Image Upsampling via Imposed Edge Statistics

Raanan Fattal*
University of California, Berkeley

Abstract

In this paper we propose a new method for upsampling images which is capable of generating sharp edges with reduced input-resolution grid-related artifacts. The method is based on a statistical edge dependency relating certain edge features of two different resolutions, which is generically exhibited by real-world images. While other solutions assume some form of smoothness, we rely on this distinctive edge dependency as our prior knowledge in order to increase image resolution. In addition to this relation we require that intensities are conserved; the output image must be identical to the input image when downsampled to the original resolution. Altogether the method consists of solving a constrained optimization problem, attempting to impose the correct edge relation and conserve local intensities with respect to the low-resolution input image. Results demonstrate the visual importance of having such edge features properly matched, and the method's capability to produce images in which sharp edges are successfully reconstructed.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithms; I.4.3 [Image Processing and Computer Vision]: Enhancement—Sharpening and deblurring

Keywords: image interpolation, image enhancement, Markov random field image modeling, super-resolution

1 Introduction

Image resizing or resampling is one of the most elementary image operation, supported by virtually all image editing software, and is used for many purposes. In the course of desktop publishing, raw images are resized, on a regular basis, to new dimensions in order to fit designated areas in documents. Low-resolution video frames from surveillance cameras are enlarged to ease the inspection of their contents. As well as the recent popularity of HDTVs brings out the need for resolution enhancement of NTSC and PAL video recordings. In 3D graphics, these interpolations are used to map image textures onto objects' surface. While satisfactory downsampled images are obtained by a proper linear pre-filtering, this is not the case for upsampling. Upsampled images usually lack small-scale texture-related features and moreover, sharp edges become blurry, original pixel grids remain noticeable (often called the '*jaggies*' artifact), and in some cases ringing appears in the vicinity of sudden transitions in intensity. Formally speaking, upsampling involves determining far more pixel intensities than the number given. This makes upsampling a particulary challenging problem and one that is highly sensitive to the additional assumptions or information needed to establish its well-posedness. Indeed, different upsampling techniques correspond to different assumptions about the nature of the upsampled image. For example, the assumption that images are smooth enough to be adequately approximated by polynomials yields analytic polynomial-interpolation formulas. On the other hand, assuming that images are limited in band yields a different family of low-pass filters. For most images, these assumptions are highly inaccurate, and as a consequence these methods produce images which suffer from excessive blurriness and the other visual artifacts mentioned earlier. We extend this discussion to more sophisticated methods in the next section.

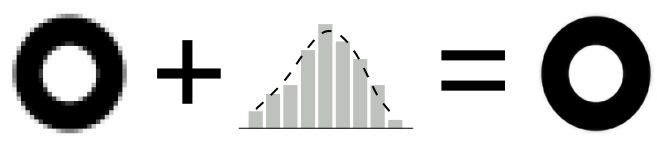

Figure 1: Sharp upsampled image resulting from a low-resolution image plus edge statistics (the right ring is an actual result from the input on the left).

In this paper we point out a unique dependency between image derivatives at different resolutions, as exhibited by real-world images; pixel differences at higher resolutions depend on their distance from an edge, the spatial distribution of that edge and the total intensity jump across it, all estimated in low-resolution. Using this non-trivial relation, we have devised a new method for upsampling images. The solution consists in promoting the predicted intensity differences in the upsampled image given the edge parameters observed at the low-resolution input. This is done while deducing absolute intensities from an 'intensity conservation' constraint that requires the total intensity in the low and high resolutions to be the same. This approach is summarized schematically in Figure 1; given a low-resolution image plus this additional parametric statistical information, sharp edges are retrieved while typical artifacts associated with upsampling are minimal.

By real-world images, we refer to scenes seen with the naked eye or more precisely, scenes captured by a photographic device (e.g. camera). This includes indoor and outdoor photos not enlarged by any digital means. Also, unless stated otherwise, we discuss only gray scale images, working with the Y channel of YUV color space. We later extend the method to handle color images as well.

This paper is organized as follows. In the next section we give a brief survey of existing work on this topic. In Section 3, we describe the statistical edge dependency, and in the following section we use it to construct the new method. In Section 5, we discuss the results obtained and compare them with results produced by other methods. Finally, in Section 6, we describe the drawbacks of this method as revealed by testing and summarize our conclusions.

