Example-Based Super-Resolution

William T. Freeman, Thouis R. Jones & Egon C. Pasztor

Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: Pixel representations for images do not have resolution independence. When we zoom into a bitmapped image, we get a blurred image. Figure 1 shows the problem for a teapot image, rich with real-world detail. We know the teapot's features should remain sharp as we zoom in on them, yet standard pixel interpolation methods, such as pixel replication (b, c) and cubic spline interpolation (d, e), introduce artifacts or blurring of edges. For images zoomed 3 octaves, such as these, sharpening the interpolated result has little useful effect (f, g).

Many applications in graphics or image processing could benefit from such pixel resolution independence, such as texture mapping, enlarging consumer photographs, and converting NTSC video content to HDTV. We don't expect perfect resolution independence—even the polygon representation doesn't have that—but increasing the resolution independence of pixel-based representations is an important task for image-based rendering. Our example-based super-resolution algorithm yields Fig. 1 (h, i).

Previous Work: Researchers have long studied image interpolation, although only recently using machine learning or sampling approaches, which offer much power.

Cubic spline interpolation [5] is a very common image interpolation function, but suffers from blurring of edges and image details. Recent attempts to improve on cubic spline interpolation [6, 8, 2] have met with limited success. Schreiber and collaborators [6] proposed a sharpened Gaussian interpolator function to minimize information spillover between pixels and optimize flatness in smooth areas. Schultz and Stevenson [7] have used a Bayesian method for super-resolution, but hypothesized the prior probability. A proprietary, undisclosed algorithm, Altamira Genuine Fractals 2.0 [1] (an Adobe Photoshop plug-in), suffers from blur in regions of texture and at fine lines.

Approach: One would expect that the richness of real-world images would be difficult to capture analytically. This motivates a learning-based approach: in a training set, learn the fine details that correspond to different image regions seen at a low-resolution; then use those learned relationships to predict fine details in other images. For the past several years [3, 4], we have been exploring this approach for enlarging images. See those publications for implementation details.

We exploit regularities over images: we use small pieces of one image, modified for generalization by the appropriate pre-processing, to create plausible image information in a second image. Without very specific training data, it is not reasonable to expect to generate the *correct* high-resolution information. We aim for the more attainable goal of generating *visually plausible* image details, such as sharp edges, and plausible looking texture.

To generate our training set, we start from a collection of high resolution images, and degrade each of them in a manner corresponding to the degradation we plan to undo in the images we later process. Typically, we blur and subsample them to create a low-resolution image of $\frac{1}{4}$ the number of original pixels.

We describe a Markov random field whos maximum probability solution corresponds to the best collection of high resolution patches from the training data to use for zooming each patch of the input image. We have explored different fast solution methods for this Markov random field problem [3, 4].

Future Work: We want to add more high level knowledge to our estimation problem, in order to remove the occasional artifacts from the algorithm outputs.

Research Support: This research was conducted at Mitsubishi Electric Research Laboratories, where WTF was a Senior Research Scientist, and TRJ and ECP were Members of Technical Staff.

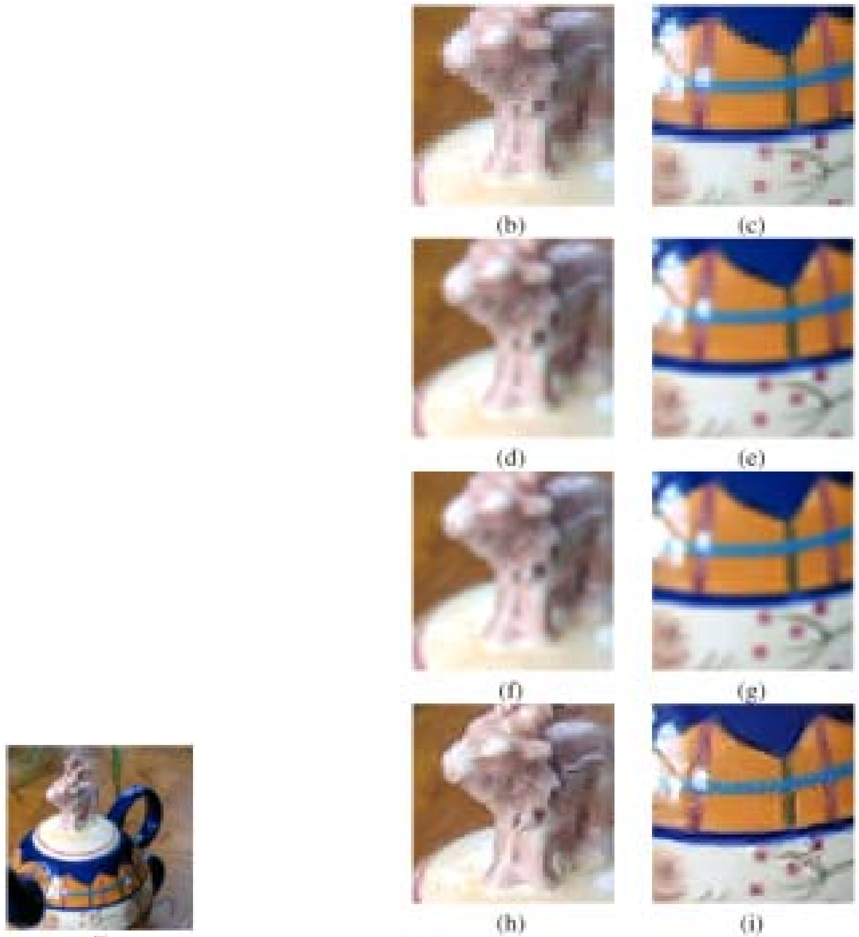

Figure 1: (a) An image (100x100) of a real-world teapot shows a richness of texture, but yields a blocky or blurred image when zoomed in by a factor of 8 in each dimension by (b, c) pixel replication or (d, e) cubic spline interpolation. (Images (b) through (i) were 32x32 pixel original sub-images, zoomed by 8 to 256x256 images). Sharpening the cubic spline interpolation may not help to increase the perceptual sharpness (f, g, using “sharpen more” in Adobe Photoshop). (h, i) show the results of our example-based super-resolution algorithm, maintaining edge and line sharpness, and plausible texture details.

References:

[1] Altamira Genuine Fractals 2.0, 2000. www.altamira.com.

[2] F. Fekri, R. M. Mersereau, and R. W. Schafer. A generalized interpolative VQ method for jointly optimal quantization and interpolation of images. In *Proc. ICASSP*, volume 5, pages 2657–2660, 1998.

[3] W. T. Freeman and E. C. Pasztor. Learning to estimate scenes from images. In M. S. Kearns, S. A. Solla, and D. A. Cohn, editors, *Adv. Neural Information Processing Systems*, volume 11, Cambridge, MA, 1999. MIT Press. See also http://www.merl.com/reports/TR99-05/.

[4] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael. Learning low-level vision. *Intl. J. Computer Vision*, 40(1):25–47, 2000.

[5] R. Keys. Bicubic interpolation. *IEEE Trans. Acoust. Speech, Signal Processing*, 29:1153–1160, 1981.

[6] W. F. Schreiber. *Fundamentals of electronic imaging systems.* Springer-Verlag, 1986.

[7] R. R. Schultz and R. L. Stevenson. A Bayesian approach to image expansion for improved definition. *IEEE Trans. Image Processing*, 3(3):233–242, 1994.

[8] S. Thurnhofer and S. Mitra. Edge-enhanced image zooming. *Optical Engineering*, 35(7):1862–1870, July 1996.