration. *Advances in Computer Vision and Image Processing*, 1:317–339, 1984.

[Irani and Peleg, 1991] Michal Irani and Shmuel Peleg. Improving resolution by image registration. *CVGIP: Graphical Models and Image Processing*, 53(3):231–239, May 1991.

[Irani and Peleg, 1993] Michal Irani and Shmuel Peleg. Motion analysis for image enhancement: Resolution, occlusion, and transparency. *Journal of Visual Communication and Image Representation*, 4(4):324–335, Dec 1993.

[Keren *et al.*, 1988] D. Keren, S. Peleg and R. Brada. Image sequence enhancement using sub-pixel displacements. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 742–746, Jun 1988.

[Peleg *et al.*, 1987] S. Peleg, D. Keren and L. Schweitzer. Improve image resolution using subpixel motion. *Pattern Recognition Letter*, pages 223–226, 1987.