2 Previous Work

The problem of image upsampling has received much attention both from the computer graphics and image processing communities. As a result, many techniques have been suggested in recent years, each with its own distinctive methodology, prior assumptions, and requirements for additional information. In this section, we discuss these related techniques by dividing them into groups of common features. We begin with the classical approach of data invariant linear filters, which are very popular in commercial software. Among these filters are the well known Nearest-Neighbor, Bilinear, Bicubic, Hann, Hamming, and Lanczos interpolation kernels. The con-

*e-mail: raananf@math.berkeley.edu

ACM Reference Format
Fattal, R. 2007. Image Upsampling via Imposed Edge Statistics. *ACM Trans. Graph. 26*, 3, Article 95 (July 2007), 8 pages. DOI = 10.1145/1239451.1239546 http://doi.acm.org/10.1145/1239451.1239546.

struction of these kernels relies strongly on the assumption that the image data is either spatially smooth or band-limited. As mentioned earlier, this is not true in general, and consequently these solutions tend to produce visual artifacts such as ringing, aliasing, blocking, and blurring. In [Thvenaz et al. 2000], a more detailed survey of these techniques and their shortcomings is given.

In recent years, more sophisticated methods, where interpolation weights adapt locally to the image data, have been developed in order to reduce blurriness and other artifacts. In [Carrato et al. 1996], the parameterization of a linear interpolation is changed according to local intensity differences allowing for the generation of faster intensity transitions. Despite this method's ability to produce sharper edges than linear filtering, the 1D treatment emphasizes vertical and horizontal edges in nonhomogeneous regions as shown in Figure 6. Also in [Su and Willis 2004], interpolation weights are adjusted locally by choosing three out of the four nearest pixels to reduce the number of variables that are averaged. This choice forms a noticeable block-like effect, showing strong continuity along one of the two diagonals. This artifact is avoided in [Li and Orchard 2001], where arbitrary edge orientation is implicitly matched by estimating local intensity covariance from the low-resolution image. As shown in Figure 6, this method is capable of generating smooth curves and of reducing jaggies. Yet, since the interpolation weights are estimated from flat regions as well, edges do not appear as sharp as they should. Also in [Ratakonda and Ahuja 1998], selective interpolation is implemented using an iterative *Projection Onto Convex Set* (POCS) scheme. Additional theoretical aspects of the POCS method concerning upsampling can be found in [Calle and Montanvert 1998].

By storing additional image data, in the form of discontinuity graphs, averaging pixels across boundaries is avoided in [Tumblin and Choudhury 2004]. This approach allows images to be enlarged while keeping boundaries infinitely sharp. Results demonstrate the effectiveness of this method on images of a strict piece-wise planar nature, *i.e.*, linear profiles separated by intensity jumps. A solution of a similar nature is reported in [Ramanarayanan et al. 2004], focusing on generating high quality texture maps.

Another way to improve the quality of upsampled images by reducing the jaggies is described in [Morse and Schwartzwald 2001]. There, after employing some given interpolation technique, the curvature of *isophotes* (curves of constant intensity) is minimized by an iterative process. In another distinct approach, reported in [Greenspan et al. 2000], the next higher harmonic of a pyramidal image representation is extrapolated such that ideal edge transitions are preserved. The method demonstrates an ability to sharpen images but suffers from some haloing artifacts due to extrapolation overshooting.

Resolution enhancement is tackled also under different settings, whereby several images, differing by some spatial transformation, are combined to create a higher resolution image. In [Capel and Zisserman 1998], an automated registration process followed by resolution enhancement is described.

Resolution enhancement was also suggested as an application of the *Image Analogies* paradigm [Hertzmann et al. 2001], and was later modified in [Freeman et al. 2002], to allow broader applicability. In the latter, an analytically interpolated image is enhanced by adding high-frequency patches from a non-parametric set of examples relating low and high resolutions. This approach sharpens edges and yields images with a detailed appearance. The use, however, of such a finite non-parametric set of examples tends to introduce some irregularities into the constructed image. A similar solution is derived from a degradation model in [Tappen et al. 2004].

The Total Variation method is used to invert a blurring process in images [Osher et al. 2003]. In this approach, the degradation is expressed by a functional which also measures the L_1 norm of the output image in order to promote regularity. This approach has been formulated for image upsampling in [Aly and Dubois 2005].

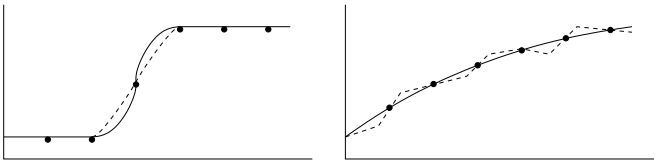

Figure 2: Possible edge profile reconstructions.

3 Edge-Frame Continuity Moduli

A sharp edge in an image corresponds to relatively large intensity gradients concentrated along the edge, while a smooth edge is composed of a more scattered set of weaker gradients. Here we upsample an image by constructing its gradient field rather than determining the pixel intensities directly. This allows us to focus on producing a properly condensed gradient field, where sharp edges can be constructed while avoiding ringing effects. To follow this approach we are required to predict the spatial intensity differences at the high-resolution based on the low-resolution input image.

A smooth intensity transition in the input image may or may not correspond to a similarly sloped and spatially distributed transition at the higher resolution. As illustrated in Figure 2, this relation is not unique and is too complicated to be quantified manually. In view of that, we resort to statistical measurements and modeling in an attempt to capture edge dependencies at different resolutions. In particular, given some local edge-related parameters extracted from the low-resolution image, such as closeness to an edge and its magnitude and scattering, we estimate the expected local intensity continuity observed at the high resolution. This relation, if found to be non-trivial with respect to these quantities, will provide a meaningful criterion for a 'correct' edge construction in the new upsampled image.

In the reminder of this section, we describe both the features which we base our predictions on, and the quantities we aim to predict at the higher resolution. Given an image I, we simulate its downsampling by means of filtering with a kernel k_d, and evenly decimating it. In order to perform the analysis at I's resolution, we upsample the decimated image by filtering with another kernel k_u to get $\tilde{I}$. Before we describe the features we extract from $\tilde{I}$, we define a few useful functions. The spatial derivatives and the gradient norm are evaluated at every pixel $\mathbf{x} = (x, y)$ as following

$$\tilde{I}_x(\mathbf{x}) = \tilde{I}(x+1, y) - \tilde{I}(x-1, y),$$
$$\tilde{I}_y(\mathbf{x}) = \tilde{I}(x, y+1) - \tilde{I}(x, y-1),$$
$$\text{and } N(\mathbf{x}) = \|\nabla\tilde{I}\|_2, \text{ where } \nabla\tilde{I} = (\tilde{I}_x(\mathbf{x}), \tilde{I}_y(\mathbf{x})).$$

The ray in the gradient direction is parameterized at each pixel by

$$\varphi_{\mathbf{x}}(t) = \mathbf{x} + t \cdot \nabla\tilde{I}(\mathbf{x})/N(\mathbf{x}).$$

The features we extract at each pixel $\mathbf{x}$ describe the shape of the nearest edges passing by $\mathbf{x}$. These quantities are computed from the gradient norm N along $\varphi_{\mathbf{x}}(t)$ when $t \in [t_1, t_2]$. This interval delimits the ridge in N which is closest to $\mathbf{x}(t=0)$ as shown in Figure 3, and is determined as follows. First, we find the point of local maxima t^* of N along $\varphi_{\mathbf{x}}$ which is closest to $t = 0$. The points t_1 and t_2 are then taken to be the points where $\varphi_{\mathbf{x}}$ ceases to decrease as we move away from t^*. As will be discussed later, the edge's statistical dependency is measured only between two resolutions differing by a factor of 2. Consequently, this search can be performed only within the limits of $|t| \leq 5$ (pixels).

Finally we are ready to define the edge features that we extract from the low-resolution image,

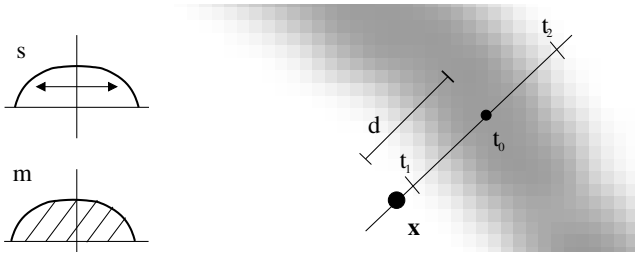

Figure 3: Illustration of d,m,s on top of $N(\mathbf{x})$, showing the segment passing through $\mathbf{x}$ along $\nabla\tilde{I}(\mathbf{x})$ passing through the edge's peak at t_0. On the left, a cross-section of the edge along this segment shows the distribution around its center and the total intensity change across it.

1. The total change in intensity across the nearest edge $m(\mathbf{x})$ is estimated from N along $\varphi_{\mathbf{x}}$ as follows

$$m(\mathbf{x}) = \int_{t_1}^{t_2} N\big(\varphi_{\mathbf{x}}(t)\big)dt.$$

2. The distance from the closest edge $d(\mathbf{x})$ is determined by the distance along $\varphi_{\mathbf{x}}(t)$ between $t=0$ and the center of mass t_0 of the closest ridge in N to $\mathbf{x}$, given by

$$t_0 = \int_{t_1}^{t_2} t\cdot N\big(\varphi_{\mathbf{x}}(t)\big)dt/m(\mathbf{x}),$$

and $d(\mathbf{x}) = |t_0|$. We also store $r(\mathbf{x}) = \text{sign}(t_0)$ which indicates on which side of $\mathbf{x}$ (along $\varphi_{\mathbf{x}}$) the edge passes.

3. The closest edge's spatial scattering $s(\mathbf{x})$ is measured by

$$s(\mathbf{x}) = \int_{t_1}^{t_2} (t-t_0)^2\cdot N\big(\varphi_{\mathbf{x}}(t)\big)dt/m(\mathbf{x}).$$

Note that these features correspond to the first three moments of the edge's cross-section as illustrated in Figure 3. We use discrete sums to approximate the integrals above.

Now, for each possible configuration of d,m,s we estimate the conditional expectation of the pixel differences, the local continuity moduli, at the high-resolution image I. We do this in a *local frame* defined by the orientation of the edge $\nabla\tilde{I}/N$. This allows us to gather dependencies which are more intrinsic to the edge (and will be shown to be significant later). This is computed using the following local transformation matrix

$$F\big(\nabla\tilde{I}(\mathbf{x})\big) = \|\nabla\tilde{I}(\mathbf{x})\|^{-1}\begin{pmatrix} r(\mathbf{x})\cdot\tilde{I}_x(\mathbf{x}) & -\tilde{I}_y(\mathbf{x}) \\ r(\mathbf{x})\cdot\tilde{I}_y(\mathbf{x}) & \tilde{I}_x(\mathbf{x})\end{pmatrix}.$$

Note that r serves to mirror the frame so that the edge will always be on the same side in the transformed coordinates, adding this additional sensitivity to the statistics. Finally we are ready to define the continuity measure between pixel $\mathbf{x}$ and its neighbor, $\mathbf{x}+\mathbf{u}$, by

$$\mathscr{C}(\mathbf{x},\mathbf{u}) = I(\mathbf{x}) - I(\mathbf{x}+F\big(\nabla\tilde{I}(\mathbf{x})\big)\cdot\mathbf{u}),$$

where $\mathbf{u}=(u,v)$ is an integer vector parameterizing a small window around $\mathbf{0}$. Under the assumption of invariance to translation and rotation, implying that there is no dependency on particular pixel position or edge orientation, we define the *edge-frame continuity modulus* (EFCM) as the conditional mean and variance of $\mathscr{C}$

$$\mu(\mathbf{u}\,|\,d,m,s) = E\big[\mathscr{C}(\mathbf{x},\mathbf{u})\;|\;d,m,s\big],$$

$$\sigma^2(\mathbf{u}\,|\,d,m,s) = E\big[\big(\mathscr{C}(\mathbf{x},\mathbf{u})-\mu(\mathbf{u}\,|\,d,m,s)\big)^2\;|\;d,m,s\big].$$

Figure 4 shows $\sigma^2(\mathbf{u}\,|\,d,m,s)$ estimated from fifteen indoor images of one mega pixels each. From these measurements a clear dependency is revealed between d,m,s and the continuity moduli exhibited by the image intensities. The continuity profiles are consistent

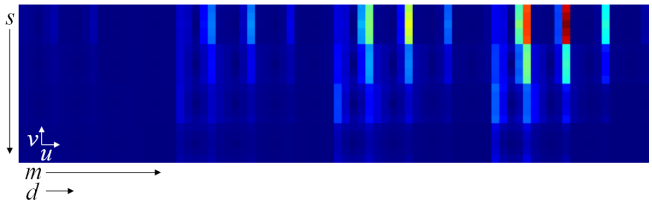

Figure 4: Displays a 5D graph spread in 2D: 5-by-5 windows of $\sigma^2(\mathbf{u}\,|\,d,m,s)$ as a function of $\mathbf{u}=(u,v)$ are shown against d,m,s. The edge scattering s increases along the vertical axis, while the distance from the edge d increases along the short horizontal axis. At every fourth window along this axis, the total intensity change across the edge m increases.

with the fact that edges pass nearby, as expected due to the image derivatives' spatial dependency, discussed in [Huang and Mumford 1999]. Moreover, as the intensity diminishes and scattering grows, the eccentricity of this profile is reduced. These measures merely quantify the anticipated local relations, allowing us to make use of them later on. The non-isotropy and non-symmetry of these graphs confirms the necessity of analyzing this relation in the local edge-frame. As stated earlier, this choice of the features d,m,s characterizes the profile of the edge cross-section. Given these features, statistics are gathered and represent a true image behavior. Therefore, the choice of these features must be judged by the their effect over the statistics and indeed Figure 4 confirms that the expected quantities are influenced greatly by conditioning with respect to these features.

Much work has been devoted in [Schaaf 1998], [Huang and Mumford 1999] and [Reinhard et al. 2004], to estimating different statistical properties of images, focusing mainly on the particular family of natural images. In these measurements, much efforts have been invested in calibrating the acquisition process so that camera-specific biases, such as nonlinear intensity transfer and particular spatial response profiles, are eliminated. Although we introduce here image statistical measurements, the aim of this work does not require them to be purely evaluated. The upsampled image need not possess any scene-intrinsic behavior; on the contrary, users expect this kind of bias to remain when enlarging an image. Also, the upsampling method we're about to describe is not restricted to any particular table of EFCM. Thus, such measurements can be performed in any suitable way.

The assumptions of spatial and rotational invariance allow us to obtain reasonable estimates of μ and σ^2 from small collections of images. Certain groups of images may show systematic statistical dependencies on different regions (such as land and sky), or specific edge orientations. Accounting for such dependencies does not introduce any principal complication to the algorithm described next, but would naturally increase the number of images needed for achieving proper statistical estimates.

4 Upsampling using the EFCM

In this section we derive a scheme for upsampling images which relies on the prior knowledge stored in the EFCM tables. Given a low-resolution image L, we compute its upsampled version $\tilde{I}$ using filter k_u (Bicubic interpolation in our implementation) in accord with the discussion of the previous section. From this temporary image, we extract the edge-related features r,d,m and s. In order to impose specific EFCM characteristics on I, we have to model the statistical behavior of the random variable $\mathscr{C}$. The samples' histogram of $\mathscr{C}$ taken from our indoor set of images bears a strong resemblance to the Normal Distribution once conditioned (on d,m,s), as shown in Figure 5. This, together with practical computational considerations, given the number of unknowns in I, leads us to model this

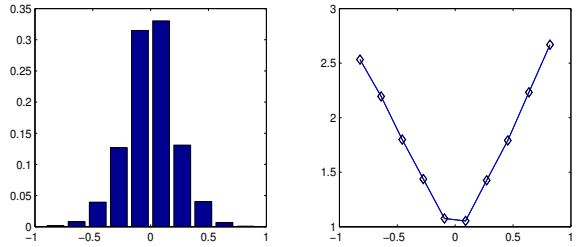

Figure 5: The left graph shows $\mathscr{C}(\mathbf{x},\mathbf{u})$ samples' histogram, H, corresponding to a particular choice of $\mathbf{u}$ and d,m,s. $\sqrt{-\log(H)}$ is plotted on the right.

distribution by the Normal distribution, characterized by μ and σ^2,

$$\mathscr{C}(\mathbf{x},\mathbf{u}) \sim \mathscr{N}\big(\mu(\mathbf{u}\,|\,d,m,s),\sigma^2(\mathbf{u}\,|\,d,m,s)\big),$$

for a given offset $\mathbf{u}$ and parameters d,m and s. This modeling of the EFCM gives rise to a *Gauss-Markov Random Field* model, reviewed in [Prez 1998], over the upsampled image space. In these terms, assuming a continuous state space for the pixel intensities and taking images as configurations, we get the following *Gibbs* distribution

$$P(I) = \frac{1}{Z}\prod_{\mathbf{x}} \exp\big(-V_{\mathbf{x}}(I)\big),$$

with the interaction potentials,

$$V_{\mathbf{x}}(I) = \sum_{\mathbf{u}} \frac{\big(\mathscr{C}(\mathbf{x},\mathbf{u}) - \mu(\mathbf{u}\,|\,d(\mathbf{x}),m(\mathbf{x}),s(\mathbf{x}))\big)^2}{\sigma^2(\mathbf{u}\,|\,d(\mathbf{x}),m(\mathbf{x}),s(\mathbf{x}))},$$

and Z is the appropriate partition function. The Markov property is evident from the limited number of pixels coupling in each of the interaction potentials. Small windows of 3-by-3 pixels involve 8 interactions similarly to a spatially varying Laplacian matrix. The resulting distribution $P(I)$ predicts the likelihood of I in terms of predicted edges behavior. Another piece of information we can use in order to reduce the space of permissable Is are the input image intensity values. This can be achieved rather naturally by introducing an intensity conservation law stating that the result by downsampling I to the input image resolution must be identical to the input image L itself. Formally, this constraint can be expressed by the following linear equation, $\mathscr{D}\cdot I = L$. Assuming that I is comprised of n pixels and L of m pixels, then $\mathscr{D}$ is a n-by-m decimation matrix whose rows are simply the downsampling kernels k_d (we used equal averages in our implementation). Thus, the upsampled image I is obtained as the solution of

$$\max_I \log\big(P(I)\big) \quad \text{s.t.} \quad \mathscr{D}\cdot I = L.$$

A similar condition was formulated in [Tappen et al. 2004] as a compatibility function which is later minimized along with another spatially-invariant regularity condition. In our formulation, this requirement is an analytical constraint, which narrows the search space. This form of constrained optimization problem can be solved using the method of Lagrange Multipliers which requires the solution of the following set of equations,

$$\frac{d\log(P)}{dI} = \mathscr{D}^\top\lambda \quad \text{s.t.} \quad \mathscr{D}\cdot I = L.$$

This is an indefinite system of linear equations known as a saddle point problem [Benzi et al. 2005]. The solver used in our implementation is equivalent to a basic *Conjugate Gradient*-based *Null Space* method as described in [Press et al. 1993] and [Benzi et al. 2005]. Also, this system is poorly conditioned due to weak interactions across edges and strong ones in flat regions. Nevertheless, simple division by the diagonal element proved to be very a effective preconditioning.

Figure 6: Top to bottom, left to right are: reference image, its downsampling by 4 (test's input pixels). Second row, Bicubic interp. and the image produced by our method. Bottom row shows the Simple Edge Sensitive and New Edge-Directed intep. results.

Color Images. The method described so far deals with gray scale images only. We apply this algorithm for color images using the assumption that color and luminance are locally related by some simple transformation. Affine transformations have been used in [Zomet and Peleg 2002] and [Levin et al. 2004]. We take a similar approach. First we upsample the luminance channel of the YUV color space. Next, for each upsampled pixel in the resulting image we compute the absolute value of its luminance difference from its four closest pixels in the low-resolution image, denoted by d_1,d_2,d_3 and d_4. Normalized weights are then computed by $w_i = d_i^{-\alpha}/w$, where $i = 1...4$, $w = \sum d_i^{-\alpha}$, and α is a positive number controlling the penalty for different in luminance. The color channels of the four neighboring pixels are linearly combined using w_i, to give the color channels of the upsampled pixel. This performs sufficiently well for images which do not contain strong changes in color unaccompanied by a change in luminance.

5 Results

In order to magnify images by a factor greater than two (along each axis), we perform the following procedure. We begin by using the input image as the low-resolution reference, from which we extract edge features. Later, for further magnification, we use the output image as the low-resolution input. All along, the con-

Figure 7: Color images magnified by a factor of 4.

strained subspace is defined by the *original* low-resolution input image. In Figure 6, by way of comparison, the same image is magnified by a factor of 4 using Bicubic interpolation, our method, the Simple Edge Sensitive [Carrato et al. 1996], and the New Edge-Directed [Li and Orchard 2001] interpolations. The main edges produced by our method are sharper and closer to the original (reference) image than the ones generated by the other methods. In Figure 8 we compare our method with the Example-Based Super-Resolution method [Freeman et al. 2002]. This comparison shows the new method's ability to generate sharper edges while producing less noise and false edge irregularities than the Example-Based Super-Resolution method. On the other hand, the images produced by the latter method appear to be somewhat more detailed. A preference between the two techniques may depend on the specific image and subjective concerns.

In Figure 7, we show color images magnified by a factor of 4, and the original input pixels. The results show that sharp edges are reconstructed and that color is successfully matched. At twice that magnification factor, as shown in Figure 9, resulting images continue to show sharpness. It appears that edges separated by one pixel show some jaggies. The chip connectors' shaded side, captured roughly by one pixel in the input, show some staircasing, while the image of the old man's face demonstrates the method's ability to construct edges of different scales. Cartoon images make good candidates for such operations because of their inherent lack of texture. The dog's outlines remain sharp, and the new image appears almost as it does at its native resolution. Magnifying higher than this factor produces less realistic images due to the absence of texture. Nevertheless, in Figure 10, we've magnified images by a factor of 16, showing that the method manages to produce sharp edges even at this high magnification factor. In Table 1 we give objective error measurements between an upsampled image and the original ground-truth image (*i.e.*, before downsampling). These differences are measured by a root mean square error (RMSE) after the images are normalized to have a zero mean and a unit variance and also by the Structural Similarity Image Quality (SSIQ) described in [Wang et al. 2004].

Figure 8: Top to bottom, left to right are: child's face upsampled by a factor of 4 using the Example-Based Super-Resolution, and using our method. Bottom row shows pixels of a teapot, their magnification by a factor of 8 using the Example-Based Super-Resolution, and the result of our method. Both the teapot original image and the images generated by the Examples-Based Super-Resolution are courtesy of William T. Freeman.

The running times of this method on a Mobile Pentium-M, running at 2.1MHz and implemented in C++ are as follows: 2 seconds to upsample an image of 128^2 pixels to twice its resolution (256^2). Upsampling four times this resolution takes 6 seconds upsampling to a resolution of 1024^2 pixels takes our implementation 22 seconds.

Figure 9: Images magnified by a factor of 8.

6 Conclusions

In this paper we presented a novel method for upsampling images. Instead of relying on inaccurate assumptions which exclude the possibly of abrupt transitions, this method is based on actual edge characteristics. The method demonstrates its ability to retrieve sharp edges even when they are poorly sampled in the low-resolution input. Artifacts, which are typical for such operations, are quite minimal in relative to the gain in clarity and sharpness of the resulting images.

The drawbacks of this method, as drawn from its testings, are the following: (i) the resulting sharp-edged images tend to emphasize lack of texture and absence of fine-details. (ii) the jaggies artifact, though minimal relative to the gain in sharpness, is sometimes noticeable in the presence of clear edges produced. (iii) acutely twisted edges, although scarce, are not captured by the EFCM and are undetected by conditioning on d, m, s alone. And (iv) the proposed method involves more computations than some of the existing techniques.

Restricting to a particular type of scenes, when collecting statistics, may produces more accurate results on images belonging to that class. Yet, as shown in Figure 12, the difference amounts to a rather subtle change in overall sharpness. The estimated tables of μ and σ^2 can thus be used with a large degree of universality.

The generic behavior of edges, as captured by the statistics, does not accurately describe every particular case. An important feature that we hold responsible for the successful edge reconstruction is the coupling *along* edges. The continuity imposed along the edges translates to strong interactions that allow neighboring edge profile 'evidence' to be propagated and combined along the edge itself. Indeed, this mechanism is detected in its absence; occasionally, in the case of long edges perfectly aligned with the grid, found in the the ring shown in Figure 12, the 2D couplings become trivial in the sense that the same feature values are found along the edge. In such cases 'generic' edges are constructed, which may lack the appropriate sharpness.

Figure 10: Child's eye and an armchair images magnified by a factor of 16.

Upsampling methods that are based on a non-parametric model, such as [Freeman et al. 2002], may suffer from a shortage of structures. Edge curves often follow a continuous set of orientations and do so at various curvatures, some of which may be absent from the set of examples. Thus the reconstructed contour may become falsely discontinuous or the edge itself may remain blurry. On the other hand, example-based techniques have shown more success in synthesizing textures than parametric approaches, see [Kwatra et al. 2005]. This suggests that a marriage between these methods and the current one may be a promising direction for further research.

JPEG and MPEG standards also follow certain band assumptions favoring lower frequencies, and hence exhibit artifacts similar to their linear interpolation counterparts (blurriness, ringing). As pointed out in [2002], this analogy suggests that achieving good results when upsampling images may offer an alternative approach to image compression. Similarly to [Simoncelli 1997], where statistical dependency between wavelet coefficients is exploited for compact image coding, more explicit edge dependencies can be used for the same purpose. For example, one can store a downsampled image, and later decompress it by upsampling. For that matter, JPEG images or MPEG frames can be taken as the 'low-resolution' images once edge relations similar to the ones used here are established. This idea can be further extended to other types of image restoration scenarios. Finally, higher-order edge properties such as curvature, if properly extracted, may improve the results in the case of highly curved edges.

The images appearing in this paper can be downloaded from:
http://www.cs.huji.ac.il/~raananf/projects/upsampling/results.html

7 Acknowledgements

We'd like to thank Dani Lischinski, Elchanan Mossel and Yossi Farjun for their helpful comments and also thank Tamar Lando for editing this text. This work was supported by the Miller Institute at the University of California, Berkeley.

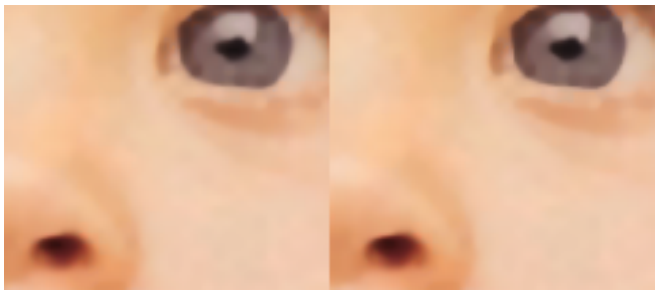

Figure 11: A comparison between two classes of images used for edge statistics estimation. Left is based on outdoor scenes and the right on indoor scenes.

Image	norm. RMSE		SSIQ	
	Our	Bicubic	Our	Bicubic
foliage1	0.226	0.278	0.946	0.924
foliage2	0.182	0.241	0.893	0.848
foliage3	0.187	0.187	0.908	0.907
foliage4	0.143	0.145	0.950	0.938

Image	norm. RMSE			SSIQ		
	Our	EBSP	GF	Our	EBSP	GF
child	0.129	0.145	0.154	0.903	0.871	0.881

Table 1: EBSR stands for Example-Based Super-Resolution method and GF for the Genuine Fractals software.

Figure 12: Figure shows a blowup of the ring from Figure 1. The horizontal edge, pointed by the red arrow, appears to be less sharp than the rest of the circle's edges.

References

ALY, H., AND DUBOIS, E. 2005. Image up-sampling using total-variation regularization with a new observation model. In *IEEE Transactions on Image Processing*, vol. 14, 1647–1659.

BENZI, M., GOLUB, G. H., AND LIESEN, J. 2005. Numerical solution of saddle point problems. In *Acta Numer.*, vol. 14, 1–137.

CALLE, D., AND MONTANVERT, A. 1998. Super-resolution inducing of an image. In *International Conference on Image Processing*, vol. 3, 232–236.

CAPEL, D., AND ZISSERMAN, A. 1998. Automatic mosaicing with super-resolution zoom. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Washington, DC, USA, 885.

CARRATO, S., RAMPONI, G., AND MARSI, S. 1996. A simple edge-sensitive image interpolation filter. In *Proceedings of the IEEE International Conference on Image Processing, 1996*, vol. 3, 711–714.

FREEMAN, W. T., JONES, T. R., AND PASZTOR, E. C. 2002. Example-based super-resolution. In *IEEE Computer Graphics and Applications*, vol. 22, 56–65.

GREENSPAN, H., ANDERSON, C. H., AND AKBER, S. 2000. Image enhancement by non-linear extrapolation in frequency space. In *IEEE Transactions on Image Processing*, vol. 9, 1035–1048–202.

HERTZMANN, A., JACOBS, C. E., OLIVER, N., CURLESS, B., AND SALESIN, D. H. 2001. Image analogies. In *ACM Transactions on Graphics*, ACM Press, New York, NY, USA, 327–340.

HUANG, J., AND MUMFORD, D. 1999. Statistics of natural images and models. In *Computer Vision and Pattern Recognition*, 1541–1547.

KWATRA, V., ESSA, I., BOBICK, A., AND KWATRA, N. 2005. Texture optimization for example-based synthesis. In *ACM Transactions on Graphphics*, ACM Press, New York, NY, USA, vol. 24, 795–802.

LEVIN, A., LISCHINSKI, D., AND WEISS, Y. 2004. Colorization using optimization. In *ACM Transactions on Graphics*, ACM Press, New York, NY, USA, vol. 23, 689–694.

LI, X., AND ORCHARD, M. T. 2001. New edge-directed interpolation. In *IEEE Transactions on Image Processing*, vol. 10, 1521–1527.

MORSE, B. S., AND SCHWARTZWALD, D. 2001. Image magnification using level-set reconstruction. In *Computer Vision and Pattern Recognition*, 333–340.

OSHER, S., SOLE, A., AND VESE, L. 2003. Image decomposition and restoration using total variation minimization and the h^{-1}. *Multiscale Modeling & Simulation 1*, 3, 349–370.

PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., AND FLANNERY, B. P. 1993. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edition ed. Cambridge University Press, Cambridge, UK.

PREZ, P. 1998. Markov random fields and images. In *CWI Quarterly*, vol. 11, 413–437.

RAMANARAYANAN, G., BALA, K., AND WALTER, B. 2004. Feature-based textures. In *Rendering Techniques*, 265–274.

RATAKONDA, K., AND AHUJA, N. 1998. Pocs based adaptive image magnification. In *ICIP (3)*, 203–207.

REINHARD, E., SHIRLEY, P., ASHIKHMIN, M., AND TROSCIANKO, T. 2004. Second order image statistics in computer graphics. In *APGV '04: Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, ACM Press, New York, NY, USA, 99–106.

SCHAAF, A. V. D. 1998. *Natural Image Statistics and Visual Processing*. PhD thesis, Rijksuniversiteit Groningen, The Netherlands.

SIMONCELLI, E. P. 1997. Statistical models for images: Compression, restoration and synthesis. In *31st Asilomar Conf on Signals, Systems and Computers*, IEEE Computer Society, Pacific Grove, CA, 673–678.

SU, D., AND WILLIS, P. 2004. Image interpolation by pixel-level data-dependent triangulation. In *Computer Graphics Forum*, vol. 23, 189–202.

TAPPEN, M. F., RUSSELL, B. C., AND FREEMAN, W. T. 2004. Efficient graphical models for processing images. In *Computer Vision and Pattern Recognition*, 673–680.

THVENAZ, P., BLU, T., AND UNSER, M. 2000. Image interpolation and resampling. In *Handbook of Medical Imaging, Processing and Analysis*, Academic Press, San Diego CA, USA, I. Bankman, Ed., 393–420.

TUMBLIN, J., AND CHOUDHURY, P. 2004. Bixels: Picture samples with sharp embedded boundaries. In *Rendering Techniques*, 255–264.

WANG, Z., BOVIK, A., SHEIKH, H., AND SIMONCELLI, E. 2004. Image quality assessment: From error measurement to structural similarity. *IEEE Trans. Image Processing 13*, 4, 600–612.

ZOMET, A., AND PELEG, S. 2002. Multi-sensor super-resolution. In *6th IEEE Workshop on Applications of Computer Vision*, 27–31.

Figure 13: A comparison between our method, Bicubic interpolation and assorted commercial softwares: Genuine Fractals (www.ononesoftware.com), FocalBlade (www.thepluginsite.com), PhotoKit (www.pixelgenius.com) and PhotoZoom (www.benvista.com).